UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYSE DES ÉMOTIONS DANS LE CONTEXTE DES RÉSEAUX SOCIAUX ET DU CHANGEMENT CLIMATIQUE

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN INFORMATIQUE COGNITIVE

PAR

DENNIS DROWN

MAI 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYZING EMOTION IN THE CONTEXT OF ONLINE SOCIAL MEDIA AND CLIMATE CHANGE

DISSERTATION

PRESENTED

AS PARTIAL FULFILLMENT

OF THE DOCTORATE IN

COGNITIVE INFORMATICS

BY

DENNIS DROWN

MAY 2023

REMERCIEMENTS

Il me semble curieux que la première section présentée dans cette thèse doctorale soit en réalité la dernière que j'écris. Il n'y a pas d'autre moyen de procéder. Autrement, comment pourrais-je être pleinement conscient du fait que je n'aurais jamais pu arriver jusque-là sans la quantité inépuisable d'aide que j'ai reçue de mes directeurs, mes professeurs, mes camarades, ma famille et mes ami·e·s ? Cet appui a pris de nombreuses formes : discuter de la théorie, contester des idées, critiquer des présentations, relire des documents, préparer de la nourriture savoureuse et prononcer des mots d'encouragement pour n'en citer que quelques-unes. Je suis incapable d'exprimer à quel point je suis reconnaissant à ceux et celles qui m'ont aidé dans cette voie.

It seems just a bit curious that the first section presented in this doctoral thesis is actually the last section that I write. It can be no other way. How else could I be so deeply conscious of the fact that I simply could never have gotten to this point without the truly endless amount of aid I received from my advisors, professors, fellow students, family and friends? This help took many forms, among these: discussing theory, challenging ideas, critiquing presentations, proofreading documents, preparing savoury sustenance, and speaking words of encouragement. I am unable to express how grateful I am to all those who have helped me on this path.

Roger Villemaire a été mon premier point de contact en ce qui concerne le programme de doctorat en informatique cognitive. J'ai parlé avec lui avant même de m'inscrire, car j'avais plusieurs questions sur la matière traitée et plus largement pour discuter de mes inquiétudes quant à l'idée de suivre un programme de 3^{e} cycle dans une université francophone. Ensuite, comme professeur de mon premier cours des concepts fondamentaux et mon professeur du cours de logique, j'ai commencé à voir Roger comme un pont stable pour moi. Il m'a aidé à faire une transition d'un monde d'informatique dans l'industrie, où je savais bien me servir de mes connaissances pour élaborer quelque chose d'utile, à l'informatique comme sujet de recherche, où le but est d'aller au-delà de ce que l'on sait afin de créer de nouvelles connaissances. Roger a participé pleinement dans la présente recherche. Il est toujours prêt à échanger des idées, à travailler pour être sûr que l'on n'a pas négligé quelque chose d'important et à discuter des possibilités liées à la recherche. J'en suis reconnaissant et ce serait difficile pour moi d'imaginer un meilleur directeur de recherche.

Serge Robert, lui aussi, a été mon professeur du cours des concepts fondamentaux et du cours de logique. Mes deux directeurs forment une excellente équipe. Or, si Roger a été un pont solide pour moi, Serge a été une falaise, mais d'une très bonne manière. Serge m'a aidé à créer un nouveau lien entre l'intelligence artificielle, comme je la connaissais dans le domaine de l'informatique, et l'IA dans un contexte qui comprend la cognition biologique en général et celle de l'être humain en particulier. Maintes fois, Serge a amené un thème que je comprenais bien (du point de vue d'un informaticien) et il l'a présenté d'une manière qui tournait à l'envers tout ce que je savais sur le sujet. Il me fait penser. Il me fait réexaminer continuellement ma perspective. Je remercie Serge du travail et de l'inspiration qu'il a contribué à ce projet de recherche.

Il n'est pas exagéré de dire que sans Mylène Dagenais, on n'aurait pas de programme de doctorat en informatique cognitive comme nous le connaissons. Mylène a été non seulement extrêmement utile pendant tout le programme, mais son aide venait toujours avec un beau sourire et un mot encourageant. Merci beaucoup, Mylène.

Estas por mi iom malfacile kompreni kiel Carlos González Avila sen-plende, dum la kvin jaroj kaj duono de tiu ĉi programo, aranĝis nian tutan vivon por ke mi estu libera koncentriĝi nur pri la doktoriĝo. Tasoj da kafo aperis; manĝoj aperis, kvazaŭ magie. Neniam mi estis bezonata por kutimaj hejmaj taskoj. Sociaj eventoj estis planitaj ĉiam en akordo kun mia universitata horaro. Li diris ade, ke mi jam havas multon por fari. Ankaŭ li faris multegon, kaj pro tio mi dankas al li plenkore.

One would never know that Norma E. Larsen does not consider herself an expert in computers as she dove in and tirelessly offered her time and energy to proofread not only this thesis, but virtually every English-language document I authored in conjunction with this PhD program. If the reader makes it to the end of this text without feeling some annoyance at how I employ commas, it will likely have been thanks to Norma's diligent efforts. She also helped me to fine-tune my delivery when I have been called upon to present this research. Perhaps most importantly, during the times when my projects were not producing the results I had expected or when the program somehow seemed overwhelming, she was always there as a mother and a friend, to listen, to remind me why I started this journey, and to let me know how she was sure that I could complete it.

Hutan do Céu de Almeida a été la première personne à qui j'ai dit que j'avais l'intention de poursuivre un programme de 3^{e} cycle à l'UQÀM. Il a été une source constante d'appui et d'encouragement. Il a relu plusieurs documents que j'ai rédigés pendant mon doctorat et j'apprécie le point de vue unique qu'il a mis sur les divers thèmes de ma recherche. En même temps, je suis reconnaissant de l'aide qu'il m'a donnée et j'ai hâte de continuer notre série de conversations intéressantes sur l'avenir et ses possibilités.

Les membres du groupe de doctorants le PhDIC, Marie-Josée Gagné, Guillaume Spalla et Jonathan Elie-Fortier, ont été une source d'aide et d'inspiration pendant ce programme universitaire. Je les remercie de tous leurs commentaires constructifs, le temps donné pour relire mes documents en langue française et surtout des moments partagés avec des amis qui comprennent tellement bien les joies et les douleurs sur le chemin d'un doctorat.

It can be difficult to maintain employment while finishing up a full-time PhD program. I am grateful to Patrick Patterson, Guillaume Amringer, and all my coworkers at Carillon Information Security for working with me to accommodate my schedule so that I would have time to finish my research, attend events associated with the university, and ultimately complete my work on this thesis. I feel fortunate to be part of a company that values education, and I hope that some of the fruits of my efforts here will lead to future endeavours with Carillon in the domain of information security.

I would like to thank Scott Helvick, Mauricio Fernández, and most recently Shawn Whitley of (defun games ()) for their support and their patience as I worked to complete this important milestone on a never-ending path towards a deeper understanding of all things that compute, reason, and seek enjoyment in life. I believe that Spycursion can only benefit from what is ultimately a new perspective on AI that I have developed throughout the course of this program.

I am also grateful for the members of my loyal class on Erlang/OTP and functional programming: Samuel Marticotte, Igor Kozlov, and Alexia Palmieri. In addition to the support and encouragement they have given me, our conversations on programming theory and more generally on just "what if" have subtly made their way into aspects of my research as well as my perspective for the future. I look forward to many more conversations with them in the years to come.

La liste ci-dessus n'est pas du tout exhaustive. Les noms mentionnés représentent les personnes qui ont joué un rôle exceptionnellement actif pour m'aider dans mon but d'obtenir un diplôme de doctorat en informatique cognitive. Toutefois, il y en a tant d'autres qui ont contribué pendant ce programme et il est impossible de les nommer toutes. C'est un privilège pour moi d'avoir eu l'opportunité de recevoir ce niveau exceptionnel de formation des professeurs à l'UQÀM et j'ai de la chance d'avoir l'appui de mes collègues, mes ami·e·s et ma famille. J'offre mes sincères remerciements.

The list above is in no way exhaustive. The names I have mentioned represent people who have taken an exceptionally active role in helping me in my goal of obtaining a doctoral degree in cognitive informatics. Yet, there are so many others who have contributed over the course of this program that it is impossible to list them all. I feel privileged to have had the opportunity to receive an exceptional level instruction from the professors at UQÀM and fortunate for the support of my colleagues, my friends, and my family. I offer my sincere thanks.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

RÉSUMÉ

Ce projet de recherche a pour but de modéliser les attitudes et les croyances sur le changement climatique des personnes dans les communautés du site de média social Twitter. Les modèles informatiques se fondent sur la série d'études sociopsychologiques, basées sur des sondages, qui s'appelle les Six Amériques. Notre recherche emploie des modèles ontologiques et d'apprentissage automatique ainsi que des méthodes de la recherche d'information. Ces composantes fonctionnent ensemble afin de créer une architecture informatique qui s'inspire de la théorie hiérarchique de la cognition biologique. Cette architecture analyse les microblogues pour identifier le sentiment et l'émotion exprimée, le niveau d'activité des utilisateurs, les liens thématiques aux questions du sondage de l'étude originale des sciences humaines, ainsi que des éléments grammaticaux des messages en ligne, y compris les dépendances syntaxiques dans un texte donné. Cette recherche représente une première étape dans le développement de cette architecture qui vise à modéliser les opinions et les points de vue des centaines de milliers, voire millions, d'utilisateurs sur les médias sociaux à partir des communications publiques qu'ils publient en ligne. Les chercheurs pourront ensuite employer ce modèle et le mettre en correspondance avec des résultats des études établies en psychologie et sociologie basées sur une méthodologie traditionnelle de sondage avec un ensemble relativement petit de sujets.

Mots-clés : analyse des sentiments, apprentissage automatique, changement climatique, Dolce+D&S Ultralite, émotions de base, logiques de description, Lucene, média social, ontologie, Plutchik, réchauffement climatique, recherche d'information, Six Amériques, traitement automatique du langage naturel, Twitter, Web Ontology Language

ABSTRACT

This research involves modelling attitudes and beliefs on climate change of people in online communities on the social media site Twitter. The informational models are based on the series of survey-based, socio-psychological studies know as the Six Americas. Our work leverages ontological and machine learning models as well as information retrieval methods functioning together to create a computing architecture inspired by the theory of hierarchical biological cognition. This architecture analyzes microblogs, identifying expressed sentiment and emotion, user activity level, topic links to survey questions from the underlying study in the human sciences, as well as grammatical elements in online posts including the syntactic dependencies within a given text. This research represents a significant first step in developing an architecture whose ultimate goal is to model the positions and perspectives of hundreds of thousands or even millions of online users on social media based on the public communications they post. Researchers may then use this model for comparison with the results of established studies in psychology and sociology based on traditional survey methods which use a relatively small set of subjects.

Keywords: basic emotions, climate change, description logics, Dolce+D&S Ultralite, emotion mining, global warming, information retrieval, Lucene, machine learning, natural language processing, ontology, Plutchik, sentiment analysis, social media, Six Americas, Twitter, Web Ontology Language

INTRODUCTION

This research project aims to leverage technologies in artificial intelligence (AI) to help provide a better understanding of how online communities view the problem of climate change. An unlikely definition from the IBM Jargon and General Computing Dictionary (Cowlishaw, 1990) gives us a somewhat odd, yet seemingly apropos, starting point for this goal:

> **artificial intelligence 1.** *n.* The opposite of natural silliness. **2.** *n.* A research topic in Computer Science. Some in the computer industry think that nothing useful can come out of artificial intelligence (but they don't trust the natural kind, either).

This is a rather droll definition we are bringing into the context of what is ultimately an exceedingly serious situation. Nonetheless, it seems curiously befitting when we consider how 97% of climate scientists are publishing that the globe is warming and human beings are causing it to do so (Cook et al., 2016); yet, many people are insisting steadfastly that all of it is nonsense. Others may believe the science but seem to feel little need to react in any significant way at the present moment (Weber, 2006).

How bad is it? Many would argue that the problem is existential and possibly represents one of the biggest challenges that humanity has ever faced. The Intergovernmental Panel on Climate Change (IPCC) asserts that the science is "unequivocal." The atmosphere, the oceans, and the land are all warming, and these changes are due to human activity. The rise in global temperatures is severely impacting the earth's climate and will continue to do so without drastic reductions to emissions of CO_2 and other greenhouse gases (IPPC, 2021).

There is remarkable agreement across vastly different countries and cultures that global warming is caused primarily by humans and that it creates extensive changes in our world, resulting in natural disasters and negative effects on health (Crona et al., 2013). However, at least in the United States, while there is clear support for a broad range of governmental policies and personal action to mitigate climate change, one also sees a wide disparity as to levels of motivation, behaviours, and policy preferences (Maibach et al., 2011).

Climate change is an important political issue with far-reaching social and economic repercussions. Various factors, e.g., political orientation, can have a major impact on a person's beliefs about climate change. Some demographic segments of the American population, notably conservative white males, are significantly more likely to deny that climate change is real (McCright & Dunlap, 2011) and that it can be tied to elevated risks for climate-related disasters such as flooding (Botzen et al., 2016). Additionally, these factors may be interrelated. For example, in the U.S. a larger income tends to have an opposite effect on Republicans and Democrats, correlating to higher probabilities that Republicans will dismiss climate change and that Democrats will rank it as the single most important environmental problem to be addressed (Bohr, 2014).

Amid this complexity the perception of environmental risk and support for a given climate policy may be influenced by the emotion and the positive or negative sentiment connected to people's values and imagery as these relate to climate change (Leiserowitz, 2006). Emotions can be a compelling factor when the goal is to modify people's behaviour to avoid risk in a dangerous situation (Weber, 2006). As we work to understand how humans reason about the existential threat that is climate change and react to the diverse forms of information they receive about it, understanding the role emotions play may ultimately prove essential (Lu & Schuldt, 2015).

0.1 Purpose

If we accept that the situation is dire, the key question with regard to this doctoral research project then becomes: How might an AI solution help? Surely, potential answers to that question are a near-endless source of research topics. For this particular thesis, however, let us return to the idea of "natural silliness," provided so pithily by IBM (Cowlishaw, 1990). Their comparison between AI and the "natural kind" of intelligence is salient. Since the late 1940s, and certainly more formally in the 1950s,[1] human cognition has been both a source of inspiration and a virtual blueprint towards the ultimate goal of creating a machine that is able to think like a human (Negnevitsky, 2011).[2] Looking at the matter from the opposite direction, AI is also often used as a research tool to gain a better understanding of biological cognition. This is immediately evident in the numerous research efforts surrounding cognitive architectures such as ACT-R[3] (Anderson, 2007), Soar[4] (Laird, 2012), and SPAUN[5] (Eliasmith, 2015).

Our purpose in the present research does not attempt to model human cognition, but we are working towards creating an architecture inspired by theories of biological cognition that will potentially serve as a research tool helping to better our understanding of why certain humans think and act as they do in a specific context. The humans we are

[1]Conferences in 1956 at Dartmouth College and the Massachusetts Institute of Technology respectively mark John McCarthy's coining of the term *artificial intelligence* and the birth of the field of cognitive science (Andresen, 2002; Miller, 2003).

[2]Note that there is a large amount of disagreement in the AI research community as to whether or not this goal is even possible.

[3]Adaptive Control of Thought–Rational: a cognitive architecture dating back to the 1970s. It has evolved and continues to do so up to the present day. The *R* in the name appeared in the 1990s with the integration of the theory of rational analysis into ACT's core design (Anderson & Lebiere, 1998).

[4]Soar is an architecture based on the *Unified Theories of Cognition* (Newell, 1990). It is a culmination of research efforts on logic and problem solving that runs back to the 1950s.

[5]Semantic Pointer Architecture Unified Network: a more recent cognitive architecture based on artificial neural networks. It maintains biological plausibility as its primary goal.

considering are those forming online communities on social media. The context could conceivably be any issue for which there is established research in the human sciences, but we have chosen the issue of anthropogenic climate change for the present research.

Our work centres around emotion. We leverage AI techniques to analyze the expression of sentiment and emotion in online communications, essentially tapping into the public conversation and then using research from psychology and sociology as a foundation on which to create a model of the community having that conversation. This model may potentially serve as an extension to the original socio-psychological study, providing a view of how online beliefs and attitudes compare with those represented in the results of more traditional survey-based research efforts.

0.2 Stylistic Considerations

I believe that it may prove beneficial to alert the reader to a few peculiarities of my writing style, especially as they relate to the present document.

0.2.1 Person and Voice: *We* and *I*

With the exception of this section on stylistic considerations, the reader will notice that I am writing the present document using the first person plural. This narrative perspective facilitates use of the active voice. I prefer to say, "we performed the experiments," rather than "the experiments were performed." Although in the second statement, written in passive voice, the reader would doubtlessly assume those experiments were performed "by us," I prefer that the writing be direct. Ignoring the possible disadvantage of sounding slightly less scientific to some readers, writing using the active voice allows me to provide the reader with more information, often using fewer words. I suspect many readers may appreciate the latter point as they make their way through the material in this text. Packing a full scientific explanation into one or two sentences can be an appreciable challenge, but that challenge is even greater when the author feels he must come across as sounding "scientific."

The use of active voice in scientific writing has become relatively common in recent work, and if this were the only stylistic peculiarity in question, I would likely have skipped writing this section. However, I have realized that my particular style with the first person plural may require further explanation. As I have conducted this doctoral research at the *Université du Québec à Montréal*, an educational institution which uses the French language, a good amount of the work and the writing I have done in conjunction with this research is in French. French is not my first language, and so I feel it especially important to have a native speaker proofread my work before I submit it. When writing scientific documents in French, I generally use a similar style to what I am employing in this thesis, but once a fellow doctoral student who was kind enough to proofread my proposal for the present research commented that she was not always certain whom I was referring to when I used the word *nous*.[6]

I imagine my style may seem a bit odd, but natural enough to a native English speaker. Yet, at that moment I understood that this style does not necessarily work quite so well when using another language. I also suspect that English will not be the first language of a good number of the people reading this document, and so I fear that there could be a similar sense of ambiguity for at least a few readers. Ultimately, her comment caused me to examine my own style. I asked myself, "Whom do I mean when I say 'we'?" The answer came quickly enough, or rather the *answers*, as I saw that I actually switch my focus according to the context. Most often *we* means "myself with the help of my advisors." I expect the reader will find this use quite common in scientific writing that uses the first person plural. However, generally when I settle into a narrative I use an inclusive *we*, imagining guiding my readers through the experiments and the observations as if we were perhaps discussing the matter in a classroom setting. Finally,

[6] *Nous* is French for "we."

I occasionally use the inclusive *we* in a more extended sense, essentially referring to humanity. Sometimes this simply reflects the sense of being part of the species *Homo sapiens*. At other times it carries a sense of reverence for the advancement of human knowledge as discovered by history's great minds. This is especially true when I find myself citing a scientist who has contributed a truly founding idea in the domain of cognitive informatics. If I may be so bold, the privilege I have in making a humble contribution to this field gives me the sense of "standing on the shoulders of giants" as Isaac Newton expressed it in 1675 in his letter to Robert Hooke (Koyré, 1952).[7]

Although *we* is doubtlessly my pronoun of choice in this thesis, there are a few places where I switch to first person singular. I do this when I am describing an observation, and I do not believe I can take it for granted that my advisors share my point of view on the matter (e.g., the present section). Most often such passages involve my thoughts as a doctoral student as I reflect on the research conducted over the course of this PhD program. I believe the reader will find the switch from *we* to *I* natural enough from the context.

On a related note, as I make liberal use of the inclusive *we* in my writing, I avoid the second person pronoun *you*. Given that I feel an author's first goal should be to write clearly, I certainly have no critique regarding the use of second person. However, switching between first person singular and plural is arguably already pushing my licence as the author of a scientific document. More importantly, I do not wish to invite further ambiguity by juggling too many pronouns. Nevertheless, there are times when I see the need to address my readers directly. In all cases I use the phrase "the reader" to express this intent.

[7]Newton is often credited with coining this phrase, but the idea predates him (Koyré, 1952).

0.2.2 Descriptive Examples: *She* and *He*

It is important, especially in this day and age, that writing be inclusive. Accordingly, when using descriptive examples in the text, this work strives to utilize the gender pronouns *she* and *he* with roughly equal frequency. Modern authors often employ expressions such as "he or she." Although I applaud the sentiment, personally I find this usage overly wordy, awkward, and unnatural. It is also common these days to utilize the pronoun *they* to refer to a single person, thus avoiding any explicit reference to that person's gender. Again, I tend to envision *they* as a reference to a group, and so this usage feels awkward to me in my writing. Therefore, in this work I simply alternate between *he* and *she*, using one through the entire explanation for a given example, and then switching to the other for the next example. This strategy has the additional benefit of a modicum of added clarity, whereby the pronouns' gender helps to distinguish "who is who" in any examples employing two hypothetical participants.

In an effort to make a tiny contribution towards making up for too long a history of gender noninclusive writing, I generally use *she* for the first participant in the first example in each section. I also tend to assign the female pronoun to participants that are demonstrating traditionally male-dominated roles, such as computer programmers and technicians[8] in the hope that the idea of a woman performing a certain role will eventually seem quite natural to readers who may automatically tend to picture a man in her place.

---

[8]This modern stereotype is particularly unfortunate due to the integral role women have played in the history of computing. Ada Lovelace's seminal work in the 1840s with algorithms destined for mechanical computing machines (Füegi & Francis, 2003) and the team of women programming the ENIAC, the world's first electronic, general-purpose computer (Light, 1999), are two out of many important examples.

0.2.3 Terminology: Climate Change and Global Warming

Terminology can be key. Back in 2002 in the United States, a consultant for the Republican party advised then president George W. Bush that his administration should use the term *climate change* for governmental documents, rather than *global warming*. The consultant believed that the former was less alarming and emotional, less likely to invoke fear at the thought of an impending disaster. Ironically perhaps, more recent research efforts indicate that *global warming* is the more divisive term (Akerlof & Maibach, 2011; Weber, 2006; Williams et al., 2015). The term *climate change* is arguably more descriptive as the changes happening on the earth inevitably take many forms including shifts of climate zones and local surface cooling effects in certain regions due to factors such as decreases in vegetation. The IPCC typically uses *climate change* (or *global climate change*) as a general "umbrella" term and *global warming* to refer more specifically to the actual warming of the planet (IPPC, 2021).

With respect to this thesis the reader may consider both terms to be synonymous. Generally, we prefer to follow the example of the IPCC, using the term *climate change* unless we are specifically discussing warming effects. However, our research is closely linked to a series of studies known as the Six Americas (Maibach et al., 2011), which has historically favoured the term *global warming*. Additionally, our research involves analyzing public communications on social media, and different groups in the online community have different preferences with regard to terminology (Williams et al., 2015). In some cases we can leverage the two terms to an extent in order to subtly distinguish between (1) *climate change* referring to what is actually happening on the earth, and (2) *global warming* referring to climate change as studied by a specific research effort (e.g., the Six Americas). Of course, the line between these two usages is often blurred, and so the reader may simply consider the terms to be synonyms throughout the present document.

0.3 The Six Americas

In the present research we look to the human sciences and to studies seeking to identify the sociodemographic groups who are more likely to believe in anthropogenic climate change and those who tend to be skeptical. Socio-psychological models from such studies may indicate who will generally support a given climate policy and who will not.

There are numerous studies investigating why some people believe the science of climate change while others do not. The model we use in this work is based on a series of studies called the Six Americas (Maibach et al., 2011), which has been running for well over a decade. The goal of these studies is to group people in the United States into six segments according to their attitudes, behaviours and political stance with respect to global warming. The value of the Six Americas goes beyond simply providing a better understanding of where adults in the U.S. stand on the subject of global warming. Governments, organizations, and researchers can use these six categories as indicators to more strategically adapt their message to an intended audience for information relating to various issues on climate change. This approach for message framing is often utilized in marketing, public health, political science, and other domains (Maibach et al., 2011; Nabi et al., 2018).

The following are the categories in the Six Americas used to describe these six segments of the U.S. population. After each segment are two numbers in parentheses. The first is the percentage of adults in the segment per the original survey [$n = 2,129$] from the Fall of 2008 (Maibach et al., 2011). The second number is the percentage according to the latest Six Americas survey [$n = 1,036$], which at the time of this writing is from December 2020 (Leiserowitz et al., 2021a).[9]

- *Alarmed (18%; 26%)*: People in the *alarmed* segment are certain that climate

[9] The Yale Program on Climate Change Communication publishes the most recent survey results on their website: https://climatecommunication.yale.edu/about/projects/global-warmings-six-americas/

change is real, that the threat is urgent, and that humans are the primary cause. They support an ambitious response from the government and make changes in their own lives in order to mitigate the danger.

- *Concerned (33%; 29%)*: People in the *concerned* segment believe in climate change and see it as a serious problem. They support a response from the government, but they are less likely to make changes in their personal lives.
- *Cautious (19%; 19%)*: People in the *cautious* segment think that climate change is a problem, through they may not be completely convinced that it is real. They generally do not see it as a threat that affects them personally and do not see any urgent need for action.
- *Disengaged (12%; 6%)*: People in the *disengaged* segment do not often think about climate change. They do not feel particularly well informed on the subject. If they do have an opinion on climate change, they will say that they could easily change their mind.
- *Doubtful (11%; 12%)*: There are three types of people in the *doubtful* segment: (1) those who believe in climate change, (2) those who do not believe in it, and (3) those who simply do not know. The people who think climate change may be real believe that it is due to natural changes in the environment. They see any threat as dubious and far in the future.
- *Dismissive (7%; 8%)*: People in the *dismissive* segment firmly deny climate change. They do not believe it exists and see no threat to humans nor to the environment. They actively participate in climate-related issues, but on the opposing side from the *alarmed* segment.

To generate the model representing the percentages of the U.S. adult population in each of the six segments, (Maibach et al., 2011) conduct Latent Class Analysis on 36 variables out of a total of 306 from the questions in the Six Americas survey. The researchers identified these 36 variables as being linked to four key focus points with respect to global warming: (1) what a person believes, (2) how involved she is in the issues, (3) her behaviour related to conservation, and (4) how she thinks society should respond.

The long-term goal of the present research endeavour is to develop a model of climate-oriented user activity on social media based on all six categories of the Six Americas. However, the scope of this doctoral research program is ultimately able to encompass only a first step in this direction. This first step involves two categories: *green* and *denier*. Table 0.1 shows the mapping from the Six Americas series of studies to the current state of our research as presented in this document. The leftmost column gives the category from the Six Americas as described above. For reference, the centre column repeats the percentages of surveyed adults in each category as of December 2020 (Leiserowitz et al., 2021a). The final column lists the mapped categories as they are modelled in the present work.

Table 0.1 Mapping the Six Americas categories to green and denier.

Six Americas	**Dec. 2020**	**Present Work**
Alarmed	26%	**Green**
Concerned	29%	
Cautious	19%	Not currently modelled
Disengaged	6%	
Doubtful	12%	**Denier**
Dismissive	8%	

People in the remaining two categories of the Six Americas (*cautious* and *disengaged*) tend not to have strong opinions on climate change and may even be unaware of it. We hypothesize that they will therefore publish little, if anything, about the subject online. Although the perceived risk and the level of engagement differ, the *alarmed* and *concerned* generally both believe that climate change is real and that it is caused by humans (Maibach et al., 2011). The present work labels both categories as green. In contrast, the *doubtful* and the *dismissive* generally do not believe in anthropogenic climate change. Again, the level of engagement of the *dismissive* is greater (broadly opposing the efforts of the *alarmed*). We label these two categories as denier.

Of course, it should be understood that in reality people have a myriad of perspectives on climate change, and reducing the analysis to those who believe the science vs. those who do not can be limiting. Subtle issues are invariably blurred such as those relating to questions about the exact nature of the problem that is anthropogenic climate change and which of countless policy choices governments and local and international organizations should choose to tackle it (Corry & Jørgensen, 2015). Labelling people in this manner may also potentially serve to dehumanize them and inadvertently emphasize certain stereotypes which may affect the analysis in any given study. This may be especially true of a label like "denier," which is not uncommon in academic papers on climate change (Howarth & Sharman, 2015). The reader should keep in mind that our choice of these two categories here has everything to do with keeping the scale of the problem at a level where we are able to obtain interesting results within the scope of a doctoral program. We fully understand that there are important complexities that we are simply not including in our model. Our work here involves handling the complexities inherent in the turbulent stream of ideas on social media that serves as the input to our model. That model's simplistic output, a binary classifier identifying a user as green or denier, is intended as a simple yet clear indicator of how well the model is handling that turbulent stream. Future research efforts shall take higher aim with regard to how our model represents the people whose online posts it is processing. Enhancing the model to represent the six categories from the Six Americas is the obvious next step for our research though many of the concerns just addressed will undoubtedly still apply.

0.4 Research Objectives

The problem we are attempting to solve in the scope of this doctoral program basically concerns the creation of an architecture which serves as an automated expert. This expert "understands" an established study in the human sciences, typically one based on traditional scientific survey methods, and is able to generate a model of the beliefs and behaviours of online communities based on what the members of those communities are posting publicly on social media. This model uses the terms and the structure of the original study in the human sciences. In this way, findings stemming from research based on the model of the online community may be compared to those of the original survey-based research.

In the present research we are concerned specifically with the analysis of emotion in the context of online social media and climate change. Essentially we are endeavouring to leverage the emotion expressed in posts on social media in order to model online communities with respect to their attitudes on anthropogenic climate change. Our research revolves around the architecture we are designing for an application potentially capable of processing millions of online posts, analyzing their content, and finally modelling the users publishing these posts in accordance with established research from the fields of psychology and sociology. For our work encompassed by this doctoral program, we are drawing from communities on the social network Twitter[10] and utilizing the Six Americas series of studies (Maibach et al., 2011) as the foundation for our model of how human beings relate to climate change.

The system we are creating is called "*Say Sila*."[11] Sila is the god of weather and spirit from Inuit mythology. Note that references to *Sila* are often mistranslated to simply mean "weather" in papers on climate research. The error seems oddly fitting given the amount of misinformation which circulates on the subject of climate change. Additionally, the concept of Sila varies somewhat from one regional group of Inuit, where Sila is a force of spirit, to another, where he is a powerful god (Leduc, 2007). The name of the application invokes a sacred element of the mythology of the Inuit with the purpose of recognizing the important role they and other indigenous peoples should have in a world needing to deal seriously with climate change. This research project is essentially

[10]https://twitter.com/

[11]Note that we have also created an ontology called "say-sila" which is incorporated into the *Say Sila* system as part of this research (see Section 2.3). To help differentiate between the two, we use a standard font for the ontology and this stylized font when referring to the architecture and the system that implements it.

about “the conversation” as it continually takes place in an online setting. The project strives to model linguistic aspects as well as important psychological and social elements as they relate to how human beings process the complicated and divisive subject of climate change. The name *Say S$\vec{\imath}$La* speaks to the input to this model, which is made up of the myriad of things people are saying (or rather tweeting) about global warming.

0.4.1 System Architecture

The research presented in this thesis is best understood in context of the architecture of the *Say S$\vec{\imath}$La* application. Throughout the course of the doctoral program, we found that we needed to concentrate our efforts on a number of specific research problems in AI relating to the functionality required by the various components in our system. We have structured the present document to reflect our work on these research problems rather than on the progressive implementation of the system itself. Nevertheless, we feel the reader will benefit from an understanding of the architecture of *Say S$\vec{\imath}$La*. This understanding should allow for a sense of coherence to what might otherwise seem to be a motley assortment of experiments in machine learning, natural language processing, and description logics. At the end of this thesis we take another look at *Say S$\vec{\imath}$La* to evaluate what we have accomplished over the course of this program and what remains to be investigated as part of our future research endeavours (see Section 8.1).

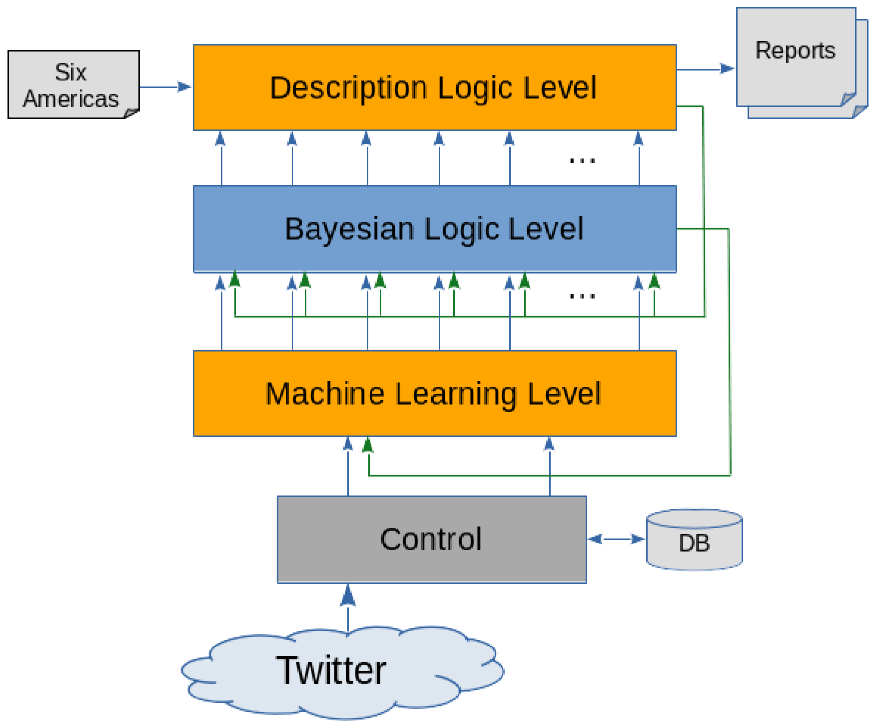
The primary goal of the system is to consume online posts from a social media network (Twitter in the present research) and to develop an ontological model of the community of users publishing these posts. This model must be founded on survey-based research from the human sciences. For our study, the ontological model is based on the Six Americas (Maibach et al., 2011); however, our intent is that the architecture shall ultimately be general-purpose. Unfortunately, the scope of this PhD program does not allow time for us to delve into a second research project from the human sciences, but we look forward to future research efforts so that the architecture may be employed not only to evaluate how other studies on *climate change* may translate when applied to online communities, but also studies on any of a vast range of important research topics.

The architecture of the application is inspired by the theory of hierarchical biological cognition. Humans beings access the outside world through the senses, but the signals associated with perception of the stimuli coming from the world are typically noisy and unclear. At times they may be altogether erroneous. Research indicates that the process of perception encodes these signals in a manner that resembles probability distributions. Hence, with regard to cognition at a computational level, perception likely functions according to biological "rules" of probabilistic inference (Knill & Pouget, 2004). According to the theory of hierarchical cognition, there are numerous levels of probabilistic models between the initial mental representation generated by the senses and the actual conceptual model of the world one holds in the mind. Furthermore, at any given stage in the hierarchy a higher level continuously transmits a signal to the level below it, predicting what the lower level should be perceiving next. This predictive signal runs down the full length of the hierarchy of models, essentially representing a progression of reduced complexity starting from higher forms of reasoning at the top levels and ending at the perception of raw stimuli at the bottom. At the same time, the parts of the brain perceiving stimuli from the outside world generate a signal that runs the hierarchy in the other direction. At any given level, this upward-bound signal serves either to confirm or to correct the probabilistic model at the next higher level (Clark, 2013).

Figure 0.1 presents the high-level design for *Say Sĭĺa*. At the bottom of the diagram, we have the online social media network Twitter. From Twitter, the blue arrow represents a stream of tweets flowing into the control module. This is essentially the "main" module of the application although it includes no AI components. The control module provides an operator interface and essentially serves as the backbone for the application. It is also responsible for collecting microblogs from the online service Twitter, handling their storage in a PostgreSQL database[12] (indicated as "DB" in the figure), and enabling their

[12]https://www.postgresql.org/

Figure 0.1 High-level design of the *Say S$\vec{\imath}$La* application.



subsequent retrieval for the experiments in our research. This was the first module we implemented in *Say S$\vec{\imath}$La*. Collecting this online data allowed us to perform an initial series of experiments called the "big player" project which we describe in Chapter 3.

The three coloured components in the middle of Figure 0.1 together form a hierarchy of models which essentially link tweets (the input) to a final representative model of their authors (the output). At the bottom of the hierarchy is the machine learning level. This level of the architecture corresponds–at least metaphorically–to layers that are lower on the hierarchy of biological cognition, layers that are associated with perception and pattern recognition in the brain. In *Say S$\vec{\imath}$La* the machine learning level analyzes online posts and their accompanying metadata with the intent to identify factors that potentially relate to a person's stance on climate change, such as gender, political orientation, position on environmental conservation, as well as the emotion expressed in the texts.

The next level up is the Bayesian logic level.[13] Continuing with our metaphor, this level corresponds to the higher layers of hierarchical biological cognition. This layer is intended to adjust the input signal to the high-level model, using probabilistic inference to enforce a "degree of belief" with respect to the data entering the system. This probabilistic inference will potentially serve to filter various forms of noise inherent in social media posts such as incorrect usage of language, sarcasm, advertisements, and off-topic messaging. The Bayesian logic level is intended to function as a collection of models of maximum likelihood estimation (Goodfellow et al., 2016). These models will take their input from the machine learning level below, and their outputs will be fed to the top layer of the architecture. The reader should note, however, that the scope of this doctoral program has allowed only for a minimal contribution with respect to the Bayesian logic level as compared to the rest of the of the architecture.[14]

At the top of the *Say Sūla* hierarchy is the description logic level. Although it is an integral part of the hierarchy in our architecture, it quite purposely departs from the metaphor of biological cognition. This level represents the system's conceptual model of the community we are studying on social media, but it is intended to be an informational model, capable of being queried and analyzed by standard computing systems (e.g., to generate a report as illustrated in Figure 0.1). In the present research this final model is based on the Six Americas (Maibach et al., 2011), indicated in the figure as an architectural input at the top of the hierarchy.

Figure 0.1 uses blue arrows to indicate the progression of the transformed data signal from the control module up to the top conceptual model. However, there are also green

[13]The name *Bayesian* refers to probabilistic logic which has roots dating back to the 18th century when the English minister Thomas Bayes developed some insightful ideas on chance. His observations were formalized decades later by the French mathematician Pierre-Simon de Laplace (Domingos, 2015).

[14]Although it is the second layer in the modelling hierarchy, the system is designed such that the machine learning level may optionally connect directly into the description logic level. The intent for our larger research effort is that we experiment first with these two layers and then insert the Bayesian logic level. Using this approach, we will be able to compare results and analyze the three-layer option using the two-layer configuration as a baseline.

arrows originating from the two higher modelling levels, each running to the next level down. These green arrows represent a predictive feedback signal. Just as in the theory of hierarchical human cognition, in our architecture the conceptual model will also serve to adjust the probabilistic models, indicating what they are likely to receive next as input. Then in turn, the probabilistic models will send an adjustment signal down to the machine learning level where this signal will be treated as another input to the data mining models on that level.

0.4.2 Technological Choices and Development

The *Say Sūla* application represents a central theme with respect to the present research. We have developed it with the constant vision that it will serve as a significant contribution to the scientific research community. To this end, we have taken care in considering the most appropriate technologies on which to base the system so that it may function as a reliable, robust, and flexible platform not only for this doctoral project but for future research as well. Of course, in light of the impressive advances in deep learning over the past decade, some readers may question our decision to create a complex architecture, based largely on ontologies and description logics to solve a problem of binary classification. It is important to keep in mind that although the initial problem we have chosen for *Say Sūla* could most likely be handled more quickly and accurately using a deep learning solution,[15] our future vision for this architecture extends well beyond basic problems of binary classification.

It is true that when it comes to pattern recognition, the performance of deep learning models readily compares and sometimes even surpasses that of human beings (He et al., 2015). An approach based on deep learning would be reasonable and certainly applicable to the field of cognitive informatics. As a ready example, the convolutional neural

[15] See (Munikar et al., 2019) for an example that leverages Google's BERT (Bidirectional Encoder Representations from Transformers) for sentiment analysis.

network, arguably the most widely recognized deep learning model, took its inspiration from biological systems for vision (Deng & Yu, 2014). Yet even considering the advances in deep learning algorithms in recent years, a number of researchers are pointing out the limitations of statistical methods and studying a variety of strategies that combine statistical and symbolic approaches to solve complex problems using AI (Domingos & Lowd, 2019; Kimmig et al., 2015). In the present research, as we look to our own brains for inspiration, we also note that dual process theories of biological cognition (Evans, 2009) indicate that hybrid architectures may provide interesting solutions for problems that call for intelligent systems (see Section 1.4), potentially allowing us to go further than we might when simply working in the framework of a single methodology.

As the present research applies to the domain of cognitive informatics, we have a particular interest in working to create an architecture that is inspired by current theories of biological cognition. We understand that we have chosen what might be considered to be a relatively simple classification problem from a standard machine learning viewpoint. We have scaled the problem down in order to focus on various aspects of the *Say S$\vec{i}$La* architecture. Our work here represents an initial endeavour in generating useful output from a system based on this architecture. Determining whether Twitter users are in a green or denier category is a readily measurable step and a building block for research involving more complex problems. It is a first step, and we are making it with the intent to continue our research, attacking increasingly difficult scenarios after the completion of this doctoral program. A number of these future studies will certainly involving replacing the baseline machine learning algorithms we are currently using in the architecture with deep learning components.

Furthermore, in addition to its use as a simple classifier, the architecture provides a valuable descriptive model of a community on social media in terms defined by an established study in the human sciences. There is distinct value in the fact that this model is easily accessible using standard technologies for the Semantic Web. In the present research the community represents Twitter users talking about climate change, and the study is the Six Americas. However, the base Say-Sila architecture is general purpose. It can be readily adapted for modelling online communities within the descriptive framework of a vast array of other areas of academic research.

Development in Erlang

The *Say Sīla* project relies heavily on the functional programming language Erlang. Erlang provides an advantage for a system such as the one we are creating for this research project as the language is designed around a fault-tolerant, highly-concurrent, distributed programming model (Armstrong, 2013). Applications in the domain of AI often require significant computational resources, and Erlang makes effective use of today's multicore processors and multiprocessor servers. Additionally, for particularly large systems, the language also facilitates spreading the load across a great many servers essentially allowing them to function as one machine.

The control module (see Figure 0.1) is programmed entirely in Erlang. In the current version of *Say Sīla*, which is v0.3, this module contains over 6300 lines of Erlang code (not including comments or blank lines) in 44 source files (including header files).[16]

Table 0.2 External Erlang libraries used in *Say Sīla*.

Library	Version	URL	Description
ecsv	0.3	https://github.com/rcouch/ecsv	CSV file parsing
epgsql	3.3.0	https://github.com/epgsql/epgsql	PostgreSQL database routines
erlsom	1.5.0	https://github.com/willemdj/erlsom	XML parsing and generation
jsx	2.11.0	https://github.com/talentdeficit/jsx	JSON formatting routines
lager	3.8.1	https://github.com/erlang-lager/lager	Server logging routines
oauth	1.6.0	https://github.com/tim/erlang-oauth	Open authorization for online access
tempo	0.4.3	https://github.com/selectel/tempo	Date/time parsing routines
yaws	2.0.7	https://github.com/klacke/yaws	Web server for the user interface

In addition to version 22 of the Erlang language and its standard library framework known as the Open Telecom Platform (OTP), the control module in *Say Sīla* makes use of a number of third-party libraries. These are listed in Table 0.2 for reference.

[16]As reported by the tokei code statistics generator: https://github.com/XAMPPRocky/tokei

Development in Clojure

As much as we benefit from Erlang as our base language for *Say Sĩla*, Erlang is not a programming language that typically targets AI systems. Therefore, we use the functional programming language Clojure for application components relating to AI. Clojure is a member of the Lisp family of programming languages that runs on the Java Virtual Machine (JVM) and allows 100% interoperability with Java packages. It is also a language that embodies a concurrent programming paradigm, facilitating the use of multiple processor cores, which is one of the benefits we mentioned for Erlang in the previous section (Hickey, 2008; Hickey, 2020). Lisps have a rich history with respect to symbolic computing in the domain of AI. Clojure is a good choice for *Say Sĩla* for this reason as well as the fact that running on the JVM provides access to a wide range of Java-based AI libraries, most notably for our research: the Weka machine learning platform (Frank et al., 2016), the Java API for OWL (Horridge & Bechhofer, 2011),[17] the HermiT OWL reasoner (Glimm et al., 2014), and the Lucene information retrieval platform from the Apache Software Foundation.

Referring back to Figure 0.1, we use Clojure for the modules associated with all of the modelling levels in *Say Sĩla*, shown as the coloured blocks in the diagram. For this research project we have written over 14,300 lines of Clojure code (excluding comments and blank lines).[18] This code constitutes 45 source files in the application.

We use Clojure version 1.10.1 for *Say Sĩla*. In addition to the base language and numerous official libraries associated with the language under the group identifier "org.clojure," we use a number of third-party libraries. These are listed in Table 0.3. Note that a number of these dependencies require importing more than a single library. We denote these cases using an asterisk (*) as the final element in the artifact identifier as it is listed in

---

[17] OWL is an oddly ordered acronym for the Web Ontology Language. The reader will find this and many other acronyms following the glossary at the end of this document.

[18] As reported by the tokei code statistics generator: https://github.com/XAMPPRocky/tokei

Table 0.3 External Clojure libraries used in *Say Sila*.

Clojars/Maven Repository	Version	Description
affective/affectivetweets	1.0.2	Sentiment & emotion processing for Twitter
clj-time	0.14.2	Date/time functionality
clojusc/wordnet	1.2.0	WordNet reference routines
defun	0.3.1	Erlang-style pattern matching
edu.cmu.cs/ark-tweet-nlp	0.3.2	Twitter NLP tools
enlive	1.1.6	HTML parsing routines
incanter/incanter-*	1.9.3	Statistics and charting library
me.raynes/fs	1.4.6	Enhanced file system routines
nz.ac.waikato.cms.weka/weka-stable	3.8.3	Weka machine learning platform
org.apache.lucene/lucene-*	7.7.3	Lucene information retrieval platform
org.erlang.otp/jinterface	1.9.1	Access to distributed Erlang nodes
net.sourceforge.owlapi/org.semanticweb.hermit	1.4.5.456	HermiT OWL reasoner
uk.org.russet/tawny-owl	2.0.3	Clojure wrappings for Java API for OWL

the Clojars[19] and Maven[20] Internet library repositories.

0.4.3 Contributions

This research project represents an integral part of the PhD program: *Doctorat en informatique cognitive (DIC)* at the *Université du Québec à Montréal (UQÀM)*. The associated work includes several contributions to the academic community:

- A functioning, open-source prototype of the *Say Sila* architecture which provides an ontological model of a study in the human sciences, populates that model using the online posts of a community on social media, provides a platform for experimental analysis of the model, and reports results from these experiments.
- An ontology called *say-sila*[21] which may be used for experimental modelling using tools for the Semantic Web for studies on climate change based on the Six Americas.
- A stand-alone ontology called *cmu-pos*[22] which may be included as part of larger

[19] https://clojars.org/

[20] https://maven.apache.org/

[21] http://www.dendrown.net/uqam/say-sila.owl#

[22] The name *cmu-pos* stands for "Carnegie Mellon University–Parts of Speech" (see Section 4.3): http://www.dendrown.net/uqam/cmu-pos.owl#

studies for natural language processing on social media.

- A relatively simple methodology (compared to graph-based approaches) for sentiment and emotion analysis on social media based on communication categories and the level of user participation.
- An analysis of the relationship between the emotion present in the microblogs of high-activity users on Twitter and that of the rest of the community for users posting tweets about global warming.
- An analysis of the stance of Twitter users on global warming based on how their tweets relate to the series of socio-psychological studies known as the Six Americas.
- A "follow the leader" method for labelling datasets of online posts for communications on climate change.
- An analysis comparing the predictive capacity of the simple co-occurrence of important concepts in microblogs vs. the predictive capacity when requiring a relationship of syntactic dependency between terms indicating these concepts.
- An analysis based on a methodology of reversed information retrieval to show levels of sentiment and emotion expressed on Twitter in posts from users for green (pro-science) or denier (skeptic) communities tweeting about global warming.
- A method of representing measured affect in online communications as visual signatures.
- An experimental analysis of how the hybrid, hierarchical architecture *Say $\vec{Si}La$* can utilize results from one pass through the modelling layers of the architecture to improve its predictive capacity for subsequent passes.
- A working example of how an iterative strategy may be used in academic research to develop a significantly large experimental architecture.

Regarding our contribution of the open source application *Say $\vec{Si}La$*, the source code is publicly available online at GitHub using the following Web link:

https://github.com/dendrown/say_sila

Furthermore, as the full development of the system does not fall into the scope of this doctoral program, we have designated a specific branch for the code repository as it pertains to the research presented in this thesis. The Web link for this branch is the following:

https://github.com/dendrown/say_sila/tree/uqam-dic

As mentioned above, we still have significant work ahead of us with *Say Sila*, and so the repository shall remain active for continued research and development for the project. However, the *uqam-dic* branch will be static and serve as a reference for the present research. At various points in this document we include URLs which point to the implementation in the code for the topic being discussed. These URLs will always reference this *uqam-dic* branch, so that the links remain stable for the reader even as the main code base continues to evolve.

We will take a brief look at the future endeavours for this project later in Section 8.1. For now *Say Sila*, as we have developed the application for this doctoral program, provides a framework on which we may model communities of people talking about climate change on social media, basing that model on established science and on the emotion and sentiment that those communities are expressing online. There is an iterative rhythm to this work. We build the application to conduct the experiments, allowing us to create a part of our model. We then use the application to evaluate that part of the model, indicating how accurate it is and what might need refinement in a subsequent stage of research. We have undertaken a rather ambitious goal in creating this system intended to extend the reach of survey-based studies in the human sciences by extracting knowledge from the realm of social media. Our research in this doctoral program represents a solid first step towards that goal. The details of this step and of the many ways it links informatics and cognitive science are described in the chapters that follow.

CHAPTER I

ANALYZING EMOTION

The analysis of sentiment and emotion may arguably be considered a smaller-but-complete version of the field of natural language processing (NLP). Although originally from the domain of computer science, it has become a point of intense interest in numerous areas including advertising and sales, management, and social science just to name a few. Sentiment analysis, frequently known as opinion mining, has been an active research area since the early 2000s. Interest has been steadily increasing since then due in large part to the vast amount of opinions publicly posted on social media. Analytical strategies span from more traditional methods based on lexica and sentence composition to statistical approaches leveraging more recent advances such as deep learning (Liu, 2015; Pozzi et al., 2017).

As would likely be expected, the field of emotion mining is closely tied to sentiment analysis. However, meanings for terms like *opinion*, *sentiment*, and *emotion* often overlap or otherwise tend to flow into each other. This can create ambiguity and sometimes confusion when considering techniques aimed at analyzing human language (Liu, 2015; Pozzi et al., 2017). People may simply express an opinion with no lexical cues as to how they feel about the matter: "*I think Senator Smith is going to be reelected.*" However, frequently they will express a certain sentiment, giving an opinion in a way that shows some degree of pleasure or displeasure: "*Smith will be getting back in, even after that awful speech in protest of the green initiative.*" People also often relay information about their inner emotional state: "*I got so angry after I heard that speech.*" In this chapter we will be taking a look at sentiment and emotion, discussing how including these elements can be useful when analyzing natural language, particularly in the context of social media and of climate change.

1.1 What is Emotion?

What is emotion? Unfortunately we will tend to get very different answers depending on whether we consult established research in philosophy, psychology, sociology, or in virtually any domain that looks into this elusive topic. Worse yet, even when coming from the same field of study, experts often disagree even as to a basic definition for *emotion*, as well as for two other terms used rather frequently in the present work: *sentiment* and *affect* (Mulligan & Scherer, 2012). In fact, Robert Plutchik, whose model of basic emotions we use in the present research, goes as far as to say, "the study of emotions is one of the most confused chapters in the social sciences" (Plutchik, 2000).

Traditionally, researchers have considered emotion more or less according to what may be called the classical view. This view goes back at least as far as Hippocrates and Aristotle and has been adopted in one form or another throughout the history of science by such profoundly influential thinkers as René Descartes, Charles Darwin, and Sigmund Freud (Barrett, 2017). The view remains prominent today in the work from respected researchers in psychology such as Paul Ekman (Ekman, 1992) and Stephen Pinker (Pinker, 2009). According to the classical view, emotions are a product of evolution, naturally selected long ago to be an integral part of human biology. Emotions are therefore universal, and allowing for very small differences due to physiology and local customs, they represent a common experience and share a common manifestation across cultures throughout the history of *Homo sapiens* (Barrett, 2017).

One of the most well known studies in support of the classical view of emotions was performed by Ekman, Friesen, and colleagues in Papua New Guinea. There, the researchers

Figure 1.1 Facial expressions for Ekman and Friesen's study of emotions.

found that the Fore people were able to match emotions and facial expressions such as those shown in Figure 1.1 for six basic emotions: anger, disgust, fear, happiness, sadness, and surprise (Ekman & Friesen, 1971). Although a number of similar studies had been performed for various cultures across the world, the Fore are notably remote and have little contact with people from the Western World (Barrett, 2017).

Alternatively, more recent research in psychology[1] counters the commonly accepted classical understanding of emotion, pointing rather to a completely different view known as the theory of constructed emotion. This theory addresses some inconsistencies in experimental data generally associated with the classical view, most notably significant variation in facial muscle activity, neural activity in the brain, and other physiological responses when emotions are experienced (Barrett, 2017). Effectively, the constructed theory considers emotions to be constructs of social reality. This means they are real

[1]Note, however, that some of the core concepts of this "recent" view date back to ancient Greece (Barrett, 2017).

only in the sense that a government or money is real. There is a necessary element of human agreement which is essential so that these constructs may exist (Searle, 1995).

According to this theory, universal emotions are a myth, and the universality of the expression of emotion is effectively an illusion brought about by shared concepts associated with the context of a given situation. For example, looking only at a human face, a scream of terror may not be so readily distinguishable from a shout of triumph. Once that face is seen in context, whether it be in a dark empty parking lot or at an important sporting event, the correct recognition of the emotion indeed becomes automatic.[2] The constructed theory argues that emotions are inextricably cultural. Rather than being innate, they come to be in a person as she develops and interacts socially with her community. The concepts she learns and the events she experiences in her young life are coupled with specific reactions in her physiology, and these "feelings" ultimately become associated with shared emotion words as they exist in the local language (Barrett, 2017).

For researchers interested in emotion, the theory of constructed emotion brings significant controversy to a quest for understanding that is already wrought with disagreement. The reader should note that the present research handles concepts relating to emotion in a way that is more readily associated with the classical view. Our intent in adopting a classical approach to modelling emotion is–more than anything else–about developing a pragmatic methodology for an analysis of sentiment and emotion in public communications on Twitter. This methodology is pragmatic in that it makes efficient use of affective NLP tools for information systems. These tools are readily available for the scientific community, and many of them have been designed specifically for use with social media. Arguably, embracing a particular theory for human emotion is not necessarily essential to creating a viable model when analyzing emotion in natural language (Liu, 2015). NLP tools associated with the classical view provide a critical element, namely

[2]The reader may wish to take another look at Figure 1.1, asking how easy it would be to identify the emotions if the faces were not labelled and if the set of emotions being considered were not known beforehand. Could these emotions be interpreted differently? Consider perhaps: loathing, contempt, grief, optimism, boredom, and interest.

a system by which we may readily quantify anger, fear, and other emotions that people are expressing online. Cultural factors which trigger these emotions or which alter the manifestation and physiological experience of the emotions are effectively included in the measurements. In the scope of the present research, we do not seek to distinguish these cultural factors or address the important underlying questions such as "what is fear?" Note, however, that we would like to make clear: our approach in this research should not be viewed as our espousing of a classical view of emotion in psychology. On the contrary, we feel it would be extremely interesting to use a system such as the one we have developed to compare these two paradigms of emotion by way of varying the design of the internal model. However, this represents a rather ambitious endeavour which shall have to wait for a future research initiative.

1.1.1 Basic Emotions

According to many researchers (oriented towards the classical view), there exists a relatively small set of basic emotions which are primary, which may correspond to specific activity in the brain and possibly specific regions of the brain, and which serve as a foundation upon which humans build more complex emotions (Barrett, 2017; Liu, 2015; Ortony & Turner, 1990). Table 1.1 lists several proposed sets of basic emotions from psychology and neurology in order of their publication within the scientific community. The reader may recognize the six basic emotions according to Ekman from Figure 1.1 in the preceding section. Plutchik's eight basic emotions are listed in bold face simply to indicate that this is the set of basic emotions we utilize in the present work. We discuss Plutchik's system more in depth in Section 1.1.4.

Basic emotions essentially serve to categorize human emotions, but as Table 1.1 shows, there is a great deal of disagreement as to what these categories should be. The use of synonyms for the names of certain emotions cuts down somewhat on the level of disparity between theories but in no way eliminates it. Furthermore, even among the

Table 1.1 Systems of basic emotions.

Researcher(s)	Basic Emotions	Source
William James	fear, grief, love, rage	(James, 1884)
William McDougall	anger, disgust, elation, fear, subjection, tender emotion, wonder	(McDougall, 1926)
John B. Watson	fear, love, rage	(Watson, 1930)
Magda B. Arnold	anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	(Arnold, 1960)
Orval Hobart Mowrer	pain, pleasure	(Mowrer, 1960)
Carroll E. Izard	anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	(Izard, 1971)
Robert Plutchik	**acceptance, anger, anticipation, disgust, fear, joy, sadness, surprise**	(Plutchik, 1980)
Paul Ekman Wallace V. Friesen Phoebe Ellsworth	anger, disgust, fear, joy, sadness, surprise	(Ekman et al., 1982)
Jeffrey A. Gray	anxiety, joy, rage, terror	(Gray, 1982)
Jaak Panksepp	expectancy, fear, rage, panic	(Panksepp, 1982)
Silvan Tomkins	anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	(Tomkins, 1984)
Bernard Weiner Sandra Graham	happiness, sadness	(Weiner & Graham, 1984)
Keith Oatley P. N. Johnson-Laird	anger, anxiety, disgust, happiness, sadness	(Oatley & Johnson-Laird, 1987)
W. Gerrod Parrott	anger, fear, joy, love, sadness, surprise	(Parrott, 2001)

Adapted from (Liu, 2015; Ortony & Turner, 1990)

researchers who have proposed sets of basic emotions, there are differing opinions as to how fundamental the theory should actually be considered with respect to human biology and psychology. For example, Izard, Plutchik, Panksepp, and Tomkins consider basic emotions to be paramount, while Mowrer and Weiner and Graham are more reserved with regard to how foundational the theory may be as they present their version in their work (Ortony & Turner, 1990).

In addition to defining a set of basic emotions, some theoretical systems continue by describing secondary emotions which may serve to categorize variations within the scope of a given primary emotion. For example, the basic emotion anger in Parrott's system encompasses a series of secondary emotions: disgust, envy, exasperation, irritability, rage, and torment. These secondary emotions may in turn encompass a number of tertiary emotions for even finer levels of emotional granularity. For example, optimism is a secondary emotion of joy, while eagerness and hope are tertiary emotions encompassed by optimism (Liu, 2015; Parrott, 2001).

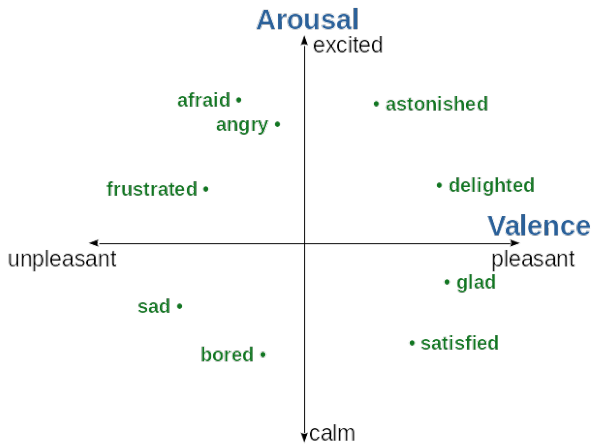
Disagreement between researchers and the substantial variations in different systems of basic emotions may certainly complicate the process of determining an effective underlying model for an emotion-based system in the field of AI. Be that as it may, a given NLP application seeking to analyze the emotion expressed in textual content is often free to ignore the question of which system of basic emotions (if any) is likely to be the more correct one. Instead, the application may simply utilize a system that covers the emotions which best relate to the subject matter of the texts to be analyzed (Liu, 2015).

1.1.2 Dimensional Models

Both in psychology and in emotion mining in NLP, models of emotion may incorporate specific dimensions which serve to identify more precisely the affective state of an individual at a given moment. Definitions using two-dimensional models typically include parameters for *valence* and *arousal* (Iglesias et al., 2017). Figure 1.2 shows an example of a two dimensional model. Note that from the viewpoint of psychology (particularly when considering the theory of constructed emotion) these dimensions represent affect or "feelings" experienced by a person and may not always correspond directly with an emotion such as those discussed in the previous section (Barrett, 2017). In emotion mining, however, the goal is generally to identify a specific emotion to associate with a given text.

Valence is a dimension which represents the range of unpleasant feelings (receiving a punch in the arm) to pleasant feelings (smelling a flower) (Barrett, 2017). The term can be associated with the word *value* in that valence refers to the psychological force by which human beings are attracted to objects they perceive as desirable or valuable and repelled by objects perceived as undesirable and devoid of value (Lewin, 1938; Mulligan & Scherer, 2012). The dimension of valence is sometimes simply called *pleasure*. It provides a basis by which to classify positive and negative emotions (Iglesias et al., 2017).

Figure 1.2 Example of a two-dimensional model of emotion.

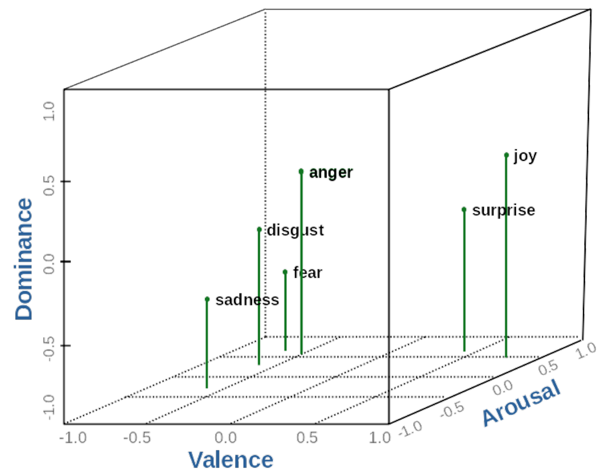


Adapted from (Russell, 1980)

Arousal is the dimension that represents the intensity of the affective feeling. The scale ranges from calm (e.g., weariness after a long day) to excited (e.g., eagerness about an upcoming event) (Barrett, 2017). In a two-dimensional model of emotion the dimension of arousal may also be called *activation* (Iglesias et al., 2017); however, some theorists argue that this name should be avoided as these two terms represent distinct phenomena in humans. Whereas *arousal* functions as described above by means of activation of the sympathetic nervous system, the term *activation* refers to the more general biological state in which an organism or a number of its subsystems are in a responsive (active) mode (Mulligan & Scherer, 2012).

There are also three-dimensional models of emotion, which include an additional axis for *dominance*, sometimes known as *control*. Dominance associates an emotion with the level of perceived control with respect to a given a situation (Bradley & Lang, 1994; Buechel & Hahn, 2017; Iglesias et al., 2017). Figure 1.3 illustrates an example of this type of model.

Figure 1.3 Example of a three-dimensional model of emotion.



Adapted from (Buechel & Hahn, 2017)

1.1.3 Sentiment Polarity

In the introduction to this chapter, we stated that sentiment analysis is often called opinion mining. The reason is that when an NLP application is processing natural language, seeking to determine the opinion a writer (or speaker) wishes to communicate on a given subject, the methodology which reveals that opinion is generally closely tied to the positive or negative sentiment expressed in the person's words. The terms *opinion* and *sentiment* can be used as direct synonyms. For example, suppose Alice says, "Climate change is real." When Bob responds, "I share that sentiment," he is referring to sharing Alice's opinion.[3] However, often the term *sentiment* refers to the feeling the words are expressing, as when Alice continues by saying, "I think climate change is just terrible." She is again giving an opinion, but now there is a negative sentiment expressed by means of the affective word *terrible*. Now the term *sentiment* is no longer completely synonymous with *opinion*, though the affect expressed continues to be directly related with the opinion conveyed (Pozzi et al., 2017).

[3]Here we mean opinion as in belief or point of view, rather than as an antonym of fact (as in knowledge based on a scientific consensus).

Sentiment in a word or a text is usually measured in terms of positive and negative polarity. When considering affect, we can think of these poles as the two ends of the dimensional axis for valence, which we saw in Figure 1.2. In fact, many researchers use a mapping from *negative* to *positive*, rather than from *unpleasant* to *pleasant* when describing affect in terms of valence (Iglesias et al., 2017; Liu, 2015).

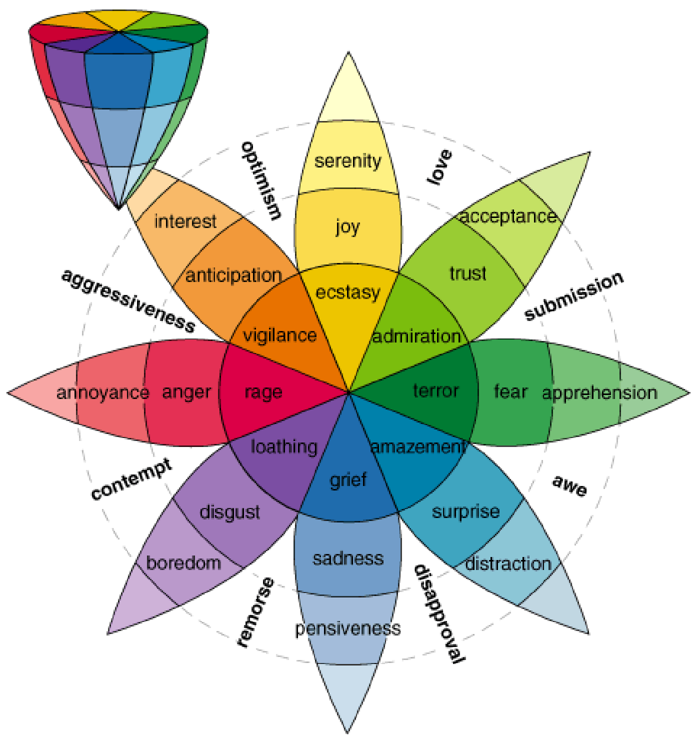
As might be expected when analyzing text, positive sentiment is embodied in words like *great*, *terrific*, and *happiness*, while examples of words expressing negative sentiment include *disagreeable*, *bad*, and *unhappy* (Liu, 2015). Tools for sentiment analysis will also typically include sentiment polarity for words that, while generally sensible, may not be immediately obvious such as: *exaggerate* (negative) and *accessible* (positive). The specific sentiment polarity values for the words listed here are taken from Bing Liu's Opinion Lexicon (Hu & Liu, 2004), which we utilize for part of the present research as described in Chapter 6. A sentiment lexicon provides mappings for a large set of words to values representing the associated sentiment. This value may be binary (i.e., *negative* or *positive*), or it may be a point on a numerical range (e.g., -1.0 to +1.0) which indicates a sentiment intensity level for the specified word. For these types of tools the actual sentiment polarity attached to any given word is decided according to the methodology used to create the tool. Such methodologies may include a statistical analysis against a corpus, human annotation, or (more recently) crowd sourcing (Bravo-Marquez et al., 2014). Also, note that in sentiment analysis it is not uncommon to utilize the concept of neutral sentiment in addition to a positive and negative polarity. Most often, neutral sentiment is associated with the absence of any expressed sentiment (Liu, 2015).

1.1.4 Plutchik's System

Of the systems of basic emotions covered in Section 1.1.1, those of Paul Ekman (Ekman et al., 1982) and of Robert Plutchik (Plutchik, 1980) are in relatively common use for the analysis of emotion in NLP (Iglesias et al., 2017). In the present work, we use

Plutchik's system by way of two emotion lexica from the National Research Council of Canada (NRC). The first is the NRC Affect Intensity Lexicon (Mohammad & Bravo-Marquez, 2017), which includes four of Plutchik's eight basic emotions (see Chapter 3). The second is the NRC Word-Emotion Association Lexicon (Mohammad & Turney, 2013), which incorporates all eight emotions as well as word mappings for positive and negative sentiment polarity (see Chapter 6).

Figure 1.4 Plutchik's eight basic emotions.



(Plutchik, 2001), reprinted per the GNU Free Documentation License 1.3

Plutchik's system is founded on the theory that emotions as we know them have developed by means of natural selection. The theory describes precursors of emotion in lower animals, a somewhat more evolved system of emotion in higher animals, and finally the complex system in humans about which so many researchers debate today. The system of emotions may be visualized as a cone as shown in Figure 1.4, where the first ring around the side at the base of the cone represents the eight basic emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. On the base itself are emotions that represent extremely intense levels of these basic emotions. Likewise, emotions on the side of the cone, closer to the vertex, are less intense versions of the basic emotions (Plutchik, 2001). For example, acceptance is now farther down the cone on the span for trust, indicating a less intense form of trust.[4]

The cone model from Figure 1.4 also serves to indicate that Plutchik's eight basic emotions are more specifically four pairs of basic emotions in opposition (i.e., fear is the opposite of anger, etc.). Finally, in the exploded view of the cone, the emotions appearing between the spans are combinations of the two basic emotions on either side. For example, the emotion awe is a mixture of fear and surprise. These combined emotions are known as *primary dyads*. The model supports further mixing as well, at any of the various levels of intensity, and so it can account for a great number of the emotions observed in human beings (Plutchik, 2000; Plutchik, 2001).

Plutchik's system is ideal for the present research in a number of ways. The organization behind the system provides an advantage with regard to both cognitive and information sciences. The system is similar to Ekman's, but the inclusion of anticipation and trust allows for the aforementioned structuring of basic emotions as pairs of opposites. This means that different emotions are quantifiable along four dimensional axes. In Section 1.1.2 we looked at dimensional models utilizing valence, arousal, and dominance. When considering these as well as other proposed dimensional models, Plutchik writes, "The identification of such axes is somewhat arbitrary and depends on the initial choices and sampling of items and on the assumptions and preferences of the investigators." He continues with the notion that the primary benefit of dimensional models, whatever the associated axes may be, is that they allow emotions to be plotted and then referenced in relation to one another (Plutchik, 2000). In our work we are considering only the basic emotions, but we still have the advantage that there is a structured relationship between them. It can be telling, for example, when we see that anger and fear are both expressed

[4]Also note that trust has taken the place of acceptance from Plutchik's original set of basic emotions (see Table 1.1).

often in tweets about climate change, while joy and sadness appear to be less prominent. We may not be working with several emotions located along the axes of anger–fear or joy–sadness; however, observing that both of the defining emotions on an axis are notably active is in and of itself an interesting finding. It may also serve to indicate what emotions (not necessarily basic ones) merit further study in future research efforts.

1.2 Analyzing Emotion on Social Media

In this work we are endeavouring to capture the affect expressed in online communications in order to reveal important information regarding the authors' beliefs and attitudes as they relate to climate change. We have created a corpus consisting of publicly-accessible microblogs, or tweets, from the online social media application Twitter. Twitter essentially makes available a vast trove of public data, potentially revealing what people think about almost any topic imaginable. According to Statistica, the service currently has well over 300 million users world-wide,[5] and it was rated fourth among popular social media sites.[6] Statistica also reports that as of January 2021 over 69 million of those users are from the United States of America, the country with more Twitter users than any other.[7] Research indicates that online invisibility and partial anonymity, among other factors, can make people feel relatively uninhibited when it comes to their online behaviour. People tend to disclose personal information more freely often expressing anger, fear and other emotions when giving their points of view (Suler, 2005). Furthermore, at least one study shows that this tendency for self-disclosure also applies to microblogs specifically such as those published on Twitter (Kaplan & Haenlein, 2011).

With sentiment analysis one is generally concerned with the polarity of natural language,

[5] https://www.statista.com/statistics/303681/twitter-users-worldwide/

[6] https://www.statista.com/statistics/248074/most-popular-us-social-networking-apps-ranked-by-audience/

[7] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

striving to determine if an author's message is positive, negative, or neutral. In our research we look at polarity, but our models also include an analysis of the emotion online users are expressing in the tweets they publish. This type of analysis is commonly called emotion mining. It could be stated more generally that in this work we focus on the affect in microblogs about climate change on social media. However, as mentioned above, the terms *sentiment*, *emotion*, and *affect* can be rather difficult to define. Researchers and practitioners in the fields of philosophy, psychology, and sociology frequently use differing definitions for these terms, often linking them with additional difficult-to-define concepts such as *mood*. Many works in the scientific arena use two or even all three terms as synonymous (Liu, 2015; Mulligan & Scherer, 2012). In the present research, we do not intend to vie for common acceptance of any given set of definitions for these terms. Nevertheless, in order to avoid ambiguity when explaining our work and our methodology, we have chosen the following usage, simple yet clear, to describe expressions in microblogs within the scope of our research:

sentiment positive or negative polarity

emotion anger, fear, joy, sadness, anticipation, surprise, disgust, or trust

affect sentiment or emotion as just described

The quest towards a better understanding of emotion is an interesting one to be sure. However, the simplification of these affective concepts allows us to concentrate our efforts on our modelling of the Six Americas and on other aspects of NLP, especially those pertaining to the inherent difficulties one encounters when analyzing natural language on social media.

1.3 The Challenge of Social Media

While social media may indeed hold a valuable trove of data that potentially represents the expressed points of view of millions of online users in the U.S. and worldwide, there is a serious challenge involved in extracting meaningful information from it. Twitter's flow of data generally takes the form of a "stream of consciousness," expressed as a series of tweets. These are short message texts with a maximum of 280 characters.[8] NLP algorithms aim to enable computers to work with human language, but traditionally they have been developed (or in many cases trained) for longer texts with a relatively proper and standardized use of language. Unfortunately, these algorithms have typically demonstrated a significant drop in effectiveness on texts from social media where formal language is notably uncommon (Farzindar & Inkpen, 2015). Online textual content is noisy from the standpoint of information processing. There is often a great deal of off-topic chatter which can make analysis difficult. This may come in the form of spam-like advertisements, or other essentially undesirable messages such as users simply saying "hello," posts with too little content to be intelligible, or texts repeating previously posted content.[9] Ongoing research aims at tackling this problem by using preprocessing methods to normalize social media texts or by developing strategies for retraining NLP algorithms to handle non-standard language (André et al., 2012; Farzindar & Inkpen, 2015).

Being a relatively new subfield of NLP, an increasing level of interest in sentiment analysis has to some extent paralleled the growing popularity of social media on the Internet. Since at least the year 2000, researchers have been looking at sentiment analysis as a tool to aid in automating the "understanding" of online content, and it has become a popular approach to accessing the massive quantity of opinion the public expresses on Twitter, Facebook, and many other online applications (Liu, 2015).

---

[8]In late 2017 Twitter increased the character limit for tweets. In their online announcement (https://blog.twitter.com/official/en_us/topics/product/2017/tweetingmadeeasier.html) Twitter reports that the results of a preliminary trial indicate that most users will likely continue sending messages near or below the original limit of 140 characters.

[9]In one study, which allowed Twitter users to give anonymous feedback, these users reported that 25% of tweets were not worth reading as they contained no novel, useful information (André et al., 2012). However, for NLP applications, tweets that are informational, yet redundant, may still potentially be useful.

Our automated analysis of these microblogs draws on various methods from the domain of NLP. Sentiment and emotion analysis are integral to the present research, but we combine these with non affect-based elements using AI methodologies centred around both machine learning and ontologies. Different methods apply to different phases of the research, and so as we present each phase in the chapters to come, we preface it with the specifics of the associated AI techniques along with the background theory and important links to cognitive science.[10]

1.4 Emotion and Climate Change

Should an emotion such as fear, or possibly anger, move human beings to act to reduce a perceived risk such as the one imposed by climate change? The dual process theory of cognition may help to provide an answer to this question. Actually, there are many dual process theories proposed in the domain of psychology, some readily interpretable in the context of a unifying theory, others perhaps less so (Evans, 2009). For our purpose when we refer to dual process theory, we are considering what is arguably a general consensus on certain properties present in several of these theories. Keeping this in mind, the answer that dual process theory provides as to whether emotion will generate action on climate change essentially presents us with a challenge–and one that is in no way trivial.

According to dual process theory, the human brain utilizes two types of reasoning processes: system 1 (S1) and system 2 (S2). S1 processes are innate and intuitive. They are much older in terms of human evolution and are shared with other species. These systems function in parallel with other activities, making spatial and temporal associations quickly. In contrast, S2 processes are acquired and require conscious control. They function in a serial manner, enabling analytical reasoning but at a much slower rate as compared to S1 processes (Epstein, 1994; Evans & Stanovich, 2013). S1 processes

[10] There is one exception to this general format: ontological modelling using description logics. As this area represents the core of the research effort for this doctoral program, we devote the entirety of Chapter 2 to the associated background theory.

are closely linked to the affective system, so they are directly associated with the anger and fear a person might feel when perceiving danger that is near or seems imminent. However, unless that person is experiencing (or has very recently experienced) a climate-related emergency, the danger represented by climate change will usually be viewed as abstract and far off in the future. This means that it is the S2 processes that will be reasoning about the danger, and this reasoning will involve little to no emotion. Furthermore, when S1 and S2 processes are in conflict (i.e., when S2 identifies a risk, but S1 does not), it is the S1 processes that will typically dominate (Weber, 2006).

At the same time, mental imagery involving emotion and positive and negative sentiment has been shown to have some influence on the perception of risk in humans and on their support for climate-related policy as well (Leiserowitz, 2006). Although S2 processes are the ones typically reasoning about climate change, emotions associated with S1 processes can motivate the S2 processes and direct them as they reason analytically (Weber, 2006). Hence, as humans decide what climate change means to them, the role of emotion can be indirect and potentially rather subtle. For this reason, research on human emotion may prove essential to our understanding of how human beings reason and respond to information about climate change (Lu & Schuldt, 2015). The present work endeavours to make a contribution towards raising this understanding.

CHAPTER II

ONTOLOGIES AND DESCRIPTION LOGICS

How might we represent concepts from the Six Americas in a computer system so that we may discern some useful information about users posting content on climate change on social media? While there are potentially many answers to that question, here we focus on one involving description logics (DL) and ontologies.

Note that we use a number of descriptive examples in this chapter, and wherever possible we take these examples from the ontological model used in the present research. In this chapter, the parts of the model we use are relatively small snippets that may be taken independently in order to demonstrate a particular construct in description logics. We cover the full model in depth in Chapter 4.

2.1 Ontologies

To understand DL, it greatly helps to first understand ontologies. This term refers to the study of *being* in philosophy or metaphysics; however, in computer science, particularly in AI, the term has a related but essentially different meaning. Defining *ontology* in this domain has historically turned out to be a rather complicated endeavour (Gómez-Pérez et al., 2004). Indeed, we acknowledge that the rather pragmatic and overly precise description we have listed in our glossary at the end of the present document may leave some theorists wanting. Perhaps the most widely accepted definition is that of Thomas R. Gruber (Gruber, 1993), refined first by (Borst, 1997) and then again by (Studer et al., 1998): "An ontology is a formal, explicit specification of a shared conceptualisation." This essentially refers to a machine readable (formal) model (conceptualization) whose concepts and constraints are defined explicitly. Furthermore, the model represents "consensual knowledge" (shared). The consensus may refer to a group of domain experts or indeed to the general population.

Given that our work mainly focuses on practical applications of theory, and our implementation is based on current research into ontological methods for AI, we shall unabashedly base our discussion on a rather pragmatic description of an ontology. Hence, we describe an ontology as referring to a formal hierarchical system of knowledge representation consisting of (1) a set of concepts, (2) individuals representing particular instantiations of those concepts, and (3) the roles or relationships between these elements. This description is closely tied to the implementation of an ontological model based on DL, more specifically on the Web Ontology Language (OWL) which we use in the present work.

The concept of an ontology as a system of classification dates back to Aristotle (Breitman et al., 2007), and indeed in this chapter we will be using the ontology we have developed for classification. The Six Americas survey uses six categories to describe people in the United States with respect to their beliefs, attitudes, and behaviours concerning climate change. We have constructed our ontology to use certain key concepts from the survey questionnaire to categorize users on Twitter who are publishing tweets relating to global warming. As mentioned in Section 0.3, our study simply incorporates two categories, green and denier, which respectively map onto two groupings of categories from the Six Americas: *alarmed/concerned* and *doubtful/dismissive*.

As we have chosen to focus on a pragmatic description of an ontology that is closely tied to the ontological model we have developed as part of the present research, it may prove more natural to continue our discussion of the details of that description in the sections that follow. In this manner we may discuss classes, concepts, individuals and their relationships as part of our presentation of description logics, OWL, and the methodology we have followed to integrate all these elements into our system.

2.2 Description Logics

The term description logic (DL) represents a family of logics[1] inspired from semantic networks and formalisms based on first-order logic to use these elements as a foundation for reasoning. Systems of logic comprising the first members of this family went by the names *terminological knowledge representation languages*, *concept languages*, or *term subsumption languages* (Gómez-Pérez et al., 2004).

Figure 2.1 Description logic constructs and associated language names.

Construct	**Syntax**	**Language**	
Concept	A		
Role name	R		
Intersection	$C \sqcap D$	$\mathcal{FL}_0$	
Value restriction	$\forall R.C$		
Top or universal concept	$\top$		
Bottom	$\bot$	$\mathcal{FL}^-$	
Limited existential quantification	$\exists R$		
Atomic negation	$\neg A$	$\mathcal{AL}$	
Negation	$\neg C$	$\mathcal{C}$	$\mathcal{S}$
Union	$C \sqcup D$	$\mathcal{U}$	$(\mathcal{ALC}_{R+})$
Existential restriction	$\exists R.C$	$\mathcal{E}$	
Number restriction	$(\geq nR)$; $(\leq nR)$	$\mathcal{N}$	
Qualified number restriction	$(\geq nR.C)$; $(\leq nR.C)$	$\mathcal{Q}$	
Nominals	$\{a_1 \ldots a_n\}$	$\mathcal{O}$	
Role hierarchy	$R \sqsubseteq S$	$\mathcal{H}$	
Inverse role	R^-	$\mathcal{I}$	

Legend:
- A: atomic concept
- C and D: general concept definitions
- R: atomic role
- S: general role definition

Adapted from (Gómez-Pérez et al., 2004)

[1]As such, the plural (description logics) is also commonly used and is arguably more correct. For this research we are primarily focusing on one system of logic ($\mathcal{SROIQ}^{(\mathcal{D})}$) from this family.

At its core DL is about knowledge represented as a set of concepts which constitute the domain being modelled. These concepts describe sets of individuals which correspond to specific instances of the various concepts in the model (Szeredi et al., 2014). For example, the ontology in this work includes the concept *OnlineUser*, and one instantiation of this concept is *IPCC_CH*, which is the online Twitter account for the IPCC.[2] Figure 2.1 presents common languages in the DL family. We see that *Concept* is the first construct listed in the figure and is one of four constructs required for the frame-based,[3] structural DL language $\mathcal{FL}_0$ (Gómez-Pérez et al., 2004). Note that this construct refers to atomic concepts (denoted by A in the figure). Atomic concepts are simply declared in an ontology and may then be used to construct composite concepts (denoted by C and D). The next construct, *Role name*, allows for binary relationships which describe and distinguish individuals by defining relationships between them. Like concepts, roles in an ontology may be atomic (denoted by R) or composite (denoted by S) (Szeredi et al., 2014). An example role from our ontology is *publishes* which forms a relationship between an instance of *OnlineUser* and an instance of *InformationObject*. This construct provides a representation of the fact that an online user (e.g., *IPCC_CH*) published an information object. For our study this will be a specific microblog, which is also an instantiated individual in the ontology.

There are two more constructs in $\mathcal{FL}_0$. *Intersection* in DL directly corresponds to the same operation on sets in mathematics.[4] The intersection of two concepts in DL implicitly creates a new concept whose associated individuals are members of both the first concept and the second concept. For example, our ontological model defines the

---

[2]The "CH" is Switzerland's Internet domain county code, as in https://ipcc.ch.

[3]Frame-based implementations group properties and constraints associated with a concept using slots in a class which represents that concept (Horrocks et al., 2003).

[4]Intersection of concept expressions in DL uses the square variant of the mathematical symbol for intersection ($\cap$) in set theory.

concept *SurveyKeyword*[5] as well as the concept *AngerToken*. When we declare:

$$\textit{SurveyKeyword} \quad \sqcap \quad \textit{AngerToken},$$

we are specifying individuals from the ontology which are both survey keywords and tokens (words in a text) that express anger. Note that it is no accidental turn of phrase that we are using the word "and" in our explanation here. Intersection in DL represents a logical *AND* operation. It is equivalent to refer to individuals resulting from the intersection of two concepts and to individuals who are members of a concept which is the conjunction of these two concepts. Finally, the last construct in $\mathcal{FL}_0$ is *value restriction*, which imposes a constraint that all individuals having the relationship R must have that relationship with individuals which are instances of concept C (Szeredi et al., 2014). For example, in our model, we define the role *denotesAffect* such that:

$$\forall\ \textit{denotesAffect.Affect},$$

where *Affect* is the concept representing sentiment polarity or emotion. Note that although it seems intuitive to a human being (and perhaps even silly) that it is affect that will be denoted when something denotes affect, the names of the properties and the concepts are simply labels. Their names give no implied meaning, and we could just as well have called them R and C or anything else. It is the formal, structured definitions we create in the DL language which allow us to represent knowledge such as that needed to model words which express emotion.

DL constructs provide the means for the "formal, explicit specification" of an ontology as defined by (Studer et al., 1998). This specification has two distinct components: the terminological box (TBox) and the assertion box (ABox). The TBox comprises intensional knowledge, formed by definitions of the concepts being modelled as well as

[5]This refers to a keyword from a research-based survey such as the Six Americas.

the relationships describing how one concept is associated with another.[6] This is the part of the ontology where meaning is assigned. In contrast, the ABox lists the extensional knowledge, which is composed of the assertions for individuals in the ontology. These are members of one or more concepts defined in the TBox. The ABox also contains assertions of the roles between two individuals as these roles are defined in the TBox. (Gómez-Pérez et al., 2004; Szeredi et al., 2014).

Continuing our traversal of the DL constructs and languages in Figure 2.1, we see that $\mathcal{FL}^-$ is syntactically a superset of $\mathcal{FL}_0$. We should point out that the present research does not make explicit use of these frame-based languages. Nevertheless, as we have chosen a pragmatic approach for this presentation of ontologies and description logics, working our way from less expressive DL languages to more powerful ones helps to provide a sensible road map to the DL we use in our model, OWL 2 DL. Notably, $\mathcal{FL}^-$ includes the *top* and *bottom* concepts. The top ($\top$) or universal concept[7] subsumes (encompasses) all concepts modelled in an ontology. Naturally, the dual of the top concept is the bottom concept ($\bot$). It is the DL equivalent of the empty set. By definition no individuals in the ontology are members of the bottom concept, and all individuals are members of the top concept. Finally, the *limited existential qualification* construct[8] may be more explicitly expressed as $\exists R.\top$. Individuals that are members of concepts constrained by this construct must have at least one assertion of the relationship R; however, the member with whom it is in relation may represent any concept in the

[6] Many ontological engineers will refer to a role box (RBox), which specifically defines the intensional knowledge of the relationships between concepts. Others consider these role axioms simply to be part of the TBox (Szeredi et al., 2014) as we have presented it here.

[7] The universal concept should be distinguished from similarly named ontological terminology, *universals*, which are the types representing the concepts in the implementation of an ontology (Arp et al., 2015). The *universal concept* is essentially a set containing all the universals for a given ontological model.

[8] This logical construct is also known as the *unqualified existential qualification* (Baader et al., 2017).

ontology[9] (Breitman et al., 2007; Szeredi et al., 2014).

Adding *atomic negation* to the constructs we have covered so far gives us $\mathcal{AL}$, the most basic of a family of DL languages known as attributive languages. Atomic negation applies only to atomic concepts (i.e., not to concepts constructed from other concepts). This DL construct functions as the complement operator in mathematical sets. Therefore, the DL expression $\neg A$ refers to the set of all individuals in the ontology which are not members of the concept A (Breitman et al., 2007).

$\mathcal{AL}$ is still a relatively simple DL language. Roles must be atomic; however, composite concepts are an important expressive component of $\mathcal{AL}$. In addition to the intersection (logical *AND*) construct mentioned above, there are two forms of concept axioms: *subsumption* and *equivalence*. The first of these, subsumption, is formally known in DL as a *concept inclusion axiom* (Szeredi et al., 2014) or a *general concept inclusion* (Baader et al., 2009). Examples in the ontological model for the present work are:

$$\begin{array}{lcl} \mathit{Token} & \sqsubseteq & \mathit{InformationObject} \\ \mathit{Text} & \sqsubseteq & \mathit{InformationObject} \\ \mathit{Tweet} & \sqsubseteq & \mathit{Text} \end{array}$$

Here we are essentially creating a hierarchy, defining the concepts *Token* and *Text* as being subsumed by the concept *InformationObject*. Looking at concepts as sets of the individuals, the subsumed concept is a subset of the concept that subsumes it. That is, all of the defined tokens and texts in the ontology are also information objects.[10] For this reason, the DL symbol for concept inclusion ($\sqsubseteq$) is a square version of the mathematical symbol for a subset. By these definitions a reasoning process will infer that any individual declared to be a tweet is also a text as well as an information object,

[9] Compare this construct to full existential restriction, which we will cover as part of the DL language $\mathcal{S}$ and which we use in the ontology for the present work.

[10] In informatics subsumption is often known as an "is-a" relationship. A tweet is a text, and a text is an information object.

and any defined roles that are associated with the latter two concepts will also apply to tweets. Another way to define composite concepts is using the equivalence axiom:

$$\mathit{OriginalTweet} \quad \equiv \quad \mathit{Tweet} \sqcap \neg \mathit{Retweet}$$

Here we define the concept *OriginalTweet* as equivalent to the intersection of the set of individuals belonging to the concept *Tweet* and the complement of the set of individuals belonging to the concept *Retweet*. This means an original tweet is a thing that is a tweet and is not a retweet. Finally, note that equivalence ($C \equiv D$) may be expressed as a pair of concept inclusion axioms ($C \sqsubseteq D$ and $D \sqsubseteq C$) (Szeredi et al., 2014).

Returning to our traversal of Figure 2.1, the next logical construct is *negation*. This construct refers to general negation of concepts, be they atomic or composite, rather than the less expressive atomic negation supported by $\mathcal{AL}$. Note that in the previous example when we defined *OriginalTweet*, the negation of *Retweet* in that definition requires only the atomic negation construct. (*Retweet* is an atomic concept.) However, if we should need to express its complement, $\neg \mathit{OriginalTweet}$, we would then require the general negation construct since *OriginalTweet* is a composite concept. (Its definition involves the conjunction of two concepts.) Negating a concept effectively refers to everything modelled which is not that concept. In the example above, $\neg \mathit{Retweet}$ encompasses all individuals in the ontology which are not members of *Retweet*. Note that this indeed includes everything else (e.g., the punned individual representing *Anger*), which is why the definition specifies the intersection of these individuals with those belonging to the concept *Tweet*. Likewise, the expression $\neg \mathit{OriginalTweet}$ covers all individuals in the ontology which are not members of the composite concept *OriginalTweet*.

Beginning with general negation, each of the remaining constructs has an associated letter which may be appended onto a given DL language acronym to indicate the constructs it comprises, such as $\mathcal{ALC}$ or $\mathcal{ALCUE}$ (Breitman et al., 2007; Gómez-Pérez et al., 2004). Negation uses $\mathcal{C}$ for *complement*. Next is *union* of concepts ($\mathcal{U}$), which again corresponds with the operation on mathematical sets, using a square version ($\sqcup$)

of the symbol for the union operator. In DL languages, union parallels intersection in that the latter functions as a logical *AND* (conjunction), while the former functions as a logical *OR* (disjunction). The final logical construct in this section of Figure 2.1 is *existential restriction* ($\mathcal{E}$). This construct is similar to limited existential quantification, mentioned above as part of $\mathcal{FL}^-$; however, rather than simply asserting that a role exists, the existential restriction construct specifies the range for the binary relation the role represents.[11] An example definition from our ontology which employs existential restriction is:

$$AngerToken \quad \equiv \quad Token \sqcap \exists denotesAffect.Anger$$

Here we define the concept of an *AngerToken*: a token that denotes anger (as a type of affect). This is the concept we use to represent words in a text that express the emotion anger.

Considering the previous three constructs as additions to the base DL language $\mathcal{AL}$, it can be shown that the DL languages $\mathcal{ALC}$, $\mathcal{ALCE}$, $\mathcal{ALCU}$, and $\mathcal{ALCUE}$ are equivalent in terms of logical expressivity. For this reason, the language with the shortest name, $\mathcal{ALC}$, is generally used to stand for all of them (Szeredi et al., 2014). $\mathcal{ALC}$ provides all the logical constructs needed for the important DL language $\mathcal{S}$, save one: transitive roles. This type of role effectively chains a given role across individuals. A common example of a transitive role is the "part of" relation. Suppose an ontology asserts that (1) an eye is part of a face; (2) a face is part of a head; and (3) a head is part of a body. If the "part of" role is transitive, then a reasoning process for the ontology can conclude that an eye is part of a body. DL languages commonly use the syntax $Trans(R)$ to declare an axiom stating that a role R is transitive. In our ontology for the present work, we use a small number of transitive roles such as:

[11] The *range* refers to the second individual (the object) in the binary relationship. The construct asserts the concept to which that individual belongs. Note that in cases where the range is the top concept ($\top$), existential restriction ($\exists R.C$) becomes logically equivalent to limited existential quantification ($\exists R$).

$$Trans(dependsOn)$$

This transitive role represents a syntactic dependency between two tokens in a tweet. These dependencies can chain over several tokens ($word_I$ depends on $word_J$, which in turn depends on $word_K$). Since our *dependsOn* role is transitive, our reasoner is able to determine that $word_I$ depends on $word_K$. The language $\mathcal{ALC}$ with transitive roles is called $\mathcal{ALC}_{R+}$, a DL language more commonly known as $\mathcal{S}$ (Gómez-Pérez et al., 2004; Szeredi et al., 2014).

Returning now to Figure 2.1, we see one last section with several logical constructs, each with an identifying letter. These letters are generally added onto the letter $\mathcal{S}$ as needed to indicate extensions to that DL language. The first two of these final constructs, *number restriction* ($\mathcal{N}$) and *qualified number restriction* ($\mathcal{Q}$), respectively parallel the constructs limited existential quantification ($\exists R$) and existential restriction ($\exists R.C$). These number restriction constructs allow constraints specifying "at least" or minimum ($\geq nR.C$) as well as "at most" or maximum ($\leq nR.C$) with respect to the number of role assertions for an individual. Again, for non-qualified number restrictions ($\mathcal{N}$) the range is the top concept ($\top$), and the axiom simply declares that a minimum or maximum number of assertions of the specified role exist for an individual with no constraint on the range (the object) in the those role assertions. Although the final ontology for the present research does not assert any number restrictions, an early-stage ontological model, based on work for the big players project (see Chapter 3), does include an example of a qualified number restriction:

$$Influencer \equiv Tweeter \sqcap (\geq 3\ isRetweetedIn.Retweet \sqcup \geq 3\ isMentionedIn.Tweet)$$

The *Influencer* concept represented potential influential users on Twitter. They were defined as a *Tweeter* who has at least three tweets that have been retweeted or who is mentioned in at least three tweets.

Continuing with these last logical constructs from Figure 2.1, we have *nominals* ($\mathcal{O}$), which assert that exactly one individual belongs to a given concept.[12] The construct *role hierarchy* ($\mathcal{H}$) allows the definition of composite roles using role inclusion axioms in a manner parallel to the way we used general concept inclusion to create a hierarchy of concepts (Baader et al., 2009; Szeredi et al., 2014). Here is an example from the ontology in the present work:

$$directlyDependsOn \quad \sqsubseteq \quad dependsOn$$

The role *directlyDependsOn* is a more specialized version of *dependsOn*. In our model it is used with texts to represent a relationship of direct dependency between two tokens in a text. In our discussion of transitive roles above, we describe a chain of syntactically dependent tokens. When processing actual tweets, we make assertions in the ABox of the form: $word_I$ directly depends on $word_J$, and $word_J$ directly depends on $word_K$. In contrast with the previous case, it is *not* true that $word_I$ directly depends on $word_K$. (The role *directlyDependsOn* is not transitive like *dependsOn*.) As the role *dependsOn* is higher in the hierarchy, a reasoner will infer that one thing depends on another thing when there is an assertion that the first thing directly depends on the second. However, it will not infer a relationship of direct dependency given an assertion of (possibly indirect) dependency.

The final logical construct in Figure 2.1 is the *inverse role* ($\mathcal{I}$). This construct allows the definition of a role which effectively represents the reversed sense of another role and swaps the domain (subject) and range (object) with respect to the original role (Antoniou et al., 2012; Szeredi et al., 2014). This may be best illustrated with an example from our ontology:

$$\begin{aligned} hasDependent \quad &\equiv \quad dependsOn^{-} \\ hasDirectDependent \quad &\equiv \quad directlyDependsOn^{-} \end{aligned}$$

[12]The concepts for affect in our ontological model (*Anger*, *Fear*, etc.) could have been declared as nominals. Instead we use an OWL 2 technique called *punning*. We discuss the differences in Section 2.3.

The upper index $^-$ indicates the inverse role construct. Just above, we looked at the *dependsOn* and *directlyDependsOn* roles. If the ontology has an assertion that $word_I$ directly depends on $word_J$, a reasoning process will infer the inverse relation: $word_J$ has the dependent $word_I$.

In our presentation of description logics in this section, the reader may have gotten the impression that DL languages and ontologies are inextricably intertwined. This need not be the case. What makes it appear as such is simply our pragmatic approach to presenting the subject in a way that readily allows it to be linked to the actual research model in the present work. DL languages do not necessarily need to be coupled with ontologies; however, this coupling provides a solution that is at once elegant and powerful for problems requiring an effective strategy for knowledge representation. This should become apparent in the following section when we look at OWL as a DL-based implementation for ontological modelling.

2.3 Web Ontology Language

Given its core emphasis on knowledge representation through concepts, DL provides a natural methodology on which to base ontologies. The Web Ontology Language (OWL) is actually a family of ontology languages, recommended and maintained by the World Wide Web Consortium (W3C) for computing applications which make use of ontologies, incorporate intelligent inference using DL, and potentially interface with the Semantic Web.

What is the Semantic Web? It intended to be "a Web of actionable information." The vision is to have a structured, machine readable, world-wide network of interlinked, symbol-driven knowledge bases for public use in a manner that parallels the tens of billions of online pages that are currently available but primarily intended for human readers. This goal was originally and officially put in place by Tim Berners-Lee in 1994 at the first World Wide Web Conference. We are still a long way from accomplishing it (Shadbolt et al., 2006). It is also possible that we will never accomplish it. The symbols driving the information in the Semantic Web are interconnected, essentially forming a shared, solid base on which to build (often complex) syntactic structures defining formalized concepts and the relationships between them as we discussed above. This interlinked base, though massive, is nevertheless floating with no grounding to non-symbolic representations, based in biological sensory and motor systems. (Harnad, 1990) argues that a formal symbolic, "top-down" approach can never provide a machine with the true sense of meaning roughly corresponding to that which humans achieve with our sensory-based, "bottom-up," biological method of understanding. Harnad declares in a blog post that the Semantic Web should in truth be called "the syntactic web."[13]

In this research we are not attempting to tackle this fundamental question of AI, which asks if a machine can truly understand the meaning behind symbols as it manipulates them in a seemingly intelligent way. Furthermore, even as the idea behind the Semantic Web seems to fit a globally-oriented version of the definition of ontology given by (Studer et al., 1998): "a formal, explicit specification of a shared conceptualisation," we are working to use the ontological techniques integrated into the Semantic Web for a simpler (perhaps more modest) goal. This goal is to extract useful information of people's attitudes towards climate change from their online posts. Certainly, efforts to have our system move beyond a syntactical to a deeper semantical "understanding" of the content of the microblogs of online users would add an extremely interesting facet to this research (even if we were ultimately unsuccessful). However, this path goes far beyond the scope of this project with respect to its incorporation in the present doctoral program and for the moment must remain an idea for this doctoral candidate to ponder.

As should be obvious from the name, OWL 2 is not the first Web Ontology Language. A brief history of the most important languages leading up to OWL 2 are (Breitman et al., 2007; Gómez-Pérez et al., 2004; Horrocks et al., 2003):

[13]https://generic.wordpress.soton.ac.uk/skywritings/2018/11/09/the-syntactic-web/

- SHOE (Simple HTML Ontology Extension) was developed in 1996 as an extension to the HTML content in online documents.
- DAML (DARPA[14] Agent Markup Language) was proposed in 1999 as a common language to support interoperability between systems incorporating ontologies. The language was released in 2000 as DAML-ONT.
- OIL (Ontology Inference Layer) was developed by the On-To-Knowledge project in Europe at approximately the same time as DAML. It provided a frame-based implementation to represent formal specifications in the $\mathcal{SHIQ}$ DL language.
- DAML+OIL was released in 2001 after the groups responsible for DAML and OIL coordinated their efforts and formed the US/EU ad hoc Agent Markup Language Committee. The language featured a syntax based on DL axioms rather than on frames.
- OWL was published in 2004 by the Web Ontology Group of the W3C. This working group used DAML+OIL as a base to begin constructing the language. OWL actually refers to three styles or sublanguages. In order of their level of capability for semantic expression, they are OWL Lite (equivalent to $\mathcal{SHIF}^{(D)}$), OWL DL (equivalent to $\mathcal{SHOIN}^{(D)}$), and OWL Full.[15] The D in the names of the equivalent DL languages indicates support for relationships with XML Schema data types (e.g., integers, strings, dates, etc.).

The present research makes use of OWL 2, which the W3C released in 2009.[16] OWL 2 is an implementation of the $\mathcal{SROIQ}^{(\mathcal{D})}$ DL language for the Semantic Web (Szeredi et al., 2014). Like the original OWL it has two sublanguages: OWL 2 DL and OWL 2

---

[14]The Defense Advanced Research Projects Agency in the United States.

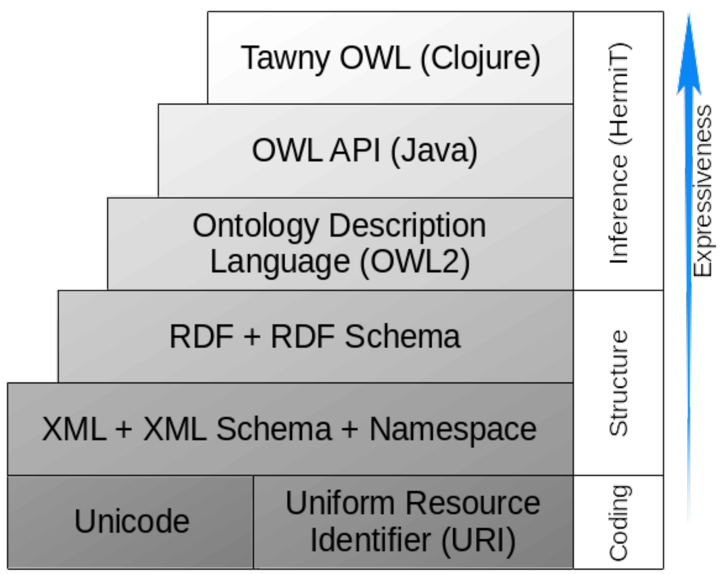
[15]OWL Full provides full compatibility with RDF and RDF Schema (see below), but it is undecidable and not well supported by way of computing applications and libraries (Hitzler et al., 2010).

[16]The second edition (2012) is the current specification as of this writing. It can be accessed at https://www.w3.org/TR/owl2-syntax/

Full.[17] Instead of an additional "Lite" sublanguage, OWL 2 DL has profiles based on OWL Lite, which increase computational performance but limit the expressivity of the DL language. Briefly, these profiles are (Antoniou et al., 2012; Horrocks, 2012):

- OWL 2 EL extends the $\mathcal{EL}$ (existential language) DL language to support ontologies with a very large TBox (many concepts and relationships) and a relatively small ABox (few individuals).
- OWL 2 QL is designed for query answering on ontologies with a relatively small and uncomplicated TBox but many assertions for individuals in the ABox.[18]
- OWL 2 RL is the subset of OWL 2 DL that may be implemented using rule-based languages. One benefit of this is that the rule-based reasoning processes can often be run in parallel.

Figure 2.2 Integration of Semantic Web architecture with Say-Sila system.



[17] Like OWL Full, OWL 2 Full is undecidable due to the fact that it provides complete compatibility with the lower-level base web technologies it incorporates (i.e., RDF and RDF Schema).

[18] As this is the case for the Twitter data modelled in the present work, we experimented with a version of our ontology which conformed to the OWL 2 QL profile. Ultimately, however, we needed to model disjunctive information (see Sections 5.1.4 and 5.1.5), which is not supported in this profile.

Figure 2.2 presents the canonical layered architecture for applications interfacing with the Semantic Web (Breitman et al., 2007)[19] along with the program libraries (top two layers) used to integrate this architecture with the larger *Say Sĩla* application, created as part of the present research. The most basic layer defines Unicode[20] as the standard character encoding for documents. It also declares the Uniform Resource Identifier (URI) as the method by which symbols are "grounded," allowing the interlinking of both concepts and individuals in different ontologies across the Semantic Web (Shadbolt et al., 2006). Note that ontologies may make use of an International Resource Identifier (IRI) instead of a URI. The two kinds of resource identifiers serve the same purpose, the difference being that by definition an IRI allows the use of non-Latin characters in the identifier (Hitzler et al., 2010). The use of URI references began with SHOE and continued with later ontology languages for the Semantic Web up to and including OWL 2. An example of a URI from the say-sila ontology[21] developed as part of this research project is:

http://www.dendrown.net/uqam/say-sila.owl#Sadness

which serves as the reference for the concept of sadness. The part of the URI up to and including the hash character (#) is known as the namespace. It is often a network location which identifies a collection of resources (Pan, 2009), in this case all those associated with the say-sila ontology. To the right of the hash character, we have the URI fragment, which may be used as a shorthand method to refer to the concept of sadness within the ontology itself. This provides a natural convention, given that the namespace also serves as the URI identifying the ontology. Also, note that XML and OWL syntax allow for the definition of a short prefix to use as a namespace identifier in place of the full URI (Hitzler et al., 2010).

---

[19] This architecture is often called "the 'layer cake' of the Semantic Web" (Antoniou et al., 2012).

[20] https://www.unicode.org/versions/Unicode14.0.0/

[21] We use the standard font "say-sila" to denote the ontology and the stylized version "*Say Sĩla*" for the application which incorporates it.

Above Unicode and URIs, the two building blocks providing the low-level coding elements, Figure 2.2 shows the eXtensible Markup Language (XML) as the first structural layer for OWL 2. Whereas SHOE was based on extensions to HTML, the DL languages OIL, DAML+OIL, and the original OWL were all based on XML, as is the current W3C standard for ontology languages, OWL 2 (Horrocks et al., 2003). Both HTML and XML are subsets of the Standard Generalized Markup Language (SGML). However in contrast to HTML, which is a markup language for handling the visual formatting of online documents, XML was created to be general-purpose. It is intended to serve as a base on which to build markup languages with very specific purposes, in this case DL-based ontology languages such as OWL 2 DL (Szeredi et al., 2014). In 1998 the W3C recommended XML as "the universal format for structured documents and data on the Web" (Arp et al., 2015).[22] Figure 2.3 presents a snippet of the XML defining the say-sila ontology. Notably, we see the *xml:base* attribute, which declares the URI for the XML document and hence the ontology. Just above that, we see the *xmlns* attribute which sets a default namespace for any non-qualified elements[23] in the XML document (Szeredi et al., 2014). The value is the W3C locator reference for the "OWL 2 Schema vocabulary."

There are quite a number of these non-qualified elements just in the short snippet given in Figure 2.3: *Ontology*, *Class*, and *SubClassOf* are among the most notable. The default OWL 2 DL namespace, using the prefix "owl," declares that *owl:Ontology* is "[t]he class of ontologies,"[24] which leads us to ask, ontologically speaking, what is a class? For the purposes of our ongoing pragmatic description, we might say that a class is the implementation of a concept in an ontology. This simplistic definition may help

[22]The quoted text is that of the XML Working Group of the W3C in the Web archive: https://www.w3.org/MarkUp/.

[23]Qualified elements have a prefix to identify the namespace to which they are associated (e.g., the "*xml*" in *xml:base*).

[24]http://www.w3.org/2002/07/owl#

Figure 2.3 The concept of Sadness in XML for the say-sila ontology.

```
<?xml version="1.0"?>
<Ontology xmlns="http://www.w3.org/2002/07/owl#"
     xml:base="http://www.dendrown.net/uqam/say-sila.owl#"
    ...
    <Declaration>
        <Class IRI="Sadness"/>
    </Declaration>
    <SubClassOf>
        <Class IRI="Sadness"/>
        <Class IRI="Affect"/>
    </SubClassOf>
    <AnnotationAssertion>
        <AnnotationProperty abbreviatedIRI="rdfs:comment"/>
        <IRI>Sadness</IRI>
        <Literal xml:lang="en">A concept which expresses the class of human affect
                               generally known as sadness according to the system
                               of base emotions by Plutchik.
        </Literal>
    </AnnotationAssertion>
    ...
    </Ontology>
```

us to understand the general idea of a class, but we immediately begin to see its limits. Perhaps even more circular than the *owl:Ontology* class is the *owl:Class* class, whose description is "[t]he class of OWL classes." The confusion can be explained by moving to the next layer up in the Semantic Web architecture shown in Figure 2.2: Resource Description Framework (RDF)[25] and RDF Schema.[26] The W3C first introduced RDF in 1997 and began officially recommending it in 1999 (Shadbolt et al., 2006). RDF essentially allows for the definition of four sets: (1) literals (2) resources, (3) properties, and (4) statements. The literals are simple character sequences. The resources represent the elements being modelled and are identified by their URIs. The properties define triples of the form $\langle subject,\ property,\ object \rangle$, where:

subject is a reference to a URI, identifying a Web resource.

property names a binary relation between the subject and the object.

object is a literal value or another URI reference.

---

[25]http://www.w3.org/1999/02/22-rdf-syntax-ns#

[26]http://www.w3.org/2000/01/rdf-schema#

The statements are a set of asserted triples using elements from the sets of literals, resources, and properties associated with a given RDF database. Such a database is commonly known as a triplestore (Arp et al., 2015; Szeredi et al., 2014). The triples are also important when building a hierarchy of concepts or roles. RDF allows the declaration of new types with a property named *rdf:type*. A class, as mentioned above, is the implementation of an ontological concept, and there is a type that denotes a class (*rdfs:Class*) defined in RDF Schema. This layer of the Semantic Web architecture became an official W3C recommendation in 2004 (Shadbolt et al., 2006). RDF Schema also designates the official property for subsumption as *rdfs:subClassOf* (Breitman et al., 2007). These elements not only allow for the creation of concepts (classes), binary relations (properties), and individuals (reified statements), all of which are needed to construct a given ontology, but they also provide the implementation of a language for the construction of ontologies in general. In other words, the ontological framework itself is using the same structural elements as the model in any given application. This is the reason why RDF Schema (as well as OWL 2) has elements such as a class of classes and a property declaring one class to be the subclass of another.

As indicated in Figure 2.2, OWL 2 builds on RDF and RDF Schema, using XML for its base language syntactical structure. Although RDF and RDF Schema serve as a relatively powerful base for ontological modelling languages, with RDF supporting binary relations and RDF Schema providing hierarchical class and role definitions, they are too limited to provide reasonable coverage of the DL constructs (Antoniou & Harmelen, 2009) as described in Section 2.2. Using RDF Schema, we could model the relationship of concept subsumption from the snippet shown in Figure 2.3:

$$\mathit{Sadness} \quad \sqsubseteq \quad \mathit{Affect}$$

However, in order to define *sila:SadnessToken*[27] we need the expressivity OWL 2 provides

[27] The say-sila ontology uses the prefix *sila* to refer to its own namespace.

to declare class equivalency (*owl:equivalentClass*) and use the DL constructs for existential quantification (*owl:ObjectSomeValuesFrom*) and logical *AND* (*owl:intersectionOf*) (Arp et al., 2015; Hitzler et al., 2010):

$$SadnessToken \equiv Token \sqcap \exists denotesAffect.Sadness$$

Figure 2.4 Affect as disjoint concepts in the say-sila ontology.

```
...
<DisjointClasses>
    <Class IRI="Anger"/>
    <Class IRI="Anticipation"/>
    <Class IRI="Disgust"/>
    <Class IRI="Fear"/>
    <Class IRI="Joy"/>
    <Class IRI="Negative"/>
    <Class IRI="Positive"/>
    <Class IRI="Sadness"/>
    <Class IRI="Surprise"/>
    <Class IRI="Trust"/>
</DisjointClasses>
...
```

OWL 2 also allows us to declare sets of classes as disjoint as shown in Figure 2.4. This means that individuals belonging to one class may not belong to any other class in the set (Antoniou & Harmelen, 2009). In order to model words which express an emotion or positive or negative sentiment, we use the property *sila:denotesAffect*, whose domain (subject) is a *Token* and whose range (object) belongs to the concept *Affect*, which is the parent class for these ten disjoint classes. An assertion declaring that a token expressed some form of affect must include an individual for the subject and the object of the *denotesAffect* property. For the token, this is straightforward enough as each word in each tweet is a named individual such as "t1268809023350063106-10" (the 10^{th} token of the tweet with the Twitter identifier 1268809023350063106). For the expressed affect, however, we must use a modelling trick because the property assertion must reference an individual, not a concept like sadness. Here we could use the DL nominals construct (*owl:oneOf* with an enumeration of only one individual). Essentially, what we need for the model is a single individual (for use in the object property) that serves only to represents the concept (used in the declaration of disjointness as well as the concept hierarchy). OWL 2 supports what is known as "punning" whereby we use the same name (e.g. "Sadness") for both the class and the individual. The context of any given assertion in the ontology makes it clear which one is used (Antoniou et al., 2012).[28] Figure 2.5 displays the snippet from the say-sila ontology involving the punning of *Sadness*. The model includes similar expressions in the XML for the other nine affective elements.

Figure 2.5 Use of punning for emotions in the say-sila ontology.

```
...
<Declaration>
    <Class IRI="Sadness"/>
</Declaration>
<Declaration>
    <NamedIndividual IRI="Sadness"/>
</Declaration>
<ClassAssertion>
    <Class IRI="Sadness"/>
    <NamedIndividual IRI="Sadness"/>
</ClassAssertion>
...
```

OWL 2 extends RDF Schema in many other ways in order to implement the $\mathcal{SROIQ}^{(\mathcal{D})}$ DL language. It includes constructs for universal quantification (*owl:allValuesFrom*); number restrictions (*owl:minCardinality* and *owl:maxCardinality*); logical *OR* (*owl:unionOf*); and logical *NOT* (*owl:complementOf*) to name a few. It also extends roles to include data properties, which specify XML Schema data types as their range; inverse roles (*owl:InverseObjectProperties*); and transitive roles (*owl:TransitiveObjectProperty*) among numerous other extensions (Antoniou & Harmelen, 2009; Arp et al., 2015).

The attentive reader will have noticed that while OWL 2 is an implementation of the $\mathcal{SROIQ}^{(\mathcal{D})}$ DL language, Table 2.1 does not include $\mathcal{R}$ among the logical constructs it presents. The $\mathcal{R}$ represents *role inclusion axioms* (Horrocks & Sattler, 2004), which create an extension to the role hierarchy construct ($\mathcal{H}$) and allow several additional DL constructs to be expressed (Hitzler et al., 2010; Horrocks et al., 2006; Szeredi et al., 2014):

[28]OWL 2 also supports using the same name for a property, but we did not need this third use for any names in our model.

Universal role: defines the role U which subsumes all other defined roles in a manner which parallels how the top concept ($\top$) subsumes all defined concepts.

Disjoint roles: $Dis(R, S)$ declares that only one of the roles R or S may be asserted between two individuals. In other words, both of the following cannot be true: $R(x, y)$ and $S(x, y)$.

Reflexive roles: $Ref(R)$ declares that $R(x, x)$ is true for all individuals in the domain defined for R.

Local reflexivity: allows concepts to be defined using the form $\exists R.Self$, which includes any individual x for which $R(x, x)$ holds true.

Irreflexive roles: $Irr(R)$ declares a relationship R which cannot be between an individual and itself. That is, an assertion of the form $R(x, x)$ would make the ontology inconsistent.

Antisymmetric role assertions: $Asy(R)$ declares that only one subject–object ordering is valid with respect to the role R. Therefore, both of the following cannot be true: $R(x, y)$ and $R(y, x)$.

Negated roles: $(x, y){:}\neg R$ asserts that it is *not* true that x and y are in relationship such that $R(x, y)$.

Composite roles: $R_1 \circ R_2$ chains roles to declare a relation $S(x, y)$, given the existence of an individual z, such that $R_1(x, z) \sqcap R_2(z, y)$.[29]

Finally, the top two layers in the architecture from Figure 2.2 depend specifically on the system that is making use of Semantic Web technologies. While we have indicated the technologies used in the *Say Sĩla* application in the figure, we shall wait to discuss these layers until Chapter 4 where we go into the details of the ontological model used in the present research.

---

[29]Note that $\mathcal{SROIQ}$ imposes constraints on roles used in composition to ensure that the TBox (or RBox) remains acyclic.

2.4 Building Blocks for a Socio-psychological Model

Part of our contribution with respect to this doctoral research is to develop a model capable of handling a rather difficult problem. Essentially, this model is tasked with describing actual online communications, the syntactic structure of those communications, the users who posted them, and finally their semantic links to content in the Six Americas survey that may be indicative of a stance with respect to climate change. On top of this, the model must carry out an analysis by which we may infer the stance on climate change of these online users. Ontologies, description logics, and specifically OWL 2 DL represent established technologies, both in academia and industry, upon which we may construct a viable tool to leverage existing studies in the human sciences to aid in research investigating online attitudes and behaviours. This is the purpose of our ontological model.

CHAPTER III

BIG PLAYERS ON TWITTER

This chapter covers the first stage of our research for this doctoral program. It does not involve ontologies or description logics as covered in the previous chapter. Neither does it involve our modelling of the Six Americas (see Section 0.3). We shall continue discussing these research efforts in Chapter 4. In this chapter, rather, we begin working with microblogs on the subject of global warming, collected from Twitter using an early version of the *Say* $\mathcal{S}\vec{\imath}La$ application described in Section 0.4. Our objectives with this project were the following: (1) begin development on the base components of the *Say* $\mathcal{S}\vec{\imath}La$ architecture; (2) become familiar with the tweets, their associated metadata, and the NLP tools we are utilizing to analyze that data; and finally (3) perform an initial analysis on the emotion expressed in tweets on the subject of climate change. As such, the "big players" project may essentially stand on its own, and we published an article on our findings (Drown et al., 2020) without any reference to *Say* $\mathcal{S}\vec{\imath}La$, nor to the larger research effort as it is laid out in the present document. However, in spite of this being an independent project in these respects, the strategy of focusing on high activity-users quickly became an integral part of the methodology for the analysis we have followed at all stages of this research project.

The big players are high-activity users in the online Twitter community tweeting on the subject of global warming. We perform an analysis of the correlation between the level of emotion these high-activity users are expressing in their tweets and the emotion level expressed in the tweets from the larger online community tweeting on this same subject. We employ an emotion lexicon to measure the levels of anger, fear, sadness, and joy expressed in the tweets tagged with the hashtag *#globalwarming* and determine the extent to which a small group of these big players may be used to predict emotion in the tweets from their larger online community.

3.1 Emotion in Twitter's Global Warming Community

We discussed in Section 1.4 how studies have indicated that emotion can be an influencing factor in people's perception of the risk associated with climate change as well as their support for policy that is intended to mitigate the problem. Additionally, emotions can effectively drive people to modify their behaviour so as to avoid risk when confronting a dangerous situation (Weber, 2006).

Microblogs from Twitter potentially represent a massive source of publicly available information in which people freely express their attitudes and talk about their behaviours with respect to climate change (or virtually any other subject). One study refers to NLP on Twitter as "an unsolicited opinion poll" (Cody et al., 2015). This study uses a sentiment analysis tool called a "Hedonometer" (Dodds et al., 2011) to measure variations in the levels of positive and negative sentiment as an indicator of happiness in an analysis of tweets on the subject of climate change. Our work here is similar in that we are measuring levels of affect in tweets, but ultimately our analysis is focused as much on the users (particularly on those tweeting extremely frequently) as it is focused on the tweets.

3.1.1 Why Look at Big Players?

The big players study is concerned with identifying users who are highly active online and who to an extent may potentially be representative of the general community with respect to the emotion expressed in their online communications on climate change.

The intent is essentially to create a methodology which serves to identify a relatively small group containing a fixed number of high-activity users. Researchers who may be studying online activity for a given purpose will then have a small number of accounts ready whose profiles and tweet activity they may consult to decide if these users are of interest as possible influencers or as prototypical examples of the community.

Often, when the goal is to identify individuals who fill a certain role in their online community, the standard approach involves the analysis of centrality using a graph-based representation of the relationships between users in that community. As research involving sentiment and emotion analysis on social media abounds, there are numerous examples using graph-based techniques. One such study endeavours to find influencers based on graph representations to describe "who follows whom" and "who reacts to whose tweets" while also measuring grade-level readability and levels of sentiment polarity in tweets (Bigonha et al., 2012). Another seeks to identify opinion leaders within Twitter communities, using a calculated competency score with respect to a given domain along with a popularity measure as determined by graphing follower relations (Aleahmad et al., 2016). A third seeks out experts based on a calculated measure for trust extended to a potential expert based in large part on the online relations for the user account (Eliacik & Erdogan, 2018).

In this study we take a different approach from the various graph-based strategies. Essentially, we put the focus on those users who set themselves apart given the exceptionally high level of online engagement they demonstrate. Our analysis involves creating affective models which serve to predict emotion in the tweets from the general community based on the emotion expressed in the tweets of these big players. Furthermore, given that these affective models take into account different types of big player activity, they may ultimately be useful for organizations seeking to evaluate various types of high-level participation, e.g. determining effective communication methods using emotion-based message framing.

3.1.2 Tweet Dataset

To represent the online community that is publishing microblogs about global warming on Twitter we created a dataset of tweets using the Twitter developer platform API.[1] The dataset consists of 414,035 tweets from 239,590 users. The tweets all incorporate the hashtag *#globalwarming* and were published between January 1, 2018 and August 31, 2019. This study tracks the hashtag *#globalwarming* as research indicates that it is used on Twitter relatively frequently by users from both the pro-science and skeptic communities, compared to other related hashtags, e.g., *#climatechange*, which tend to be used more often by climate activists than by skeptics or deniers in their tweets (Williams et al., 2015).

All tweets in the dataset selected for this study are in English. The metadata which accompanies each tweet collected via the Twitter API provides additional information about that tweet such as the language of the text. We use this metadata to identify the language and filter the tweet for inclusion in the dataset.

3.2 Tools for Natural Language Processing

We use a number of NLP tools to process the tweet text and extract the information used in the affective models. This section describes these tools and how we have incorporated them into our methodology.

3.2.1 A Lexicon of Emotion

In Section 1.1.4 we discussed the system of eight basic emotions according to Plutchik (Plutchik, 2001). For the big players study, we employ the NRC Affect Intensity Lexicon (NRC-AIL) from the National Research Council of Canada, which includes four of the

[1] https://developer.twitter.com/

basic emotions from Plutchik's model: anger, fear, sadness, and joy (Mohammad & Bravo-Marquez, 2017). The lexicon was manually created for communications on Twitter using a technique known as Best-Worst Scaling (Louviere et al., 2015), which has a number of people select the two items out of a set that are most and least representative of a specific property. For the lexicon, three project annotators plus a large number of crowd-sourced individuals were asked to choose the best and worst tweets with respect to how well they typify one of the four basic emotions. The tweets were grouped in sets of four examples from a dataset of emotion-laden tweets collected specifically for the project. The Best-Worst Scaling process allows for a calculated value of the emotional intensity attached to words, rather than simply categorically indicating the presence of an emotion with respect to a given word. In our study, we use this lexicon to provide an intensity level for tweets based on the words they contain.

We access the lexicon via the Affective Tweets plugin (Bravo-Marquez et al., 2019)[2] for the Weka platform for machine learning (Frank et al., 2016). It reportedly contains "close to 6,000 entries" (Mohammad & Bravo-Marquez, 2017), but the implementation used in this project does not provide access to all of these entries. Although version 0.5 (used in the present research) contains 5814 words, the supposedly equivalent lexicon for the Affective Tweets package for Weka provided as an Attribute-Relation File Format (ARFF) file, contains only 4192 entries.[3]

Our system creates an emotion vector for each tweet, treating the text essentially as a bag of words and using the Weka plugin as a library to calculate floating-point values for the four emotions covered by the NRC-AIL lexicon. The Affective Tweets plugin sums intensity levels across these emotions for all words in the text that appear in the lexicon. Together, these values represent the emotion intensity for the whole tweet for the four

---

[2]https://github.com/felipebravom/AffectiveTweets

[3]We are unclear as to why. The two formats have the same header information, including the lexicon version. Words in the ARFF file start with *A* and continue to *Z*, and spot-checking the entries shows the same intensity levels for the various emotions. Our work uses the shorter ARFF lexicon.

basic emotions we are using in the study. For example, this tweet:

> *The New Coal Crisis - #EndFossilFuels destroying the environment, boosting #globalwarming & threatening our health https://t.co/Ig1sCBpEFS*

generates the following emotion vector:

$$\langle anger{=}1.726, fear{=}2.485, sadness{=}0.734, joy{=}0.493\rangle$$

As of this writing, we are pleased to note that the NRC has recently released a new version (v1.0) of this lexicon, now called the NRC Emotion Intensity Lexicon (NRC-EIL).[4] This release includes all eight of the basic emotions in Plutchik's system (Mohammad, 2018). Unfortunately, this newer version was not available when we conducted the experiments in the present work, which uses the NRC-AIL as described above.

We should note briefly that although the analysis of sentiment and emotion in the present research is based on lexica, these are not the only tools available. There has been extensive research into machine learning methods for emotion classification in text. These include well-known supervised learners, such as decision trees, naive Bayes, support-vector machines (Alm, 2008; Chaffar & Inkpen, 2011; Lin et al., 2007; Liu, 2015; Mohammad, 2012) as well as deep convolutional neural networks (Yang et al., 2022). There is also some work involving unsupervised learning with hybrid neural networks (Wang et al., 2017). We shall be interested in looking further into machine learning strategies for the detection of emotion as part of our continued research. Incorporating such learners into the machine learning level of the $\mathcal{S}ay\ \mathcal{S}\vec{\imath}\mathcal{L}a$ architecture will allow for an extensive comparison of these techniques against our initial methodology using lexica as well as an opportunity for experimenting with hybrid solutions.

---

[4]http://www.saifmohammad.com/WebPages/AffectIntensity.htm

3.2.2 Preprocessing Tweet Text

In addition to providing program access to the NRC-AIL lexicon, the Affective Tweets library (Mohammad & Bravo-Marquez, 2017) makes available a number of NLP tools. We use its tokenizer, which is designed specifically for texts on Twitter (Gimpel et al., 2011). The tokenizer delimits words, punctuation, emojis, and other symbols in the text. It also standardizes usernames and web URLs so that they do not interfere with the calculation of emotion intensity levels for the tweet.

The plugin allows a program to specify a stop list so that common words, generally devoid of any analytical value (e.g., *a* and *of*), are removed from the text before determining the associated emotion values. It also enables the use of a stemming algorithm, which reduces each word to its lexical stem (e.g., "warms," "warmed," and "warming" all become "warm"). As part of our initial research efforts for the big players project, we experimented with affective models created with the use of Apache Lucene's[5] list of English stop words and also the Snowball Porter stemmer (Porter, 2006) for English.[6] However, incorporating these two tools did not improve results substantially, and so this chapter presents the results from models with no stop list and without stemming.[7]

Although stop lists and stemming are common practice in NLP applications, we do not find it extremely surprising that these preprocessing steps did not lead to affective models that consistently performed better in our study. Traditional NLP tools tend to suffer when analyzing human language as it is generally found on social media. The authors of these tweets are not professional writers. They will often express their ideas with minimal consideration given towards clearly organizing the intended content. Slang and

---

[5]https://lucene.apache.org/

[6]http://snowball.tartarus.org/

[7]We do use stop word lists and stemming when processing tweets for the ontological model (see Section 5.1.2); however, we do not perform a comparison of results with and then without these preprocessing steps.

abbreviations are common, as are spelling errors. While our analysis makes use of tools intended for social media, the state of the art for processing natural language online still lags behind what is currently possible for traditional, more formal texts (Farzindar & Inkpen, 2015).

3.3 Affective Models

The affective models in this study use emotion expressed in big player tweets to predict levels of emotion in tweets from the "regular players," the rest of the Twitter community participating in the online conversation about global warming. In order to determine which of the top high-activity users form the set of big players for a model, we observe the correlation between the emotion measured in the tweets of a tentative set of big players and the emotion measured in the tweets the regular players are publishing.

This section describes the different ways Twitter users can demonstrate high activity as defined by our methodology in this study. It also covers how we group candidate sets of big players and how we use these elements in linear regression models to predict emotion in tweets from the regular players. The dataset used to train and evaluate these models is described in Section 3.1.2. In the experimental process we use only a portion of the full dataset at any one time. This portion represents the tweets published during a fixed period of time. We call this data subset a *tracking run*. It is effectively the tweet data for a time slice from the year and eight months covered by the full dataset.

3.3.1 Communication Categories

A Twitter user can demonstrate a high level of online activity in a number of different ways. For the purpose of creating affective models, we define four *communication categories* and rank users in terms of tweet count for each one. The communication categories are defined as follows (with the associated code for the affective model in parentheses):

1. **Original Tweeters** (*oter*): Users publishing messages that they have personally authored.
2. **Retweeters** (*rter*): Users retweeting (republishing tweets originally written by someone else) using the Twitter syntax "*RT @author* ...."[8] The act of retweeting may be likened to the concept of forwarding an email for others to read (Boyd et al., 2010).
3. **Retweeted Authors** (*rted*): Users who publish tweets that are then retweeted by other users. Retweeted authors originate the tweets that players in the previous category are retweeting.
4. **Mentioned Authors** (*tmed*):[9] Users who are directly mentioned in another user's tweet by means of Twitter's "*@user*" notation. This construct in a tweet is generally used to address specific users or attract their attention (Honeycutt & Herring, 2009).

It is worth noting that somewhat frequently a single tweet will be accounted for multiple times. For example, if the user *alice* publishes tweet T_1, which she has personally authored and which mentions user *realdonaldtrump*, then T_1 contributes to *alice*'s participation in the *oter* category and also to *realdonaldtrump*'s participation in *tmed*.

Similarly, as illustrated in Figure 3.1, if the user *bob* sends a retweet T_2, originally published by *alice* (T_1), and which mentions user *realdonaldtrump* then T_2 contributes to *bob*'s participation in *rter*, to *alice*'s participation in *rted*, and to *realdonaldtrump*'s participation in *tmed*. As we see, each of the communication categories corresponds to a distinct type of participation. In this second example, for a given experiment, *alice*'s

[8]Although this syntax is the most common, the format is not standardized, and there are alternatives. Note that the metadata, which the Twitter API includes with each tweet, also identifies messages as retweets and specifies the retweeted author.

[9]Please note that due to an effort to keep the category names parallel in the system, the *tmed* code for the model is backwards: "*m*ention*ed* in *t*weet."

Figure 3.1 Communication categories for tweet activity.

original tweet (T_1) would not be considered if it were sent before the beginning of the tracking run. For the same reason, only T_2 will contribute to *realdonaldtrump*'s emotion vector for the *tmed* category if *alice*'s T_1 tweet came before the start of the tracking run (as opposed to contributions from both T_1 and T_2 in the case where T_1 was published during the period covered by the tracking run.)

Of course, there is a distinct relation between big players who are retweeters and big players who are retweeted authors since the exact same set of retweets are used to determine who is a big player. This in no way means that we will identify the same big players in each category. If it helps to clarify the distinction, the reader may keep these two questions in mind with respect to all the microblogs tagged as retweets: which users are sending most of the retweets (*rter*), and which users are most often retweeted (*rted*)?

We can also look at *oter* and *rter* as communication categories representing active participation. Similarly, we can think of *rted* and *tmed* as categories which represent forms of passive participation. To a point, it may be useful to consider the concept of passive big players; yet, this analogy is not valid in every sense. Users on Twitter may often seek rather actively to be retweeted (Boyd et al., 2010), and frequently users will mention each other reciprocally in their tweets when establishing a communication thread (Honeycutt & Herring, 2009).

3.3.2 The Top N Big Players

For a given tracking run we identify a set number of Twitter users who rank highest with respect to each of the four communication categories described in the previous section. These are the “Top N” players for the tracking run. As such, we are working with what are essentially four types of high-activity users. Note that it is possible for one user to be in two categories.[10]

Our methodology involves evaluating a series of big player groups whose size ranges from 5 to 25. As mentioned in Section 3.1.1, we would like to keep the size relatively small so that researchers interested in utilizing a big player group for a given project will not be overwhelmed if the research calls for a manual examination of the users' profiles or their tweets. Ideally, we are seeking a group size N which demonstrates a significant correlation between the levels of emotion measured in the big player tweets and the levels found in the general community.

We define the regular players simply as the users (1) who have published at least one original tweet with the hashtag *#globalwarming* over the course of the specified tracking run and (2) who are not in any of the four big player communication categories. The affective models aim to predict emotion levels in these original tweets published by the regular players. We take it as a reasonable assumption that the original tweets authored and published by the *#globalwarming* community on Twitter likely represent a fair reflection of what that community is feeling.[11]

---

[10]For the series of experiments conducted in this study, only five out of the $239,590$ users showed up in two categories. In each case the first communication category was either *oter* (original tweeter) or *rter* (retweeter), while the second category was always *rted* (retweeted author).

[11]Note that we are not examining regular players as retweeters nor as retweeted or mentioned authors

3.3.3 Linear Regression

In this study we use linear regression models to predict the level of emotion in the regular players' tweets. This basic machine learning algorithm generates a function of the form:

$$\hat{y} = \sum_{j=0}^{k} w_j a_j = w_0 a_0 + w_1 a_1 + \ldots + w_k a_k. \tag{3.1}$$

The function maps a vector of k elements to a scalar, our target prediction (Witten & Frank, 2005). For each data instance, a_j represents the value for a given attribute, and w_j is the weight. A higher weight value w_j indicates that the attribute a_j has greater significance when predicting the target value, whereas a negative coefficient demonstrates an inverse relationship between the attribute and the target. The zeroth term a_0 is always set to 1, and the corresponding weight w_0 is the function's intercept, a bias term indicating the output of a linear regression model when there is no input data. Mathematically speaking, this is an affine function for any non-zero intercept; however, the description "linear" is common in machine learning (Goodfellow et al., 2016). Note that the intercept generally appears last.

Linear regression determines values for the weights by minimizing the sum of the mean squared error across the instances used to train the model:

$$min \sum_{i=1}^{n} \left(y^{(i)} - \hat{y}^{(i)}\right)^2 \tag{3.2}$$

where $y^{(i)}$ is the true target value for instance i, and $\hat{y}^{(i)}$ is the corresponding prediction.

As part of our initial work on the *Say Sĩla* architecture we created a Clojure-based testing implementation which utilizes the *LinearRegression* Java class provided by the Weka machine learning platform, which we access as a programming library to create

---

in the way we do for the big players.

the regression model.[12] We use Weka's default hyperparameters for the learner. One of these is the ridge parameter, which works as a penalty term as the algorithm minimizes the sum of the mean squared error. This term serves to regularize the attribute weight coefficients, keeping them from growing too large and thereby limiting the complexity of the model and its tendency to overfit the training data (Goodfellow et al., 2016; Pereira et al., 2016).[13]

Default hyperparameters for Weka's *LinearRegression* also include the elimination of collinear attributes as well as the use of the M5′ algorithm when selecting which data attributes to use in a linear regression model. M5′ (Wang & Witten, 1997) is an improvement on J. Ross Quinlan's M5 algorithm (Quinlan, 1992).[14] M5′ is a tree-based algorithm for continuous attributes. It works to reduce the number of attributes in the model by identifying the attributes that reduce the variance in the target class. The algorithm uses a metric of standard deviation reduction, adjusted to handle missing values, and those attributes which maximize this metric are selected (in our case) for the regression model.

In addition to being a standard machine learning algorithm, regression has been used in research to analyze emotion in texts published on social media. For example, (Preoţiuc-Pietro et al., 2015) use linear and non-linear regression models, whose attributes include sentiment polarity and emotion, to predict the income of users on Twitter. While we recognize that other machine learning algorithms may result in more accurate models, often the improvements are relatively minor, and many of these algorithms require that we sacrifice the visibility that regression-based models provide (Weisberg, 2014). As we

---

[12] Weka's fully qualified class name is *weka.classifiers.functions.LinearRegression*. The following links reference our code where we (1) invoke this learner as well as a number of others and (2) automate the training and evaluation of the model:
(1) https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/weka/core.clj#L160
(2) https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/weka/tweet.clj#L380

[13] Weka's default value for the ridge parameter is 10^{-8}. See Section 7.1.1 for more details on ridge regularization.

[14] Quinlan also created the well-known C4.5 decision tree, discussed in Section 7.1.2.

are seeking to analyze online content from the big players which serves to predict emotion in the general community, a black box model is of somewhat limited use, no matter how accurate it may be. Additionally, a rule of thumb in machine learning is to start first with simpler and speedier learning algorithms before moving to more complex strategies (Domingos, 2012). Linear regression is an established statistical method that performs well in many applications. Moreover, even for non-linear problems, results from linear regression can be a valuable indicator of what to expect from more elaborate learners (Witten & Frank, 2005). Given the associated methods of analysis of variance to evaluate how well a model fits the underlying data, linear regression seems an excellent approach towards creating a clear initial view as to who is representative of online communities with respect to how they feel about global warming.

Note that we also performed the experiments for the big player project with a number of other machine learning models available using Weka (Witten & Frank, 2005) including Gaussian processes, decision lists using separate-and-conquer (M5Rules), random forests, and support-vector machines with first and second degree polynomial kernels (SMOreg using PolyKernel with exponents set to 1 and 2). In each case we used Weka's default settings associated with the specified learner. Linear regression consistently showed better results than these other models while maintaining the advantage of being a "clear box" modelling technique that allows us to understand how the affective model determines its predicted values. Hence, in this chapter we are reporting only the results from the linear regression models.

3.3.4 Linear Regression and Emotion in Tweets

We are using these linear regression models to predict the variation from one week to the next in the levels of emotion as measured in tweets from the regular players in the *#globalwarming* community on Twitter. For a given tracking run, we create four models, one for each of the emotions covered by the NRC-AIL lexicon: anger, fear, sadness, and joy. Each model has 16 attributes which represent the variation, week by week, in the average levels of all four emotions for the tweets from the big players across the four communication categories described in Section 3.3.1. These attributes are the same for each of the four models.

The name of each big-player attribute, as well as the dependent (target) variable used to predict average emotion levels for the regular players, all follow a three-part naming convention which indicates (1) the community group, (2) the communication category, and (3) the expressed emotion:

$$\binom{big}{reg}\_\begin{pmatrix} oter \\ rter \\ rted \\ tmed \end{pmatrix}\_\begin{pmatrix} anger \\ fear \\ sadness \\ joy \end{pmatrix}$$

As one example, the attribute big_rter_fear represents the variation in the measured level of fear in retweets (via "RT @author") from users in the big retweeters group. As another, the attribute big_tmed_joy gives the variation in joy as measured in tweets that contain frequently-mentioned authors (via "@username"). Note that all attributes except the target start with *big* as these values correspond to the variation in emotion from the big player tweets. In contrast, the target attribute in each model always begins with *reg* since we are predicting variation in emotion levels for the regular players. Additionally, the second component in the name of the target attribute is always *oter* as our prediction is always with regard to the original tweets of these users.

The following linear equation presents an example of a typical affective model, which predicts the variation in the level of anger expressed in the regular players' tweets for a given week with respect to the previous week:

$$\begin{aligned} reg\_oter\_anger = \quad & 0.261 \times big\_oter\_anger \\ & + 0.167 \times big\_oter\_joy \\ & + 0.067 \times big\_rter\_fear \\ & - 0.059 \times big\_oter\_fear \\ & + 0.030 \times big\_rted\_joy \\ & - 0.001 \end{aligned} \tag{3.3}$$

This equation might usually be presented along with our results. It is the anger model for the Top 12 big players in the experiment representing our first tracking run (2018-01-01 up to 2018-10-01). However, the full series of experiments generates 720 of these models (4 emotions × 20 Top-N levels × 9 tracking runs).[15] It would be interesting to attempt to determine an all-encompassing interpretation for all these models, but as our purpose here is to create a base architecture capable of running machine learning experiments on Twitter data, we are simply using this single result as an example in the current discussion. In this example, the last term is the intercept. This small, negative number indicates that if there is no change in any of the big player emotions in the model for a given week during the period, then we can expect a very slight drop in the level of anger measured in the original tweets of the regular players. Assuming that there *is* some measured variation in the emotion expressed by the big players, the attribute weights appearing before the intercept allow us to calculate a prediction for the variation in anger expressed by the regular players. The *big_oter_anger* attribute has the largest positive coefficient. Hence, we should expect that when big players express more (less) anger in their original tweets than they did in the previous week, then given relatively little change in the other attributes,[16] we will see an increase (decrease) in

[15]The architecture also creates 14,400 reference models for comparison with the 720 based on the big players.

[16]A thorough interpretation of these models would include an analysis of collinearity between at-

the anger from the regular players in their original tweets. The same is true, but to a decreasing extent, for joy in original tweets from the big players, fear in their retweets, and joy when they are retweeted.[17] The final term, *big_ oter_fear* indicates a negative correlation, whereby more (less) fear in the big players' original tweets means we will predict less (more) anger in original tweets from regular players.

When considering these models, we should note that we have no particular reason to expect a linear relationship between the emotional levels measured in the communications of the regular players and those of the big players. Nevertheless, generating a regression model that explains a major portion of the variance seen in the data is valuable because it clearly indicates the factors involved in producing the predicted value for a given emotion.

3.3.5 Data Preparation for a Tracking Run

We have developed a multi-step methodology to prepare the raw tweet data to train the regression models associated with a given tracking run. Our system implements the following procedure:

1. **Apply emotion vectors**: We run the tweets tagged with *#globalwarming* through the Affective Tweets Weka filter to generate an emotion vector for each tweet. For every user we keep four running averages as we process the tweets. The running averages are also emotion vectors, and there is one for each of the four communication categories. We apply each tweet's emotion vector either as an original tweet (*oter*) or as a retweet (*rter*). If it is a retweet, we also apply the emotion vector to the user who originally authored the tweet (*rted*). Lastly,

---

tributes.

[17] It is likely that many of the microblogs represented by the *big_ rted_joy* term are retweeted versions of the microblogs represented by *big_ oter_joy*.

we determine if the tweet text contains any user mentions. If so, we apply the emotion vector to every user mentioned in the tweet (*tmed*). When the system applies an emotion vector to a user, it effectively incorporates the emotion levels for the tweet being processed into an emotion vector which represents the running averages for the four emotions on the day the tweet was published for the appropriate communication category. The system also increments a tweet counter for the user for that communication category. The counter is used to determine the user's ranking in the category with respect to the other users in the tracking run.

2. **Rank and group players**: Once it has processed all the tweets, the system determines who the big players are. For values of N ranging from 5 to 25, we generate player activity rankings for each of the four communication categories and select the N highest ranked users for inclusion in a big player group for that category. Occasionally there is a tie. For example, when determining the Top 10 players in a given communication category, the users ranked #10 and #11 may both have the same number of tweets. When such a tie happens, we include both users, and the Top 10 group for that communication category will actually have 11 players. After we have determined who is in the big player groups for each of the four communication categories, all the remaining users are grouped together to represent the regular players. (The models will be predicting the variations in emotion from one week to the next for the tweets from these users.)
3. **Convert day vectors into week vectors**: Now that we have identified the four groups of big players and the one group of regular players, the system calculates average emotion vectors for every group across all the users in the group for each day of the tracking run. It sorts the day groupings (by calendar order) and then bundles them into week groupings (i.e., seven consecutive day groupings). Lastly, the system reduces all the emotion vectors in each week grouping to create a single vector that represents the average emotion expressed in the tweets for all players in a community (i.e., *big* or *reg*) and communication category (i.e., *oter*, *rter*, *rted*, or *tmed*) over the course of that week.

4. **Compute weekly variation**: We begin this final step with what is essentially a set of data instances, where each individual instance represents the emotion intensity for one week of tweets. The system takes each of these instances and subtracts the associated attribute values from those of the previous instance, thus obtaining the variation in emotion intensity from one week to the next. That is, for an attribute A_Δ the system computes the variation: $A_\Delta = A_i - A_{i-1}$, where A_i is the emotion level for the current week, and A_{i-1} is the level for the previous week.[18]

After completing the final step, the system will have calculated the weekly variations in the average levels for the big and regular players for each communication category and for all four emotions. There are 16 attributes for the big players (e.g., *big_oter_anger*) which represent data values to be substituted for a_1 through a_k in the linear function shown in Equation 3.1. There are also four target attributes (y in Equation 3.2) representing the week-to-week variations for the four emotions as expressed in the original tweets of the regular players (e.g., *reg_oter_fear*). Each serves as the dependent variable for a separate linear regression model. Thus, for every tracking run we create four models, one for each emotion. The 16 big-player attributes are constant across each of these four models for a given tracking run. The target attribute is all that changes between each predictive model.

3.4 Experimentation, Results, and Analysis

We now have an understanding of the structure of the affective models intended to predict the variations in expressed anger, fear, sadness, and joy in regular player tweets from one week to the next as these players discuss global warming. This understanding

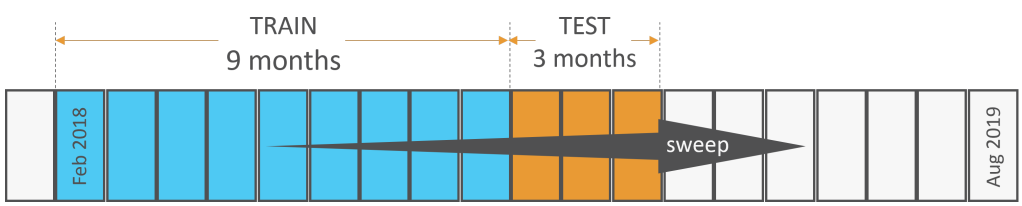
[18]Some readers may think it better to calculate the relative change instead, dividing A_Δ by A_{i-1} which results in a percentage. This approach was not possible here as occasionally for low values of N, a big player group would not express one of the four emotions during a full week of tweets, resulting in a relative change which is undefined.

includes the format of the dataset for training and testing these models with the 16 big-player attributes (4 communication categories × 4 emotions) representing big player activity and one dependent attribute (an emotion) for four separate models which together represent the regular player activity. In this section we present the procedure used to run the experiments for the big player project, report the results from those experiments, and discuss our findings.

3.4.1 Tracking Run Sweeps

Conducting an experiment involves the creation of a series of affective models for each of the four basic emotions for groups of Top N players with N ranging from 5 to 25. As mentioned in Section 3.1.2, our dataset contains tweets with the hashtag *#globalwarming*, published between January 1, 2018 and August 31, 2019. The dataset therefore represents 20 full months of tweets.

Figure 3.2 Sweeping a tracking run window for an experiment.



For an experiment we run 9 tracking runs. Every run uses 12 months of data, beginning on the first day of month M_i at midnight, where i ranges from 1 to 9, and ending on the last day of month M_{i+11} at 23:59:59.999. Each of the 9 tracking runs shifts the start month by one, which means that we are effectively working with a 12-month window, which we sweep across the 20 months of data collected from Twitter. Figure 3.2 illustrates this part of the experimental methodology. To better clarify the process, the figure is displaying the second tracking run of an experiment. There is one tracking run that has already occurred, the run indicated in the figure, and then seven additional runs are needed to reach the end of the data and complete the experiment.

We have chosen a 12-month data window because as each instance in the dataset represents one full week of activity on Twitter, we need to cover a relatively large time span in order to have enough data to train the models. Additionally, the one-year window represents a natural block of time with respect to the calendar. For each 12-month tracking run we designated the first nine months (75%) as the training period and used the last three months (25%) to test the model. We report our results using the average Pearson correlation coefficient (PCC) across all 9 tracking runs. The PCC metric ranges between -1 and 1, where a 1 would indicate an exact positive correlation between the prediction of the affective model and the values measured from the test data.[19] Table 3.1 presents these results for the linear regression models for each of the four emotions.

For each value of N, in order to determine if the group of big players is significant with respect to its capacity to predict the emotion expressed in the general community, we compare the results for these big players with results for reference groups of N members. However, to build a meaningful model we need to ensure that users in these reference groups show a minimal level of participation in terms of the number of tweets they have published during the tracking run. To this end, we select random users to create four groups, one for each communication category (*oter*, *rter*, *rted*, and *tmed*).[20] The size of each reference group is the same as that of the big players (N), and all its members must have published at least 40 tweets in the communication category for the group. We also require that users in the reference groups not be big players themselves. This constraint aims to avoid biased results whereby one or a few big players end up making the major contribution for a reference group. Once we have created the reference groups, we process

---

[19]Please consult the glossary for a more complete description of the PCC.

[20]Our code for the random selector utilizes the Erlang *rand* library's implementation (*exrop*) of the Xoroshiro116+ pseudorandom number generator (Vigna, 2016) with 58 bits of precision and a period of 2^{116-1}.

the tweets of their members and create the linear regression models as described above, only replacing the four sets of big players with the four reference groups. For all tracking runs performed and for each group size of N reference users (replacing the Top N players), we repeat this procedure 20 times.

Table 3.1 Correlation (PCC) for models predicting emotion in the last 3 months.

	Anger		**Fear**		**Sadness**		**Joy**	
N	**BIG**	**ref.**	**BIG.**	**ref.**	**BIG**	**ref.**	**BIG**	**ref.**
5	0.1376	*****	0.0129	*****	-0.0992	*****	0.1130	*****
6	0.1660	*****	-0.0818	*****	-0.1369	*****	0.1473	*****
7	0.1465	0.3301	0.0509	0.2432	-0.1539	-0.2223	0.1680	-0.0852
8	0.3415	0.0751	0.0429	-0.1945	-0.0758	-0.3385	0.0170	0.1602
9	0.2727	0.0090	-0.0337	-0.1281	-0.1429	0.2542	0.1183	-0.0504
10	0.2294	-0.0753	0.0204	0.1587	-0.1407	-0.0119	0.1291	0.1278
11	0.2534	0.0951	0.0031	0.0112	-0.2119	0.0155	0.1066	0.0938
12	0.3189	0.0595	-0.0009	-0.0340	-0.2380	-0.0701	0.1058	0.0816
13	0.3035	0.0144	0.0197	0.0425	-0.1542	0.0436	0.0935	0.0766
14	0.3024	-0.0786	0.1568	0.0128	-0.1876	0.0682	0.0854	-0.0219
15	0.2793	-0.0599	0.2358	0.0334	-0.1989	-0.0854	0.0115	-0.0606
16	0.2770	-0.0229	**0.2972**	-0.0332	-0.1221	0.0127	0.0632	-0.0007
17	**0.3869**	0.0139	0.2566	0.0099	-0.1812	0.0468	0.0945	0.0639
18	0.3787	0.0580	0.1282	0.0530	-0.2420	-0.0236	-0.0110	0.0829
19	0.3402	-0.0351	-0.0134	0.0053	-0.1857	-0.0609	0.2517	-0.0044
20	0.2764	0.0046	0.0389	0.0204	-0.1725	-0.0063	0.0016	0.0721
21	0.2917	0.0233	0.0914	0.0677	-0.0572	0.0015	0.0531	-0.0348
22	0.2042	0.0524	0.0934	-0.0289	-0.1444	0.0032	0.0942	-0.0314
23	0.1969	-0.0066	0.1732	0.0245	-0.1179	0.0222	0.1064	-0.0126
24	0.1816	-0.0012	0.1331	0.0177	-0.0533	0.0185	0.1061	0.0595
25	0.1201	0.0552	0.0594	0.0241	-0.0897	-0.0442	0.0522	0.0115

The “ref” columns in Table 3.1 report the average PCC for each value of N over 180 reference model runs (20 reference groups per tracking run × 9 tracking runs). When a field contains asterisks (*****), this indicates that for every model in the 180 runs, there was at least one week during which a reference group did not publish any tweets with the specified emotion for one or more communication categories. This issue primarily occurs with small values of N when there are fewer reference players contributing. Their level of activity may not be consistent, or the variety with respect to tweet content may not be sufficient to cover every week in the tracking run.[21] We should keep in mind that

[21]The problem also occurs when the training period is longer (i.e., 12 months instead of 9) as there is a higher probability of a “slow” week.

the reference users are necessarily less active than the big players (or else they would be big players themselves).[22]

When we compare the columns for weekly variation in anger in Table 3.1, we note that the PCC from the big players (BIG) is closer to 1 and higher than the average PCC across the 180 reference groups (ref). Following N as we consider these columns, it is interesting that the PCC for the big players generally increases as the size of the group increases until we obtain a value of 0.3869. Then, as we continue with ever larger groups of big players, the PCC falls off. A group of around 17 big players produces the best results.

As we analyze the predicted variations in fear from Table 3.1, we see that although our models for fear do not always outperform the reference models, they show an improved linear correlation when N is between 14 and 18, and we observe the highest PCC, 0.2972, at $N = 16$. Big players score better than the reference groups towards the middle of our selected range for N, but not for models representing smaller groups nor larger groups. These results indicate that when considering anger and fear, the big players are indeed significant as a group. Additionally, for both of these emotions the PCC is maximal when the group size is between 16 and 18. This finding is noteworthy as the general idea behind the big player technique is to allow researchers working on their own projects involving social media to identify just a reasonable number of users who may be representative of the general community. These researchers may then focus their efforts on these high-activity users.

The results in Table 3.1 for sadness and joy are not so clear. The models for sadness show no significant level of linear correlation, neither for the big players, nor for the reference groups. The results for joy are not much better. There is one notable PCC of 0.2517 at $N = 19$, but this result seems anomalous rather than significant.

---

[22] We could effectively avoid this problem by raising the minimum tweet limit for inclusion in a reference group past 40; however, going too high results in a smaller pool of viable users to fill the reference groups when N is relatively large.

Given these results, it seemed important to verify that the observed lack of correlation with sadness and joy was not due to any inherent limitation of linear regression, As mentioned in Section 3.3.3, we reran the experiments with a number of other machine learning algorithms including Gaussian processes, decision lists, random forests, and support-vector machines with first and second degree polynomial kernels. Notably, none of these learners surpassed the results obtained from the linear model.

As mentioned above and as illustrated in Figure 3.2, this stage of experimentation used the tweets from the final three months of each tracking run as the test dataset. Essentially, the affective models were using the first nine months of the period to predict variations in emotion for the last three months. However, we also created linear regression models to predict the weekly variations in emotion *during* the 12-month period. Instead of using an independent test dataset created from the tweets occurring in the three-month span immediately following the training period, we used 10-fold cross-validation on a full 12 months of training data.[23]

As can be seen in Table 3.2 the experiments using cross-validation markedly improved the PCC scores for the big player models for all emotions across all values of N. However, the tests also significantly improved the results of the reference models. Indeed, the scores for these baseline models were high enough to preclude the hypothesis that the Top N big players were a better choice than random users (with the 40-tweet minimum level of activity) for the problem of predicting variation in emotion levels during the actual period used to train the model. Looking down the rows of the table, we do not see the PCC tending towards a peak at a certain value for N like we did for anger and fear when testing using the final three months of the tracking run. Therefore, for researchers using this methodology the task of picking a good value for N becomes rather problematic. Much more significantly, the PCC scores for the big player and the reference groups are comparable, indicating that there may likely be no advantage in using the big players as

[23]Please see the glossary for a description of the process of cross-validation.

Table 3.2 Correlation (PCC) for models evaluated using 10-fold cross-validation.

	Anger		Fear		Sadness		Joy	
N	**BIG**	**ref.**	**BIG.**	**ref.**	**BIG**	**ref.**	**BIG**	**ref.**
5	0.3521	*****	0.0721	*****	0.3203	*****	0.1580	*****
6	0.3620	*****	0.2865	*****	0.3309	*****	0.1559	*****
7	0.4067	*****	0.3893	*****	0.3138	*****	0.2326	*****
8	0.4127	-0.0685	0.4050	0.1910	0.2369	0.1227	0.2317	0.5470
9	0.4576	0.3461	0.3075	0.2587	0.2970	-0.0342	0.2418	0.3372
10	0.4545	0.1041	0.2906	0.2317	0.3216	0.1892	0.2215	0.2865
11	0.4322	0.2651	0.3815	0.2895	0.3102	0.2942	0.2428	0.3487
12	0.4327	0.1883	0.3369	0.2678	0.3000	0.1924	0.1974	0.3406
13	0.3800	0.1766	0.3189	0.1528	0.2297	0.1536	0.2233	0.3262
14	0.4299	0.2770	0.3341	0.3137	0.2917	0.4133	0.2367	0.2675
15	0.4364	0.2358	0.3238	0.3180	0.2610	0.3765	0.2432	0.2508
16	0.4098	0.2712	0.3625	0.2686	0.2416	0.3788	0.1705	0.3115
17	0.3740	0.2361	0.2448	0.2638	0.2651	0.2903	0.1393	0.2719
18	0.3823	0.3284	0.2670	0.2845	0.2697	0.2582	0.1425	0.2410
19	0.3550	0.2701	0.3248	0.2804	0.2871	0.2609	0.1194	0.2724
20	0.3955	0.2926	0.2482	0.3146	0.2509	0.2686	0.0858	0.2840
21	0.4236	0.3049	0.3583	0.2927	0.3052	0.3601	0.1160	0.3011
22	0.3716	0.2828	0.3044	0.2811	0.2900	0.3102	0.1401	0.2419
23	0.4306	0.3191	0.3106	0.2623	0.3086	0.3263	0.1364	0.3048
24	0.4248	0.3189	0.3235	0.3178	0.3080	0.2934	0.1673	0.3004
25	0.4201	0.3382	0.2678	0.2700	0.2740	0.3054	0.1909	0.2849

representative of the larger community with respect to predicting the emotion expressed in their tweets. This finding stands in stark opposition to the predictive value of the big players for the final three months of a period.

Figures 3.3 and 3.4 graph the PCC (y-axis) respectively for anger and fear as we vary the number of users in the big player groups (x-axis). Both figures display two graphs. In each case the underlying data for the upper graph is in Table 3.1. These are the results from the experiments where for each tracking run we predict variation in emotion at the end of the period, using the first nine months of data for training the model and the last three months for testing it. In the lower graph of each figure the data is from Table 3.2. These results are for the experiments using 10-fold cross-validation where the model is effectively predicting the variation in emotion over the full period of the tracking run. In each of the four graphs the dotted line represents the average PCC for the 180 reference models measured for each value of N. The peaks for maximum predictive capability are

Figure 3.3 Correlation for models predicting anger.

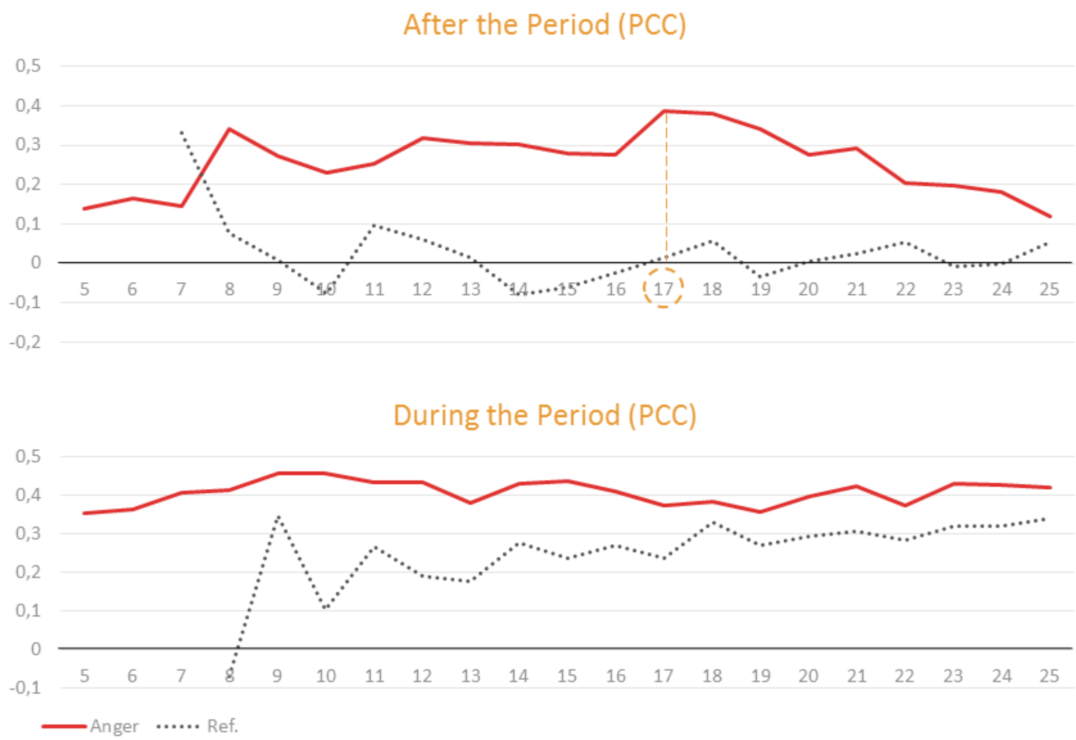
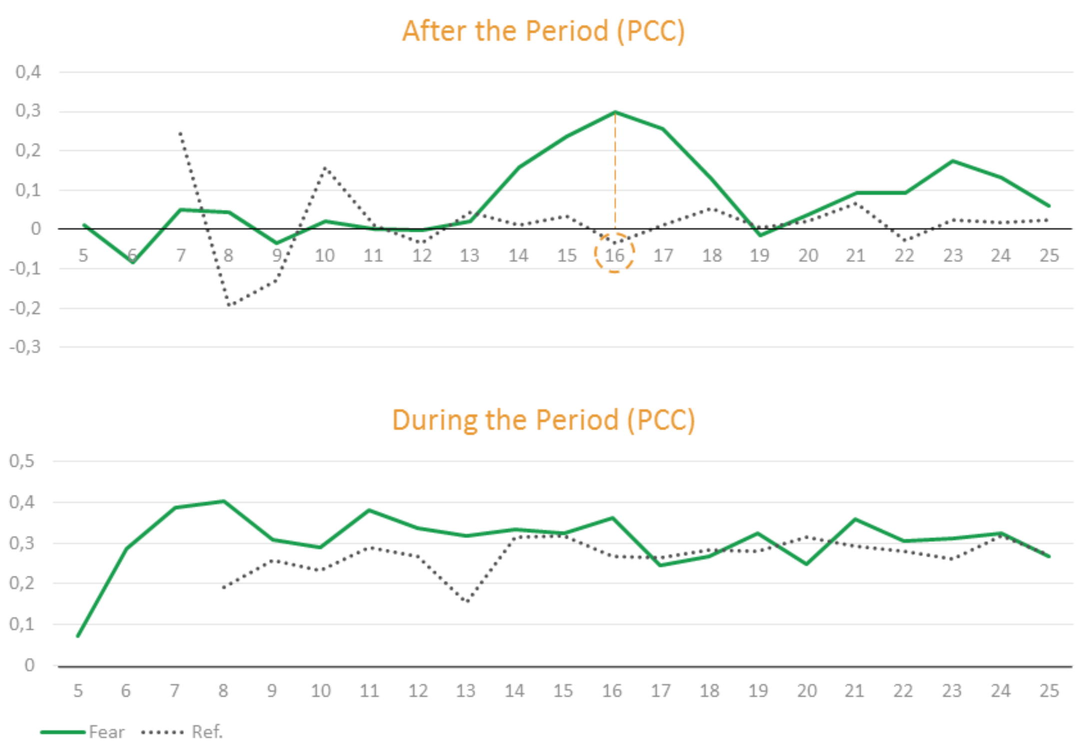


Figure 3.4 Correlation for models predicting fear.



readily seen in the top graphs at big player group sizes of 17 for anger and 16 for fear. In the lower graphs, however, there is no peak, and with fear especially, the big players show no increase in predictive capacity compared to the reference groups. Clearly the affective models perform better when predicting variation in emotion in the months following the training period.

With the goal of investigating this interesting finding further, we repeated the experiment for prediction during the tracking run period. However, this time instead of using cross-validation, we randomly selected a number of data instances corresponding to three month's time from throughout a given 12-month period. We removed the instances from the training data and used them to create an independent test dataset for the tracking run. In this way we were able to employ a method which effectively parallels that of cross-validation, yet which is distinctly different, and still predicts variation in emotion during the course of each 12-month period. The experimental results using the independent test datasets for each tracking run were comparable to those using 10-fold cross-validation as reported above. Again, the affective models did not show predictive capability during the time span in the way they did when predicting at the end of it.

3.4.2 Discussion

We would like to offer three hypotheses on how one might interpret the results from this study of the big players on Twitter:

1. Anger and fear as expressed in big players' tweets is representative of anger and fear within the larger online community.
2. Anger and fear in tweets from the big players have an influence on the emotional state of the community. This may be the desired goal for some big players as high levels of activity on Twitter can indicate the intent to gain an online presence or communicate a specific message to a perceived audience (Boyd et al., 2010; Marwick & Boyd, 2011).

3. Big players and the general community are each reciprocally influencing each other's expressed emotional state with regard to anger and fear. If this hypothesis is accurate, the results may be a reflection of how users tend to interact online mainly with other like-minded individuals (Fersini, 2017; Williams et al., 2015).

The big players may potentially be of interest for researchers in each of these three cases. Our methodology can be used to identify a relatively small number of representative accounts, reducing the amount of time needed to manually evaluate multiple user profiles and to read through numerous tweets. Furthermore, in addition to the set of basic emotions, the affective models incorporate different categories of big player activity. Hence, such models may ultimately prove useful for research seeking to evaluate communication methods representing different types of online participation involving emotion-based framing on social media.

Our results indicate that anger and fear may be two emotions of particular interest for research on climate change. Indeed these two may arguably be the most pertinent of all the emotions we consider in the present work. In Chapter 6 we will see that anger and fear are also meaningful when analyzing communications about global warning on Twitter using our ontological model of key concepts from the Six Americas series of studies.

Affective elements like anger and fear have traditionally been rare in cognitive models for decision making and risk analysis, both of which are central themes when studying how human beings relate to climate change (Leiserowitz, 2006). However, methodologies may be changing somewhat. Recently there have been a number of studies in psychology and sociology that are indeed looking at emotion. More recent, related research (Leiserowitz et al., 2016) makes use of a number of emotions including anger and fear to model peoples' attitudes and beliefs about climate change in the United States. Another example is the work of (Haltinner & Sarathchandra, 2018) which studies how fear can invoke denial and skepticism when people receive information on climate change. A third example is a study by (Nabi et al., 2018), which evaluates the efficacy of messages intended to promote climate change advocacy when these messages are framed around content based on fear (as opposed to hopeful or emotionally neutral content).

Regarding sadness and joy, our models do not show any significant correlation between the big players and the larger community for these emotions. As we repeated the experiments using a number of other machine learning algorithms, the lack of interesting results for these emotions does not appear to be an artifact of linear regression. Continued research regarding the analysis of online expression for these emotions is certainly warranted. Sadness may indeed be a particularly salient emotion as previous research has pointed to it as an important aspect of human reaction to climate change and likely a factor which can influence how international climate policy is shaped (Farbotko & Mcgregor, 2010).

Undoubtedly, affective models for sadness and joy need to incorporate factors which we have not included in this study. These emotions likely represent a more difficult research problem as one might expect that people do not typically express them in a clear, unblurred manner in messages they post about climate change. Irony is one factor which should be examined in the scope of a continued research effort as one study on online communications has demonstrated sadness to be a common emotion in tweets expressing irony (Sulis et al., 2016). The same study finds that joy is frequently expressed in tweets linked to a second complicating factor: sarcasm. We can find an example from a big player in our dataset for this project:

> Greenhouse gasses are good. Just look how plants grow inside a greenhouse. It's wonderful! #ClimateChange #GlobalWarming

Of course, the corresponding emotion vector shows only joy:

$$\langle anger{=}0.0, fear{=}0.0, sadness{=}0.0, joy{=}1.652\rangle$$

It is not hard to imagine how a big player posting such tweets in large quantity will make for a more complicated modelling problem than one where all big players are expressing their emotions in a direct manner in the online content they post. So as our results indicate, we might expect that sadness and joy will prove to be emotions that are relatively difficult to model for the online topic of global warming. For the present study, we are left to speculate as to the extent that ironic and sarcastic microblogs have influenced our affective models.

3.5 Limitations

As always, when working with statistical models we need to remember that finding a given correlation does not mean we understand the causes behind the phenomenon we are investigating. Elements not included in the model may ultimately be responsible in great part for the observed results. For example, the affect expressed on Twitter may fluctuate with climate-related events as well as with factors such as the geographic location, level of education, and even the general state of health of the users (Mitchell et al., 2013). This means that while the big players project represents a means by which to predict emotion levels in tweets from the global warming community on Twitter, we cannot say that the big player activity we have incorporated into the affective models is the cause of any emotion expressed online.

Additionally, as we discussed in Section 1.3, standard NLP tools often have difficulty when tasked with analyzing content from social media. Microblogs are inherently noisy. They are filled with informal, non-standard language that can be problematic for traditional NLP techniques (Farzindar & Inkpen, 2015). Users also may not follow the rules a researcher is expecting them to follow. They may employ the "@" sign, but they are not mentioning another user. They may publish the same tweet repeatedly, or they may retweet another user, but change the original author's message (Boyd et al., 2010). These complexities and many others, known or unknown, should remind us that it is important to exercise a measure of restraint when interpreting the results of an analysis such as the present one.

In addition to any improvements we may make to our NLP methodology, we might also take some additional steps to refine our general strategy with respect to machine learning. As with all the experiments in the present research, our focus has been primarily on the *Say Sũ̃La* architecture and only secondly on the problem we are tasking it to solve. In the introduction to this chapter, we stated that one of the purposes of the big players project was to familiarize ourselves with the tweets we were collecting from Twitter as well as the copious metadata that comes with those tweets. In an initial phase of this project, before we incorporated the methodology of sweeping the 12-month window across 20 months of Twitter data, we conducted an initial analysis of this data. Before running any learner algorithms, we looked over means and standard deviations for the attributes in our initial datasets and created graphs to visually inspect possible trends based on player activity. It is reasonable to expect that by conducting a new analysis which takes into account the structural changes involved in an experiment now covering 320 big-player attribute groupings (4 emotions $\times$ 4 communication categories $\times$ 20 Top-N levels) and their changes over nine tracking runs, we could improve upon the results reported in this chapter. This approach for feature engineering is a good strategy in the general sense: taking an iterative approach to a given machine learning problem whereby we work to understand our data, clean it up and structure it such that a learner can make best use of the inherent data features, run experiments with the learner, and then use the results of the experimentation to repeat the full process and improve on those results (Domingos, 2012). However, we did not follow this iterative strategy to completion for the big player project itself since the goals of this particular study were primarily aimed at (1) getting the architecture to a point where it could collect tweets and use them to perform an experiment with a basic version of its machine learning level and (2) performing a preliminary analysis in order to better understand the data coming from Twitter as well as general tendencies in the expression of emotion from the online community tweeting about climate change.

A final limitation of these experiments, and one that is perhaps more fundamental for any type of analysis of what people are publishing online, is that we are reducing the intricacies and subtle nuances of human expression to a small and relatively simple set of numeric metrics. As we study people through the emotion they express in their online communications, we need always to remember that human cognition is complex, and emotions represent but a part of it. When following research goals such as how to effectively reach people on social media to inform them on the subject of climate change, we should always remain aware that models are necessarily simplifications. The processes of human understanding and human behaviour that we are modelling are ever more complex (Chapman et al., 2017).

3.6 Contributions and Continued Research

This work essentially represents an alternative strategy for emotion mining with regard to the conversation about global warming on Twitter. In some ways this methodology for identifying key users is simpler than the graph-based approaches which are relatively common in current research for social media. Our methodology defines specific categories of online participation and selects users who are most active with respect to these types of activity. The emotion expressed via these communication categories provides a structure for tweet data that may then be analyzed with standard machine learning techniques. Here we have concentrated on linear regression, which has the advantage of being a relatively fast algorithm, and more importantly, one that delivers a model clearly describing the essential elements used to predict (in this case) the emotion expressed online about global warming.

The big players project demonstrates that very high-activity users do not show a universal predictive capacity with respect to the emotion expressed in tweets from the larger global warming community on Twitter. Yet, for the emotions anger and fear, two emotions that are notably relevant to the subject of climate change, these high-activity accounts do indeed exhibit a distinct predictive capability when predicting for the three month span immediately following the period used to train the model. Moreover, the highest correlation is found in groups of users small enough that researchers may manually study these accounts individually as needed for a given research endeavour. For sadness and joy, however, groups of the "Top N" players do not show this same predictive capacity, suggesting that there are additional underlying complexities involved when modelling these two emotions for an analysis of microblogs discussing global warming on social media.

At the beginning of this chapter we mentioned that the big players project was in many ways an independent study since it stands apart from any endeavour to use our architecture to model the activity of a Twitter community in the framework of the Six Americas. The project essentially represents our initial work on *Say* $\mathcal{S}\vec{\iota}$*La*, aimed at creating and validating the base architecture to ensure a sound foundation on which to continue our research. Of course, high-activity users and their online behaviour is an interesting research topic in its own right. When it comes to communicating a given message on an important subject such as climate change, it would undoubtedly be valuable to have a good understanding of the effectiveness associated with proper framing and selection of content versus simply acting as a big player, publishing online posts again and again, possibly with no particular concern as to the quality of the message. Which method best invokes emotion, makes people listen, and gets them engaged? Possibilities abound for continued research in this direction.

In addition to its contribution as an independent project, our work with the big players has also had an influence on our remaining research endeavours. In fact, looking to the high-activity users became an essential strategy throughout the research effort. We shall see examples of this in the chapters to come. Furthermore, the project readily integrates itself into the *Say* $\mathcal{S}\vec{\iota}$*La* application as it represents an initial design and implementation for an emotion module in the machine learning layer of the *Say* $\mathcal{S}\vec{\iota}$*La* architecture as described in Section 0.4. In our continued research we strive to build upon our findings from the big players project and integrate them into the larger scope of this work.

CHAPTER IV

THE SAY-SILA ONTOLOGY

This chapter describes the principle elements of the ontology which we have created to model users and their online communications and which will ultimately serve to aid us in determining the stance on climate change taken by these users as they post their thoughts on the subject on social media. While our preliminary work with the "Big Players" (see Chapter 3) was centred around machine learning techniques, here we use an ontology and description logic to model the conversation about *#globalwarming* on Twitter.

Back in Chapter 2 we presented some of the basic concepts of our research model as part of our general discussion on ontologies and description logics (DL). Here, we build on that foundation, starting by presenting the final two layers on the top of the Semantic Web technology stack shown in Figure 2.2: Tawny OWL and the Java API for OWL. We then begin to cover our ontological model much more in depth, beginning with the top-level ontology upon which it is founded. The ontological model essentially performs two functions. Firstly, it describes the syntactic and affective structure of the tweets we have collected from Twitter on the subject of global warming. It describes the online users who publish these tweets, and it describes important concepts from the Six Americas survey, which may be included in the textual content of the tweets. Secondly, the ontology is used for an analysis of the modelled data so that we may classify users according to their stance with respect to global warming. In this chapter we present the modelling methodology for the tweets and online users. In the following chapter, we shall incorporate the concepts from the Six Americas and the associated analytical elements from the ontology.

4.1 Implementation in Clojure

The top two layers in the architecture in Figure 2.2 depend on the application making use of Semantic Web technologies. The system we have designed for the present research, *Say Sīla*, uses a library known as Tawny-OWL[1] (Lord, 2013) for the Clojure programming language. Tawny OWL builds on the Java API for OWL[2] (Horridge & Bechhofer, 2011). Example code for the *Text* and *OnlineAccount* classes from our ontological model is shown in Figure 4.1. This is the actual programming syntax we use to generate the ontological expressions in XML such as those shown in the examples from Chapter 2. Inside the `defclass` expression in the code, the `:super` keyword allows us to build a model using a hierarchy of classes and will create the necessary *owl:SubClassOf* constructs at the XML level. The *dul* prefix[3] for the *InformationObject* and *SocialObject* classes refers to the Dolce+D&S Ultralite (DUL) ontology, which we discuss in Section 4.2. We may also declare classes as disjoint using the `:disjoint` keyword, which will serve to generate *owl:DisjointClasses* XML expressions such as those shown in the listing in Figure 2.4. Note that the *pos* prefix for the *Token* class refers to a smaller ontology called cmu-pos,[4] developed as part of the present research and imported into the larger say-sila ontology (see Section 4.3). The cmu-pos ontology defines classes for concepts relating to parts of speech for words, emojis and other tokens commonly used in online platforms such as

---

[1]https://github.com/phillord/tawny-owl

[2]https://owlcs.github.io/owlapi/

[3]The identifier to the left of the slash (/) actually refers to the namespace identifying a Clojure module. When programmatically defining an ontology, the Clojure namespace works as an abstraction, providing access to the ontology's namespace (separated by a colon [:] in the XML). At the highest level they both essentially represent the same thing.

[4]http://www.dendrown.net/uqam/cmu-pos.owl#

Twitter. Finally, Tawny-OWL's `defclass` call can also generate XML expressions for a human-readable label, a descriptive comment for the class, as well as other RDF Schema elements.

Figure 4.1 Using Tawny-OWL to define classes in the say-sila ontology.

```
(defclass Text
  :super    dul/InformationObject
  :disjoint pos/Token
  :label    "Text"
  :comment  "An Information Object consisting of text.")
(as-disjoint Text pos/Token)

(defclass OnlineAccount
  :super    dul/SocialObject
  :label    "Online Account"
  :comment  "A user account for an online service.")
```

Modelling using Tawny-OWL is a powerful strategy when creating systems based on ontologies. The developer has access to the design abstractions available in a high-level, functional programming language like Clojure. For example, when creating several classes which all have the same structure, we use Lisp macros[5] to create a `defemotion` call,[6] which generates Tawny-OWL code for the class representing a single emotion. At a higher level, this macro is invoked by a `defemotions` call[7] which generates the code for a full set of emotions such as the base emotions according to Plutchik (Plutchik, 2001) which we are using in this study. We are also free to utilize the lower-level Java API routines when necessary. Figure 4.2 gives an example function from the *Say Sila* application which loads an ontology into the system from disk. The call to Tawny-OWL's *owl-ontology-manager* returns the Java API's *OWLOntologyManager*,[8] which we use to load an ontology document into Java instances usable by the API. Other Java classes

---

[5] Clojure is a member of the Lisp family of programming languages.

[6] https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/sila.clj#L291

[7] https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/sila.clj#L307

[8] *org.semanticweb.owlapi.model.OWLOntologyManager*

like IRI[9] and OWLOntologyID[10] are used seamlessly in the Clojure code where they are needed.

Figure 4.2 Using the Java API for Java with Tawny-OWL in Clojure.

```
(defn load-ontology
  "Reads an ontology from the file system."
  [^String iri
   ^String fpath]
  (let [id  (OWLOntologyID. (IRI/create iri))
        rsc (io/as-file fpath)
        man (owl-ontology-manager)]
    (remove-ontology-maybe id)
    (.loadOntologyFromOntologyDocument man (IRI/create rsc))))
```

As mentioned above, the say-sila ontology imports other ontologies. This allows us to build our model on formalized conceptualizations that essentially represent a submodel that is complete in and of itself (e.g., the part-of-speech concepts in cmu-pos) or a foundational ontological structure which forms the basis for a domain-level ontology such as say-sila. Dolce+D&S Ultralite is such an ontology. In the next section we cover its use with regard to the present research.

4.2 Dolce+D&S Ultralite

As any given ontology is ideally intended to be "a formal, explicit specification of a shared conceptualisation" (Studer et al., 1998), it readily follows that, as much as possible, ontologies should *share* a common foundation. Certainly a level of basic interoperability will exist for ontologies created with a common implementation language, such as OWL 2 DL, but the implementation language does nothing to ensure (or arguably even encourage) a shared representation of the actual knowledge being modelled. Even across various projects from a common domain in research or industry, the ontologies associated with these projects are often incompatible. The information being modelled

---

[9] *org.semanticweb.owlapi.model.IRI*

[10] *org.semanticweb.owlapi.model.OWLOntologyID*

cannot be easily shared, and projects that might readily benefit from an exchange of data must function as essentially independent silos. This is where a top-level ontology comes in. A top-level ontology[11] strives to define the most basic, elemental concepts, which are known to be sound philosophically, so that domain and application ontologies may then define their specific concepts in terms of the fundamental concepts modelled in the top-level ontology (Arp et al., 2015; Borgo & Masolo, 2009).

In an endeavour to facilitate efforts towards consistency and reuse among ontology-based sociological studies on social media, we have created the say-sila ontology using Dolce+D&S Ultralite (DUL)[12] as its top-level ontology. As its name suggests, DUL was derived from the DOLCE ontology (Gangemi et al., 2002), which stands for "Descriptive Ontology for Linguistic and Cognitive Engineering." DOLCE is based on a modal logic of possible worlds and has often been used for modelling in the domain of biology and the social sciences. The ontology resulting from its implementation in OWL is called DOLCE Lite.[13] DUL is essentially an adaptation of DOLCE Lite in order to employ ontological terminology more in line with what is commonly used for projects linked to the Semantic Web. DUL is also "lighter" than DOLCE Lite. Branding it as "ultralite," its creators[14] endeavoured to incorporate only the minimum axiomization necessary to express the principle concepts modelled by DOLCE (Arp et al., 2015; Presutti & Gangemi, 2016).

Additionally, the Dolce+D&S Ultralite ontology integrates a number of other upper-level ontologies that have been established as sound foundations when representing certain common ontological design patterns. The most notable of these is the Description and Situation ontology, indicated by the "D&S" in the name, which serves to model concepts

[11]In the literature these ontologies are also known as *foundational ontologies*, *formal ontologies*, or *upper-level ontologies*.

[12]http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#

[13]http://ontologydesignpatterns.org/ont/dlp/DOLCE-Lite.owl

[14]The Lifecyle of Networked Ontologies (NeOn) EU FP7 Integrated Project.

that may have an inherent level of ambiguity depending on a given situation or context. Other upper-level ontologies incorporated into DUL are: the Plan ontology, the Information Objects ontology, and the Collection ontology (Presutti & Gangemi, 2016). Of these upper-level ontologies integrated into DUL, only the Information Objects ontology is used as part of the present research.

In the remainder of this section we take a look at the parts of DUL and the Information Objects ontology which serve as the foundation for concepts modelled in the say-sila ontology as well as our smaller cmu-pos ontology, which say-sila imports. When the *Say Sĭla* application utilizes these ontologies, it has two modes by which to bring in DUL as the top-level ontology for say-sila and cmu-pos. The first is by way of an *owl:Import* expression, which Tawny-OWL will add as part of the initial information in the extensive *owl:Ontology* node that is the ontological model expressed in XML. Essentially, this one line in the XML causes the system to load the full DUL ontology from the online resource identified by the URI, and all the axioms defined in DUL are available to build upon.

The second mode for bringing in the DUL ontology effectively makes the "Ultralite" ontology even lighter by extracting only the parts of DUL which we utilize in the say-sila and cmu-pos ontologies. These parts include the classes and the object properties (binary roles) needed to describe the structural elements of the tweets and the users who publish them as well as the analytical elements needed for classification and analysis. We are utilizing only 9 of the 76 classes defined in DUL and 5 of the 107 defined object properties.[15] This is a similar strategy to the one taken for the bio-zen plus ontology for medical research, which incorporated only the needed elements of DOLCE into an OWL DL top-level ontology for the project (Samwald & Adlassnig, 2008). After some preliminary testing to verify that the say-sila and cmu-pos ontologies using the minimal DUL base were functionally equivalent to the ontologies when importing all of DUL, we

[15] The Clojure source code that switches between the full import mode and the extraction mode, which recreates only the needed parts of Dolce+D&S Ultralite, can be accessed here: https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/dolce.clj

proceeded using only the minimal (extracted) top-level ontology.[16]

4.3 The CMU-POS Ontology

The cmu-pos ontology is a relatively small ontology we created which models the concepts of tokens and parts of speech adapted for online communications according to the standards associated with the TweetNLP project[17] at Carnegie Mellon University (CMU). In the present research we use cmu-pos only in the context of the main say-sila ontology, which imports it and builds upon it. However, cmu-pos is an independent ontological model which could be used in conjunction with other domain ontologies in other research projects. Like the say-sila ontology, it uses DUL as its top-level ontology.

Arguably the most important class in cmu-pos is the *Token* class, which we use to model the words and other symbolic entities (e.g., punctuation or emojis) in the text of a microblog, which we have collected from Twitter and included in our dataset (see Sections 5.1.1 and 5.2). Figure 4.3 presents the class hierarchy for *Token*.[18] The class *owl:Thing* represents the top concept ($\top$) in OWL 2 DL. The arcs with the arrows correspond to *owl:SubClassOf* declarations. Effectively, these are the DL concept inclusion

---

[16]The verification process involved ensuring that an inductive learning tool known as DL-Learner (Bühmann et al., 2016) generated identical learned class expressions for test runs using for the two sets of ontologies after populating them with individuals (see Section 5.1.1).

[17]https://www.cs.cmu.edu/~ark/TweetNLP/

[18]The visual representations of the ontology were created using Stanford's OntoGraph plugin for the Protégé ontology editor. Note that OntoGraph uses a "has subclass" relation to denote class hierarchy, and so the arrows point in the opposite direction as is customary in UML (Unified Modelling Language) and as would be indicated by the *rdfs:subClassOf* role.

OntoGraph: https://protegewiki.stanford.edu/wiki/OntoGraf
Protégé: https://protege.stanford.edu/

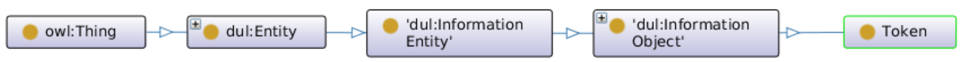
axioms to define *Token*:

$$dul{:}InformationEntity \sqsubseteq dul{:}Entity \tag{4.1}$$

$$dul{:}InformationObject \sqsubseteq dul{:}InformationEntity \tag{4.2}$$

$$Token \sqsubseteq dul{:}InformationObject \tag{4.3}$$

Figure 4.3 Class hierarchy for the *Token* class in the cmu-pos ontology.



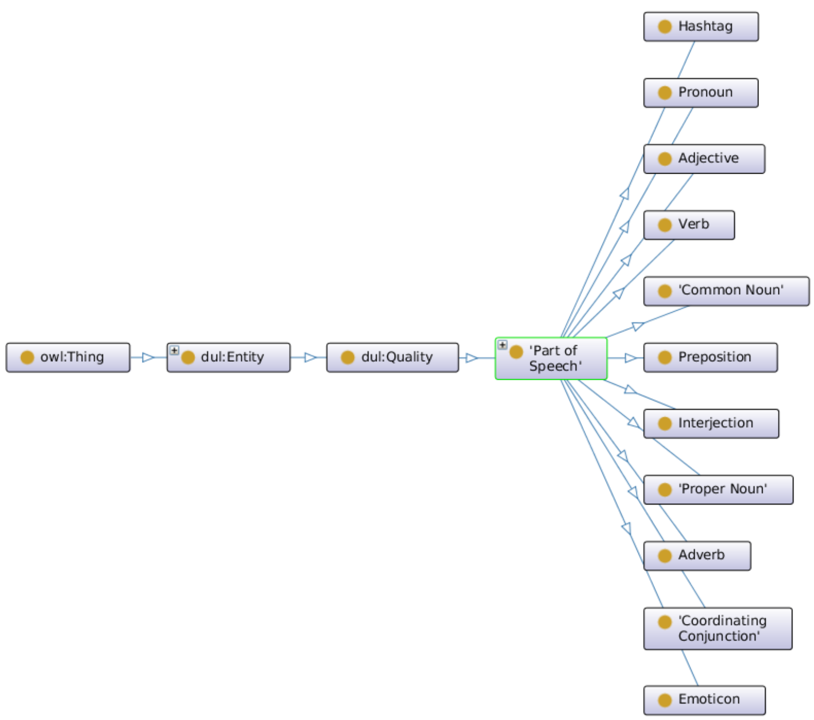
All concepts modelled using DUL are subsumed by *dul:Entity*.[19] The DUL ontology describes an entity as "[a]nything: real, possible, or imaginary, which some modeller wants to talk about for some purpose." Many of the things we are modelling in this work are information objects. As mentioned above, the Information Objects ontology is included as part of DUL. The concept of an information entity encompasses forms of knowledge and communication be they physical or digital, words or byte code, drawings or photographs (Gangemi & Peroni, 2016). The list of possible examples may be endless, but in our work we are modelling the essential elements of microblogs. In this case the element is a token in an online text, which is an information object. This parent class, *dul:InformationObject*, simply represents something that holds information with no specification as to what information is held, nor how it is held. Child classes such as *Token* establish these types of details.

Our model also builds on *dul:Quality*, which represents characteristics of entities which are not an actual part of the entity, but which also do not exist in and of themselves without the entity (Presutti & Gangemi, 2016). In the cmu-pos ontology we define one direct child class of *dul:Quality*, which is *PartOfSpeech*. This class in turn subsumes

[19]The concept of an entity is the root concept for the full DUL ontology, not just the portion we extracted for our model.

several classes representing the different parts of speech that may be associated with a given token. Figure 4.4 shows a subset of these classes in relation to their place in the concept taxonomy of the cmu-pos ontology. The parts of speech in the model correspond to those created specifically for NLP on Twitter as part of the TweetNLP project (Owoputi et al., 2013) at Carnegie Mellon University. The name of the ontology, cmu-pos, comes from this system of parts of speech (POS) from this NLP project at CMU. Table 4.1 presents the full set of part of speech classes defined in the ontology. Note that in addition to standard grammatical parts of speech, the set of modelled concepts includes online-specific elements such as web URLs, email addresses, emoticons, and hashtags.

Figure 4.4 Class hierarchy for classes representing parts of speech in cmu-pos.



The relationship between an individual token and its part of speech is modelled by the *isPartOfSpeech* role, which is a subproperty of the DUL role which relates an entity and

Table 4.1 Parts of speech concepts from the cmu-pos ontology.

Class	Tag	Description
Address	U	URLs and email addresses
Adjective	A	adjectives
Adverb	R	adverbs
AtMention	@	at-mentions (indicating another user in the text of a tweet)
CommonNoun	N	common nouns
Continuation	˜	multi-tweet text indicator
CoordinatingConjunction	&	coordinating conjunctions
Determiner	D	determiners
Emoticon	E	emoticons
ExistentialPlusVerbal	Y	usage of existential *there* with a verbal
ExistentialThere	X	usage of existential *there*, predeterminers
Hashtag	#	a hashtag to set the topic in a tweet
Interjection	!	interjections
NominalPossessive	S	usage of nominal + possessive
NominalVerbal	L	usage of nominal + verbal
Numeral	$	numerals
Other	G	abbreviations, foreign words, possessive endings, symbols, etc.
Preposition	P	pre- or postpositions and subordinating conjunctions
Pronoun	O	personal pronouns
ProperNoun	ˆ	proper nouns
ProperNounPlusPossessive	Z	proper nouns with a possessive
ProperNounPlusVerbal	M	proper nouns with a verbal
Punctuation	,	punctuation
Verb	V	verbs including copula and auxiliaries
VerbParticle	T	verb particles

Adapted from (Gimpel et al., 2011)

a quality:

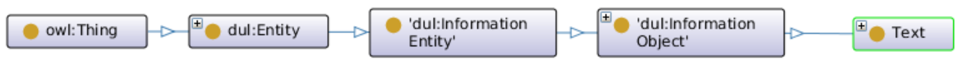
$$isPartOfSpeech \sqsubseteq dul{:}hasQuality \tag{4.4}$$

As might be expected, the domain (subject) of *isPartOfSpeech* is *Token*, while the range (object) is *PartOfSpeech*.

4.4 The Say-Sila Ontology

The say-sila ontology is the main model for the present research. It builds upon DUL and imports the cmu-pos ontology described above. This ontology is relatively large, compared to cmu-pos. We have already seen many of the basic concepts from this ontology back in Chapter 2 when we discussed the theory of ontologies and description logics. In this section we present these concepts more completely in order to explain how the say-sila ontology builds on DUL and cmu-pos. Note that there are other concepts modelled in this ontology, which are more intricately tied to our analysis of the Twitter data. We will describe these later in Section 5.1.

Figure 4.5 Class hierarchy for the *Text* class in the say-sila ontology.

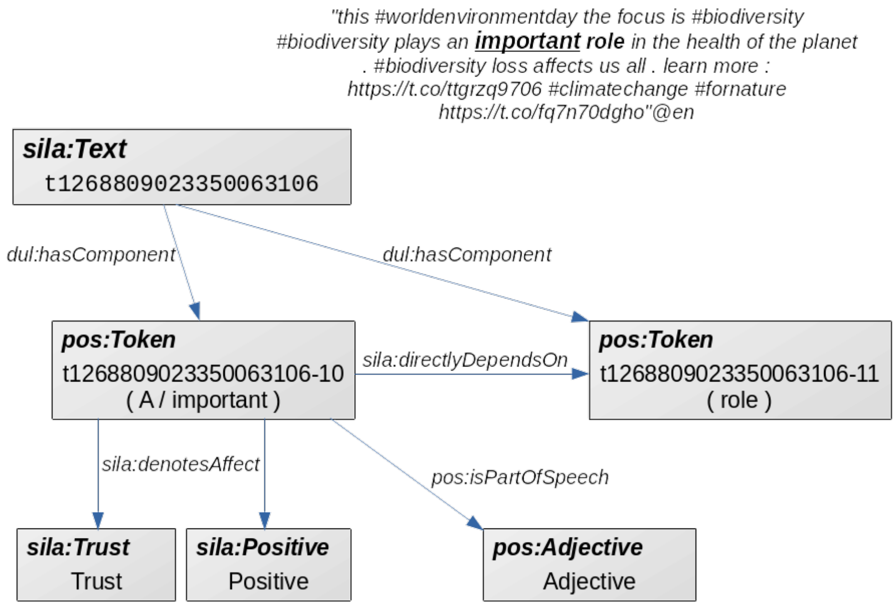


The *Text* class is central to the say-sila ontology. Instances of tweet texts are members of this class.[20] Figure 4.5 shows the hierarchy for the *Text* class. Like *pos:Token*, it is an information object. *Text* and *pos:Token* are declared disjoint in the ontology (see Figure 4.1). To model the relationship between a tweet and the tokens it contains, we use the relationship *dul:hasComponent*. In DUL both the domain and the range of this role are declared to be of type *dul:Entity*. When we model a tweet, we declare several instantiations of this role where each time the subject is the tweet text entity (labelled according to its Twitter identifier), and the objects are the tokens making up the textual content. Figure 4.6 shows the DL relationships in the model between an individual tweet (Twitter ID 1268809023350063106) from the IPCC and the 10$^{\text{th}}$ ("important") and 11$^{\text{th}}$ ("role") tokens in its textual content. We can also see how the ontology models that "important" is an adjective using the *pos:isPartOfSpeech* role.

In Figure 4.6 the class names are in the upper left of the boxes, and the individuals' names are centred toward the bottom of the boxes. Note that for *Adjective* in the cmu-pos ontology, we are using OWL 2 DL punning so that *Adjective* will be interpreted as a class or as an individual as is appropriate for a given DL construct. We also use punning for affect in the say-sila ontology as evidenced in the objects (*Trust* and *Positive*) to the *sila:denotesAffect* role. The full hierarchy for *Affect* in the ontology is given in Figure 4.7.

[20] An initial version of the ontology included child classes for *Tweet* and *Retweet*, which were subsumed by *Text*. We removed them in the current version in an effort to lighten the ontology. As we are not modelling other forms of text in the present research, this level of distinction was not necessary. However, these child classes may be reintegrated into the ontology as part of our continued research.

Figure 4.6 Portion of model for an individual tweet in the say-sila ontology.

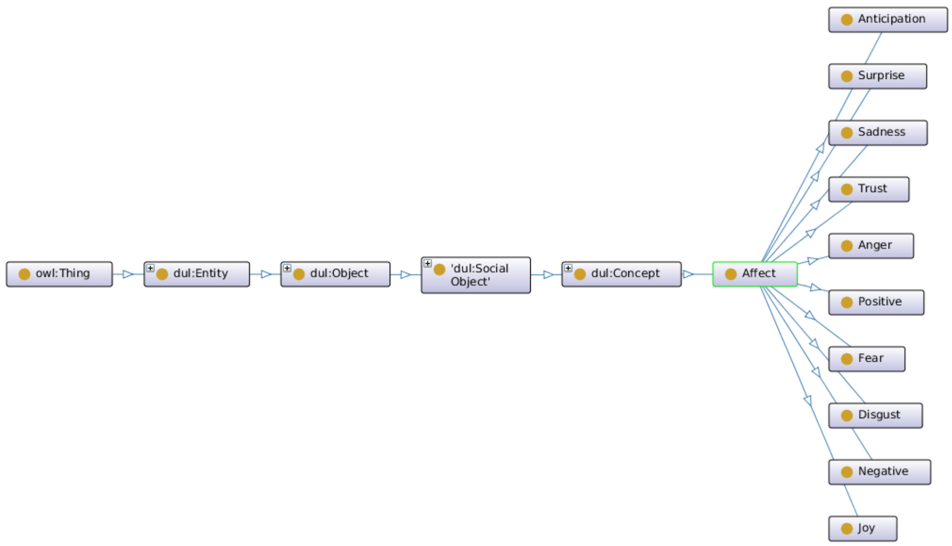


We have modelled the concept of affect building off of *dul:Concept*.[21] Travelling down the chain of the hierarchy from the top DUL class *dul:Entity* (anything modelled), we have *dul:Object*, which includes "[a]ny physical, social, or mental object, or a substance" and *dul:SocialObject*, which is an object that requires some form of communication to exist (Presutti & Gangemi, 2016). Subclasses for *Affect* include concepts for positive and negative sentiment polarity, as well as the full set of basic emotions from Plutchik's system (Plutchik, 2001).

Not least in importance, the say-sila ontology also includes a class for the users on social media who have authored the microblogs in the model (e.g., the Twitter account for the IPCC). Figure 4.8 presents the hierarchy for the *OnlineAccount* class. Note that although DUL includes classes for agents and persons, we chose to model online accounts as social objects so that in future research we may differentiate the account from the person or the organization who controls it. Finally, *OnlineAccount* includes two child classes, *GreenAccount* and *DenierAccount*, whose members are the user accounts identified as

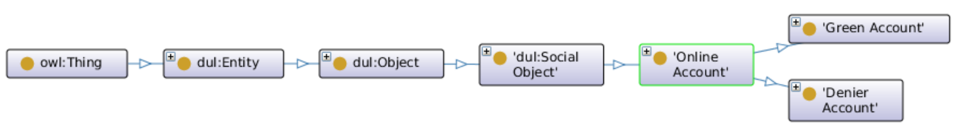
[21] The *dul:Concept* class is often used in conjunction with the Descriptions and Situations portion of DUL; however, our current model is not using this design pattern.

Figure 4.7 Class hierarchy for the *Affect* class in the say-sila ontology.



belonging respectively to the green (*alarmed/concerned* from the Six Americas) or the denier (*doubtful/dismissive*) categories as described in Section 2.1.

Figure 4.8 Class hierarchy for the *OnlineAccount* class in the say-sila ontology.



The classes described above form the basic conceptual elements of the say-sila ontology. Note, however, that there are several additional classes in the ontology which are intended for the analysis of data collected from Twitter. We will look at these parts of the model in the next chapter.

4.5 Towards a Descriptive and Analytical Ontological Model

Here we have covered the portion of our ontological model intended to describe the online users and their communications on Twitter which we will be analyzing as part of this research effort. In essence what we have created up to this point is similar to existing research utilizing ontologies for the automated analysis of sentiment in text (Salguero & Espinilla, 2016; Salguero & Espinilla, 2018), specifically those analyzing sentiment in microblogs on Twitter (Cotfas et al., 2015). Our model differs from these examples in that it builds on an established top-level ontology, Dolce+D&S Ultralite, which ensures a sound ontological foundation and also promotes reuse of both the cmu-pos and the say-sila ontologies, particularly with regard to other research efforts in the human sciences.[22] In the next chapter we will take our ontological model a step further, using it to link concepts from a socio-psychological study, the Six Americas, to the modelled online communications we have collected from Twitter. We will also discuss the analytical elements defined in our ontology by which we may use description logic to infer the stance of the users with respect to climate change.

[22]Note that the model created by (Cotfas et al., 2015) makes use of W3C-recommended ontologies for the Semantic Web. This methodology also facilitates reuse.

CHAPTER V

THE SIX AMERICAS ON TWITTER: ONTOLOGICAL MODEL

Our discussion of the say-sila ontology spans three chapters in the present work. In Chapter 2 we got a preliminary glimpse of the ontology as we looked at the general theory of description logics. Later, in Chapter 4 we covered the basic elements of the say-sila ontology as well as the cmu-pos ontology it imports so that we may model users on Twitter and the content of their tweets. In the present chapter we delve into elements in the ontology which allow for an analysis of these modelled tweets and online users using important concepts[1] from the Six Americas survey.

As mentioned previously, we have two objectives with respect to our ontological model. The model must be descriptive such that it describes users on Twitter and their tweets, and it must be analytical such that it maps these users to specific categories based on the Six Americas project. Modelling conceptual elements related to the Six Americas survey in the say-sila ontology essentially represents a bridge between these two objectives. In this chapter we cover elements in the model describing syntactic and semantic dependencies occurring in the microblogs collected from Twitter. We are endeavouring to use these modelled dependencies to map social media texts to specific concepts covered in the Six Americas survey questionnaire (Maibach et al., 2011).

---

[1] In this chapter to avoid ambiguity between a *concept* as content from the survey and a *concept* in description logic, our general use of the word *concept* means the former. We shall use the term *class* (the implementation of a DL concept [here in OWL 2]) when referring to our ontological model.

5.1 Survey Concepts and Analytical Elements

Our underlying methodology for this stage of the research is largely an iterative process. To an extent, this process has been inspired by the methodology behind the open-source, climate-change toolkit of the EU Decarbonet project (Maynard & Bontcheva, 2015) in that the analysis of tweets and the comparison to an authoritative reference is partially algorithmic and partially manual.

We create the elements of syntactic and semantic dependency in the ontological model by iteratively comparing the textual content of the tweets (to determine what is being said) to the questions of the Six Americas survey (to determine what indicates a given stance on climate change). What exactly does it mean to iteratively compare microblogs to survey questions? In the scope of the present research it involves modelling the occurrence and the co-occurrence of key words in tweets as well as the syntactic dependency between the words if such a dependency exists.[2] These key words are words which are indicative of opinions on climate change. We must select words representing important concepts from the survey questions in the Six Americas study (Leiserowitz et al., 2010), choosing only the concepts that are actually being discussed frequently on Twitter. We also incorporate WordNet (Fellbaum, 1998) so that the model includes words from tweets which are synonymous with the words being used to represent our targeted concepts. In essence we must determine what part of the tweets and what part of the Six Americas we need to model in order to infer a user's position on the subject of climate change.

For the present research our iterative process entails that data is required to complete the ontological model. We have compiled a dataset from collected tweets, which we have named the *survey concept dataset*. Its purpose is to aid us in determining what concepts from the Six Americas are being talked about on Twitter and which should therefore be modelled in the say-sila ontology. The general idea is that the tweets and consequently

[2]We are using relationships of syntactic dependency according to the output of the TweeboParser NLP tool for social media (Kong et al., 2014), which we describe in Section 5.1.2.

the users publishing them will fall into specific categories of survey concepts, e.g., a user account publishing one or more tweets expressing the concept “human-caused.” These survey concept categories may then serve as indicators in order to infer a user’s stance on climate change.

Note that the survey concept dataset is used only as the final part of the process of model creation. Afterwards, we use a separate dataset with the same attribute characteristics to perform the experimental analysis (see Section 5.2). As our methodology is an incremental one, before looking at the DL definitions of the elements describing survey concepts and the analytical parts of the model, it is worthwhile to understand how the ontology is populated with the individual texts, tokens, and online users that will later be analyzed. The survey concept dataset is an integral part of this process.

5.1.1 Creating a Survey Concept Tweet Dataset

What we have designated the survey concept dataset has a very specific purpose. It is not intended to capture the general online conversation about climate change.[3] The objective behind the survey concept dataset is essentially to link the conversation on Twitter with key concepts from the Six Americas. We created the dataset using the Twitter developer platform API,[4] collecting English-language tweets with the hashtags *#climatechange* and *#globalwarming* published between October 1, 2019 and June 30, 2020. Although other hashtags may also be pertinent to this this topic, (Williams et al., 2015) found these to be among the most general and commonly used in a study comparing 27 hashtags relating to climate. This effort yielded 1,452,241 raw tweets from 403,748 users.

How do we use this data to link the online conversation to the Six Americas? Ultimately

[3]This will be the purpose of the “2019 *#globalwarming*” tweet dataset used for analysis as described in Section 5.2.

[4]https://developer.twitter.com/

our goal is to determine which of the Six Americas categories best classifies a given user on Twitter. As explained in Section 0.3, to remain within the scope of the present research we map six categories onto two (green and denier) and limit ourselves to determining which of these two best represents a user with respect to her stance on climate change. Thus, our task is to find tweets in the dataset where users may be indicating which stance they have taken. We look to the questions themselves in the Six Americas survey (Leiserowitz et al., 2010) for key concepts that may serve to indicate a person's stance on climate change. Helping to determine people's views on this divisive issue is, after all, the primary purpose of the survey.

Of course, these concepts must be part of the conversation on Twitter to be useful for identifying the stance of its users on global warming. The survey concept dataset is intended to reveal the concepts being discussed, but the tweets it contains should ideally be those most indicative of a person's stance on climate change. Therefore, we choose to focus specifically on the subject of environmental conservation as this theme was found to be the most significant predictor of a person's stance with respect to climate change based on a meta-analysis by (McCright et al., 2016), which analyzes 87 peer-reviewed empirical studies from 1998 to 2016. Accordingly, we filtered the tweets using a caseless search with a regular expression[5] compared against the full text of the tweet. We formed this regular expression using stemmed prefixes of the keywords "environmental" and "conservation" as well as pertinent synonyms and sister terms (the immediate hypernym and its immediate hyponyms) as reported by WordNet (Fellbaum, 1998).[6] This process led to an intermediate, unlabelled dataset of 162,274 tweets from 62,555 users (11.2% of the full set of collected tweets and 15.5% of the associated users).[7]

---

[5]This is the PCRE (Perl Compatible Regular Expression) we used for filtering the tweets: `"environment|[[:<:]]env[[:>:]]|[[:<:]]conserv|[[:<:]]preserv|[[:<:]]sav(e|es|ing)[[:>:]]"`.

[6]https://wordnet.princeton.edu/

[7]Note that tweet collection for our test dataset (see Section 5.2) does not include this filter for the topic of environmentalism. While we argue that the filtering is specifically intended to be part of the methodology for creating the ontology, we do recognize that there is an inherent lack of parallelism at

As we go through the details of the methodology for the creation of the survey concept dataset, it might be helpful to recall that in addition to describing users and microblogs on Twitter, the ontological model has a second objective, which is to analyze the stance of online users with respect to climate change. Specifically, we would like to determine if someone tweeting about global warming should be represented in the ontology by a *GreenAccount* (corresponding to the *alarmed/concerned* from the Six Americas) or a *DenierAccount* (corresponding to the *doubtful/dismissive*). Unfortunately, the tweets we have collected are unlabelled with respect to these account categories, and to our knowledge no dataset is available for an analysis of this type from within the scientific community. In order to evaluate how well the model is performing the analytical task of classifying user stance on climate change, we need a method of labelling our dataset.

Our process leverages two independent methods of determining which users are in the green category and which are in the denier category. One allows us to label users, and the other represents the ontological model's method of classification. The first method involves follower relationships on Twitter. We call this method "follow-the-leader," and we employ it as a ground truth with respect to a user's stance on climate change. We describe the method in this section as it is an integral part of the methodology used to create the survey concept dataset. The second method to determine user category involves logical inference using weak and strong indicators from the model. We cover the second method in Sections 5.1.4 and 5.1.5.

For the follow-the-leader method we again used Twitter's development API to identify which users are following a set of known "leader" accounts with a high in-degree (number of followers) and a clear stance on the subject of climate change. To choose leader accounts we manually examined whom the highest-activity users are following on Twitter. The 269 high-activity users selected for this purpose (4.7% of the users in the survey

---

the testing stage. We discuss this further in Section 5.4.1.

concept dataset) are those publishing a minimum of 20 tweets in the dataset.[8] We selected users whose tweets were easily discerned by a human evaluator as green or denier and examined the accounts these users were following, searching for a small set of leader accounts run by known organizations or individuals with 10,000 followers or more.

Why do we consider these leader accounts important for determining the stance of the online users who follow them? Research indicates that humans take cues as to their opinions about a controversial subject such as climate change from others who are considered to be in an elite position and whom they trust (Bohr, 2014; Krosnick, 2000). The online leader accounts represent an established stance on climate change. They are the accounts of organizations, politicians, and other well-known entities who are widely followed on Twitter and who make public their views with regard to the science of climate change. Additionally, a leader account's position on climate change is trivially verifiable by having a human examine the account's profile and the tweets it publishes.[9] Table 5.1 shows our chosen leader accounts and their approximate number of followers as of November 2020. This general method of using leader accounts in order to determine a position for users of social media on a given theme has been used in previous studies. Examples include identifying Christians and atheists on Twitter by utilizing followers of public figures known to be Christian or atheist (Ritter et al., 2014) and also identifying Republicans and Democrats in the United States by including followers of well-known politicians registered to one of those parties (Sylwester & Purver, 2015).

Ideally, we would label as many users as possible for the survey concept dataset. Doing so is not practical, however, due to rate limitations for queries against the Twitter

---

[8] As checking followers was a largely manual, labour-intensive process, we had to limit ourselves to a relatively small subset. Accordingly, we simply chose the group with the highest minimum participation level that we consider in the study.

[9] What we mean by "trivially verifiable" is that there is text in the account's profile or in several tweets in line with these examples: "Standing with women and girls on the frontlines of the climate emergency and working towards climate justice for all" (CAREClimate: green) or "Climate Change/Global Warming has NOTHING to do with Man Made CO2" (ClimateRealists: denier).

Table 5.1 Green and denier leader accounts.

	Account	Name	Followers
Green	UNFCCC*	UN Climate Change	758.9k
	350	350 dot org	392.5k
	IPCC_CH*	IPCC	223k
	FAOclimate*	FAO Climate Change	143.8k
	ClimateChangeUS	ClimateChangeUS	17.9k
	CBI_CC*	CBI climate change	13k
	UKClimateEnvoy	Nick Bridge	13.1k
	CAREClimate	CARE Climate Change	11.4k
Denier	BreitbartNews	Breitbart News	1.6M
	EPAAWheeler	EPA Administrator Andrew Wheeler	77.7k
	ClimateRealists	Climate Realists (Ugly Folk)	48.5k
	ClimateDepot	Marc Morano	25.7k
	HeartlandInst*	The Heartland Institute	15.8k
	CFACT	CFACT	15.k

* denotes accounts in the survey concept dataset

developer API and the large number of users publishing only a single tweet. For these reasons and taking into account the indications we found that relatively active users may be more representative of their general community on Twitter (see Chapter 3), we limit the users considered for inclusion in the survey concept dataset to those who have published two or more tweets from the filtered dataset. Hence, we removed the 45,027 users who published only one tweet (72.0% of the filtered dataset) and proceeded to query the Twitter API to determine which leaders the remaining users were following. We discarded 11,649 users (18.6%) who were not following any of the leader accounts or whose account had been cancelled or suspended by Twitter. Additionally, there were 151 users following one or more leader accounts from both the green and the denier list. We manually checked each user's profile description and several of their tweets. When it was trivially verifiable by a human evaluator that the user belonged to the green or denier category,[10] we assigned the user to the appropriate category. When there was ambiguity, we discarded the user from the dataset. Of the users following both types of leader accounts, 27 were green, 5 were deniers, and 119 were discarded (0.2%).[11]

[10]Please refer to note 9 for an explanation of what we mean by "trivially verifiable"

[11]Although we included these 32 users in the final survey concept dataset, this small number of accounts should be negligible. For future research it may be beneficial to skip the work involved in

The final survey concept dataset contains 44,038 tweets from 5,760 users (27.1% of the filtered tweets from 9.2% of the associated users). As might be expected, given that we are working with a theme of environmentalism here, our resulting dataset is heavily imbalanced. Of these final users 5,337 (92.7%) are in the green category, and only 423 (7.3%) are in the denier category. It may be interesting to note that by labelling our dataset according to who is following the leader accounts, we are essentially performing an analysis of the followers of the leaders. These leader accounts do not necessarily have tweets on the subject of environmental conservation in the survey concept dataset, but the accounts marked with an asterisk (*) in Table 5.1 have at least two tweets in the dataset.

Although creating a fully annotated dataset is beyond the scope of the present research, we did perform an initial verification of our follow-the-leader method by randomly sampling 200 active accounts from the 2019 *#globalwarming* test dataset (described in Section 5.2) and having a human annotator classify these users as being in the green or the denier category based on their Twitter profile and original tweets they have published. Of these 200 samples, the human and our system's results from the follow-the-leader method were in agreement for 195 (97.5%) of the accounts. Of course, given this high level of accuracy for the automated follow-the-leader analysis, one might question why we do not simply use this as our final strategy for determining the stance of Twitter users, rather than creating a full architecture based on an ontology and description logics to make the same determination. The answer is that the goal of our architecture goes beyond the green–denier classification problem that we have chosen for this doctoral program. This research represents our first steps towards a generalized architecture that is capable of modelling other studies in the human sciences, studies for which a follow-the-leader methodology will not necessarily generate the answer we are seeking. Furthermore, the ontological model in *Say $S\vec{\imath}la$*'s description logic level can be used in the scope of other research efforts which utilize Semantic Web technologies, allowing the

manually checking so many accounts in order to focus efforts on other aspects of the study.

results of an analysis using our architecture to be available to applications beyond the original project for which the ontology was created. With these long-term objectives in mind, we are using the follow-the-leader method here to generate the category labels needed to validate our approach for the Twitter users considered in this study.

5.1.2 Natural Language Processing Tools

Our goal is to create an ontological model so that we may perform a limited semantic analysis of the tweets in the dataset described above. We process the tweet text using the Affective Tweets library (Mohammad & Bravo-Marquez, 2017), available as a plugin for the machine learning platform Weka (Frank et al., 2016). This NLP tool suite includes a tokenizer and part of speech (POS) tagger (Gimpel et al., 2011), trained on Twitter data and specifically intended for working with tweets. The tags for the parts of speech correspond to those described in Section 4.3. The plugin can facilitate the preprocessing of tweets by removing tokens according to a list of stop words (common words devoid of analytical value).[12] It also allows the selection of a stemming algorithm to convert words to their grammatical roots. When comparing tweet content to ontological concepts, we stem tokens using the Snowball Porter stemmer (Porter, 2006) for the English language.[13]

We also make use of the TweeboParser (Kong et al., 2014) to determine dependencies between the tokens in the tweet text. TweeboParser is a graph-based syntactic dependency parser, based in turn on TurboParser (Martins et al., 2013), which employs an integer linear programming model to score candidate first-order parse trees for a given sentence. In a first-order parse tree the arcs factor over two consecutive tokens only. Second- and third-order trees allow multiple siblings and grandparent/grandchild relations between tokens; however, generating higher-order, non-projective parse trees is

---

[12]We use the English stop word list from Apache's information retrieval package Lucene: https://lucene.apache.org/

[13]http://snowball.tartarus.org/

generally an NP-hard problem. Non-projective parsers do not require dependency arcs to be nested, and so they are necessary for languages where word order is relatively flexible. TurboParser extends its first-order models using a method known as alternating directions dual decomposition (AD^3) (Martins et al., 2011), which approximates an ideal parse tree with second- and third-order features by splitting the higher-order parsing task into local subproblems and using an augmented Lagrangian optimization technique to solve them.

TweeboParser (Kong et al., 2014) seeks to augment TurboParser's effectiveness for NLP in the world of social media, specifically when one is working with text messages published on Twitter. One of the changes that TweeboParser incorporates is the use of hierarchical clusters of words, known as Brown clusters, to aid in part-of-speech tagging (Brown et al., 1992; Owoputi et al., 2013). Note that TweeboParser uses the same coarse POS mappings as the Affective Tweets (Mohammad & Bravo-Marquez, 2017) library for Weka. Another notable change is that in an effort to minimize any issues stemming from a new parsing methodology, TweeboParser uses information adapted from the established Penn Treebank (Marcus et al., 1993) to be a factor in the parsing of a tweet.

While the Penn Treebank represents a well-accepted example of a large, high-quality syntactically-annotated corpus,[14] (Kong et al., 2014) argue that the standardized language of the texts serves to limit its effectiveness for training parsers intended to process language as it is commonly used on social media. The extent to which the Penn Treebank data influences how tweets are parsed is determined by TweeboParser's learning algorithm. The published training dataset is made up of a relatively small set of tweets (929 tweets containing 12,318 tokens), the goal being to keep up with the dynamic nature of Twitter communications by allowing "imperfectly-trained" teams of annotators to quickly produce small training datasets. The annotation style for TweeboParser training

[14]The texts making up the extensive Penn Treebank corpus undergo an initial automated parsing phase, the results of which are then corrected by trained human annotators.

input data is generally based on the conventions proposed by (Schneider et al., 2013), which closely follow the Yamada-Matsumoto conventions (Yamada & Matsumoto, 2003) used in their work on statistical dependency parsing. These conventions also aimed for rapid creation of training datasets using annotators who are not necessarily experts in linguistics. Tweebo-style annotations (and thus parse trees) do not include punctuation, topic-indicator hashtags, and other tokens not contributing to the syntactic structure of the text.

The annotation conventions presented by (Schneider et al., 2013) and used as the base style for TweeboParser focus on multi-word lexical units.[15] A given text may not necessarily be parsed into a tree with a single root, but rather a set of tree fragments in which multiple roots give access to the different lexical units. The annotation system aims to be lightweight while handling dependency relations such as coordinating conjunctions and conjuncts[16] as well as representing relationships between anaphora and their antecedents. Generally, the idea is to "underspecify" syntactic information during the annotative process, thereby allowing for some level of uncertainty from the annotator, ambiguity in the text, or flexibility in the constraints of a given project.

As Tweebo is a graph-based dependency parser, it does not parse texts according to a formal grammar. Rather, it follows a statistical methodology, using machine learning to create a scoring function based on the human-annotated training corpus known as the "Tweebank" (Kong et al., 2014).[17] In essence this scoring function is used to generate probability scores for candidate parse trees in order to find the tree which most closely aligns with the dependency representations inherent in the training corpus (Nivre, 2010). It should be noted with respect to our current work that we would likely

[15]The say-sila ontology models the reported multi-word lexical units; however, they are not currently used in the analysis of tweets.

[16]These elements are also included in the ontological model to capture the syntactic structure of a tweet, but as of yet are not utilized when performing an analysis.

[17]The Tweebank corpus is available at http://www.ark.cs.cmu.edu/TweetNLP.

expect some level of improvement in our results by training the TweeboParser with a human-annotated corpus created from tweets which specifically discuss the topic of climate change. However, for our initial analysis we are using the default TweeboParser, trained on the Tweebank corpus. We leave for future research efforts the questions concerning experimentation using a TweeboParser more finely tuned for communications about global warming on social media.

Tweebo outputs dependency information as tab-separated values in the format associated with the dependency parsing shared task from the tenth Conference on Computational Natural Language Learning (CoNLL-X) (Buchholz & Marsi, 2006). To aid our iterative process, our system converts the TweeboParser output into a text-based visual tree. Figure 5.1 shows this tree for an example tweet:[18]

> *@BreitbartNews It's about the fools errand of SAVING the PLANET from itself and it's closed co2 loop million year ClimateCycle. #ClimateChange*

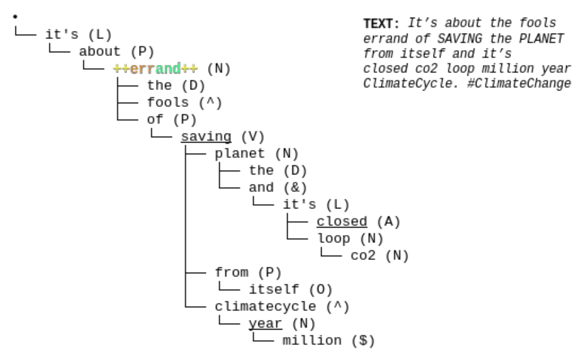


Figure 5.1 Dependency tree for a tweet from an account in the denier category.

Back in Section 2.2 we described the ontological roles *dependsOn* and *directlyDependsOn* as defined in the say-sila ontology. We use these roles in the ontological model to describe the syntactical dependencies between the tokens in the text (see Figure 4.6) as determined by the TweeboParser output for a given tweet. At the bottom of Figure 5.1, for

[18]Underlined tokens are keywords in the Six Americas survey. The codes in parentheses indicate the token's part of speech. The colours denote emotion (anticipation and trust), while the plus signs (++) indicate positive sentiment.

example, "million" directly depends on "year," and "year" directly depends on "climatecycle." The two roles are hierarchical with *dependsOn* subsuming *directlyDependsOn*, and so the model also represents the knowledge that the former token depends on the latter in both cases. Of course, *dependsOn* is a transitive role, while *directlyDependsOn* is not, Therefore, we can use the model to infer that "million" depends on "climatecycle" (but not directly).

5.1.3 Survey Concept Rules

Our goal is to use the syntactic dependencies identified by the Tweebo Parser to create a number of semantic indicators[19] in the model which function as markers for talk about certain concepts related to climate change. To this end we turn to the questions from the Six Americas survey (Leiserowitz et al., 2010). Our first step is to determine which parts of the questionnaire Twitter users are talking about frequently. Starting with the full text of the tweets, we removed stop words and created a list in descending order of usage counts for the stemmed tokens in the intermediate (unlabelled) tweet dataset. Table 5.2 shows the tokens from the list that are used at least 150 times.[20] The tokens displayed in bold face represent concepts found in the questions from the Six Americas survey which we have selected for the ontological model in this research. Note that we have removed tokens relating to terms used to filter for tweets on the subject of environmental conservation. We also disregard words which necessarily refer to a very large number of questions in the survey such as *climate*, *change*, *global*, *warming*, as well as *people*.

We are aiming to identify concepts that are discussed often, but in the say-sila ontology we will be modelling the co-occurrence of words and relations of syntactic dependency

[19] As discussed in Section 2.3, we realize some researchers may object to the term *semantic* in reference to Semantic Web technologies. An alternative term might be "topic indicators."

[20] The count for a given token includes both its normal and its #hashtag form.

Table 5.2 Stemmed tokens by decreasing frequency.

Tokens	Frequency
energi	7643
natur	4889
new	4395
world	4371
pollut	3728
compani	2811
us	2562
news	2498
carbon	2408
daili	2389
govern	2374
help	2096
support	1941
protect	1709
year	1880
polit	1806
water	1720
futur	1704
plant	1674
emiss	1648
time	1617
recycl	1601
action	1579
human	1525
medium	1489
share	1319
now	1316
peopl	1298
renew	1296
blog	1226
solar	1218
take	1205
inform	1193
scientist	1095
interest	1046
electr	1027
work	989
issu	980
oil	966
economi	957
build	943
power	933
econom	925
research	897
report	895
reduc	862
home	780
india	775
democrat	768
caus	765
speci	765
polici	736
air	735
commun	726
anim	711
money	705
industri	702
effect	689
respons	688
today	688
watch	660
state	636
show	612
light	565
event	563
increas	554
point	552
wind	542
american	538
gas	535
nation	532
record	531
organ	527
bill	528
countri	528
problem	523
stori	523
start	522
clean	514
job	505
visit	507
follow	497
weather	486
singl	485
leader	481
intern	475
expert	443
effort	432
sign	423
import	485
activ	435
develop	423
small	379
discuss	367
onlin	369
produc	369
public	363
movement	355
generat	348
fund	387
china	366
trust	321
friend	320
winter	320
realli	319
cost	318
step	312
coast	397
hope	309
happen	296
meet	298
contribut	290
provid	290
children	366
vote	379
product	365
greenhous	316
tax	311
panel	268
bad	263
buy	257
u.s.	251
trash	244
sourc	242
drive	223
presid	221
local	219
media	217
current	207
letter	200
journal	209
area	195
differ	193
requir	191
high	191
never	190
growth	199
regist	196
rain	194
urg	172
parti	169
view	169
listen	167
person	163
network	163
york	163
prioriti	157
unit	154
post	154

between two words. Therefore, we are ultimately looking for questions in the Six Americas survey that refer to a pair of concepts, both of which are invoked by words used relatively frequently in the tweets from the survey concept dataset.[21] Although the Porter stemmer (Porter, 2006) is generally quite consistent as it reduces various word forms to their corresponding linguistic stems, we can see in the case of "economy" (stem: *economi*) and "economic" (stem: *econom*) that discrepancies can indeed occur.[22] This is not especially problematic for this stage of the present research as we are manually looking through the list generated by the system for pairs of concepts that relate to a question from the Six Americas. In the case of *economi* and *econom*, we simply consider these two stems to be one concept. In our case, we found a survey question in which they may be paired with *growth* (though this concept is tweeted less frequently) to form a pair we call "economic–growth." Note that many stems from Table 5.2, although popular in tweets and present in the survey, did not readily pair up with a second stem. For the present research it is important that the concepts be paired. Effectively, we are seeking a number of these concept pairs that may indicate either a green or a denier stance when incorporated into the ontological model.

We build on the *dul:Concept* class from the Dolce+D&S Ultralite (DUL) top-level ontology (Presutti & Gangemi, 2016) to model what we are calling a survey concept rule (SCR). The term "rule" simply indicates that the concept must be put to use, in our case by its inclusion in a tweet. We declare an SCR for important concepts from the Six Americas survey (Leiserowitz et al., 2010) being discussed in the tweets from the intermediate dataset, identified by the token frequency list (Table 5.2).

[21]The procedure involved manually consulting the list in Table 5.2 while searching for pairs of these terms in the questions from the Six Americas survey (Leiserowitz et al., 2010). Standardizing this process, generalizing it for other studies in the human sciences, and ideally automating it would constitute a worthwhile project for our continued research.

[22]Replacing stemming with lemmatization as a preprocessing step would help to minimize this issue at the cost of increased processing time and computational resources.

$$SurveyConceptRule \sqsubseteq dul{:}Concept \tag{5.1}$$
$$SurveyConceptRule\text{–}HUMAN \sqsubseteq dul{:}SurveyConceptRule \tag{5.2}$$
$$SurveyConceptRule\text{–}NATURE \sqsubseteq dul{:}SurveyConceptRule \tag{5.3}$$
$$SurveyConceptRule\text{–}CAUSE \sqsubseteq dul{:}SurveyConceptRule \tag{5.4}$$
$$SurveyConceptRule\text{–}CO2 \sqsubseteq dul{:}SurveyConceptRule \tag{5.5}$$
$$SurveyConceptRule\text{–}CUT \sqsubseteq dul{:}SurveyConceptRule \tag{5.6}$$
$$SurveyConceptRule\text{–}ECONOMIC \sqsubseteq dul{:}SurveyConceptRule \tag{5.7}$$
$$SurveyConceptRule\text{–}GROWTH \sqsubseteq dul{:}SurveyConceptRule \tag{5.8}$$

In the say-sila ontology we define tokens that indicate a given SCR using an *indicatesRule* object property, which is a subrole of DUL's *expresses* and has a domain of *pos:Token* and a range of *SurveyConceptRule*. Concept tokens are named simply using the main word representing an SCR concept (e.g., *HumanToken*).

$$indicatesRule \sqsubseteq dul{:}expresses \tag{5.9}$$
$$HumanToken \equiv pos{:}Token \sqcap \exists indicatesRule.SurveyConceptRule\text{–}HUMAN \tag{5.10}$$
$$CauseToken \equiv pos{:}Token \sqcap \exists indicatesRule.SurveyConceptRule\text{–}CAUSE \tag{5.11}$$

We define the rest of the concept tokens in the same manner. Note that when processing a dataset and deciding whether a given input token (i.e., the actual word from a tweet's text) should be added as an individual of a concept token class, our system consults WordNet (Fellbaum, 1998) and includes the synonyms reported for the appropriate synsets.[23] The synsets we chose for our four survey concepts are listed in Table 5.3. There were two concept tokens where we found that the synsets defined in WordNet did not produce synonyms that adequately covered words likely to be used in tweets. The

[23]The *Say Sila* application uses the Clojure WordNet library to access semantic relations: https://github.com/clojusc/wordnet

Note that WordNet also allows us to find other semantic relations, such as those for "type of" (hyponym/hypernym) and "part of" (meronym/holonym) (Bratko, 2012), which we hope to investigate in the scope of our continued research. Our code which pulls synonyms from WordNet is linked here: https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/wordnet.clj#L124

Table 5.3 WordNet synsets for detecting Six Americas survey concepts

Concept	Synset	Gloss
cause	SID-07341157-N	events that provide the generative force that is the origin of something; “they are trying to determine the cause of the crash”
	SID-00007347-N	any entity that produces an effect or is responsible for events or results
	SID-01649143-V	give rise to; cause to happen or occur, not always intentionally; “cause a commotion”; “make a stir”; “cause an accident”
	SID-00772482-V	cause to do; cause to act in a specified manner; “The ads induced me to buy a VCR”; “My children finally got me to buy a computer”; My wife made me buy a new sofa”
human	SID-02474924-N	any living or extinct member of the family Hominidae characterized by superior intelligence, articulate speech, and erect carriage
	SID-02754015-A	characteristic of humanity; “human nature”
	SID-02754145-A	relating to a person; “the experiment was conducted on 6 monkeys and 2 human subjects’"
	SID-01261689-A	"having human form or attributes as opposed to those of animals or divine beings; “human beings”; “the human body”; “human kindness”; “human frailty”
nature	SID-09389659-N	the natural physical world including plants and animals and landscapes etc.; “they tried to preserve nature as they found it”
CO_2	SID-14821139-N	a heavy odorless colorless gas formed during respiration and by the decomposition of organic substances; absorbed from the air by plants in photosynthesis
carbon	SID-14657384-N	an abundant nonmetallic tetravalent element occurring in three allotropic forms: amorphous carbon and graphite and diamond; occurs in all organic compounds
cut	SID-00430013-V	cut down on; make a reduction in; “reduce your daily fat intake”; “The employer wants to cut back health benefits”
	SID-00244786-V	reduce in scope while retaining essential elements; “The manuscript must be shortened”
regulate	SID-00300122-V	fix or adjust the time, amount, degree, or rate of; “regulate the temperature”; “modulate the pitch”
	SID-02517217-V	bring into conformity with rules or principles or usage; impose regulations; “We cannot regulate the way people dress”; “This town likes to regulate”
	SID-00702806-V	shape or influence; give direction to; “experience often determines ability”; “mold public opinion”
	SID-00235689-V	restrain the emission of (sound, fluid, etc.)
economic	SID-02727475-A	of or relating to an economy, the system of production and management of material wealth; “economic growth”; “aspects of social, political, and economical life”
	SID-02587892-A	concerned with worldly necessities of life (especially money); “he wrote the book primarily for economic reasons”; “gave up the large house for economic reasons”; “in economic terms they are very privileged”
growth	SID-13518338-N	a process of becoming larger or longer or more numerous or more important; “the increase in unemployment”; “the growth of population”

Extracted from WordNet 3.1 (Fellbaum, 1998) first of these is CO_2, which we combined with the appropriate synset for *carbon*.[24] The second concept is *cut*, which we have combined with *regulate*.[25]

When the architecture populates the say-sila ontology, it compares the grammatical stem of the input token with those of the concept word and its synonyms. If there is a match, we insert an individual of the appropriate concept token type into the ontology. Otherwise, we insert an individual of the general type *pos:Token*.

5.1.4 Weak Indicator Accounts

With the concept tokens defined, we now have what we need in the ontology to define the first set of what we call indicator classes: the weak texts. A weak indicator text is a tweet that has two concept tokens as components, indicating a pair of survey concept rules. The pairs represent concepts that are intrinsically linked in the questionnaire for the Six Americas (Leiserowitz et al., 2010) to indicate important issues (e.g., “economic growth”) related to a person’s stance on climate change. Table 5.4 lists the SCRs we are using in the present work. The classes tagged with a dagger (†) indicate issues that may more likely be embraced by people in categories of the Six Americas which lean towards climate change denial. The last column denotes the reference table in the description of the Six Americas questionnaire (Leiserowitz et al., 2010) which includes the specified pair of concepts.

The definition in the ontology for texts that contain a weak indicator that there should be a reduction (cut) of CO_2 is as follows:

---

[24] WordNet’s synonyms for “CO2” are “carbonic acid gas” and “carbon dioxide.” The synonyms for “carbon” (per synset SID-14657384-N) are “C” and “atomic number 6” (though our purpose was simply to include “carbon.”)

[25] Synonyms for *cut* using our chosen synsets are “trim,” “shorten,” “abbreviate,” “foreshorten,” “trim down,” “cut down,” “trim back,” “abridge,” “contract,” “reduce,” “bring down," and “cut back." Synonyms for *regulate* are “baffle,” “influence,” “shape,” “mold,” “modulate,” “determine,” “order,” “regularise,” and “govern.”

Table 5.4 Text classes based on pairs of survey concept rules from the Six Americas.

Class	Concept Token A	Concept Token B	Six Americas Survey Ref.
$WeakHumanCauseText$	$HumanToken$	$CauseToken$	T5
$WeakNatureCauseText^{\dagger}$	$NatureToken$	$CauseToken$	T5
$WeakCO2CutText$	$CO2Token$	$CutToken$	T25
$WeakEconomicGrowthText^{\dagger}$	$EconomicToken$	$GrowthToken$	T11

† denotes a denier SCR pair

$$CO2Token \equiv pos{:}Token \sqcap \exists indicatesRule.SurveyConceptRule\text{–}CO2 \tag{5.12}$$

$$CutToken \equiv pos{:}Token \sqcap \exists indicatesRule.SurveyConceptRule\text{–}Cut \tag{5.13}$$

$$\begin{aligned} WeakCO2CutText \equiv\ & Text \\ & \sqcap \exists dul{:}hasComponent.CO2Token \\ & \sqcap \exists dul{:}hasComponent.CutToken \end{aligned} \tag{5.14}$$

The other weak indicator texts are defined in the same manner. Note that for these weak indicator texts, there is no requirement for a relationship of syntactic dependency between the two concept tokens. They merely have to both be components of the tweet text. This is the reason we call them *weak* indicator texts.

Continuing along this line, we may now define a set of classes known as weak indicator accounts. Simply put, a weak indicator account represents the account of an online user who has published one or more weak indicator texts for a given SCR pair. In the say-sila ontology we define an online account as a subclass of *dul:SocialObject*. We also define the role *publishes* with a domain of *OnlineAccount* and a range of *dul:InformationObject*. As an example, here is the definition of the weak indicator account for the “human–cause” SCR pair.

$$OnlineAccount \sqsubseteq dul{:}SocialObject \tag{5.15}$$

$$\begin{aligned} WeakHumanCauseAccount \equiv\ & OnlineAccount \\ & \sqcap \exists publishes.WeakHumanCauseText \end{aligned} \tag{5.16}$$

Again, the weak indicator accounts for the remaining SCR pairs in Table 5.4 are defined according to the same pattern. Now we are at the point where we can group the weak indicator accounts according to the users' green or denier stance, based on which of the leader accounts they were following on Twitter. As it identifies an account's stance via this follow-the-leader method, the system creates each individual online account in the ontology explicitly as a *GreenAccount* or a *DenierAccount* (one individual for each of the 5,760 in the survey concept dataset) according to whom these users are following.

$$GreenAccount \sqsubseteq OnlineAccount \tag{5.17}$$
$$DenierAccount \sqsubseteq OnlineAccount \tag{5.18}$$
$$GreenWeakHumanCauseAccount \equiv GreenAccount \sqcap WeakHumanCauseAccount \tag{5.19}$$
$$DenierWeakHumanCauseAccount \equiv DenierAccount \sqcap WeakHumanCauseAccount \tag{5.20}$$

Individuals inferred to be in the "green weak human–cause account" class (1) have been found to be in the green stance category based on the leader accounts they follow and (2) have published one or more tweets which contain tokens related to the concepts "human" and "cause" with no requirement that the tokens have any syntactic dependency relationship beyond their both being components of the microblog. Likewise, individuals inferred to be in the "denier weak human–cause account" class have published weak human–cause texts, but they have been found to be in the denier category due to to the leader accounts they are following. Figure 5.2 displays the part of the ontology relevant to the weak human–cause accounts.[26] Of course, there are corresponding fragments in the ontology for the weak indicator accounts representing each of the SCR pairs from Table 5.4.

Finally, we define two types of weak classes to identify accounts with an inferred green stance on climate change. These two weak indicator classes are defined as follows:

---

[26]The figure includes hierarchical subsumption up to the first class defined in DUL.

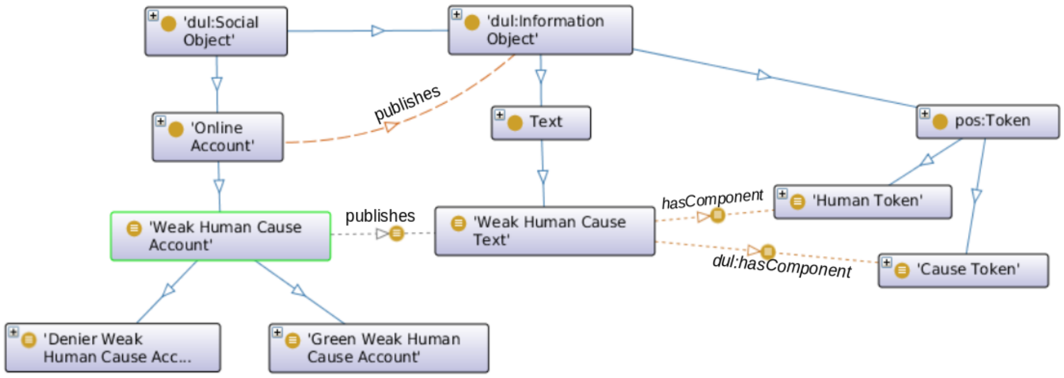


Figure 5.2 Say-sila ontology fragment for the weak human–cause indicator.

$$\begin{aligned} WeakInferredGreenAccount \equiv\ & WeakHumanCauseAccount \\ & \sqcup WeakCO2CutAccount \end{aligned} \tag{5.21}$$

$$\begin{aligned} WeakInferredGreenAccountPlus \equiv\ & WeakHumanCauseAccount \\ & \sqcup WeakNatureCauseAccount \\ & \sqcup WeakCO2CutAccount \\ & \sqcup WeakEconomicGrowthAccount \end{aligned} \tag{5.22}$$

The *WeakInferredGreenAccount* class represents the union of accounts indicating the SCR pairs from this study which describe concepts more often believed by people in the green categories in the Six Americas (the SCR pairs without a dagger (†) in Table 5.4). The *WeakInferredGreenAccountPlus* is simply the disjunction for all the SCRs considered in the study.[27] The "plus" in the name essentially means that the class also encompasses indicator accounts associated with non-green (denier) SCR pairs. In Section 5.3 we explain our reasons for including this "plus" variant for weak inferred green accounts in our model.

[27]Note that while we do not expressly declare it, *WeakInferredGreenAccount* is a subclass of *WeakInferredGreenAccountPlus*. An OWL reasoner such as HermiT (Glimm et al., 2014) will infer this relation.

5.1.5 Strong Indicator Accounts

Up until now we have been describing weak indicators, but as might be expected, for each weak indicator text and account class, our ontology includes corresponding strong indicator classes. To be a weak indicator text, as explained above, a tweet need only have the two SCR concept tokens as components with no requirement for any relationship of syntactic dependency between them. However, for every tweet in the dataset, we run a dependency analysis using the TweeboParser NLP tool (Kong et al., 2014). To be a strong indicator text for a given SCR pair, one concept token must depend on the other as declared by the output dependency tree created by this parser. We represent these dependencies in the ontology by defining the role *dependsOn* (the domain and range of which are both *dul:Entity*).

Recall that Table 5.4 listed A and B concept tokens for each of the subclasses of *Text* associated with an SCR pair. When inserting the individuals representing the two tokens in a dependency relationship, the system declares one token (usually token A) using a new type which specifically represents the dependency, while the other token simply has the basic concept token type as described above. For example, if token A is "carbon" and token B is "reduce" (synonyms of "CO_2" and "cut" respectively), and the TweeboParser has found that token A depends on token B, then token A will be of type *CO2CutTokenAB*. Token B will simply be typed as *CutToken*. Here are the token dependency definitions for the SCR pair "CO_2–cut" which we have been using as an example. The definitions associated with the other SCR pairs follow the same pattern.

$$CO2CutTokenAB \equiv CO2Token \sqcap \exists dependsOn.CutToken \tag{5.23}$$

$$CO2CutTokenBA \equiv CutToken \sqcap \exists dependsOn.CO2Token \tag{5.24}$$

In other words, the *CO2CutTokenAB* and *CO2CutTokenBA* classes essentially represent a sort of specialty token whose existence indicates that another token in the text exists and, according to the parser output, creates a dependency relationship with respect to a given SCR pair. It may be worth noting that TweeboParser dependencies ("carbon" depends on "reduce") do not necessarily reflect word order in the tweet ("we must reduce our carbon footprint"). In Table 5.4 and in the ontology, we have ordered the concepts A and B such that "A depends on B" represents normal language usage (rather than "B depends on A") for each SCR pair.[28] A strong indicator text class then simply describes a text which has one of these dependency tokens as a component. That token implies the existence of the other token on which it is dependent. Continuing with the example of the "CO_2–cut" SCR pair, the strong indicator text class is defined as follows:

$$\begin{aligned} StrongCO2CutText \equiv\ & Text \\ & \sqcap\ dul{:}hasComponent.(CO2CutTokenAB \sqcup CO2CutTokenBA) \end{aligned} \tag{5.25}$$

The classes for strong indicator texts for the remaining SCR pairs in Table 5.4 are defined in the same manner.

As mentioned above, the definitions for the strong indicator accounts directly parallel those for the weak indicator accounts. The same is true of the definitions used to group the strong indicator accounts by stance category (green or denier), determined by checking which leader accounts a given user is following. The definitions in the ontology which model these account classes for the "CO_2–cut" SCR pair are:

$$\begin{aligned} StrongCO2CutAccount \equiv\ & OnlineAccount \\ & \sqcap\ \exists publishes.StrongCO2CutText \end{aligned} \tag{5.26}$$

$$GreenStrongCO2CutAccount \equiv GreenAccount \sqcap StrongCO2CutAccount \tag{5.27}$$

$$DenierStrongCO2CutAccount \equiv DenierAccount \sqcap StrongCO2CutAccount \tag{5.28}$$

[28] In the course of this research we did separately compare results for strong indicator texts when token A depends on token B (most common) as well as when B depends on A (relatively rare). However, in the present document we only present results for texts with a strong relation as a dependency in either direction.

Again, these classes for the other SCR pairs in this study are defined in the same manner. Figure 5.3 presents the associated part of the ontology.

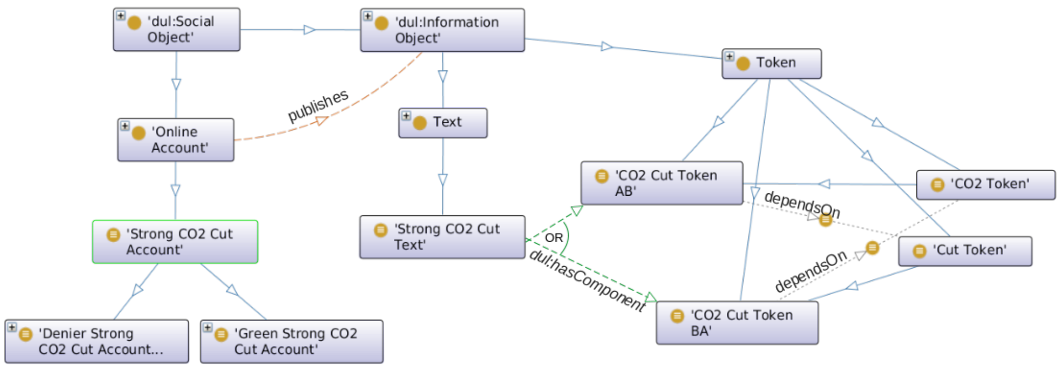


Figure 5.3 Ontology fragment for the strong CO2–cut indicator.

The definition of the final strong inferred account classes also directly parallels the weak inferred account classes. *StrongInferredGreenAccount* is a disjunction of only the indicator accounts directly corresponding to a green stance in the Six Americas survey (Leiserowitz et al., 2010), while *StrongInferredGreenAccountPlus* is a disjunction of all the indicator accounts examined in the present research.[29]

$$\begin{aligned} StrongInferredGreenAccount \equiv\ & StrongHumanCauseAccount \\ & \sqcup StrongCO2CutAccount \qquad (5.29) \\ StrongInferredGreenAccountPlus \equiv\ & StrongHumanCauseAccount \\ & \sqcup StrongNatureCauseAccount \\ & \sqcup StrongCO2CutAccount \\ & \sqcup StrongEconomicGrowthAccount \qquad (5.30) \end{aligned}$$

In the following section we describe how we can use this model to compare the individuals covered by the inferred accounts with those in the *GreenAccount* and *DenierAccount*

[29]Again, Section 5.3 explains why we are using both green and denier indicators to infer which accounts are in the green category.

classes, declared explicitly according to the leader accounts a given user is following.

5.2 The 2019 Global Warming Tweet Dataset

As discussed in Section 5.1.1, we generated two datasets containing tweets collected from Twitter. Whereas the first, the survey concept dataset, was used to identify the survey concept rules for the analytical elements of the say-sila ontology, the second dataset contains the microblogs representing the general online conversation about global warming on Twitter. We use this dataset to perform the actual analysis for this study. We call it the "2019 *#globalwarming*" dataset as it includes tweets with the hashtag *#globalwarming* collected for the full year from January 1 to December 31, 2019. The structure of this dataset is identical to that of the survey concept dataset. However, in addition to covering a different time frame, our criteria for tweet inclusion is slightly different. Specifically, while the survey concept dataset included tweets with the hashtags *#climatechange* and *#globalwarming* that passed a filter for the topic of environmental conservation (an indicator of one's stance on climate change), the 2019 dataset for our analysis simply includes all tweets tagged with the hashtag *#globalwarming* over that year.

We recognize that a number of readers may argue that the inclusion criteria should have been the same for both the dataset used to create the analytical elements of the model and the one used to perform the analysis. We shall address these concerns directly in Section 5.4.1. However, the reason for the difference in criteria stems from the fact that there are two separate objectives for which we created these two datasets. For the survey concept dataset the focus is on the completion of the final model. The intent is to identify aspects of the Six Americas survey that may be used to help distinguish Twitter users in the green and denier categories. We are essentially seeking to extract the threads of the online conversation on climate change which target these distinguishing factors. For the 2019 *#globalwarming* dataset the focus is on the online users themselves. We want a dataset that better reflects the general conversation on global warming as well as a dataset that is relatively balanced with respect to the green and denier user categories. Hence, we include tweets with the *#globalwarming* hashtag for the same reason we did for our analysis of the "Big Players" in Section 3.2. This hashtag has been shown to be more commonly used by both the pro-science (green) and skeptic (denier) communities on Twitter (Williams et al., 2015).

As was the case with the survey concept dataset, the 2019 global warming dataset includes only users who have published at least two tweets (with the hashtag *#globalwarming*) over the course of the year. When performing the analysis for this part of the research, we repeat experiments multiple times, each time using a portion of the 2019 global warming dataset wherein included users publish a minimum of N tweets with N ranging from 2 to 20 (see Section 5.3). When $N = 2$, we are using the entire dataset. This approach has two benefits. First of all, it allows us to progressively analyze only the most active users in the online discourse and potentially identify trends as the minimum level of participation varies. Secondly, there is a practical benefit to progressing "backwards," starting with an N of 20 and working our way down to 2. This strategy allows us to begin experimentation using a relatively small subset of the full dataset, representing the users with the highest activity levels. We can begin running the experiments for these users, even as the time-intensive process of dataset creation continues for the rest of the users who are less active. Then incrementally, we repeat the analysis, each time with more users. The time and computing resources needed to produce the TweeboParser dependency output for each tweet is significant.[30] Additionally, Twitter limits the rate at which one may query its developer API,[31] and so checking leader accounts for thousands of users also takes a substantial amount of time. By starting with a subset of users who have published at least 20 tweets from the dataset and

[30]On an AMD FX-8370 server running at 4.0 GHZ the process requires approximately 10 GB of memory and parses about 2-4 tweets/minute, depending on the number of tokens.

[31]See: https://developer.twitter.com/en/docs/twitter-api/rate-limits

then repeating the analysis for successively larger groups, we can essentially verify that our process is functioning as expected using a relatively small subset of users. Once we have confirmed the model and the methodology are correct, we continue, progressively considering larger sets of users who are tweeting less frequently.

Table 5.5 Tweet and user counts for 2019 dataset.

Min. Tweets	2019 Counts					
	Tweets	Users	Greens		Deniers	
2	20,856	2,376	1529	64.35%	847	35.65%
3	18,816	1,347	886	65.78%	461	34.22%
4	17,550	927	612	66.02%	315	33.98%
5	16,662	705	457	64.82%	248	35.18%
6	15,947	561	365	65.06%	196	34.94%
7	15,401	468	305	65.17%	163	34.83%
8	14,869	392	253	64.54%	139	35.46%
9	14,485	345	223	64.64%	122	35.36%
10	14,107	302	195	64.57%	107	35.43%
11	13,807	271	173	63.84%	98	36.16%
12	13,466	240	155	64.58%	85	35.42%
13	13,178	216	142	65.74%	74	34.26%
14	12,931	198	132	66.67%	66	33.33%
15	12,735	183	123	67.21%	60	32.79%
16	12,465	166	110	66.27%	56	33.73%
17	12,241	151	100	66.23%	51	33.77%
18	12,003	137	94	68.61%	43	31.39%
19	11,877	130	89	68.46%	41	31.54%
20	11,668	119	82	68.91%	37	31.09%

Table 5.5 shows the number of users for each minimum tweet count as well as the number of tweets these users have published in the dataset. It also records how many of the users have been labelled as green or denier accounts according to the follow-the-leader labelling strategy. Figure 5.4 gives a graphical representation of the user (left y-axis) and tweet (right y-axis) counts at each minimum tweet level. Figure 5.5 shows the user counts again along with the number of accounts labelled as being in the green or denier categories.

For each activity level we find roughly two thirds of the accounts in the green category and one third in the denier category. The minimum green and maximum denier percentages occur at $N = 11$ (63.8% vs. 36.2%), while the maximum green and minimum denier percentages are at $N = 20$ (68.9% vs. 31.1%). We note that for high levels of N, the tendency is to have fewer accounts labelled as denier. While still unbalanced, the

Figure 5.4 User and tweet counts for the 2019 dataset.

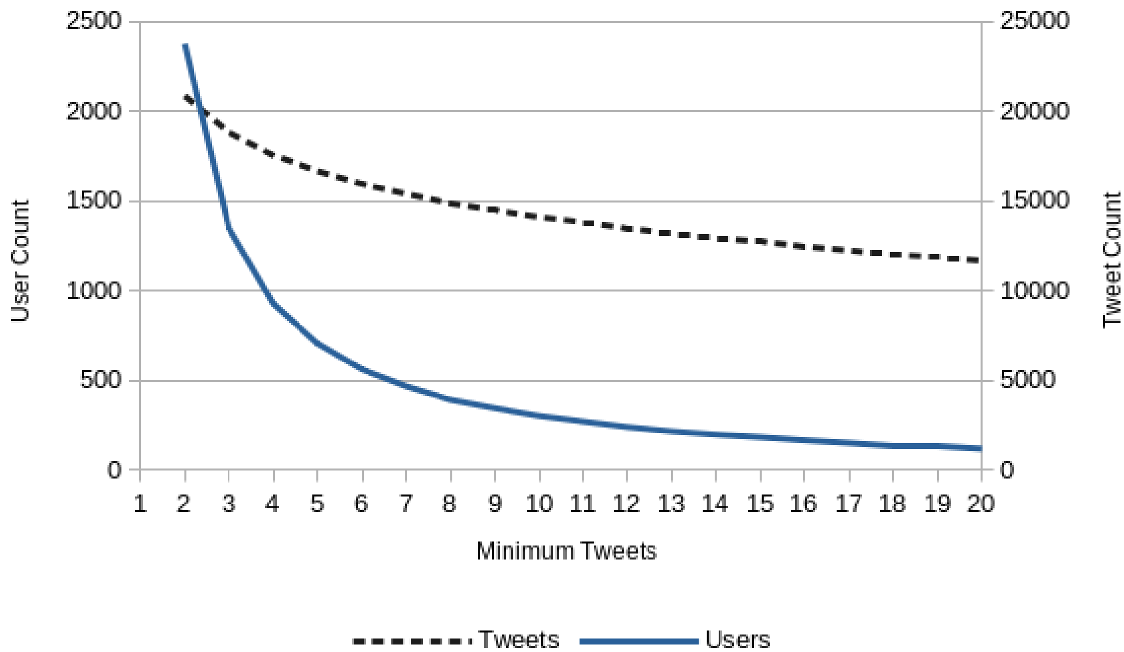
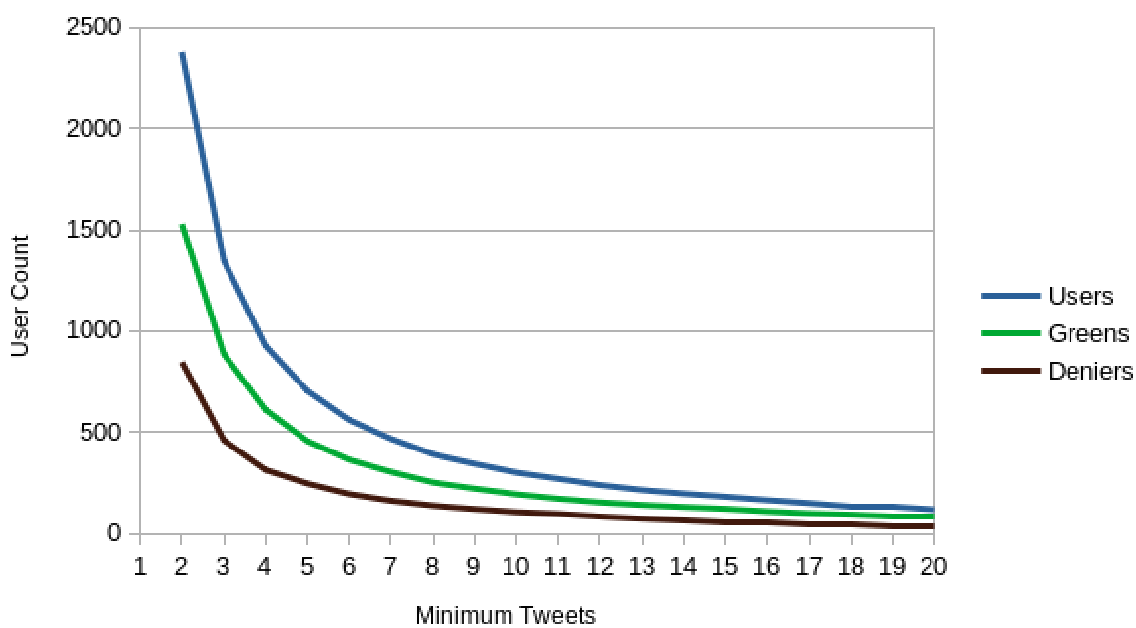


Figure 5.5 User counts for green and denier categories in the 2019 dataset.



ratio is not nearly as extreme as was the case with the survey concept dataset where the number of accounts labelled as denier never surpassed 7.4% at any activity level.[32] Perhaps more importantly, given that we are working with an online community which has not been adjusted for any set of demographics, the proportion of accounts for the green and denier categories in our 2019 *#globalwarming* dataset is arguably a fair reflection of the percentages found in the December 2020 results of the Six Americas survey (Leiserowitz et al., 2021a) as reported back in Table 0.1. In the actual survey, the *alarmed* and *concerned* categories (which together map to our green category) totalled 55% of the respondents, whereas the *doubtful* and *dismissive* categories (which map to our denier category) totalled 20%.[33]

5.3 Inferring Stance from Tweets Linked to Survey Concepts

We performed a series of experiments to investigate to what extent we can determine the category of users (green or denier) using the weak and the strong inferred account indicator classes in the say-sila ontological model. Once again, for the purposes of this analysis we model an account's stance according to which of the known green or denier leader accounts the user is following as described in Section 5.1.1. In the sections which follow we use precision and recall to quantify how well the weak (co-occurrence of survey concepts) and strong (syntactical dependency of survey concepts) indicators perform at determining the stance category of users with respect to each of the four pairs of survey concept rules from Table 5.4.

Precision and recall are information metrics used to measure how well a model performs

---

[32]While our methodology defined two distinct objectives for the two datasets, we recognize that the imbalance in the survey concept dataset suggests some level of bias with respect to the SCR pairs incorporated into the model. In Section 5.4.1 we discuss the question of parallel datasets for the creation of the model and for the analysis stage.

[33]The remaining 25% from the Six Americas December 2020 study comes from the *cautious* and the *disengaged* categories, which we do not model in the present research.

the task of classifying a set of examples. If we are interested in identifying users in the green category, then to calculate the precision and recall, we must consider the following sets of individuals as generated by our model:

1. *true positives* (TP) users correctly inferred to be in the green category
2. *false positives* (FP) users the model inferred to be green but who are really in the denier category.
3. *false negatives* (FN) users the model did not infer to be green but who *are* actually in the green category.

Precision is then defined as:

$$precision := \frac{TP}{TP + FP} \tag{5.31}$$

while recall is defined as:

$$recall := \frac{TP}{TP + FN} \tag{5.32}$$

Recall is also known as the *true positive rate*, which makes sense when we note that $TP + FN$ simply represents all the individuals who are in the green category (Witten & Frank, 2005).

With these metrics in mind, our hypotheses at the start of this work are the following:

- **1A**: the green SCR pairs, "human–cause" and "CO_2–cut," will serve to identify users in the green category.
- **1B**: the denier SCR pairs, "nature–cause" and "economic–growth," will serve to identify users in the denier category.
- **2A**: weak indicators will identify more users of a given category but with decreased accuracy, resulting in relatively higher recall and lower precision
- **2B**: strong indicators will infer user categories with better accuracy, but will miss many users, resulting in relatively higher precision and lower recall.

It became evident quite early in the experimental process that hypothesis 1B was incorrect. We mention this finding now because it makes the presentation of our results much more intuitive if we simplify the methodology such that we are always using the model to identify users in the green category.[34] The data as presented will still show coverage of the accounts labelled as denier for each SCR considered, but the calculations for precision and recall will be with respect to green accounts. We also provide the calculated F1 measure, which is the harmonic mean of precision and recall (Ye et al., 2012). It represents a single metric by which to score an SCR indicator for a given minimum activity level. The F1 measure is calculated as follows:

$$F1 := \frac{2 \times recall \times precision}{recall + precision} \tag{5.33}$$

Accordingly, for each SCR pair we run the experiment with minimum activity levels ranging from 2 to 20 tweets.

5.3.1 Human–Cause Indicators

An experimental run involves using the say-sila ontological model to determine the presence of select concepts from the Six Americas survey (Leiserowitz et al., 2010) in the tweet texts. In this first case the targeted SCR pair is "human–cause," which is intended to test hypothesis 1A. Tweets which contain tokens for both concepts are inferred to be members of the *WeakHumanCauseText* class. If there is a syntactic dependency between the tokens in a text, then it is inferred to be a member of the *StrongHumanCauseText* class.

Table 5.6 gives the counts and percentages of accounts in the say-sila ontology, populated with individuals representing the data in the 2019 *#globalwarming* dataset, that have

[34] Of course, as we are modelling only two stances, the remaining users must be assumed to be in the denier category. A model that explicitly detects the denier category would arguably be more valuable as this is the minority label (Weiss, 2004). Further research is certainly warranted in this direction.

published one or more weak (co-occurring concepts) human–cause indicator texts. Each line in the table gives results for a minimum tweet activity level from 2 to 20. The "No. Users" column gives the number of users in the dataset who have posted at least the minimum number of tweets over the course of the year in 2019. The three double-columns each give the counts (left) and percentages (right) of the weak human–cause indicator accounts. The first double-column shows all the weak human–cause accounts, regardless of the stance on global warming. The second and third double-columns list the number of these accounts for which the follow-the-leader method found them to be respectively in the green or the denier category. In each row the "Green" and "Denier" counts sum to the full count of weak human–cause accounts. The percentages in all cases represent the coverage with respect to all the users in the dataset for a given minimum activity level.

Table 5.6 Coverage of green and denier weak human–cause accounts.

Min. Tweets	**No. Users**	**Weak Human–Cause Accounts**					
				Green		**Denier**	
2	2376	156	6.57%	105	4.42%	51	2.15%
3	1347	122	9.06%	83	6.16%	39	2.90%
4	927	101	10.90%	65	7.01%	36	3.88%
5	705	89	12.62%	54	7.66%	35	4.96%
6	561	81	14.44%	46	8.20%	35	6.24%
7	468	76	16.24%	43	9.19%	33	7.05%
8	392	71	18.11%	41	10.46%	30	7.65%
9	345	66	19.13%	38	11.01%	28	8.12%
10	302	61	20.20%	36	11.92%	25	8.28%
11	271	58	21.40%	34	12.55%	24	8.86%
12	240	52	21.67%	31	12.92%	21	8.75%
13	216	48	22.22%	27	12.50%	21	9.72%
14	198	44	22.22%	24	12.12%	20	10.10%
15	183	43	23.50%	24	13.11%	19	10.38%
16	166	43	25.90%	24	14.46%	19	11.45%
17	151	40	26.49%	23	15.23%	17	11.26%
18	137	35	25.55%	22	16.06%	13	9.49%
19	130	33	25.38%	21	16.15%	12	9.23%
20	119	29	24.37%	19	15.97%	10	8.40%

Figure 5.6 presents a graphical representation of the percentage of coverage shown in Table 5.6 for (1) all the weak human–cause indicator accounts [WHCA], (2) the green weak human–cause indicator accounts [GWHCA], and (3) the denier weak human–cause

indicator accounts [DWHCA]. This SCR pair has relatively good coverage as a weak indicator. However, we must note that even though "human–cause" is a green indicator, the number of users in the denier category who are demonstrating this indicator is not as low as one might expect, compared to the coverage shown by users in the green category.

Table 5.7 Coverage of green and denier strong human–cause accounts.

Min. Tweets	**No. Users**	**Strong Human–Cause Accounts**					
				Green		**Denier**	
2	2376	55	2.31%	32	1.35%	23	0.97%
3	1347	39	2.90%	21	1.56%	18	1.34%
4	927	35	3.78%	17	1.83%	18	1.94%
5	705	32	4.54%	14	1.99%	18	2.55%
6	561	30	5.35%	12	2.14%	18	3.21%
7	468	29	6.20%	12	2.56%	17	3.63%
8	392	28	7.14%	11	2.81%	17	4.34%
9	345	26	7.54%	10	2.90%	16	4.64%
10	302	24	7.95%	10	3.31%	14	4.64%
11	271	24	8.86%	10	3.69%	14	5.17%
12	240	21	8.75%	9	3.75%	12	5.00%
13	216	20	9.26%	8	3.70%	12	5.56%
14	198	19	9.60%	8	4.04%	11	5.56%
15	183	18	9.84%	8	4.37%	10	5.46%
16	166	18	10.84%	8	4.82%	10	6.02%
17	151	16	10.60%	7	4.64%	9	5.96%
18	137	15	10.95%	7	5.11%	8	5.84%
19	130	13	10.00%	6	4.62%	7	5.38%
20	119	11	9.24%	5	4.20%	6	5.04%

Of course, while weak indicators identify texts with a simple co-occurrence of the concepts in an SCR pair, strong indicators identify those where one of the tokens representing the concepts has a syntactic dependency on the other (see Section 5.1.5). In this case these tokens are varying grammatical forms of "human" and "cause" as well as their synonyms as given by the appropriate synsets in WordNet. Table 5.7 lists the corresponding counts and percentages of coverage for users who have tweeted strong (syntactically dependent) human–cause indicator texts. The table layout is the same as described for the weak indicator texts except that the three double-columns give the counts for the strong indicator accounts. The coverage in the double-columns is also presented graphically in Figure 5.7, where the blue, green, and dark magenta lines respectively represent (1) all the strong human–cause indicator accounts [SHCA], (2) the green strong human–cause indicator accounts [GSHCA], and (3) the denier strong human–cause indicator accounts [DSHCA]. We note that overall coverage increases at higher minimum activity levels until we hit a minimum-tweet level of 16-18, after which it begins to taper off. Comparing Figures 5.6 and 5.7, We do not see this tapering off when looking only at users in the *GreenWeakHumanCauseAccount* class. However, we do see it for the users in *DenierWeakHumanCauseAccount* where the coverage begins to wane after the 16-tweet minimum.

Notably, at a minimum-tweet level of 4 or greater, there are more users in the *DenierStrongHumanCauseAccount* class than there are in the *GreenStrongHumanCauseAccount* class. This essentially means that while more users labelled as green are including the concepts "human" and "cause" in their tweets (resulting in weak indicators), more users in the denier category are using expressions that capture a syntactic dependency between these concepts in theirs (e.g., "humans cause," "caused by humanity," "human induced," etc.). This result is surprising and stands somewhat in contrast to our expectations per hypothesis 1A when considering the strong green indicators.

Table 5.8 allows us to examine this finding more closely. Instead of the coverage across all participating users, the table reports the ratios of green and denier human–cause indicator accounts with respect to all the human–cause indicator accounts. On the left side of the table we list the weak type, based on a simple co-occurrence of the terms in a tweet. Here we see that at the lower and higher ends of the range of activity levels, the model finds these users to be approximately two-thirds green vs. one-third denier, similar to the ratio for the full dataset (see Section 5.2). As we move up or down the table towards a minimum-tweet level of 14, the ratio approaches 55% green to 45% denier. However, for the strong (syntactic dependency) indicator class on the right side of the table, we see that the model selects a smaller percentage of strong green and a larger percentage of strong denier indicator accounts than it does for weak indicator accounts at every activity level.

Figure 5.6 Percent coverage of green and denier weak human–cause accounts.

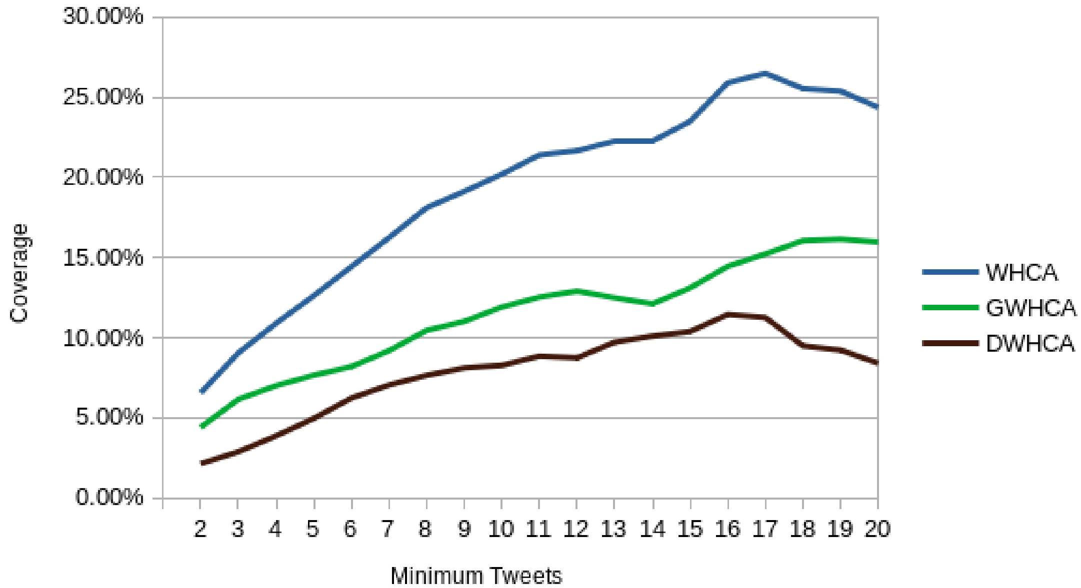


Figure 5.7 Percent coverage of green and denier strong human–cause accounts.

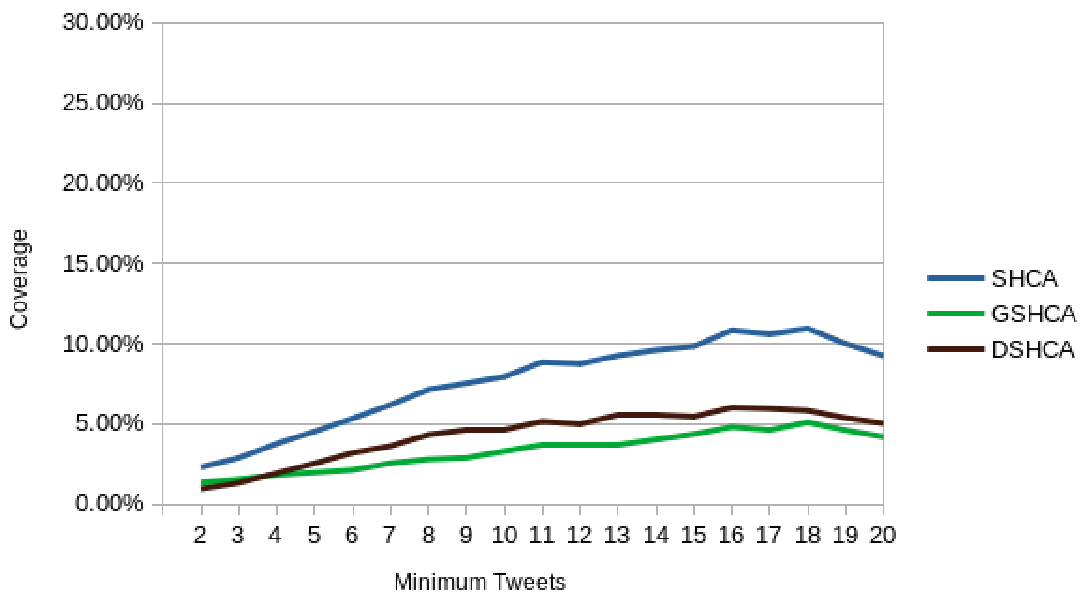


Table 5.8 Green and denier ratios for weak and strong human–cause accounts.

Min. Tweets	**Weak Human–Cause Account**					**Strong Human–Cause Account**				
	Total	**Green**		**Denier**		**Total**	**Green**		**Denier**	
2	156	105	67.31%	51	32.69%	55	32	58.18%	23	41.82%
3	122	83	68.03%	39	31.97%	39	21	53.85%	18	46.15%
4	101	65	64.36%	36	35.64%	35	17	48.57%	18	51.43%
5	89	54	60.67%	35	39.33%	32	14	43.75%	18	56.25%
6	81	46	56.79%	35	43.21%	30	12	40.00%	18	60.00%
7	76	43	56.58%	33	43.42%	29	12	41.38%	17	58.62%
8	71	41	57.75%	30	42.25%	28	11	39.29%	17	60.71%
9	66	38	57.58%	28	42.42%	26	10	38.46%	16	61.54%
10	61	36	59.02%	25	40.98%	24	10	41.67%	14	58.33%
11	58	34	58.62%	24	41.38%	24	10	41.67%	14	58.33%
12	52	31	59.62%	21	40.38%	21	9	42.86%	12	57.14%
13	48	27	56.25%	21	43.75%	20	8	40.00%	12	60.00%
14	44	24	54.55%	20	45.45%	19	8	42.11%	11	57.89%
15	43	24	55.81%	19	44.19%	18	8	44.44%	10	55.56%
16	43	24	55.81%	19	44.19%	18	8	44.44%	10	55.56%
17	40	23	57.50%	17	42.50%	16	7	43.75%	9	56.25%
18	35	22	62.86%	13	37.14%	15	7	46.67%	8	53.33%
19	33	21	63.64%	12	36.36%	13	6	46.15%	7	53.85%
20	29	19	65.52%	10	34.48%	11	5	45.45%	6	54.55%

For the human–cause indicator at least, the strong class appears more selective than the weak with respect to the denier category. That is to say, taking syntactic dependency into account demonstrates an interesting denier-oriented bias for this concept pair. Certainly, it would be beneficial to identify factors which tend to make the model more discriminating towards one category or the other, and so it will be interesting to see if this pattern holds true for the remaining indicator classes.

In addition to analyzing the coverage of modelled users from the dataset and the corresponding green-to-denier ratios, we can evaluate how effective the human–cause SCR pair is for identifying the stance of these users on global warming. Table 5.9 shows the precision, recall and F1 scores with respect to users in the green category for both the weak and strong indicator accounts across the range of minimum tweet activity levels. These metrics indicate how well individuals who are members of the *WeakHumanCause-Account* and *StrongHumanCauseAccount* classes correspond with the members of the *GreenAccount* class. Specifically, in this case the precision represents:

$$precision := \frac{GreenWeakHumanCauseAccount}{WeakHumanCauseAccount} \tag{5.34}$$

Likewise, the recall represents:

$$recall := \frac{GreenWeakHumanCauseAccount}{GreenAccounts} \tag{5.35}$$

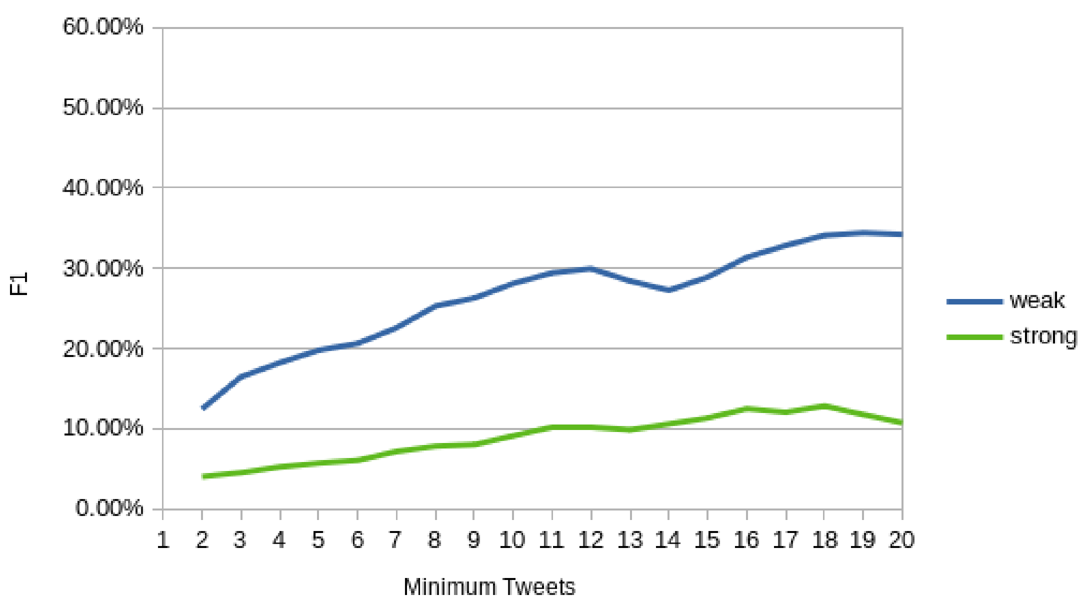
As mentioned above, the F1 measure effectively combines these two metrics into a single score. Figure 5.8 gives a graphical representation of the F1 measures reported in Table 5.9. With regard to our hypothesis 1A, presented at the beginning of this section, we observe that we have a fairly low level of predictive capability with a single SCR pair. Note, however, that we are first presenting the results individually, and the final predictive analysis will employ a disjunction of SCR pairs. Recall is low, which is to be expected, given that we are targeting a single concept across all topics being discussed with the hashtag-identified theme of global warming. The recall is indeed higher for the weak indicator accounts as per hypothesis 2A. However, regarding precision, we do not see results corresponding to hypothesis 2B since the strong indicator accounts show lower, rather than higher, precision compared to the weak indicator accounts across all minimum-tweet activity levels.

To add a bit of perspective, we can juxtapose this analysis of online Twitter communications with actual findings from the survey-based experiments of the Six Americas in 2010. We see that a strong majority of subjects in the alarmed (91%) and concerned (75%) categories from the "Beliefs" section of the Six Americas survey (Leiserowitz et al., 2010, T5) assert that "[a]ssuming global warming is happening," they believe that it is "[c]aused mostly by human activities." As we are mapping these two categories onto our green class of Twitter users, we expect to have a higher percentage of green accounts among the users publishing human-cause indicator texts. Table 5.8 shows this to be the case for the weak indicator, but not for the strong, except at the lowest levels of

Table 5.9 Predictive scores for weak and strong human–cause accounts.

Min. Tweets	**Weak Human–Cause Account**			**Strong Human–Cause Account**		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
2	67.31%	6.87%	12.46%	58.18%	2.09%	4.04%
3	68.03%	9.37%	16.47%	53.85%	2.37%	4.54%
4	64.36%	10.62%	18.23%	48.57%	2.78%	5.26%
5	60.67%	11.82%	19.78%	43.75%	3.06%	5.73%
6	56.79%	12.60%	20.63%	40.00%	3.29%	6.08%
7	56.58%	14.10%	22.57%	41.38%	3.93%	7.19%
8	57.75%	16.21%	25.31%	39.29%	4.35%	7.83%
9	57.58%	17.04%	26.30%	38.46%	4.48%	8.03%
10	59.02%	18.46%	28.13%	41.67%	5.13%	9.13%
11	58.62%	19.65%	29.44%	41.67%	5.78%	10.15%
12	59.62%	20.00%	29.95%	42.86%	5.81%	10.23%
13	56.25%	19.01%	28.42%	42.86%	5.81%	10.23%
14	54.55%	18.18%	27.27%	42.11%	6.06%	10.60%
15	55.81%	19.51%	28.92%	44.44%	6.50%	11.35%
16	55.81%	21.82%	31.37%	44.44%	7.27%	12.50%
17	57.50%	23.00%	32.86%	43.75%	7.00%	12.07%
18	62.86%	23.40%	34.11%	46.67%	7.45%	12.84%
19	63.64%	23.60%	34.43%	46.15%	6.74%	11.76%
20	65.52%	23.17%	34.23%	45.45%	6.10%	10.75%

Figure 5.8 F1 scores (green) for weak & strong human–cause accounts.



minimum participation.[35] For the segments from the Six Americas that we have mapped to our denier category, the doubtful (10%) and the dismissive (2%), generally do not report believing that climate change is caused by humans.

5.3.2 Nature–Cause Indicators

In contrast to the human–cause SCR pair, we hypothesized that the nature–cause (or natural cause) SCR pair would be a denier indicator in that it represents a topic centred around the supposition that if climate change is occurring, it is not due to human activity. As a denier indicator, the nature–cause pair was originally intended as a test for hypothesis 1B. Tweets containing a co-occurrence of words linked to each of the two concepts are represented by the *WeakNatureCauseText*. Likewise, tweets are modelled by the *StrongNatureCauseText* class when they contain words linked to these concepts, and these words have a relation of syntactic dependency.

We analyze this SCR pair in the same manner as we did for the human–cause pair in the preceding section. Table 5.10 reports the counts and percentage of coverage of individuals inferred to be members of the *WeakNatureCauseAccount* class with respect to the user base for the 2019 *#globalwarming* dataset at minimum levels of participation ranging from 2 to 20 tweets. Figure 5.9 presents the coverage data in graphical form. The blue, green, and dark magenta lines on this chart respectively indicate the percentages of the base user set covered by individuals in the weak account indicator classes: *WeakNatureCauseAccount* [WNCA], *GreenWeakNatureCauseAccount* [GWNCA], and *DenierWeakNatureCauseAccount* [DWNCA].

Keeping in mind that this SCR pair and that of human–cause share one concept (i.e., "cause"), when we compare the percentage of coverage here to the results for the human–cause SCR pair (Table 5.6 and Figure 5.6), we see that many more users are publishing

[35] Note that the counts for the strong indicator account are always less, as the individuals that are members of this class represent a subset of those of the weak indicator.

Table 5.10 Coverage of green and denier weak nature–cause accounts.

Min. Tweets	**No. Users**	**Weak Nature–Cause Accounts**		**Green**		**Denier**	
2	2376	54	2.27%	35	1.47%	19	0.80%
3	1347	48	3.56%	33	2.45%	15	1.11%
4	927	40	4.31%	27	2.91%	13	1.40%
5	705	34	4.82%	24	3.40%	10	1.42%
6	561	28	4.99%	20	3.57%	8	1.43%
7	468	25	5.34%	17	3.63%	8	1.71%
8	392	24	6.12%	17	4.34%	7	1.79%
9	345	24	6.96%	17	4.93%	7	2.03%
10	302	22	7.28%	17	5.63%	5	1.66%
11	271	20	7.38%	16	5.90%	4	1.48%
12	240	19	7.92%	15	6.25%	4	1.67%
13	216	19	8.80%	15	6.94%	4	1.85%
14	198	18	9.09%	14	7.07%	4	2.02%
15	183	17	9.29%	13	7.10%	4	2.19%
16	166	17	10.24%	13	7.83%	4	2.41%
17	151	17	11.26%	13	8.61%	4	2.65%
18	137	14	10.22%	11	8.03%	3	2.19%
19	130	13	10.00%	10	7.69%	3	2.31%
20	119	10	8.40%	8	6.72%	2	1.68%

tweets referring to the concept "human" than to the concept "nature." *WeakHumanCauseAccount* and *WeakNatureCauseAccount* both peak at an activity level of 17 tweets. However, while the former reaches 26.49% coverage of the user base, the latter only hits 11.26%. Both SCR pairs show similar graphs in that the coverage increases as the minimum activity level increases up to an activity level of 16 or 17 tweets. After this the percentage of coverage falls off for higher levels of minimum activity.

The corresponding coverage data for users whose tweets contain strong nature–cause indicators is displayed in Table 5.11. Figure 5.7 graphs the data from this table with the blue, green, and dark magenta lines respectively representing the percentage of user coverage for the classes: *StrongNatureCauseAccount* [SNCA], *GreenStrongNatureCauseAccount* [GSNCA], and *DenierStrongNatureCauseAccount* [DSNCA]. As with the human–cause SCR pair, the coverage for the nature–cause pair is much lower for strong than for weak indicators. Of course, we expect a lower coverage because strong indicators represent a subset of weak indicators with the additional constraint of syntactic dependency between the word references to the two concepts in the pair. Nevertheless,

Table 5.11 Coverage of green and denier strong nature–cause accounts.

Min. Tweets	**No. Users**	**Strong Nature–Cause Accounts**					
				Green		**Denier**	
2	2376	11	0.46%	8	0.34%	3	0.13%
3	1347	10	0.74%	7	0.52%	3	0.22%
4	927	9	0.97%	7	0.76%	2	0.22%
5	705	8	1.13%	7	0.99%	1	0.14%
6	561	8	1.43%	7	1.25%	1	0.18%
7	468	6	1.28%	5	1.07%	1	0.21%
8	392	5	1.28%	5	1.28%	0	0.00%
9	345	5	1.45%	5	1.45%	0	0.00%
10	302	5	1.66%	5	1.66%	0	0.00%
11	271	5	1.85%	5	1.85%	0	0.00%
12	240	4	1.67%	4	1.67%	0	0.00%
13	216	4	1.85%	4	1.85%	0	0.00%
14	198	4	2.02%	4	2.02%	0	0.00%
15	183	3	1.64%	3	1.64%	0	0.00%
16	166	3	1.81%	3	1.81%	0	0.00%
17	151	3	1.99%	3	1.99%	0	0.00%
18	137	3	2.19%	3	2.19%	0	0.00%
19	130	3	2.31%	3	2.31%	0	0.00%
20	119	3	2.52%	3	2.52%	0	0.00%

the coverage is lower than we might like to see, given that our eventual goal is to use these indicators to predict a user's stance on climate change.

As mentioned at the beginning of Section 5.3, it very quickly became obvious that our hypothesis 1B was not valid and that the denier SCR pairs would not serve to identify users in the denier category. The issue with this hypothesis is surprising, yet exceedingly clear, when looking at the user coverage for *DenierStrongNatureCauseAccount* (Table 5.11 and Figure 5.10). There are many more users labelled green who are tweeting about the concepts of "nature" and "cause." This may stand to reason for the weak (co-occurrence) indicators; however for strong (syntactical dependency) indicators, one might expect a significant number of messages with expressions such as "caused by nature" from users in the denier category. Yet, the results show that very few users labelled as deniers are using these expressions, and this number falls to zero once we hit a minimum activity level of 8 tweets. Finally, we note that the number of *StrongNatureCauseAccount* individuals does not taper off at the very highest activity levels as has been the tendency up to now. The same three green users remain active at a minimum of 15 tweets and

Figure 5.9 Percent coverage of green and denier weak nature–cause accounts.

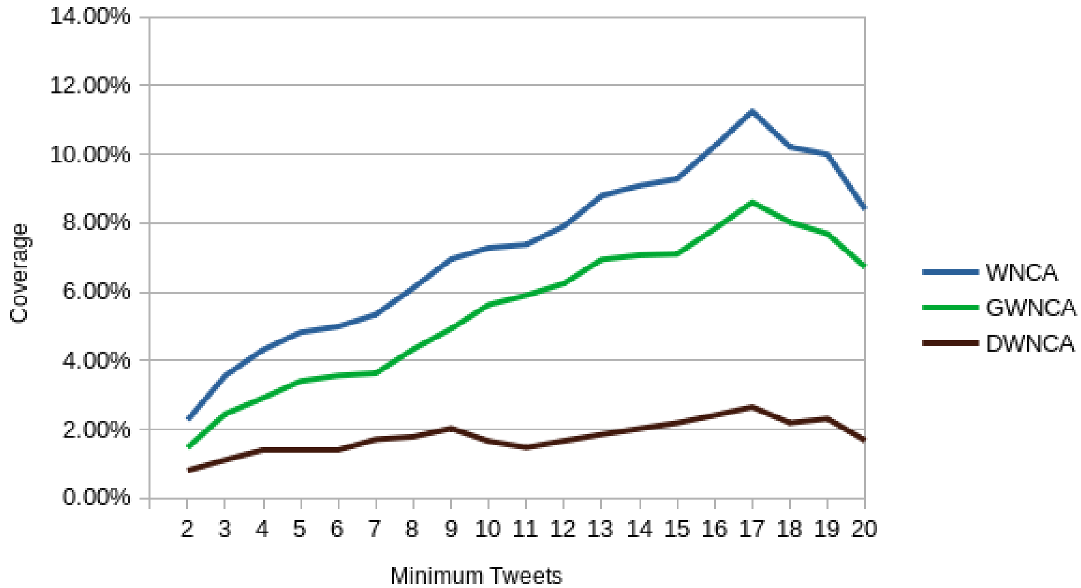
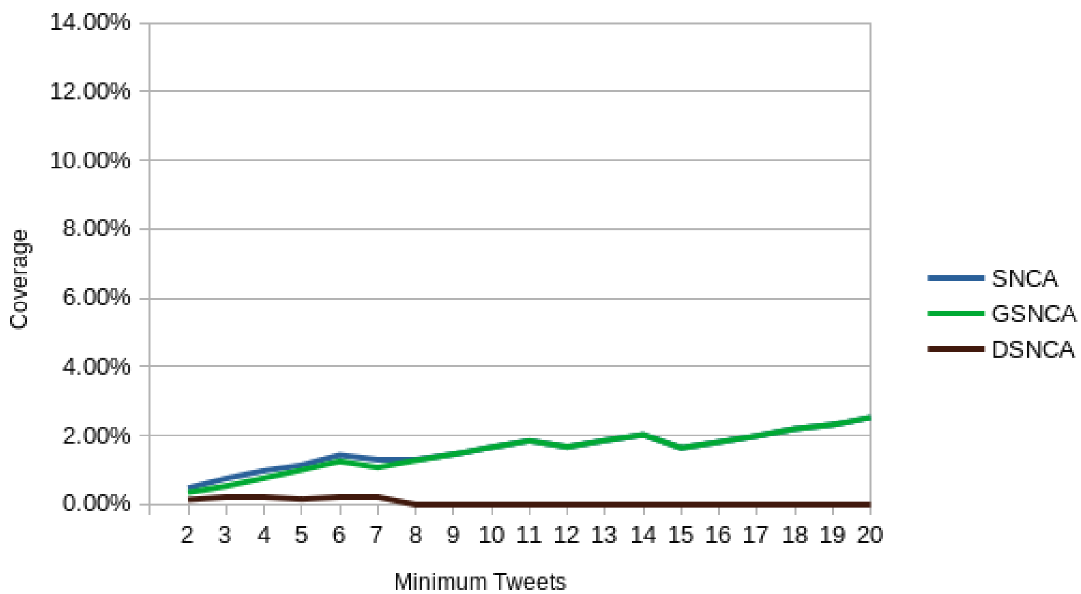


Figure 5.10 Percent coverage of green and denier strong nature–cause accounts.



above, representing an increasing percentage of coverage as the user base grows smaller (see Table 5.11).

The inability of this SCR pair to identify users in the denier category is perhaps even more evident in Table 5.12, which lists the green and denier ratios for nature–cause indicator accounts. For weak nature–cause accounts (left side) at a 2-tweet minimum, we see the same split from the dataset of roughly two-thirds green and one-third denier, but as we raise the minimum activity level, the distribution falls even more towards the green accounts until it reaches 80% green and 20% denier. For the strong indicator accounts (right side), we do not see any bias in the model towards the denier category as we did with the human–cause SCR pair (see Table 5.8). Rather, the indicated accounts are mostly green when considering low levels of minimum participation and always green for a minimum activity level of 8 or higher. Again, the results for this SCR pair do not support our hypothesis 1B.

Table 5.12 Green and denier ratios for weak and strong nature–cause accounts.

Min.	**Weak Nature–Cause Account**					**Strong Nature–Cause Account**				
Tweets	**Total**	**Green**		**Denier**		**Total**	**Green**		**Denier**	
2	54	35	64.81%	19	35.19%	11	8	72.73%	3	27.27%
3	48	33	68.75%	15	31.25%	10	7	70.00%	3	30.00%
4	40	27	67.50%	13	32.50%	9	7	77.78%	2	22.22%
5	34	24	70.59%	10	29.41%	8	7	87.50%	1	12.50%
6	28	20	71.43%	8	28.57%	8	7	87.50%	1	12.50%
7	25	17	68.00%	8	32.00%	6	5	83.33%	1	16.67%
8	24	17	70.83%	7	29.17%	5	5	100.00%	0	0.00%
9	24	17	70.83%	7	29.17%	5	5	100.00%	0	0.00%
10	22	17	77.27%	5	22.73%	5	5	100.00%	0	0.00%
11	20	16	80.00%	4	20.00%	5	5	100.00%	0	0.00%
12	19	15	78.95%	4	21.05%	4	4	100.00%	0	0.00%
13	19	15	78.95%	4	21.05%	4	4	100.00%	0	0.00%
14	18	14	77.78%	4	22.22%	4	4	100.00%	0	0.00%
15	17	13	76.47%	4	23.53%	3	3	100.00%	0	0.00%
16	17	13	76.47%	4	23.53%	3	3	100.00%	0	0.00%
17	17	13	76.47%	4	23.53%	3	3	100.00%	0	0.00%
18	14	11	78.57%	3	21.43%	3	3	100.00%	0	0.00%
19	13	10	76.92%	3	23.08%	3	3	100.00%	0	0.00%
20	10	8	80.00%	2	20.00%	3	3	100.00%	0	0.00%

Continuing on to the predictive capability of the nature–cause SCR pair, Table 5.13 gives the precision, recall, and F1 scores for both the weak and the strong indicator accounts for this pair. Figure 5.11 plots these F1 scores graphically. Again, we are predicting with respect to the green label. According to our hypothesis 1B, the nature–cause pair should be more suitable for predicting the denier class, but the analysis of coverage shows that this is not the case. The experiments provide more evidence for hypotheses 2A and 2B as we have higher recall with the weak indicator and higher precision with the strong indicator. However, the low coverage and consequently low recall make for extremely poor F1 scores for *StrongNatureCauseAccount*. The *WeakNatureCauseAccount* class provides better predictive capability, peaking at an F1 measure of 22.22% at a minimum activity level of 17 tweets.

We should perhaps remind the reader that the say-sila model will be using indicator classes for SCR pairs together (in disjunction). Each indicator is intended to hit separate themes in a continuing online conversation. We have chosen concept pairs that have been shown to be important with respect to the Six Americas series of studies, but we fully expect that users will be discussing some themes more often than others.

Looking once again at the "Beliefs" section of the Six Americas survey (Leiserowitz et al., 2010, T5),[36] we find that the majority of subjects in the doubtful (80%) and dismissive (64%) categories, which we map to the denier category in our model, respond that they believe that global warming is "[c]aused mostly by changes in the environment."[37] Not surprisingly, zero percent of the subjects in the alarmed and concerned categories, which we map to green accounts, answered that they believe environmental changes are the primary cause of global warming. As noted above, given how the subjects answered in the Six Americas survey, we expected to see a much larger ratio of users from the denier category. Yet, our experiments show that it is primarily green users publishing tweets

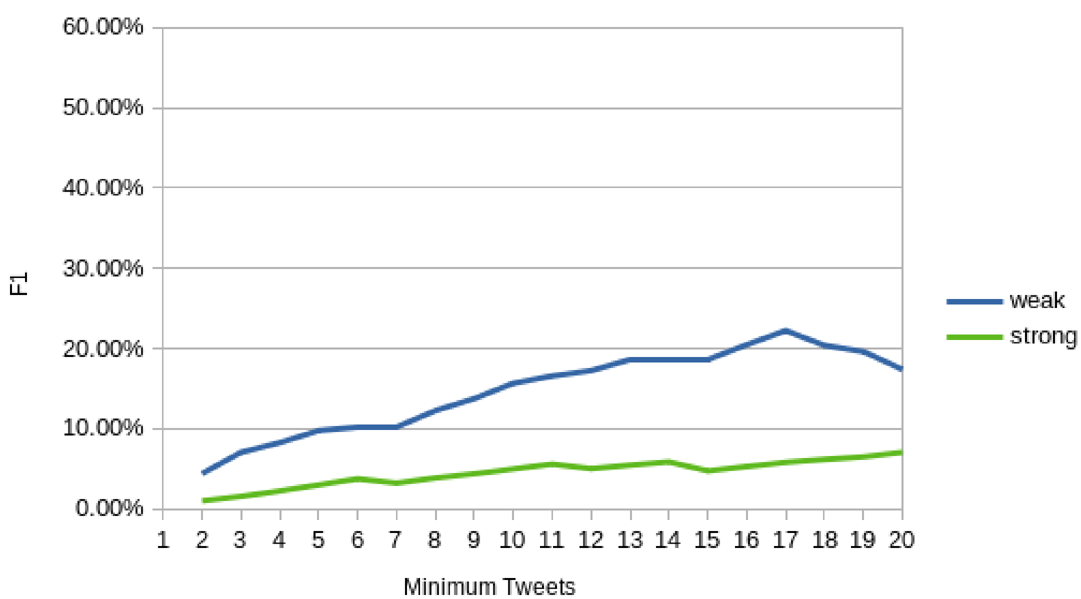
[36]The "Beliefs" section is for the same Six Americas question (T5) that we considered for the human–cause SCR pair.

[37]When considering the lower percentage for the dismissive (who are at the far end of the belief spectrum for climate change, opposite the alarmed (Maibach et al., 2009)), it should be noted that 30% from this category answered, "None of the above because global warming isn't happening," even though the question clearly states that they are to assume it *is* happening.

Table 5.13 Predictive scores for weak and strong nature–cause accounts.

Min.	Weak Nature–Cause Account			Strong Nature–Cause Account		
Tweets	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
2	64.81%	2.29%	4.42%	72.73%	0.52%	1.04%
3	68.75%	3.72%	7.07%	70.00%	0.79%	1.56%
4	67.50%	4.41%	8.28%	77.78%	1.14%	2.25%
5	70.59%	5.25%	9.78%	87.50%	1.53%	3.01%
6	71.43%	5.48%	10.18%	87.50%	1.92%	3.75%
7	68.00%	5.57%	10.30%	83.33%	1.64%	3.22%
8	70.83%	6.72%	12.27%	100.00%	1.98%	3.88%
9	70.83%	7.62%	13.77%	100.00%	2.24%	4.39%
10	77.27%	8.72%	15.67%	100.00%	2.56%	5.00%
11	80.00%	9.25%	16.58%	100.00%	2.89%	5.62%
12	78.95%	9.68%	17.24%	100.00%	2.58%	5.03%
13	78.95%	10.56%	18.63%	100.00%	2.82%	5.48%
14	77.78%	10.61%	18.67%	100.00%	3.03%	5.88%
15	76.47%	10.57%	18.57%	100.00%	2.44%	4.76%
16	76.47%	11.82%	20.47%	100.00%	2.44%	4.76%
17	76.47%	13.00%	22.22%	100.00%	3.00%	5.83%
18	78.57%	11.70%	20.37%	100.00%	3.19%	6.19%
19	76.92%	11.24%	19.61%	100.00%	3.37%	6.52%
20	80.00%	9.76%	17.39%	100.00%	3.66%	7.06%

Figure 5.11 F1 scores (green) for weak & strong nature–cause accounts.



with this SCR pair (see Table 5.12).

5.3.3 CO_2-Cut Indicators

The CO_2–cut SCR pair[38] is another green indicator from the Six Americas survey. It involves reducing carbon emissions. Like the human–cause pair, CO_2–cut is a green indicator, intended to represent a topic line that serves to identify users in the green category as stated in hypothesis 1A. In the same manner as the previous SCR pairs, tweets containing words that are linked to these two concepts are modelled by the *WeakCO2CutText* class. If the words are syntactically dependent, as reported by the TweeboParser tool (Kong et al., 2014), then they are modelled by the *StrongCO2CutText* class.

Table 5.14 Coverage of green and denier weak CO_2–cut accounts.

Min. Tweets	No. Users	Weak CO_2–Cut Accounts					
				Green		Denier	
2	2376	78	3.28%	58	2.44%	20	0.84%
3	1347	63	4.68%	47	3.49%	16	1.19%
4	927	55	5.93%	39	4.21%	16	1.73%
5	705	49	6.95%	33	4.68%	16	2.27%
6	561	45	8.02%	29	5.17%	16	2.85%
7	468	40	8.55%	24	5.13%	16	3.42%
8	392	38	9.69%	23	5.87%	15	3.83%
9	345	38	11.01%	23	6.67%	15	4.35%
10	302	37	12.25%	23	7.62%	14	4.64%
11	271	34	12.55%	20	7.38%	14	5.17%
12	240	33	13.75%	19	7.92%	14	5.83%
13	216	29	13.43%	16	7.41%	13	6.02%
14	198	27	13.64%	15	7.58%	12	6.06%
15	183	27	14.75%	15	8.20%	12	6.56%
16	166	26	15.66%	14	8.43%	12	7.23%
17	151	23	15.23%	13	8.61%	10	6.62%
18	137	21	15.33%	13	9.49%	8	5.84%
19	130	21	16.15%	13	10.00%	8	6.15%
20	119	20	16.81%	13	10.92%	7	5.88%

Table 5.14 lists the counts and percentages of coverage of the user base for individuals in the *WeakCO2CutAccount* class for our standard set of minimum activity levels ranging

[38] Although "cut-CO_2" may seem more intuitive, the order of the concepts reflects the most common syntactic dependency in tweet texts: $dependsOn(CO2Token, CutToken)$ (see Section 5.1.5).

from 2 to 20. As before, we have the number of users participating at each level and then counts and percentages for the number of those who have at least one tweet with the CO_2–cut SCR pair in the lot of *#globalwarming* tweets they published in 2019. Finally out of these users, we show the number who have previously been labelled as part of the green or denier categories according to the follow-the-leader method (see Section 5.1.1). The percentages in all cases are with respect to the full user base for a given activity level. Figure 5.12 plots these percentages with the blue, green, and dark magenta lines respectively representing coverage of the users for a given activity level by individuals in the following classes: *WeakCO2CutAccount* [WCCA], *GreenWeakCO2CutAccount* [GWCCA], and *DenierWeakCO2CutAccount* [DWCCA].

In the results from the previous SCR pairs, we noted that coverage of the user base by indicator accounts generally peaks at a minimum activity level of 16-18 tweets. Here, however, we see this peak only for the denier weak CO_2–cut indicators accounts (7.23% at 16 tweets). The green indicator accounts keep rising for the highest minimum activity levels, and this keeps the general indicator account coverage (green + denier) rising as well. Referring back to Table 5.14, which shows the underlying data for the chart, we see that it is the same 13 accounts from the green category which continue to make the minimum-tweet cut all the way to the last minimum participation level of 20 tweets. Meanwhile, as the minimum number of tweets rises for each experiment, the associated user base gets smaller since there are fewer users publishing so many tweets. Therefore, these 13 users in the *GreenWeakCO2CutAccount* class cover an increasingly large percentage of this shrinking user base.

We see a similar situation for the strong CO_2–cut indicator accounts. Table 5.15 lists the counts and percent coverage for users publishing at least one *StrongCO2CutText*. The associated chart is shown in Figure 5.13 with the customary blue, green, and dark magenta lines respectively representing coverage by these indicator classes: *StrongCO2CutAccount* [SCCA], *GreenStrongCO2CutAccount* [GSCCA], and *DenierStrongCO2CutAccount*

Figure 5.12 Percent coverage of green and denier weak CO_2–cut accounts.

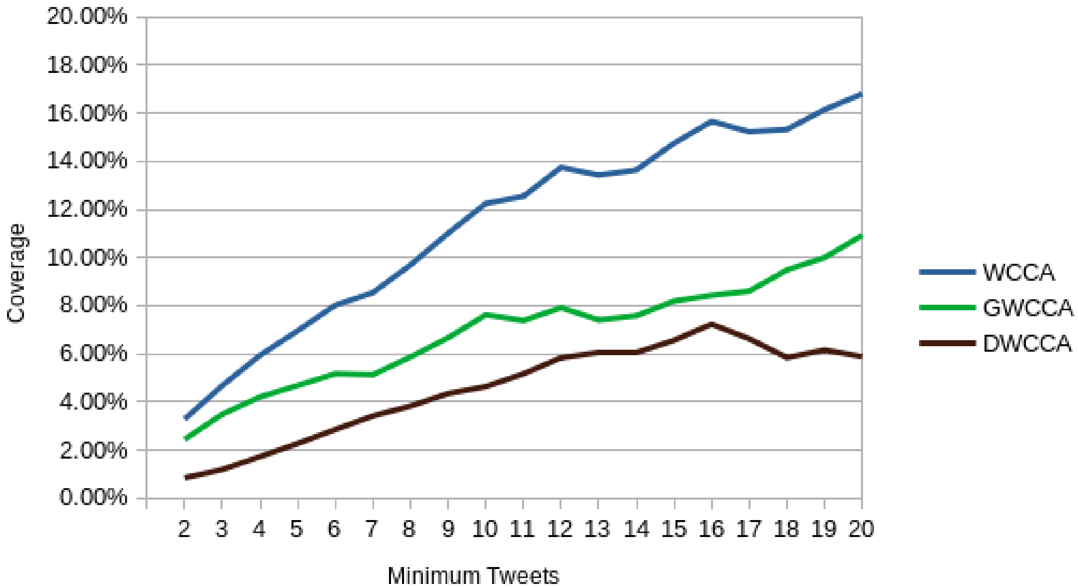


Figure 5.13 Percent coverage of green and denier strong CO_2–cut accounts.

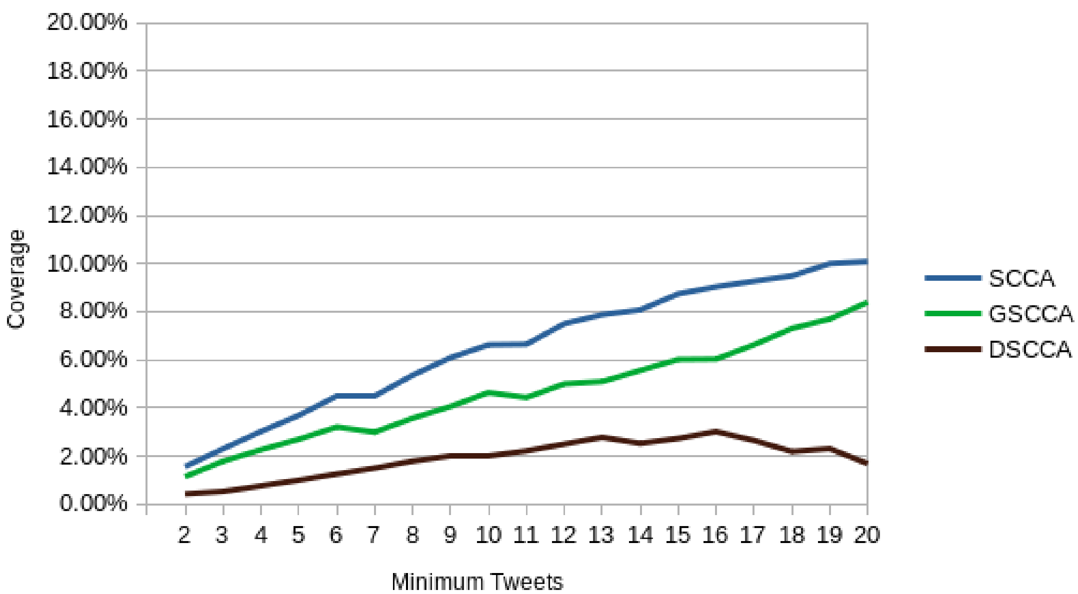


Table 5.15 Coverage of green and denier strong CO_2–cut accounts.

Min. Tweets	**No. Users**	**Strong CO_2–Cut Accounts**		**Green**		**Denier**	
2	2376	37	1.56%	27	1.14%	10	0.42%
3	1347	31	2.30%	24	1.78%	7	0.52%
4	927	28	3.02%	21	2.27%	7	0.76%
5	705	26	3.69%	19	2.70%	7	0.99%
6	561	25	4.46%	18	3.21%	7	1.25%
7	468	21	4.49%	14	2.99%	7	1.50%
8	392	21	5.36%	14	3.57%	7	1.79%
9	345	21	6.09%	14	4.06%	7	2.03%
10	302	20	6.62%	14	4.64%	6	1.99%
11	271	18	6.64%	12	4.43%	6	2.21%
12	240	18	7.50%	12	5.00%	6	2.50%
13	216	17	7.87%	11	5.09%	6	2.78%
14	198	16	8.08%	11	5.56%	5	2.53%
15	183	16	8.74%	11	6.01%	5	2.73%
16	166	15	9.04%	10	6.02%	5	3.01%
17	151	14	9.27%	10	6.62%	4	2.65%
18	137	13	9.49%	10	7.30%	3	2.19%
19	130	13	10.00%	10	7.69%	3	2.31%
20	119	12	10.08%	10	8.40%	2	1.68%

[DSCCA]. Of the 13 users discussed above, 10 are green strong indicator accounts,[39] and their coverage with respect to the user base becomes greater as that base shrinks with each increase in the level of minimum activity. Also, as we saw with the weak denier indicator accounts for this SCR pair, the coverage of the strong denier indicator accounts does indeed fall off after it peaks at a minimum-tweet level of 16. These results are not surprising given that CO_2–cut is a green SCR pair, and hypothesis 1A states that identifying green concept pairs will aid in identifying green users.

Table 5.16, shows the green and denier ratios for the CO_2–cut indicator accounts. As with the previous SCR pairs, the weak (co-occurrence) type is on the left and the strong (syntactic dependency) type is on the right. For this concept pair the model does not appear to discriminate towards the denier category as we saw with the strong human–cause indicator accounts. On the contrary, for strong CO_2–cut indicator accounts around

[39]Recall that a strong indicator is effectively subsumed by the corresponding weak indicator in the ontological model.

the mid-range of minimum activity, the distribution falls fairly close to the base dataset with two-thirds green and one-third denier. Additionally, for both low and high levels of minimum activity, any bias would be towards the green accounts.

Table 5.16 Green and denier ratios for weak and strong CO_2–cut accounts.

Min.	Weak Human–Cause Account					Strong Human–Cause Account				
Tweets	Total	Green		Denier		Total	Green		Denier	
2	78	58	74.36%	20	25.64%	37	27	72.97%	10	27.03%
3	63	47	74.60%	16	25.40%	31	24	77.42%	7	22.58%
4	55	39	70.91%	16	29.09%	28	21	75.00%	7	25.00%
5	49	33	67.35%	16	32.65%	26	19	73.08%	7	26.92%
6	45	29	64.44%	16	35.56%	25	18	72.00%	7	28.00%
7	40	24	60.00%	16	40.00%	21	14	66.67%	7	33.33%
8	38	23	60.53%	15	39.47%	21	14	66.67%	7	33.33%
9	38	23	60.53%	15	39.47%	21	14	66.67%	7	33.33%
10	37	23	62.16%	14	37.84%	20	14	70.00%	6	30.00%
11	34	20	58.82%	14	41.18%	18	12	66.67%	6	33.33%
12	33	19	57.58%	14	42.42%	18	12	66.67%	6	33.33%
13	29	16	55.17%	13	44.83%	17	11	64.71%	6	35.29%
14	27	15	55.56%	12	44.44%	16	11	68.75%	5	31.25%
15	27	15	55.56%	12	44.44%	16	11	68.75%	5	31.25%
16	26	14	53.85%	12	46.15%	15	10	66.67%	5	33.33%
17	23	13	56.52%	10	43.48%	14	10	71.43%	4	28.57%
18	21	13	61.90%	8	38.10%	13	10	76.92%	3	23.08%
19	21	13	61.90%	8	38.10%	13	10	76.92%	3	23.08%
20	20	13	65.00%	7	35.00%	12	10	83.33%	2	16.67%

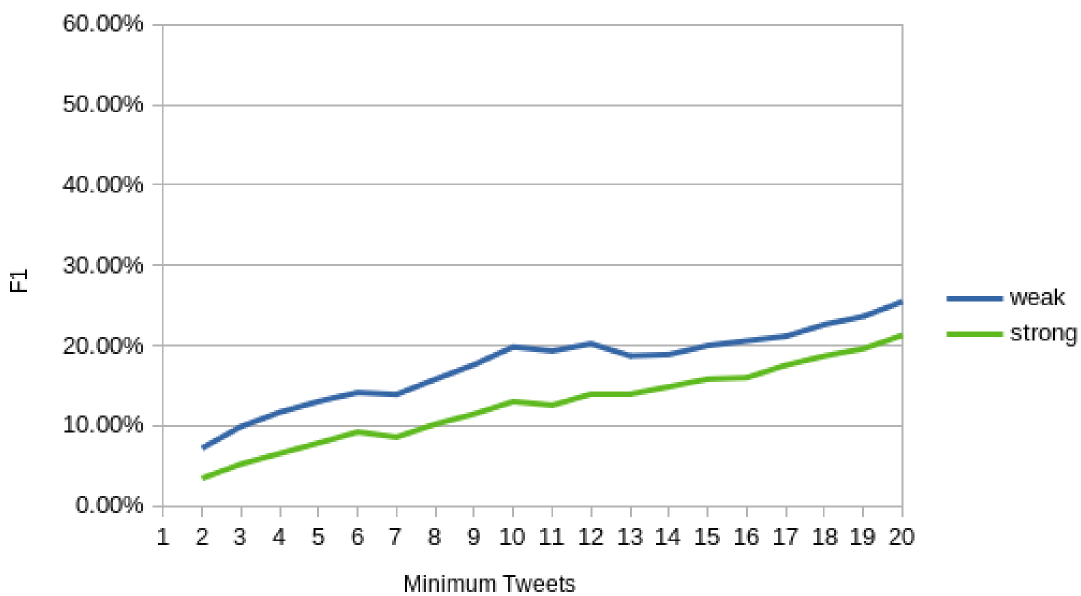
We now turn our attention to the predictive capability of the CO_2–cut SCR pair. Table 5.17 lists the precision, recall, and F1 scores for the weak and strong CO_2–cut indicator accounts for each of the minimum tweet activity levels we are considering. The F1 scores are presented graphically in Figure 5.14. There are a number of interesting results for this concept pair. First we see that the levels of precision of weak CO_2–cut indicators are greater for lower levels of minimum-tweet activity than we saw for the other green SCR pair, weak human–cause (Table 5.9). These values even out at a minimum-tweet level of about 11, after which the two SCR pairs show similar precision scores. For the strong indicators, however, the precision metric is consistently higher for CO_2–cut than it is for human–cause for all considered activity levels.[40] Recall is better for the

[40]We suspect the reason for these findings may have something to do with the generality of the synonyms for "cause" (as given by WordNet) compared to those of the other modelled concepts; however,

Table 5.17 Predictive scores for weak and strong CO_2–cut accounts.

Min. Tweets	Weak CO_2–Cause Account			Strong CO_2–Cause Account		
	Precision	Recall	F1	Precision	Recall	F1
2	74.36%	3.79%	7.22%	72.97%	1.77%	3.45%
3	74.60%	5.30%	9.91%	77.42%	2.71%	5.23%
4	70.91%	6.37%	11.69%	75.00%	3.43%	6.56%
5	67.35%	7.22%	13.04%	73.08%	4.16%	7.87%
6	64.44%	7.95%	14.15%	72.00%	4.93%	9.23%
7	60.00%	7.87%	13.91%	66.67%	4.59%	8.59%
8	60.53%	9.09%	15.81%	66.67%	5.53%	10.22%
9	60.53%	10.31%	17.62%	66.67%	6.28%	11.48%
10	62.16%	11.79%	19.83%	70.00%	7.18%	13.02%
11	58.82%	11.56%	19.32%	66.67%	6.94%	12.57%
12	57.58%	12.26%	20.21%	66.67%	7.74%	13.87%
13	55.17%	11.27%	18.71%	64.71%	7.75%	13.84%
14	55.56%	11.36%	18.87%	68.75%	8.33%	14.86%
15	55.56%	12.20%	20.00%	68.75%	8.94%	15.83%
16	53.85%	12.73%	20.59%	66.67%	9.09%	16.00%
17	56.52%	13.00%	21.14%	71.43%	10.00%	17.54%
18	61.90%	13.83%	22.61%	76.92%	10.64%	18.69%
19	61.90%	14.61%	23.64%	76.92%	11.24%	19.61%
20	65.00%	15.85%	25.49%	83.33%	12.20%	21.28%

Figure 5.14 F1 scores (green) for weak & strong CO_2–cut accounts.



weak human–cause indicator in all cases, which makes sense as that indicator has better coverage than the weak CO_2–cut. The same cannot be said for the strong indicators as the CO_2–cut indicator has more similar coverage to the strong human–cause and better recall at most levels of minimum activity. Finally, we note that the F1 scores do not peak at a minimum activity level of 16, 17, or 18 as they have tended to do with indicators for the other SCR pairs. Rather, the F1 measure continues to rise through to the end of the range at a 20-tweet minimum.

The concepts for the CO_2–cut SCR pair come from two Six Americas survey questions on the topic of "Support for National Response: Specific Climate and Energy Policies" (Leiserowitz et al., 2010, T25). When asked about a policy to "[r]egulate carbon dioxide (the primary greenhouse gas) as a pollutant," the majority of subjects in the alarmed (97%) and concerned (95%) categories respond that they either "strongly support" or "somewhat support" such a policy. Again, we are mapping these categories to our green accounts in the ontology, and we consider this SCR pair to be a green indicator. In contrast, the Six Americas categories that we map to the denier accounts responded with a much lower level of support for this type of policy: doubtful (56%) and dismissive (23%). In the same section of the Six Americas survey, another question asks the subject if she would support a policy to "[s]ign an international treaty that requires the United States to cut its emissions of carbon dioxide 90% by the year 2050." The majority of subjects in the green categories, alarmed (97%) and concerned (87%), once again said they would either "strongly support" or "somewhat support" this policy, while relatively few subjects in the denier categories said they would support the policy: doubtful (30%) and dismissive (15%).

---

additional research is needed on this point.

5.3.4 Economic-Growth Indicators

The economic–growth SCR pair is the second denier concept pair we consider in this study. The Six Americas survey first queries a subject as to whether he believes efforts towards environmental conservation tend to "improve" or "reduce" growth in the economy (or neither). A follow-up part to the question asks the subject to decide if economic growth should be prioritized over protecting the environment when the two issues are in conflict (Leiserowitz et al., 2010, T11). Although the first part potentially allows for a green perspective regarding economic growth, we chose this concept pair as a denier indicator due to how the second part frames the issue as choice between environmentalism and the economy. Thus, our initial intent was that the economic–growth indicator be a test for hypothesis 1B.

Table 5.18 Coverage of green and denier weak economic–growth accounts.

Min. Tweets	**No. Users**	**Weak Economic–Growth Accounts**					
				Green		**Denier**	
2	2376	17	0.72%	17	0.72%	0	0.00%
3	1347	12	0.89%	12	0.89%	0	0.00%
4	927	11	1.19%	11	1.19%	0	0.00%
5	705	10	1.42%	10	1.42%	0	0.00%
6	561	8	1.43%	8	1.43%	0	0.00%
7	468	8	1.71%	8	1.71%	0	0.00%
8	392	6	1.53%	6	1.53%	0	0.00%
9	345	5	1.45%	5	1.45%	0	0.00%
10	302	5	1.66%	5	1.66%	0	0.00%
11	271	5	1.85%	5	1.85%	0	0.00%
12	240	5	2.08%	5	2.08%	0	0.00%
13	216	5	2.31%	5	2.31%	0	0.00%
14	198	4	2.02%	4	2.02%	0	0.00%
15	183	4	2.19%	4	2.19%	0	0.00%
16	166	4	2.41%	4	2.41%	0	0.00%
17	151	4	2.65%	4	2.65%	0	0.00%
18	137	4	2.92%	4	2.92%	0	0.00%
19	130	3	2.31%	3	2.31%	0	0.00%
20	119	3	2.52%	3	2.52%	0	0.00%

Table 5.18 lists the counts and percentages of coverage for the weak economic–growth account indicators for the range of minimum-tweet activity levels considered in this series of experiments. We see that this indicator has the lowest coverage of the base user

set out of all the SCR pairs presented. Moreover, we expected this concept pair to serve in identifying users in the denier category, but the coverage for accounts labelled as denier is zero at all activity levels. Figure 5.15 shows this data graphically with the blue, green, and dark magenta lines respectively representing the classes: *WeakEconomicGrowthAccount* [WEGA], *GreenWeakEconomicGrowthAccount* [GWEGA], and *DenierWeakEconomicGrowthAccount* [DWEGA]. Note that the blue line is completely covered by the green as all the weak economic–growth indicator accounts are in the green category. Accordingly, the dark magenta line simply traces the x-axis as there is 0% coverage for the accounts labelled denier. We should also note that, as is often the case, we see a peak in coverage at a minimum-tweet level of 18 for the weak indicator. Moreover, looking at Table 5.19, we find the same peak for the strong indicator.

Table 5.19 Coverage of green and denier strong economic–growth accounts.

Min. Tweets	**No. Users**	**Strong Economic–Growth Accounts**					
				Green		**Denier**	
2	2376	12	0.51%	12	0.51%	0	0.00%
3	1347	9	0.67%	9	0.67%	0	0.00%
4	927	9	0.97%	9	0.97%	0	0.00%
5	705	8	1.13%	8	1.13%	0	0.00%
6	561	6	1.07%	6	1.07%	0	0.00%
7	468	6	1.28%	6	1.28%	0	0.00%
8	392	4	1.02%	4	1.02%	0	0.00%
9	345	4	1.16%	4	1.16%	0	0.00%
10	302	4	1.32%	4	1.32%	0	0.00%
11	271	4	1.48%	4	1.48%	0	0.00%
12	240	4	1.67%	4	1.67%	0	0.00%
13	216	4	1.85%	4	1.85%	0	0.00%
14	198	3	1.52%	3	1.52%	0	0.00%
15	183	3	1.64%	3	1.64%	0	0.00%
16	166	3	1.81%	3	1.81%	0	0.00%
17	151	3	1.99%	3	1.99%	0	0.00%
18	137	3	2.19%	3	2.19%	0	0.00%
19	130	2	1.54%	2	1.54%	0	0.00%
20	119	2	1.68%	2	1.68%	0	0.00%

Given that strong (syntactic dependency) indicators are subclasses of their corresponding weak (co-occurrence) indicators, it is not surprising that the results for strong economic–growth indicators parallel those of the weak indicators, especially given that we are considering only green accounts with this concept pair. Data for the counts and percentage

Figure 5.15 Percent coverage of green and denier weak CO_2–cut accounts.

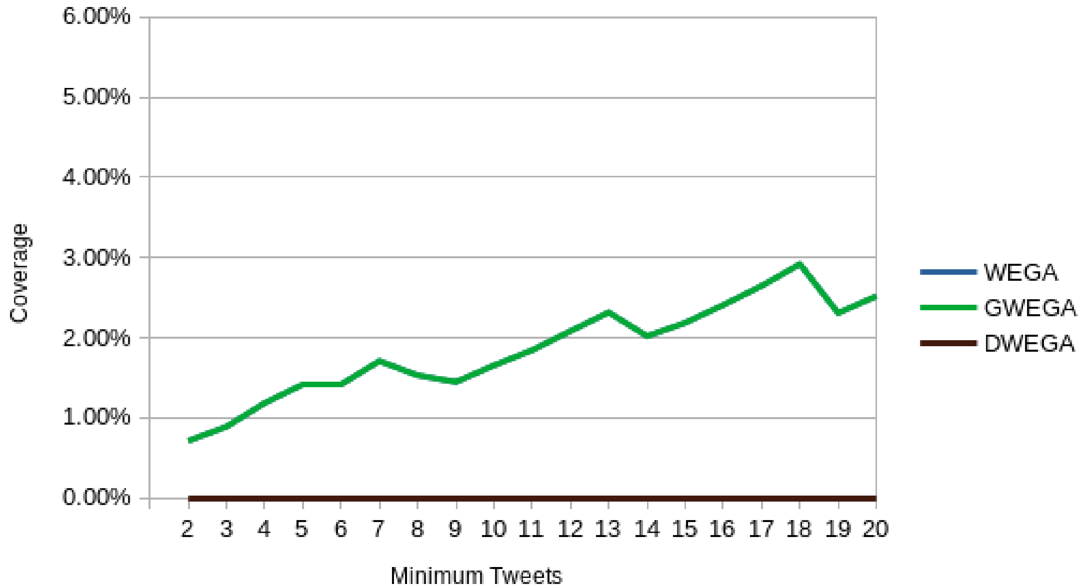
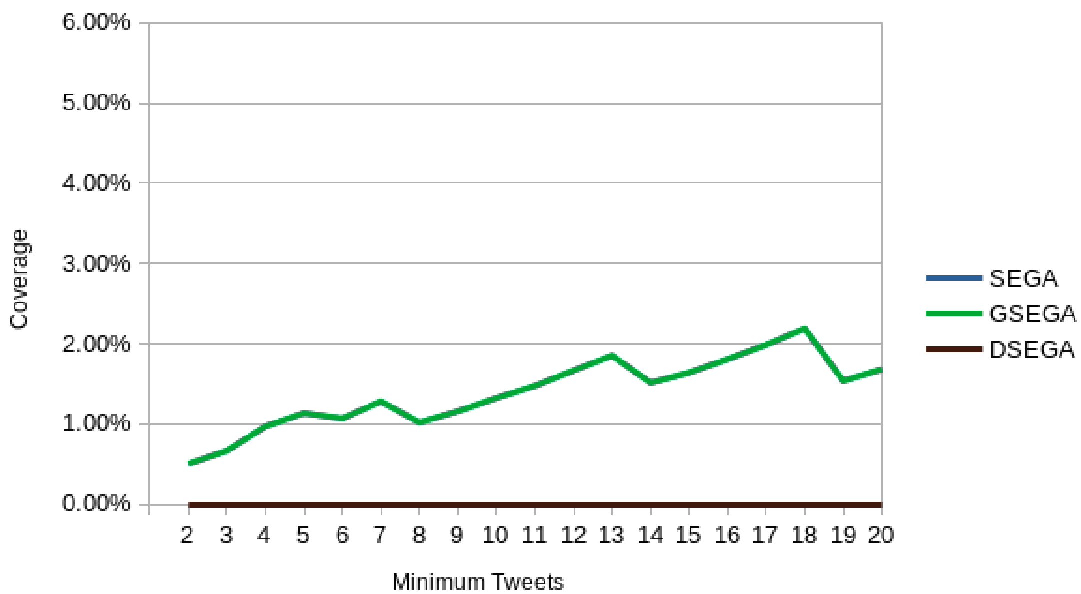


Figure 5.16 Percent coverage of green and denier strong CO_2–cut accounts.



of base user coverage are shown in Table 5.19. For the strong indicator, coverage of users labelled green basically follows that of the weak with slightly lower values in each case. Of course, once again the coverage for users in the denier category is 0%. These results may be visualized more clearly in Figure 5.16, where the blue, green, and dark magenta lines respectively represent the classes: *StrongEconomicGrowthAccount* [SEGA], *GreenStrongEconomicGrowthAccount* [GSEGA], and *DenierStrongEconomicGrowthAccount* [DSEGA]. Once again, the green line follows the blue exactly and covers it. The dark magenta line for the denier indicator simply runs the axis at 0%.

Table 5.20 Green and denier ratios for weak and strong economic–growth accounts.

Min.	**Weak Economic–Growth Account**					**Strong Economic–Growth Account**				
Tweets	**Total**	**Green**		**Denier**		**Total**	**Green**		**Denier**	
2	17	17	100.00%	0	0.00%	12	12	100.00%	0	0.00%
3	12	12	100.00%	0	0.00%	9	9	100.00%	0	0.00%
4	11	11	100.00%	0	0.00%	9	9	100.00%	0	0.00%
5	10	10	100.00%	0	0.00%	8	8	100.00%	0	0.00%
6	8	8	100.00%	0	0.00%	6	6	100.00%	0	0.00%
7	8	8	100.00%	0	0.00%	6	6	100.00%	0	0.00%
8	6	6	100.00%	0	0.00%	4	4	100.00%	0	0.00%
9	5	5	100.00%	0	0.00%	4	4	100.00%	0	0.00%
10	5	5	100.00%	0	0.00%	4	4	100.00%	0	0.00%
11	5	5	100.00%	0	0.00%	4	4	100.00%	0	0.00%
12	5	5	100.00%	0	0.00%	4	4	100.00%	0	0.00%
13	5	5	100.00%	0	0.00%	4	4	100.00%	0	0.00%
14	4	4	100.00%	0	0.00%	3	3	100.00%	0	0.00%
15	4	4	100.00%	0	0.00%	3	3	100.00%	0	0.00%
16	4	4	100.00%	0	0.00%	3	3	100.00%	0	0.00%
17	4	4	100.00%	0	0.00%	3	3	100.00%	0	0.00%
18	4	4	100.00%	0	0.00%	3	3	100.00%	0	0.00%
19	3	3	100.00%	0	0.00%	2	2	100.00%	0	0.00%
20	3	3	100.00%	0	0.00%	2	2	100.00%	0	0.00%

For completeness, Table 5.20 lists the green and denier ratios for the economic–growth indicator accounts. If there is discriminatory bias for this indicator, it is certainly towards the green category as 100% of the members of both the weak (left side) and the strong (right side) indicator classes are green accounts. This surprising result is similar to what we found for the other denier indicator, nature–cause (Table 5.12) except that the lack of any capacity to identify accounts in the denier category is even more striking. In Section 5.3.5 we will see if these denier indicators might actually serve to help identify green accounts.

With such low coverage of the user base, we do not expect a high level of predictive capability for the economic–growth SCR pair. Table 5.21 reports the precision, recall, and F1 scores for both the weak and strong indicator accounts. Note again that predictions are with respect to the green label even though the economic–growth pair was originally intended to serve as a denier indicator. The results for recall support hypothesis 2A as the weak indicator performs better than the strong. However, little may be said with regard to precision and hypothesis 2B. All individuals who are members of the *WeakEconomicGrowthAccount* class are labelled green, which implies the same is true for the subset of individuals belonging to the *StrongEconomicGrowthAccount* class. Therefore, in all cases the precision is 100%. Of course, the poor scores for recall bring down the F1 to the lowest of all the other SCR pairs for weak indicators. For the strong indicators, this concept pair is roughly comparable to nature–cause, and both concept pairs score lower F1 measures than either of the green pairs. Figure 5.17 charts the F1 scores for economic–growth. Both the weak and the strong F1 scores peak at a minimum activity level of 18 with scores of 8.16% and 6.19% respectively.

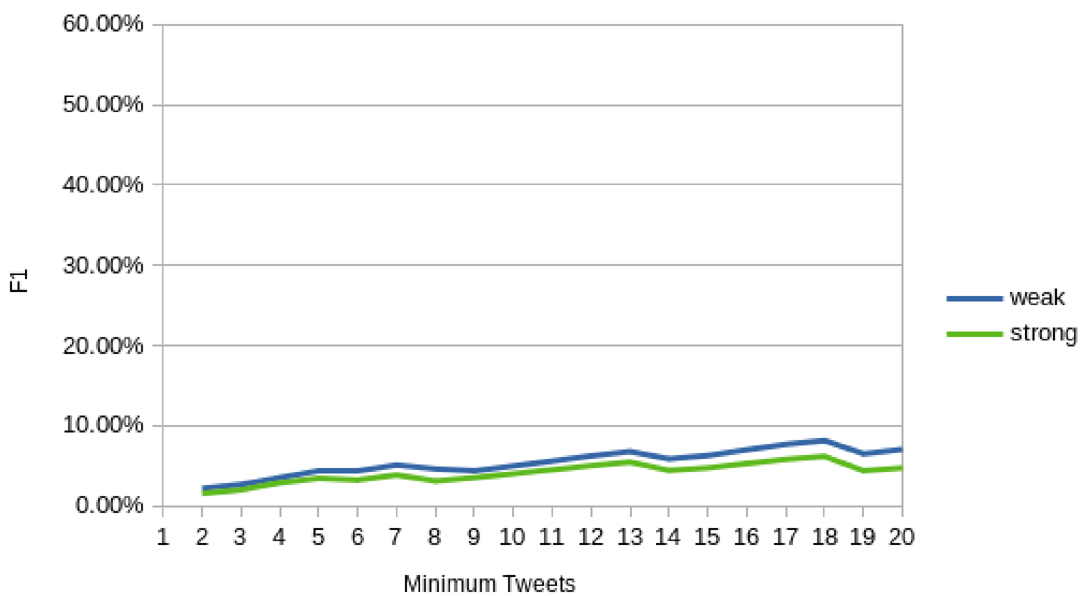
If the nature–cause SCR pair made for a poor denier indicator, the economic–growth pair performs even worse in this regard. As mentioned above, it was originally chosen as such because in the section on "Environmental protection and economic growth," the Six Americas survey (Leiserowitz et al., 2010, T11) asks the question, "When there is a conflict between environmental protection and economic growth, which do you think is more important?" The majority of subjects in the denier categories, doubtful (60%) and dismissive (74%), answered, "Economic growth, even if it leads to environmental problems." Unsurprisingly, very few subjects in the green categories gave this answer: alarmed (3%) and concerned (16%). Nevertheless, it appears primarily to be users in the green category who are speaking their mind on economic growth with tweets such as this one from the survey concept dataset:

Deutsche Bank report highlights sustainable growth: The bank asks if humans will sac-

Table 5.21 Predictive scores for weak and strong economic–growth accounts.

Min. Tweets	Weak Economic–Growth Account			Strong Economic–Growth Account		
	Precision	Recall	F1	Precision	Recall	F1
2	100.00%	1.11%	2.20%	100.00%	0.78%	1.56%
3	100.00%	1.35%	2.67%	100.00%	1.02%	2.01%
4	100.00%	1.80%	3.53%	100.00%	1.47%	2.90%
5	100.00%	2.19%	4.28%	100.00%	1.75%	3.44%
6	100.00%	2.19%	4.29%	100.00%	1.64%	3.23%
7	100.00%	2.62%	5.11%	100.00%	1.97%	3.86%
8	100.00%	2.37%	4.63%	100.00%	1.58%	3.11%
9	100.00%	2.24%	4.39%	100.00%	1.79%	3.52%
10	100.00%	2.56%	5.00%	100.00%	2.05%	4.02%
11	100.00%	2.89%	5.62%	100.00%	2.31%	4.52%
12	100.00%	3.23%	6.25%	100.00%	2.58%	5.03%
13	100.00%	3.52%	6.80%	100.00%	2.82%	5.48%
14	100.00%	3.03%	5.88%	100.00%	2.27%	4.44%
15	100.00%	3.25%	6.30%	100.00%	2.44%	4.76%
16	100.00%	3.64%	7.02%	100.00%	2.73%	5.31%
17	100.00%	4.00%	7.69%	100.00%	3.00%	5.83%
18	100.00%	4.26%	8.16%	100.00%	3.19%	6.19%
19	100.00%	3.37%	6.52%	100.00%	2.25%	4.40%
20	100.00%	3.66%	7.06%	100.00%	2.44%	4.76%

Figure 5.17 F1 scores (green) for weak & strong economic–growth accounts.



> *rifice economic growth to halt environmental damage. #Sustainability #ClimateChange #Davos2020 @DeutscheBank*

5.3.5 Indicator Ensembles

We have analyzed the four SCR pairs separately; however, our intention when designing this experimental model was to use them together as a tool to predict the stance on climate change for Twitter users in the dataset. We selected two green indicators (human–cause and CO_2–cut) and two denier indicators (nature–cause and economic–growth). The green indicators are intended to test hypothesis 1A, which states that they may be used to identify users in the green category. In Section 5.1.4 we discussed definitions in the say-sila ontology for two weak inferred-green account classes. We repeat the definition of the first one (Equation 5.21) here for reference:

$$\begin{aligned} WeakInferredGreenAccount \equiv\ & WeakHumanCauseAccount \\ & \sqcup WeakCO2CutAccount \end{aligned} \tag{5.36}$$

Likewise, in Section 5.1.5 we presented a strong inferred-green account class (Equation 5.29), which we also repeat for reference:

$$\begin{aligned} StrongInferredGreenAccount \equiv\ & StrongHumanCauseAccount \\ & \sqcup StrongCO2CutAccount \end{aligned} \tag{5.37}$$

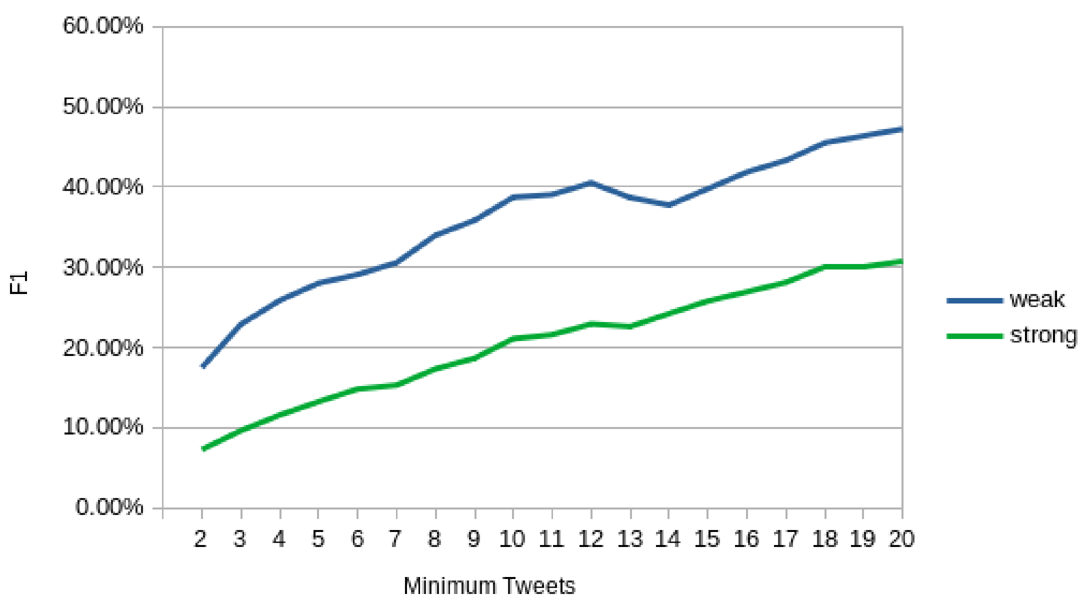
The reader may recall that there are two types of weak and two types of strong inferred-green classes in the ontology (regular and “plus”). The regular types are the disjunction of the two (weak or strong) green concept indicators, human–cause and CO_2–cut, as presented in Equations 5.36 and 5.37.

Because we are simply using a disjunction of the two account indicator classes to define

Table 5.22 Predictive scores for weak and strong inferred-green accounts.

Min.	**Weak Inferred-Green-1 Account**			**Strong Inferred-Green-1 Account**		
Tweets	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
2	69.86%	10.01%	17.51%	66.29%	3.86%	7.29%
3	70.76%	13.66%	22.89%	67.65%	5.19%	9.64%
4	67.59%	16.01%	25.89%	63.93%	6.37%	11.59%
5	64.06%	17.94%	28.03%	60.71%	7.44%	13.26%
6	60.34%	19.18%	29.11%	58.49%	8.49%	14.83%
7	58.88%	20.66%	30.58%	56.25%	8.85%	15.30%
8	60.00%	23.72%	33.99%	55.32%	10.28%	17.33%
9	60.00%	25.56%	35.85%	55.56%	11.21%	18.66%
10	61.80%	28.21%	38.73%	59.52%	12.82%	21.10%
11	60.24%	28.90%	39.06%	57.50%	13.29%	21.60%
12	61.04%	30.32%	40.52%	59.46%	14.19%	22.92%
13	58.57%	28.87%	38.68%	57.14%	14.08%	22.60%
14	57.81%	28.03%	37.76%	60.61%	15.15%	24.24%
15	58.73%	30.08%	39.78%	62.50%	16.26%	25.81%
16	58.06%	32.73%	41.86%	61.29%	17.27%	26.95%
17	59.65%	34.00%	43.31%	64.29%	18.00%	28.13%
18	64.71%	35.11%	45.52%	69.23%	19.15%	30.00%
19	65.31%	35.96%	46.38%	70.83%	19.10%	30.09%
20	66.67%	36.59%	47.24%	72.73%	19.51%	30.77%

Figure 5.18 F1 scores for weak & strong inferred green accounts.



the inferred-green class, we do not explicitly present results for the dataset coverage nor the green and denier ratios for *WeakInferredGreenAccount*. These values simply represent the union of the individuals who are members of the weak (or the strong) human–cause account indicator classes from Table 5.6 (or Table 5.7) or else the weak (or the strong) CO_2–cut account indicator classes from Table 5.14 (or Table 5.15). Of greater interest undoubtedly is the predictive capability for this ensemble class. Table 5.22 lists the scores for precision and recall for the weak and strong inferred-green account classes. For these regular (not "plus") inferred-green account classes, as expected, we see that the recall scores are higher for the weak ensemble indicator than they are for the strong. For precision, however, we note that the weak indicator actually performs better than the strong up to a minimum activity level of 13, after which the strong indicator scores higher. These results support hypotheses 2A and 2B with respect to recall but support them for precision only at higher levels of minimum activity. This finding is not too surprising given that results for the tweet dataset are generally more stable at higher minimum-activity levels, and we did see that the weak human–cause indicator reports better precision than the associated strong indicator when considering that SCR pair separately. Additionally, Table 5.22 reports the F1 scores for the weak and strong ensemble indicator classes, and Figure 5.18 charts these scores for the range of minimum activity levels being considered. We note immediately that the F1 scores are visibly higher than they were for the individual account indicators (see Figure 5.8 for human–cause and Figure 5.14 for CO_2–cut).

These results show a level of predictive capability for the green SCR pairs and therefore lend support for hypothesis 1A. Yet, as disclosed at the beginning of this section, it became immediately obvious during the early test runs that the SCR pairs we are considering to be denier indicators were not useful at all for identifying users in the denier category. This initial finding made it clear that hypothesis 1B was not valid (at least not with respect to our chosen denier concept pairs). However, we did see that the nature–cause and economic–growth SCR pairs showed some predictive capability for users in the green category. Therefore, we added these indicators to the disjunctive series of account indicator classes to define the "plus" inferred-green account classes. We repeat their definitions (Equations 5.22 and 5.30) here for the reader's convenience:

$$\begin{aligned} WeakInferredGreenAccountPlus \equiv\ & WeakHumanCauseAccount \\ & \sqcup WeakNatureCauseAccount \\ & \sqcup WeakCO2CutAccount \\ & \sqcup WeakEconomicGrowthAccount \end{aligned} \tag{5.38}$$

$$\begin{aligned} StrongInferredGreenAccountPlus \equiv\ & StrongHumanCauseAccount \\ & \sqcup StrongNatureCauseAccount \\ & \sqcup StrongCO2CutAccount \\ & \sqcup StrongEconomicGrowthAccount \end{aligned} \tag{5.39}$$

Simply put, the inferred-green account "plus" classes take into account all the SCR pairs we are considering in this study. A user who publishes a tweet that can be modelled by any one of the four text indicator classes is a member of either the *WeakInferredGreenAccountPlus* class (if the tweet has a co-occurrence of tokens linked to a pair of concepts) or the *StrongInferredGreenAccountPlus* class (if those tokens have a relation of syntactic dependency in the text). Again, we do not explicitly present the base user coverage nor the green and denier ratios for these ensemble indicators, noting rather that they are the union of the members from the weak or strong account indicator classes for the four SCR pairs: human–cause, nature–cause, CO_2–cut, and economic–growth. Table 5.23 lists the precision, recall, and F1 scores for these weak and strong inferred-green account "plus" classes. As we saw with the regular ensemble classes, while recall is indeed better for the weak inferred-green class as opposed to the strong, precision is improved for the strong classes only at a minimum-tweet level of 14 or higher. These findings support hypotheses 2A and 2B only for those users participating at higher levels of minimum activity and suggest that an analysis of the larger community on Twitter may not be so direct (see Section 5.4.3). The F1 scores are the highest we have found and show improvement over those of the regular indicator ensembles (Table 5.22) for all levels of minimum activity.

Table 5.23 Predictive scores for weak and strong "plus" inferred-green accounts.

Min. Tweets	**Weak Inferred-Green-2 Account**			**Strong Inferred-Green-2 Account**		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
2	71.32%	12.03%	20.59%	69.44%	4.91%	9.16%
3	72.28%	16.48%	26.84%	69.88%	6.55%	11.97%
4	69.19%	19.44%	30.36%	68.00%	8.33%	14.85%
5	66.89%	22.10%	33.22%	66.18%	9.85%	17.14%
6	63.36%	22.74%	33.47%	63.49%	10.96%	18.69%
7	61.67%	24.26%	34.82%	60.71%	11.15%	18.84%
8	62.73%	27.27%	38.02%	59.62%	12.25%	20.33%
9	62.50%	29.15%	39.76%	60.00%	13.45%	21.98%
10	64.95%	32.31%	43.15%	63.83%	15.38%	24.79%
11	63.74%	33.53%	43.94%	62.22%	16.18%	25.69%
12	64.71%	35.48%	45.83%	63.41%	16.77%	26.53%
13	62.82%	34.51%	44.55%	61.54%	16.90%	26.52%
14	62.50%	34.09%	44.12%	63.89%	17.42%	27.38%
15	62.86%	35.77%	45.60%	64.71%	17.89%	28.03%
16	62.32%	39.09%	48.04%	63.64%	19.09%	29.37%
17	64.06%	41.00%	50.00%	66.67%	20.00%	30.77%
18	67.86%	40.43%	50.67%	71.43%	21.28%	32.79%
19	68.52%	41.57%	51.75%	73.08%	21.35%	33.04%
20	69.39%	41.46%	51.91%	75.00%	21.95%	33.96%

Figure 5.19 F1 scores for weak & strong "plus" inferred green accounts.

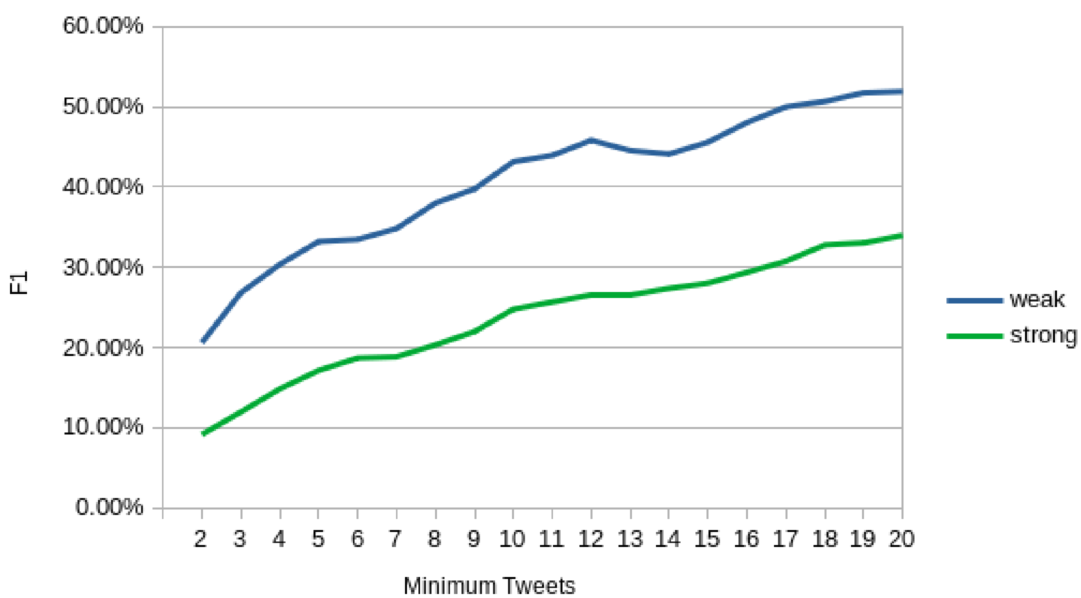


Figure 5.19 presents the F1 scores in graph form for the weak and strong green-inferred "plus" indicator classes. The weak indicator ensemble outperforms the strong as we have seen for every indicator class in the study. Note that we do not see a peak in the F1 scores for the ensemble as we did with both the (supposed) denier concept pairs when tested individually. Although there is a slight dip in the prediction performance at a minimum activity level of 14,[41] generally the more published tweets needed to include users in an experimental run, the better the model is at determining whether those users are in the green or the denier category. This makes sense as the whole collected set of a user's tweets are used for the purpose of determining the category representing her stance on global warming.

This series of experiments with the say-sila ontological model supports hypothesis 1A in that the SCR pairs provide a level of predictive capability for identifying users in the green category. However, we were forced to abandon hypothesis 1B early in the process. Moreover, when we changed what were thought to be denier concept pairs into additional green concept pairs (modelled using the "plus" inferred-green indicator ensemble classes), the predictive capability improved. One might speculate, therefore, that by adding more concept pairs to the ensemble, we could achieve better scores. One problem with this strategy is that the SCR pairs in this work were chosen because we were able to identify concepts from the Six Americas study which potentially reflect frequently discussed content from the general *#globalwarming* conversation on Twitter (see Section 5.1.3). Additional concepts from the Six Americas that do not show up in tweets will be of limited use. However, we did have two additional concept pairs that were a very large part of the textual content in the survey concept tweet dataset: "energy–conservation" and "environment–protect." We did not include these SCR pairs in the general model because the associated concepts were part of the filtering process used to generate this dataset. However, while these indicators may not be appropriate

---

[41]Notably, 14 is the level of minimum activity where the strong indicator ensembles begin to have higher precision than the weak.

for the official series of experiments, they can serve to give us a rough idea of how adding more indicators may help to improve the model. When analyzed separately, these pairs showed very modest predictive capability for both the weak and strong account indicator classes. When added as part of new ensemble classes (i.e., "plus-plus"), they resulted in an additional 1-2% for the F1 scores over the "plus" indicator ensemble classes. Our interpretation of these informal results is that additional SCR pairs in indicator ensemble classes will likely yield diminishing improvements to the results once one has identified the primary relevant concepts that show up frequently in the microblogs.

5.3.6 Overall Observations

We feel that the most striking outcome of this set of experiments is that while the green concept-pair indicators demonstrate a modest level of predictive capability with regard to classifying green users, the denier indicators did not demonstrate any predictive capability whatsoever for users in the denier category. On the contrary, these denier indicators actually served to help predict green users as we saw when using the "Plus" ensemble to classify users in the final experiments in this chapter.

Our experiments support our hypothesis that using weak indicators (a co-occurrence of paired concepts) serves to increase the recall when attempting to identify green users as weak indicators led to better recall than strong indicators across all experiments. However, we found only partial support for our hypothesis that strong indicators (a syntactic dependency between the words expressing paired concepts) would lead to stronger precision in our experiments. Rather, we found this to be the case only at higher levels of minimum user participation. For minimum tweet levels lower than around 14 tweets, we were surprised to see the weak indicators generally giving higher scores for precision.

Weak indicators are clearly the best performer among the ontological entities we have examined. The concept pair with the highest predictive capability is human–cause, followed in order by CO2–cut, nature–cause, and finally economic–growth. Using these indicators together in ensemble classes provides an even higher predictive capability in terms of the F1 scores when classifying green users.

5.4 Limitations

As we developed the say-sila model and performed the experiments covered in this chapter, we noted a number of limitations with respect to the research efforts we have made. We list these here so they may be duly noted and consulted for future research endeavours.

5.4.1 Consistency between the SCR and Test Datasets

As described in Section 5.1.1, we selected the survey concept rules (SCR) by finding terms from the Six Americas questionnaire (Leiserowitz et al., 2010) which were frequently used in a collection of tweets filtered for textual content dealing with environmental conservation. (We ignored concepts related to the filter keywords.) We chose this selection method because it has the advantage of targeting tweets that are discussing a topic which has been demonstrated to be a primary predictor of a person's stance on climate change (McCright et al., 2016). Along with this advantage, however, we realize that the strategy creates an inherent inconsistency between the SCR dataset and the test dataset used to conduct our analysis. Essentially, we are incorporating existing research on the psychology of climate change to construct our ontological model and then using the general 2019 *#globalwarming* conversation on Twitter to populate that model.

Our aim in the present research is to use established findings on climate change from the human sciences to predict the attitudes and beliefs of general online users who are publishing microblogs on the many facets of global warming. Therefore, we stand by the methodology we have elected for the experimentation as laid out in this chapter. That said, we recognize that continued research would be invaluable to compare these findings to the results generated when more parallel methods of tweet selection are employed for dataset creation.

5.4.2 Redundancy in Tweets

We mentioned in Section 5.1 that the creation of the say-sila ontological model was an iterative process. As such, we spent an appreciable amount of time looking at individual tweets either to evaluate how well the automated model creation processes were working or to determine why a given set of results was showing something that was perhaps unexpected. Given that we consistently see that the most conclusive results are those from users participating at a higher activity level (i.e., publishing more tweets), one pattern in the tweets that we saw and deemed noteworthy is a high level of redundancy in frequent tweets from a number of online users. This is a known problem with online NLP analysis and is often considered a form of noise (Farzindar & Inkpen, 2015). Some users may be advertising a product or simply sending the same (or similar) text multiple times, possibly via an automated process (i.e., a "bot"). The scope of this doctoral research has not allowed for an in-depth look at these individuals, but continued research efforts are certainly warranted on this topic. They may serve not only to refine the capabilities of research models such as this one, but also to better understand the nature of online communications in their own right.

5.4.3 A Need for Big Players

In the present work, we have found numerous occasions where looking at a subset of higher activity users on Twitter may allow us to identify trends and patterns more clearly than is possible when analyzing the full set of users in an online community (e.g., those tweeting about *#globalwarming*). Indeed from our initial experiments described in Chapter 3, grouping users according to a minimum level of participation has become an essential part of our methodology for this analysis.

NLP applications processing online communications generally have to deal with a large amount of noise (Farzindar & Inkpen, 2015). The fact that we produce less conclusive findings for lower levels of minimum participation may likely be due at least in part to an increase in noise in the text of the tweets being analyzed. In Chapter 3 we found some indications that emotion in high-activity users may be representative to an extent of emotion in the larger community. Although we have not tested it explicitly for our ontological model, we have proceeded with some underlying expectation that our results will be more interesting and likely more stable when analyzing groups of users that are publishing more tweets than the others in their community.

Yet, we must take note of the results from our experiments with the ensemble indicators (Section 5.3.5) where hypotheses 2A and 2B proved valid for the higher activity levels, but not for the lower ones. These results are markedly different from those where higher activity users simply score higher, and may indicate underlying considerations significantly more complex than just "less noise." The redundancy issue described in the previous section may also be at play here. Further research may help to explain the different factors emphasized by various subsets of the larger online community and how they may affect a given research model that is intended to represent them.

5.4.4 Integration with Semantic Web Ontologies

In many ways this chapter represents the culmination of a discussion of ontologies which began back in Chapter 2. There we cited the commonly-accepted definition that an ontology is a "shared conceptualisation" (Studer et al., 1998). We have notably embraced the idea of a shared model in our choice of DUL (Presutti & Gangemi, 2016) as the top-level ontology for the say-sila model. However, there are concepts modelled in our ontology which have been used in other research projects. We considered a number of these for incorporation into our model but in the end chose not to include them.

An important example is the EmotionsOnto ontology (Gil et al., 2015). Emotions are an important theme in the present research, and although they show up minimally in our final experiments for predicting the stance of online users on climate change, they were still part of the larger experimental procedure covered in this chapter. Moreover, they form an integral part of the say-sila ontological model. On one hand, modelling affect and emotion using the EmotionsOnto ontology would be ideal since it uses DOLCE as a top-level ontology. On the other hand, its use is essentially geared towards emotion from a biological standpoint. Although this makes it appropriate for applications in psychology, we found it did not integrate seamlessly with the elements in our ontology intended to model the required aspects of communications on social media.

Of course, there are ontologies modelling the various facets of social media, which are common on the Semantic Web and recommended by the W3C. Most notably, FOAF (Friend of a Friend)[42] defines standard elements concerning people, including basic information, e.g., age, gender, phone number, as well as account data pertinent to a person's online presence (Hitzler et al., 2010). Also, SIOC (Semantically-Interlinked Online Communities)[43] builds on FOAF to model users on social media, specifically with respect to their online activities, forums, their relationships with other users, and the online communities they form (Gandon et al., 2012). While in many ways these ontologies would be ideal for our use in the present research, they are not built on DUL nor on any other established top-level ontology, and we found it somewhat awkward to force them into a common ontology based on DUL. Nevertheless, *Say Sila* does support a configuration option that declares that a number of classes from FOAF and SIOC are equivalent with the associated classes in the say-sila ontology.[44] However, the ontological model, as presented in this chapter, has that configuration option off.

One study by (Cotfas et al., 2015) takes an approach employing these ontologies for

---

[42] http://xmlns.com/foaf/0.1/

[43] http://rdfs.org/sioc/ns#

[44] https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/sila.clj#L1582

modelling social media and online activities. The research team created a set of ontologies for emotion mining in Twitter communications. Concepts in these ontologies build on the FOAF and SIOC ontologies where appropriate.[45] For the present project we ultimately chose a solid integration with a top-level ontology over an integration with the W3C recommendations for the Semantic Web. Nevertheless, we believe it would be well worth the effort to continue our research efforts with the goal of integrating the say-sila model more closely with established Semantic Web ontologies while keeping a firm foundation on an upper ontology like DUL.

5.5 Contributions and Continued Research

This chapter describes the multiple uses of the say-sila ontology for our analysis in the present research effort. We use the ontology to denote each Twitter user's stance (green or denier) with respect to climate change as a function of which known leader accounts the user is following. We also utilize the ontology for the goal of inferring a user's category (green or denier) based on the tweets she has published in 2019 on the topic of global warming. Finally, we use the ontology itself to find when the stance inferred using description logic matches the stance indicated by checking whom the user is following.

In the previous section, we noted a number of limitations for this study, and these begin to pave a road towards continued research. In addition to addressing these limitations, however, there are several positive findings on which to build as we continue to use the say-sila ontology as a research tool to analyze online attitudes towards climate change. Perhaps the most interesting finding is the fact that results when modelling using weak indicators were much better than results from strong indicators. While we generally saw that weak indicators scored higher for recall, and strong indicators scored higher for

[45] The article states that the ontology is available at https://github.com/liviucotfas/TweetOntoSense; however, the website simply gives instructions about whom to contact for more information.

precision, supporting our hypotheses 2A and 2B,[46] the F1 measures were notably better for weak indicators. This observation holds for all SCR pairs and for the ensembles at all the considered levels of minimum participation.

This is an important finding which leads to a crucial question when examining possible directions for continued research. From a certain point of view, a good deal of time and computing resources are needed to determine the syntactic dependency of the tokens in each tweet so that we may model these relationships when populating the ontology. Furthermore, the additional ontological constructs needed to model the strong indicators make for a "heavier" ontology for a given number of individuals, increasing the time a DL reasoner needs to run so that it can make the inferences we are seeking. As a doctoral candidate performing the research discussed above, I can recount that throughout the iterative process of model creation and evaluation, I had an ongoing goal of hitting an acceptable level of predictive capability so that we could then move to the next level of sophistication for the model and thereby improve upon that capability. However, it was the less "sophisticated" modelling technique which demonstrated the best results time and again. The co-occurrence of a key pair of concepts in a text were more indicative of user stance on climate change than a relationship of syntactic dependency between those concepts. We could interpret our results as a reminder that often the simple solution is the ideal one. In future research, relying solely on the weak indicators may allow a more focused effort towards evaluating the most important concepts from the Six Americas and perhaps other socio-psychological studies on climate change.

Yet, if we look at the weak and strong models from a different point of view, we are essentially left here with a mystery to solve. Certainly, concept pair indicators based on syntactic dependency did not lead to models with increased scores for precision as we expected. But why not? One might consider a strong model to be rather similar to its weak counterpart, just with an added constraint. After all, strong account indicators do

[46]The human–cause indicator was an exception with respect to hypothesis 2B as the strong indicators did not show higher precision compared to the weak indicators.

represent subclasses of the weak account indicators in the ontological model. However, our results seem to indicate that the strong models are different from the weak in a way that is somehow more fundamental. When looking at the human–cause SCR pair, it seemed that the strong models might be discriminatory towards the denier category. This particular observation did not hold true for the other SCR pairs we considered; however, something is indeed different about the strong models. A closer examination of syntactic dependency in online posts could shed light on the more surprising results found in the present research. We have mentioned previously how natural language as it exists on social media is inherently noisy and how its analysis represents a sizable challenge (see Section 1.3). It may be enlightening to compare how our results in this study differ from those produced when the model is based on a corpus of more formally written documents on global warming. We have also discussed how the usage of irony and sarcasm can potentially interfere with models of online conversations (see Section 3.4.2). It would be interesting to bring these types of expressive elements into consideration in order to determine to what extent they may be tied to relations of syntactic dependency between tokens in a microblog.

Of course, as we look forward, our biggest challenge is perhaps figuring out how to directly identify users in the denier category. Our methodology demonstrated general support for hypothesis 1A (green concept pairs allow us to identify users in the green category).[47] However, we were unable to demonstrate any support for hypothesis 1B (denier concept pairs allow us to identify users in the denier category). Rather, the denier SCR pairs actually proved to be an additional means by which to find those same green users. The denier category represents the rare class among the online users we are considering, and so at least from the standpoint of a machine learning problem, identifying these users is the more desirable goal. Working towards achieving this goal is an obvious next step in our ongoing research.

---

[47] We noted one exception, however, for users in the *DenierStrongHumanCauseAccount* class (see Section 5.3.1).

Even as we anticipate continuing our research endeavours and refining the say-sila ontology, it is worth reflecting momentarily on our model as it currently stands and on the contributions it represents for the scientific community. We have developed an iterative methodology which essentially links online communications about a given subject to a survey-based research effort in the human sciences. In our case the communications were microblogs on the subject of global warming, and the study was the Six Americas. Yet, this same process may be readily adapted to virtually any online conversation as it relates to various important research efforts. Additionally, we have taken an established method for determining ground truth for online users with respect to their position on a given subject, adapted this method for the subject of climate change, and integrated it into an ontological model. This feature allows the model to be used not only for logical inference, but also as an evaluation method for the results it produces. The say-sila ontology can augment the excellent, survey-based research associated with the Six Americas, essentially expanding its reach to enable the study of online communities of hundreds of thousands of users or more. Our model effectively represents a tool by which research in cognitive science may be automated to generate discoveries about people online that may then be used to corroborate findings made using the traditional, survey-based experiments. Finally, knowledge gained from studying humans via their online activity may serve as a significant contribution back to the original cognitive study which is the foundation of the model. We look forward to coming full-circle in this way with this ontological model and indeed with all our research for the *Say Sila* project.

CHAPTER VI

THE SIX AMERICAS ON TWITTER: AFFECTIVE SIGNATURES

In the previous chapter we modelled several important concepts from the Six Americas survey (Leiserowitz et al., 2010) as an aid in determining the stance (green or denier) of users on Twitter with respect to their beliefs and attitudes about climate change. Now we will look in more detail at the questions from the survey. Specifically, we are interested in the affect expressed in tweets that may be related to each question as well as the affect expressed by the users in all their tweets about *#globalwarming* when these users can be linked to a given question from the survey.

6.1 Information Retrieval

Twitter is generally a platform of free expression, and obviously the users publishing their thoughts on climate change are not, as a matter of course, intentionally aligning their texts to survey questions from the Six Americas. However, it is reasonable to assume that the topics discussed in tweets with the hashtag *#globalwarming* will often hit some point covered in the survey. Although it is possible that the subject of a tweet may be concerned with multiple questions, the short length of these texts should tend to make this multiple-hit occurrence less frequent. For our analysis, we are stipulating that a tweet may be linked to only one question at most. (It may be linked to none.) Of course, in cases where a tweet's subject matter is in reality related to more than one question, we would like to determine which one represents the "best fit" and link the tweet to that question.

To make this determination, we turn to another subdomain of NLP: information retrieval (IR). Two common examples of IR-based applications are (1) document retrieval by author, title, or keywords in the online catalogue for a university or public library and (2) website retrieval by keywords in Internet search engines. In both cases the goal is essentially to index a corpus of documents and then later query against that index to retrieve documents relevant to a search. In the first example, the application creates an indexed representation of a large corpus of books, periodicals and associated articles, as well as other informational media. In the second, the index must allow speedy look-ups of the vast number of sites on the World Wide Web.

IR techniques have also been used to search large corpora of collected microblogs. One example is a study by (Makki et al., 2015) which seeks to augment accuracy of the tweets returned for a given request by actively involving the user in the retrieval process. Another is given by (Chen et al., 2013) who propose a hybrid IR model that uses dynamic temporal profiles for queries to boost the standard technique of pseudo-relevance feedback (automatically repeating searches with an expanded query using terms from the documents retrieved by the initial query).

However, in our case we do not wish to create queries with a certain subject matter and then search for the relevant microblogs, but rather the reverse. We already have the tweet, and we want the subject matter. The documents we need to retrieve are the questions from the Six Americas survey. In lieu of a traditional search by keywords to find tweets, we use the text of the tweet itself as our search request to find the desired document, the most pertinent survey question.

Note that ours is arguably not a standard use case for IR. Our document corpus is very small, just 31 survey questions all touching on a related theme. Thus, we might not expect the weighted term “signatures” to be sufficiently different from one document to another, which potentially may negatively affect our results. Supervised machine learning methods to classify tweets as one of 31 question categories would be an alternate approach. The objective here is analogous to entity extraction in sentiment analysis, where hidden Markov models (Rabiner, 1989) and conditional random fields (Lafferty et al., 2001) have proven to be successful techniques (Liu, 2015). Of course, supervised learning techniques will not work for us directly as we do not have a pre-labelled dataset with known tweet-to-question relationships. We could attempt to incorporate few-shot or zero-shot learning techniques into a supervised learning methodology (Romera-Paredes & Torr, 2015). Alternatively, we could investigate unsupervised techniques for topic modelling, such as Latent Dirichlet Allocation (Blei et al., 2003) or Probabilistic Latent Semantic Analysis (Hofmann, 1999). These all represent interesting avenues for future projects aimed at improving the baseline for the *Say S$\vec{\imath}$La* architecture that we are establishing in the scope of the present doctoral program. For this research, we chose our current direction because term-frequency-based approaches, similar to those used in IR, are often employed in sentiment analysis when trying to determine the subject of a given text (Ku et al., 2006). When considering strategies based on frequency, terms are often linked with parts of speech (Hu & Liu, 2004; Scaffidi et al., 2007), which we model in the *Say S$\vec{\imath}$La* architecture's description logic level, even if this part-of-speech information is not included as part of the IR-based methodology explored in the present research.

Note also that as our goal is to retrieve at most a single question for a given tweet, we are not employing IR query expansion techniques such as the use of synonyms, hypernyms, hyponyms, or pseudo-relevance feedback to increase the number of documents returned in a search. In other words, we wish to maximize precision, and we are less concerned with recall when linking a tweet to a survey question. Query expansion techniques are generally used to enhance recall. Furthermore, they are most effective when used in conjunction with short queries (Voorhees, 1994). Despite the name and the application-imposed character limit, microblogs are typically longer than the short queries of a few keywords that these IR techniques are meant to augment. Finally, in this study we are seeking to understand something about the users on Twitter, and the insight revealed by an analysis considering *all* the tweets that a given user publishes about global warming[1] can potentially be compared favourably to extending the information retrieved by automated query expansion techniques.[2]

6.1.1 Leveraging Lucene

Lucene is an open source IR platform that is part of the Apache Software Foundation.[3] It was originally written by Doug Cutting, who named it after his wife[4] and her grandmother. As an IR tool, Lucene enables the search for documents, document passages, or meta-data relating to documents from what would typically be a large corpus (McCandless et al., 2010).

As mentioned above, the corpus for this work is relatively small, consisting of 31 questions taken from the third update to the extended series of studies, generally known as the Six Americas. These studies began in 2008 (Maibach et al., 2009) and have been performed twice yearly on a representational sample in the United States up to the present day (Leiserowitz et al., 2021b). The paper describing the third update (Leiserowitz et al., 2010) includes the full text for each question in the survey. We have saved these texts in a series of simple ASCII documents for indexing with Lucene. The authors present the questions (along with their survey results) in Tables 2 through 31 of their paper. Accordingly, in our study we refer to these questions as T2 through T31.[5]

---

[1]As indicated by inclusion of the hashtag *#globalwarming* within the time period covered by the dataset.

[2]The scope of the present research does not include such a comparison, but it is an interesting direction for future research.

[3]https://lucene.apache.org/

[4]Lucene is her middle name.

[5]Henceforth, we refrain from any explicit use of the word *table* when referring to the survey questions

Using Lucene to run queries to search a given corpus requires a number of configuration decisions. These define how both the corpus documents (the survey questions) and the query text (the tweets) are preprocessed with respect to similarity weighting, tokenization, stop words, and stemming.[6] Lucene is highly configurable with respect to these kinds of options; however, for a basic discussion of possible configurations we need to take a look at similarity scoring and analyzers.

Similarity Scoring

Our goal is to index the corpus of questions from the Six Americas and subsequently to query against it using the tweets in the dataset and retrieve the question relevant to each tweet (if any). The key is to score the similarity between a vector in the index, which represents the content of a document (question), and a vector representing the search terms in a query (tweet). When Lucene creates an index for a corpus, it computes a term vector for each document it contains. Afterwards, when we perform a query, Lucene creates a similar vector, usually from a set of keywords, but in our case from the author's words in her tweet. Lucene determines the similarity between the query vector and the term vectors for the documents in the corpus. The documents with term vectors most similar to the query vector are returned as "hits" for the request.

The elements in these vectors are weighted values, and the weighting algorithm is one of the configuration choices in Lucene. Historically, the most well-known algorithm for this purpose is term frequency × inverse document frequency (TF*IDF). As indicated by the name, the algorithm will weigh a document's relevance more heavily with respect

---

in an effort to avoid ambiguity with our tables in the present work. Note also that the study splits the question on "Risk Perceptions" into two tables: 8a for "Who is at Risk" and 8b for "When Harm Will Occur." We do likewise in our work, referring to these questions as T8a and T8b respectively.

[6] Lemmatization may be used in lieu of stemming. Generally this choice results in some improvement in precision at the cost of an increase in computational resource requirements. Our research focuses on stemming.

to a given term when that term is found more frequently in the document. However, when a term occurs frequently in many documents across the corpus, this serves to decrease the relevance of those documents with respect to that term. This similarity algorithm is accessed in Lucene using an instantiation of the class *ClassicSimilarity*.[7] It is "classic" because previously it was the default similarity for Lucene. The associated algorithm computes the cosine similarity between document and query vectors, weighted using a form of the standard TF*IDF metric for each term. Additionally, Lucene's implementation adds boost and normalizing factors to speed up the query and give higher scores to shorter document fields.

The current default similarity algorithm for Lucene and the one used in our research is Okapi BM25 ("Best Match 25"). Developed as part of the Okapi IR system in conjunction with TREC (Text REtrieval Conference),[8] it employs a probabilistic model to rank documents in order of probable relevance to a given query. The model is largely based on TF*IDF but demonstrates significantly improved performance based on trials at a number of TREC conferences. It imposes a saturation point which limits the contribution of a term's frequency to the relevance score of a document, and it normalizes document length across the corpus (Robertson et al., 1995; Robertson & Zaragoza, 2009).

Analyzers

Lucene uses an *Analyzer* class to handle a number of NLP preprocessing steps, which are generally necessary both for indexing documents in a corpus and for executing queries. Lucene offers several *Analyzer* child classes, which are ready for use in an application. In our research we primarily use the *EnglishAnalyzer* class. However, it may be useful to look briefly at a few of Lucene's major analyzer classes in order of increasing functionality

[7] It derives from the abstract class *TFIDFSimilarity*. For details, consult the javadocs for the Core Lucene API at https://lucene.apache.org/core/7_7_3/core/.

[8] https://trec.nist.gov/

to better understand the work that *EnglishAnalyzer* is doing:[9]

Analyzer: This analyzer represents the baseline of our IR functionality linking tweets to survey questions. The only preprocessing it does is to tokenize the input based on Unicode standards, delimiting words based on whitespace, punctuation, control codes, etc.

StandardAnalyzer: This is Lucene's general-purpose analyzer. After tokenizing as above, it converts words to lower case and removes stop words. The analyzer has a default stop list for English, but a developer may specify her own list when instantiating the analyzer.

Porter Stemmer: Our system derives this analyzer by extending Lucene's base *Analyzer* class. This analyzer tokenizes, converts to lowercase, removes stop words, and then applies the Porter stemming algorithm (Porter, 2006) to use the roots of input words as the final terms.[10]

EnglishAnalyzer: This Analyzer comes from Lucene's "analyzers-common" API (rather than the core API). This extension to Lucene provides effective, pre-configured analyzers for a large number of languages. Browsing the source code for this API, we can see that after tokenization, the analyzer removes *'s* endings from possessives before converting the input to lowercase and removing stop words. Then, prior to running the same Porter stemmer, the analyzer identifies specific keywords in an attempt to minimize over-aggressive stemming.

Note that all these analyzers use Lucene's default BM25 similarity class (*BM25Similarity*) unless specifically configured for TF*IDF (*ClassicSimilarity*) or any of several other similarity scoring algorithms available in Lucene.

---

[9]Our system supports these analyzers as well; however, all the results we report in this chapter use the *EnglishAnalyzer*.

[10]This is essentially the analyzer our system *would* use if only Lucene's core API were available. However, the Lucene "analyzers-common" API allows us to incorporate additional functionality.

Question Hits

For this portion of the research we continue using the 2019 *#globalwarming* dataset as described in Section 5.2. As a first step, we created an index of the ASCII documents containing the Six Americas survey questions using Lucene's *EnglishAnalyzer*. We then used each of the 97,666 tweets (prepared by the *EnglishAnalyzer*) as an individual query against this IR index. Table 6.1 lists the hit counts reported by the query run for each question in the survey in the original order as published in (Leiserowitz et al., 2010). Lucene returns the top N hits for a query for any configured value of N; however, for this initial study we are using $N = 1$. Of course, it is entirely possible that despite the short format, some tweets may be linked to multiple questions, but here we are seeking only the best hit (closest similarity) to match a tweet to a survey question.[11] Figure 6.1 presents the same results as the table in graphical form, but note that the questions (x-axis) are ordered by descending number of tweets (y-axis) which refer to (or hit) a given question.

6.1.2 Adding Affect

Now that we have linked the tweets in the 2019 *#globalwarming* dataset to specific questions from the Six Americas, we have a means by which to quantify the affect that the associated community on Twitter is expressing with respect to the individual questions. Note that while *Say S$\vec{\imath}$La* uses Lucene to link the tweets to the questions as described in the preceding section, for the analysis of affect the system once again leverages the Affective Tweets library (Mohammad & Bravo-Marquez, 2017), specifically for access to its sentiment and emotion lexica.

In Section 3.2.1, when we modelled emotion for the "big players" in the *#globalwarming*

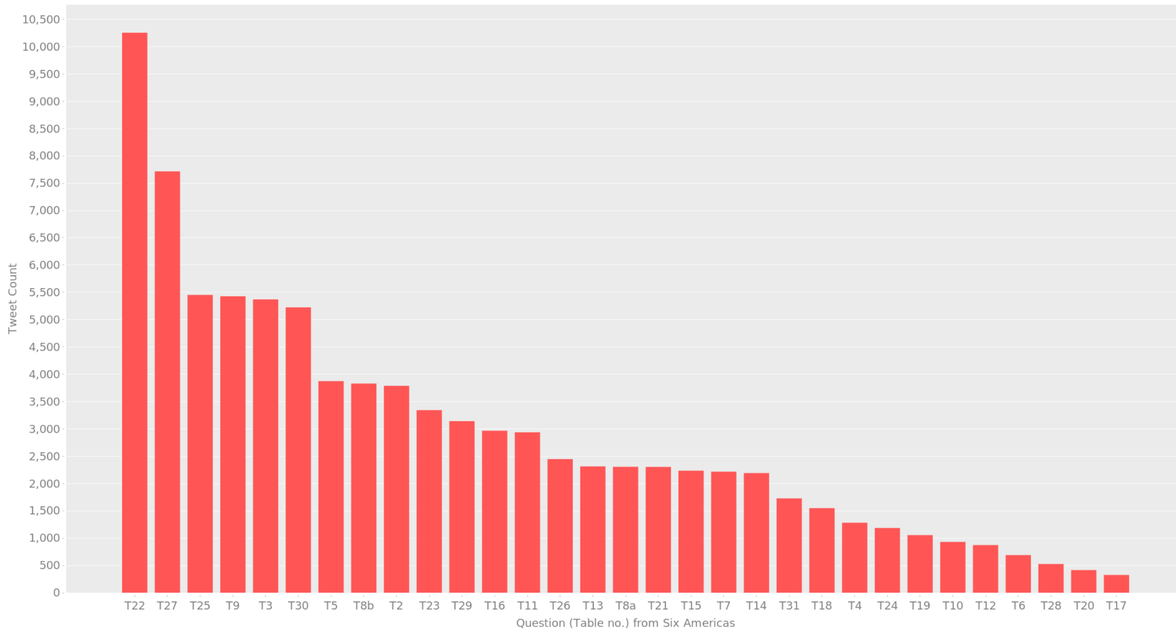
[11]See Section 6.4.2 for an explanation of why using $N >= 2$ might allow us to improve upon the results presented in this chapter.

Table 6.1 Number of tweets relating to individual survey questions.

Question	Tweets	Six Americas Title
T2	3789	Attitudinal Certainty and Need for Additional Information to Decide
T3	5372	Questions About Global Warming
T4	1282	Information-Seeking and Attention
T5	3875	Beliefs
T6	689	Emotions
T7	2217	Issue Involvement
T8a	2306	Risk Perceptions: Who Is at Risk
T8b	3831	Risk Perceptions: When Harm Will Occur
T9	5429	Perceptions of Weather and Climate
T10	930	Impact of the Economic Downturn
T11	2938	Environmental Protection and Economic Growth
T12	872	Conservation Actions
T13	2314	Conservation Intentions
T14	2191	Perceived Impact of Own Actions
T15	2235	Consumer Activism
T16	2967	Political Activism
T17	325	Perceived Importance of Conservation Behaviors
T18	1549	Perceptions of Social Norms
T19	1055	Interpersonal Communication
T20	414	Family Communication
T21	2304	Opinion Leadership
T22	10256	Outcome Expectations
T23	3344	Support for a National Response: Conditions for & Magnitude of Action Desired
T24	1186	Issue Priority
T25	5454	Support for National Response: Specific Climate and Energy Policies
T26	2447	Attention and Response to Climategate
T27	7717	Attention and Response to IPCC Errors
T28	526	Trust in Information Sources
T29	3142	Media Preferences
T30	5226	Attention to Specific Programs and Media Sources
T31	1727	Party Identification, Political Ideology, and Voter Registration
	7757	*Tweets not referring to a question*
	97666	**Total Tweets**

community on Twitter, we used the NRC Affect Intensity Lexicon (NRC-AIL) (Mohammad & Bravo-Marquez, 2017). Here, for our analysis of the questions from the Six Americas we use a different lexicon, also from the National Research Council of Canada, called the NRC Word-Emotion Association Lexicon (NRC-10) (Mohammad & Turney, 2013). This 14,182-word, human-annotated lexicon was created as a crowd-sourcing project using Amazon's Mechanical Turk. The words are linked to one or more of the full set of eight basic emotions in Plutchik's model. These emotions function as four pairs of opposites: anger–fear, sadness–joy, surprise–anticipation, and disgust–trust (Plutchik, 2001). The lexicon also indicates if a given word is associated with a positive or a negative sentiment, giving a total of 10 possible affect associations, which we use as the base of our analysis of the survey questions. For this reason the lexicon is called

Figure 6.1 Tweet counts for questions in the Six Americas survey.



the NRC-10.[12] Note that although this lexicon provides affective attributes to fully represent Plutchik's model along with the sentiment polarity of a word (rather than just four emotions as with the NRC-AIL), we no longer have an indicator of the strength of a given attribute. For example, while the NRC-AIL reports the word "change" to have a level of fear of 0.198 on a scale of 0 (no emotion) to 1 (highest expressed emotion), the NRC-10 simply gives a binary value (1 of {0,1}) to indicate that fear is associated with this word.

As we will see later in this chapter, our initial analysis using the NRC-10 revealed an interesting tendency relating to sentiment polarity (positive/negative) from Twitter users in the green and denier categories with respect to their tweets linked to the Six Americas survey questions via Lucene. To investigate this tendency further, we conduct a parallel analysis on the same tweets using a second sentiment lexicon commonly known as Bing Liu's Opinion Lexicon (Hu & Liu, 2004). This lexicon was also created manually, and it

[12]We will commonly refer to the lexicon using this short name in an effort to avoid confusion and distinguish it well from the NRC-AIL that was part of the experiments in Chapter 3.

includes slang, misspellings, and morphological variations (Bravo-Marquez et al., 2014). This feature potentially makes the lexicon an invaluable tool for sentiment analysis on social media sites such as Twitter where this type of language use is exceedingly common (Farzindar & Inkpen, 2015). The current version of this lexicon comprises 2006 positive words and 4783 negative words.

Both of these lexica are available as part of the Affective Tweets plugin (Bravo-Marquez et al., 2019), for the Weka machine learning platform (Frank et al., 2016), which we incorporate programmatically into our system.

Affect Signatures

Whereas Figure 6.1 showed the number of tweets which the Lucene-based analysis linked to each survey question, Figure 6.2 shows the affect levels by question for these same tweets. The level of affect (y-axis) represents the number of words in the tweets linked to an emotion or a sentiment polarity per the NRC-10 lexicon. It is interesting to note here that although the order of the questions (x-axis) still runs from left to right according to the descending number of linked tweets, just as shown in Figure 6.1, the peaks of the bars are jagged now. Even though a higher tweet hit count will generally mean more words and therefore potentially more words expressing affect, several questions with a lower hit count are showing more sentiment and emotion than those with a higher count. A reasonable hypothesis would be that some questions invoke more emotion than others.

Simply charting the raw count of affect words for the survey questions is of limited value, especially when looking at the questions towards the right with lower tweet hit counts and, therefore, a small bar on the chart whose tiny subdivisions are difficult to distinguish. Figure 6.3 looks at this same data in terms of the percentage in which each affect attribute is represented in the set of tweets corresponding to each question. Here, the tweet hit count for a given question is represented only by its position on the x-axis, and the height of each bar is always 1, representing 100% of the affect expressed for a

Figure 6.2 Affect in tweets by Six Americas survey question.

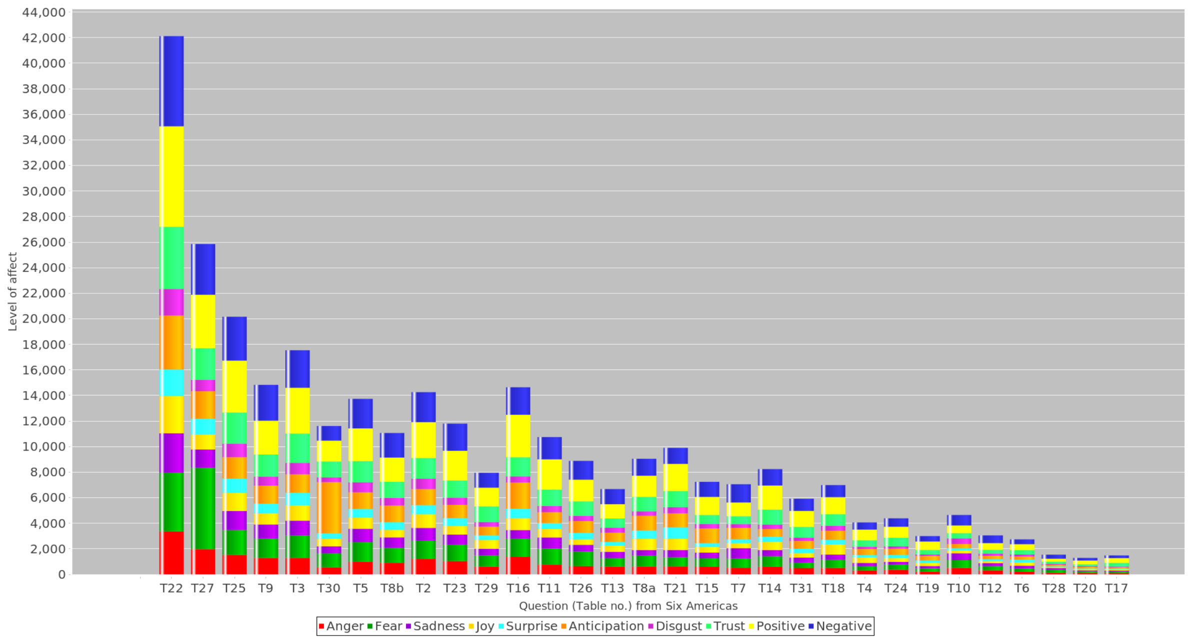


Figure 6.3 Affect signature as percentages for questions from the Six Americas.

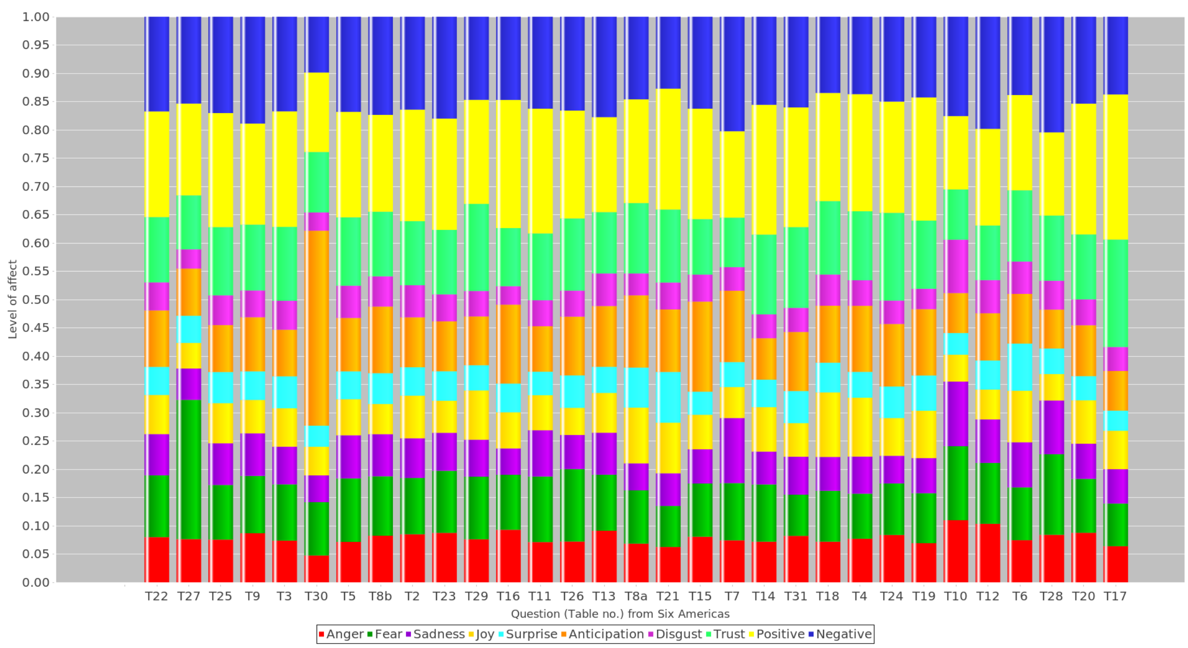


Table 6.2 Affect word counts and percentages for questions from the Six Americas.

Quest.	Anger		Fear		Sadness		Joy		Surprise		Anticipation		Disgust		Trust		Positive		Negative	
T2	1213	8.5%	1418	9.9%	1002	7.0%	1074	7.5%	716	5.0%	1259	8.8%	808	5.7%	1617	11.3%	2811	19.7%	2341	16.4%
T3	1294	7.4%	1745	9.9%	1170	6.7%	1190	6.8%	993	5.7%	1442	8.2%	899	5.1%	2292	13.1%	3584	20.4%	2933	16.7%
T4	314	7.7%	324	8.0%	267	6.6%	423	10.4%	186	4.6%	475	11.7%	184	4.5%	496	12.2%	842	20.7%	557	13.7%
T5	982	7.1%	1544	11.2%	1047	7.6%	876	6.4%	676	4.9%	1297	9.4%	784	5.7%	1660	12.1%	2560	18.6%	2311	16.8%
T6	205	7.5%	256	9.3%	219	8.0%	250	9.1%	229	8.3%	241	8.8%	157	5.7%	346	12.6%	462	16.8%	380	13.8%
T7	525	7.4%	713	10.1%	813	11.5%	386	5.5%	312	4.4%	890	12.6%	294	4.2%	618	8.8%	1078	15.3%	1428	20.2%
T8a	619	6.8%	853	9.4%	433	4.8%	890	9.8%	641	7.1%	1155	12.8%	349	3.9%	1127	12.5%	1658	18.3%	1322	14.6%
T8b	913	8.3%	1162	10.5%	824	7.4%	590	5.3%	602	5.4%	1303	11.8%	592	5.4%	1263	11.4%	1894	17.1%	1920	17.4%
T9	1288	8.7%	1504	10.1%	1118	7.5%	875	5.9%	750	5.1%	1418	9.6%	703	4.7%	1726	11.6%	2652	17.9%	2800	18.9%
T10	511	11.0%	608	13.1%	531	11.4%	221	4.8%	178	3.8%	328	7.1%	439	9.4%	413	8.9%	603	13.0%	816	17.6%
T11	766	7.1%	1246	11.6%	881	8.2%	664	6.2%	448	4.2%	863	8.0%	496	4.6%	1270	11.8%	2370	22.0%	1748	16.3%
T12	316	10.3%	329	10.8%	235	7.7%	161	5.3%	157	5.1%	255	8.3%	179	5.9%	295	9.7%	522	17.1%	605	19.8%
T13	610	9.1%	662	9.9%	496	7.4%	470	7.0%	309	4.6%	717	10.7%	385	5.8%	725	10.9%	1122	16.8%	1186	17.7%
T14	593	7.2%	834	10.1%	479	5.8%	648	7.9%	399	4.8%	605	7.3%	348	4.2%	1163	14.1%	1891	22.9%	1283	15.6%
T15	585	8.1%	681	9.4%	438	6.0%	443	6.1%	296	4.1%	1153	15.9%	346	4.8%	711	9.8%	1414	19.5%	1178	16.3%
T16	1363	9.3%	1421	9.7%	685	4.7%	934	6.4%	744	5.1%	2042	13.9%	480	3.3%	1503	10.3%	3318	22.7%	2153	14.7%
T17	95	6.4%	112	7.5%	90	6.1%	101	6.8%	53	3.6%	104	7.0%	63	4.2%	282	19.0%	381	25.7%	204	13.7%
T18	502	7.2%	629	9.0%	419	6.0%	797	11.4%	367	5.3%	703	10.1%	386	5.5%	907	13.0%	1336	19.1%	941	13.5%
T19	208	6.9%	265	8.8%	186	6.2%	251	8.4%	187	6.2%	352	11.7%	108	3.6%	362	12.1%	653	21.8%	428	14.3%
T20	113	8.8%	123	9.5%	80	6.2%	99	7.7%	55	4.3%	116	9.0%	59	4.6%	148	11.5%	298	23.1%	198	15.4%
T21	620	6.3%	718	7.2%	573	5.8%	886	8.9%	889	9.0%	1093	11.0%	472	4.8%	1279	12.9%	2117	21.4%	1259	12.7%
T22	3369	8.0%	4600	10.9%	3070	7.3%	2919	6.9%	2087	5.0%	4210	10.0%	2087	5.0%	4859	11.5%	7866	18.7%	7052	16.7%
T23	1033	8.8%	1295	11.0%	794	6.7%	667	5.7%	616	5.2%	1042	8.8%	564	4.8%	1343	11.4%	2323	19.7%	2125	18.0%
T24	368	8.4%	400	9.1%	215	4.9%	293	6.7%	245	5.6%	485	11.0%	182	4.1%	681	15.5%	863	19.7%	659	15.0%
T25	1524	7.6%	1948	9.7%	1488	7.4%	1430	7.1%	1111	5.5%	1673	8.3%	1061	5.3%	2433	12.1%	4065	20.2%	3435	17.0%
T26	641	7.2%	1139	12.8%	536	6.0%	426	4.8%	510	5.7%	922	10.4%	410	4.6%	1135	12.8%	1694	19.1%	1474	16.6%
T27	1977	7.6%	6368	24.6%	1430	5.5%	1172	4.5%	1238	4.8%	2157	8.3%	881	3.4%	2468	9.5%	4193	16.2%	3970	15.4%
T28	130	8.4%	220	14.2%	147	9.5%	72	4.7%	70	4.5%	106	6.9%	79	5.1%	178	11.5%	227	14.7%	316	20.5%
T29	606	7.6%	881	11.1%	520	6.5%	694	8.7%	354	4.4%	686	8.6%	359	4.5%	1225	15.4%	1464	18.4%	1169	14.7%
T30	553	4.8%	1094	9.4%	548	4.7%	589	5.1%	433	3.7%	4004	34.5%	375	3.2%	1240	10.7%	1633	14.1%	1144	9.9%
T31	487	8.2%	431	7.3%	398	6.7%	352	5.9%	336	5.7%	618	10.4%	253	4.3%	846	14.3%	1253	21.2%	950	16.0%
MEAN		7.9%		10.5%		6.9%		6.9%		5.2%		10.7%		4.8%		12.1%		19.1%		16.0%
STD-DEV		1.19		3.05		1.64		1.76		1.19		4.90		1.13		2.09		2.86		2.28

question. The coloured subdivisions of each bar give the percentage which each affect characteristic contributes towards the sum of all the expressed affect in the tweets for that question. In essence we are stretching out the bars from the count-based graph in order to get percentage-based affect signatures for these questions. Table 6.2 shows the data used to generate both Figures 6.2 and 6.3. For each column representing an affective element, the numbers on the left of the column are the counts of words expressing the emotion or sentiment according to the lexicon (graphed in Figure 6.2). The numbers on the right of the column give the percentage score for the affect characteristic with respect to all the expressed affect (graphed in Figure 6.3).

We will take a look at percentages more closely for a number of survey questions later in this chapter when we analyze the activity of users on Twitter presupposed to be in either the green or the denier category. For the moment, however, with this rough look at the affect expressed in the unlabelled dataset, we start to see a number of interesting points emerge:

- There is generally more positive sentiment (mean: 19.1%) than negative (mean: 16.0%) for most questions. Obvious exceptions are:
  - T7 ("Issue Involvement") at 20.2% negative vs. 15.3% positive
  - T10 ("Impact of the Economic Downturn") at 17.6% negative vs. 13.0% positive
  - T12 ("Conservation Actions") at 19.8% negative vs. 17.1% positive
  - T28 ("Trust in Information Sources") at 20.5% negative vs. 14.7% positive

  Questions T8a, T8b, and T13 also express more negative sentiment than positive; however, the difference does not exceed one percentage point. This leaves 24 questions expressing greater positive sentiment than negative. One question (T27) is borderline at only 0.8% more expressed positive affect, while the rest range from 1.2% up to 12% higher.
- Tweets linked to T27 ("Attention and Response to IPCC Errors") express more

fear (24.6%) compared with other questions (mean: 10.5%).

- Tweets referring to T30 ("Attention to Specific Programs and Media Sources") express much more anticipation (34.5% vs. a mean of 10.7%). and decidedly less anger (4.8% vs. a mean of 7.9%).
- Tweets referring to questions T10 ("Impact of the Economic Downturn") and T12 ("Conservation Actions") express more anger (11.0% and 10.3% respectively vs. a mean of 7.9%).

Polarity Signatures

Figure 6.3 shows positive (yellow) and negative (blue) sentiment polarity to be important elements among the affect attributes covered by the NRC-10 lexicon. For this reason, we repeat the analysis described above using Liu's Opinion Lexicon, which being purely a sentiment lexicon, includes the two polarity characteristics only. Figure 6.4 presents the level of sentiment in the tweets for each question in order of the number of linked tweets (refer to Figure 6.1). This chart parallels the one in Figure 6.2 except that the only two affect characteristics are positive and negative sentiment. Once again we see the peaks of the bars are jagged, indicating that a greater number of tweets for a given question does not necessarily mean that the level of sentiment expressed in those tweets will be greater as well. Also, we should note that the size of Liu's sentiment lexicon is significantly smaller than that of the NRC-10, and so we can reasonably expect the count of sentiment words identified in the tweets to be smaller as well. We see this is indeed the case, looking at the reported level of affect (y-axis) in Figure 6.4.

Once again, we propose that it is advantageous to represent the levels of sentiment as percentages as shown in Figure 6.5. The stretched bars, show the percentages of positive and negative sentiment expressed for each of the survey questions in order of descending tweet hit count. Table 6.3 gives the corresponding counts and percentages in tabular form for Liu's Opinion Lexicon.

Figure 6.4 Sentiment in tweets by Six Americas survey question.

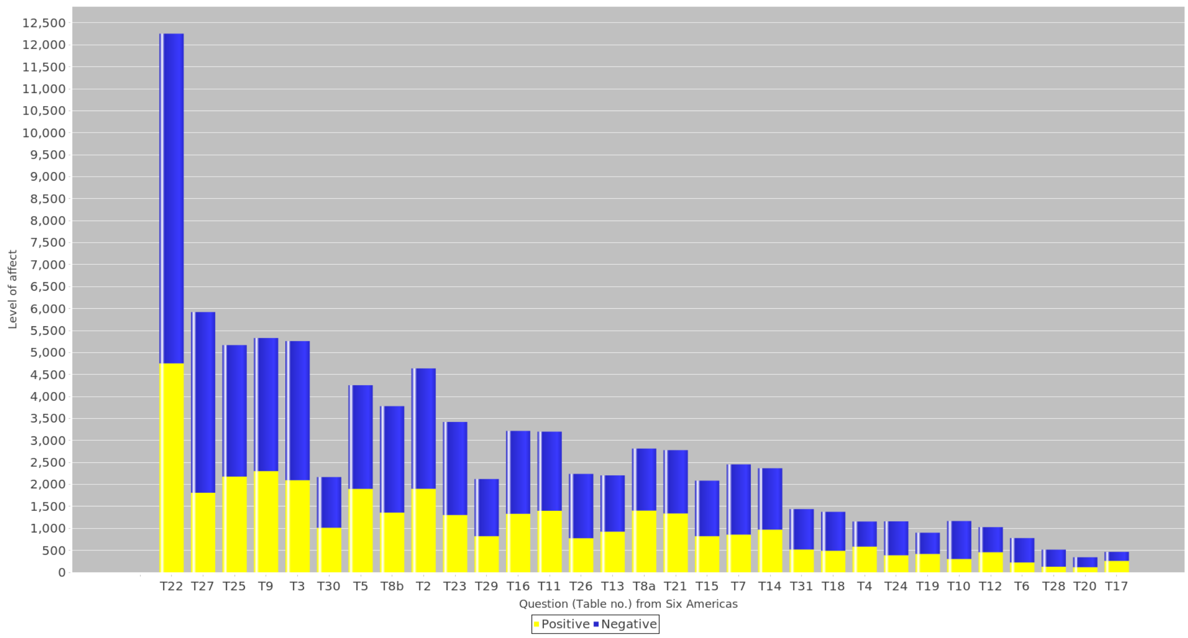


Figure 6.5 Sentiment signature as percentages for Six Americas questions.

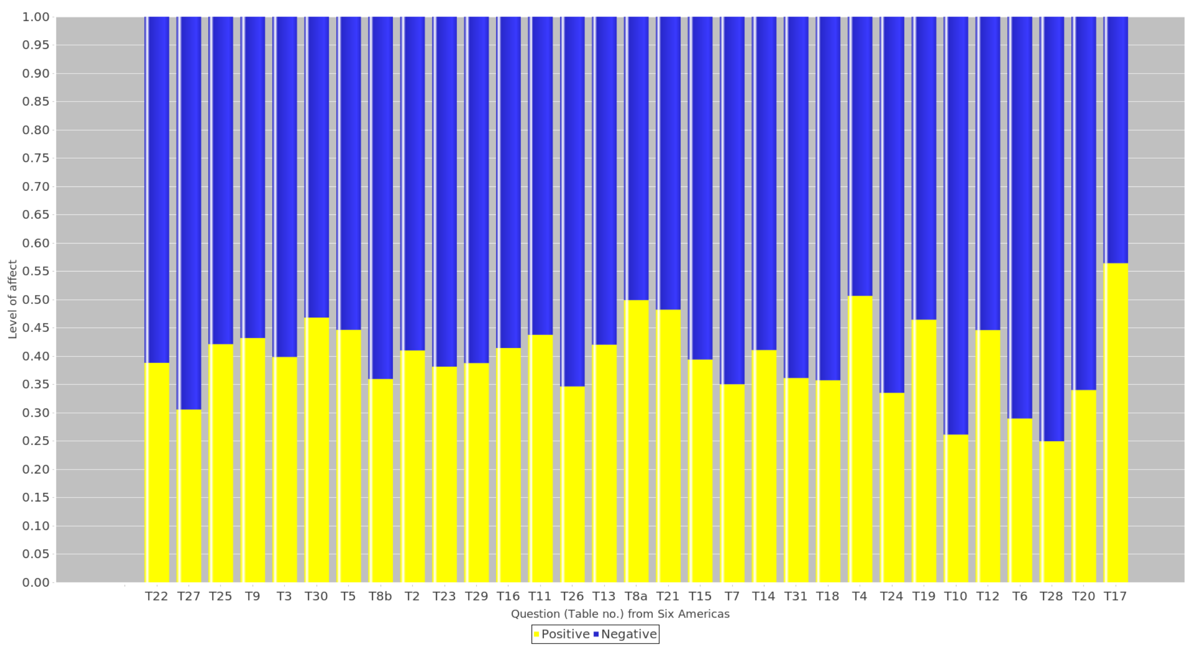


Table 6.3 Sentiment word counts and percentages for questions from the Six Americas.

Question	**Positive**		**Negative**	
T2	1903	41.0%	2738	59.0%
T3	2097	39.9%	3164	60.1%
T4	586	50.6%	571	49.4%
T5	1900	44.6%	2356	55.4%
T6	226	29.0%	554	71.0%
T7	860	35.0%	1595	65.0%
T8a	1406	49.9%	1411	50.1%
T8b	1360	36.0%	2421	64.0%
T9	2304	43.2%	3028	56.8%
T10	305	26.1%	862	73.9%
T11	1401	43.8%	1800	56.2%
T12	459	44.6%	570	55.4%
T13	927	42.0%	1279	58.0%
T14	973	41.1%	1396	58.9%
T15	822	39.4%	1264	60.6%
T16	1333	41.4%	1884	58.6%
T17	263	56.4%	203	43.6%
T18	492	35.7%	885	64.3%
T19	419	46.5%	483	53.5%
T20	117	34.0%	227	66.0%
T21	1341	48.2%	1440	51.8%
T22	4756	38.8%	7498	61.2%
T23	1305	38.1%	2117	61.9%
T24	389	33.5%	772	66.5%
T25	2178	42.1%	2992	57.9%
T26	776	34.6%	1464	65.4%
T27	1810	30.6%	4111	69.4%
T28	129	25.0%	388	75.0%
T29	823	38.8%	1300	61.2%
T30	1015	46.8%	1153	53.2%
T31	520	36.1%	919	63.9%
MEAN		39.8%		60.2%
STD-DEV		7.15		7.15

We see immediately that we do not have the same general trend of more positive sentiment than negative with Liu's lexicon. A likely explanation is that this lexicon has over twice as many negative words as positive words, while the NRC lexicon has a somewhat smaller ratio of approximately 3 negative to 2 positive.[13] As we note differences in composition of the lexica, we should also take into consideration reported findings that people may tend to express positive sentiment more often than negative and that this tendency extends to what they publish on social media (Guerra et al., 2014).

Figures 6.6 and 6.7 present a comparison of positive and negative sentiment respectively between the NRC-10 lexicon and Liu's Opinion Lexicon for tweets linked to the Six

[13]The Liu lexicon has 4783 negative and 2006 positive words. For the NRC lexicon, the ratio is 3316 negative to 2313 positive words.

Figure 6.6 NRC-10 vs. Liu for positive sentiment in survey questions.

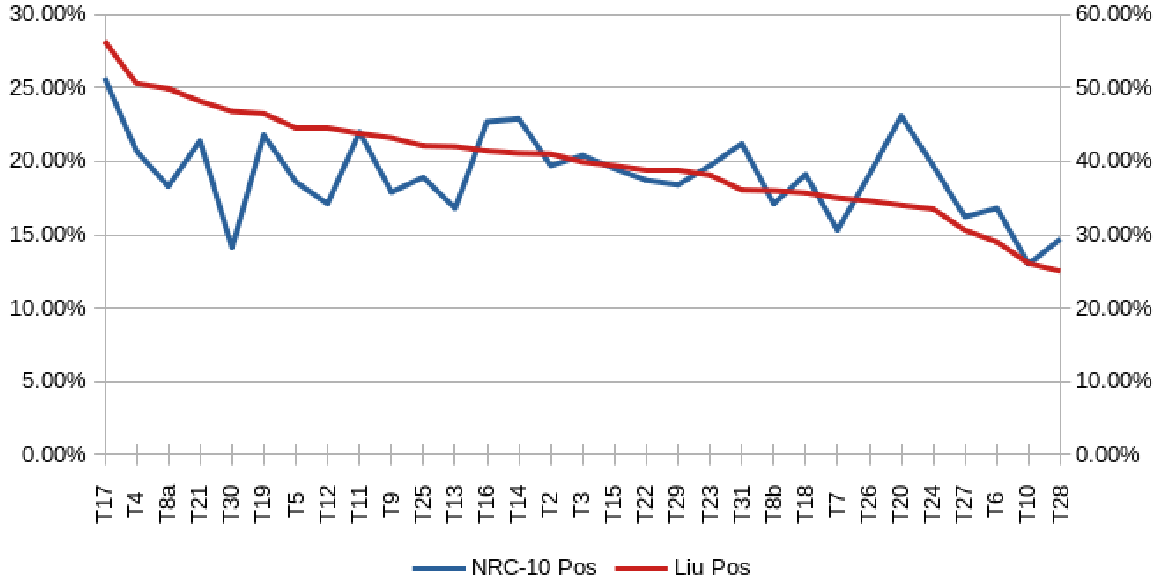
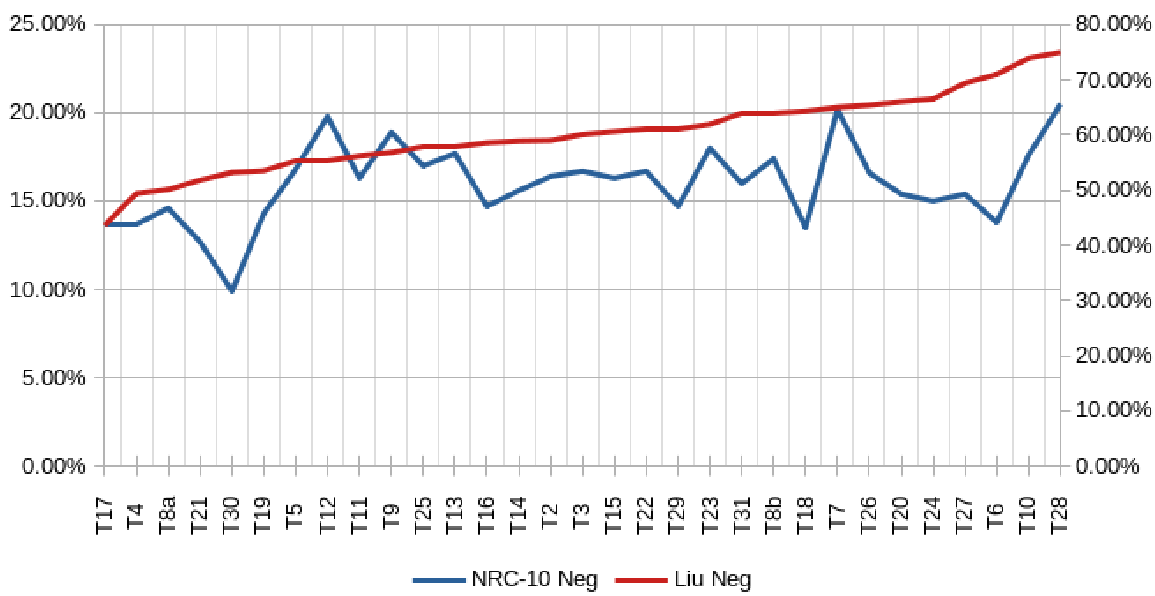


Figure 6.7 NRC-10 vs. Liu for negative sentiment in survey questions.



Americas survey questions. The questions (x-axis) are listed in order of decreasing positive (Figure 6.6) and increasing negative (Figure 6.7) sentiment according to the Liu lexicon.[14] The left y-axis represents the percentage of positive (Figure 6.6) or negative (Figure 6.7) sentiment for NRC-10 (out of 10 affect characteristics). The right y-axis represents the same for the Liu lexicon (out of two characteristics). The values plotted in these charts are from the "Positive" and "Negative" columns in Table 6.2 for NRC-10 and Table 6.3 for Liu's lexicon. The difference in the number of affect characteristics between the two lexica is the main reason for the difference between the ranges (percentages) shown on the left and right y-axes. (The sentiment polarities are two out of ten "slices" for NRC-10, while they represent the whole pie for Liu.)

Figures 6.6 and 6.7 take another step towards explaining why with Liu's Opinion Lexicon we do not see the same trend of increased positive sentiment in the tweets with respect to the negative sentiment that we see with NRC-10. We should not expect that sentiment intensity will be the same, given the number of attributes in each lexica as well as the number of positive and negative terms actually defined. However, we might reasonably expect that an analysis with one lexicon will more-or-less follow the same ordering of questions as the other with respect to increasing positive or negative sentiment.[15] Ideally, we would see *both* curves generally fall for positive and rise for negative sentiment for the question ordering in these charts. Looking at tweet counts for the various questions (see Table 6.1 and Figure 6.1), we note that sometimes the positive sentiment curve juts back up at questions where we have fewer associated tweets (T20, T28, T19, and perhaps arguably T31 as well). However, there are other questions where a similar

---

[14] Although we are using the NRC-10 as the primary tool for our affective analysis in this chapter, using the Liu lexicon as the base for this comparison allows us to keep the same question order for both positive and negative sentiment. Because the NRC-10 includes 10 affect characteristics, a greater percentage of positive sentiment will not necessarily entail a lesser percentage of negative sentiment (and vice versa).

[15] The reader should take caution when considering the charts in Figures 6.6 and 6.7. At a casual glance it may appear that Liu's lexicon results in a smooth and proper ordering, while the NRC-10 flits all around. This is only because we chose Liu's lexicon as the base for the graph. If we had chosen NRC-10, then that lexicon would appear to be the stable one (at a casual glance).

jutting occurs, but the questions have better coverage in terms of linked tweets. We see a similar situation for negative sentiment, where the rising curve juts back down for a number of questions with lower tweet counts (T6 and T18), but again the same happens for other questions with more reasonable coverage.

As both lexica were created manually, we should account for the subjective nature of the task given to the human annotators. Decisions as to the affect characteristics of a given word may vary significantly in one context or another and from one person to the next. For this reason (Bravo-Marquez et al., 2014) recommend verifying how well attributes representing sentiment or emotion in a given lexicon correspond to known target data examples. The presupposed use case for this advice is a data mining application, which is closer to the final phase of our research as described in Chapter 7. Here, however, we have an integral limitation in that we are using AI tools to study data from social media with little ground certainty as to what we should find. We discuss this further in Section 6.4, but we can say briefly here that as we use the tools to explore the data, we are also evaluating the tools themselves for this particular research problem. What we may conclude at this point is that we must be cautious when identifying trends for expressed sentiment, noting where results differ between the two lexica. We also recognize that as we do not have results with a second lexicon for the emotions (as opposed to sentiment polarity) covered by NRC-10, we should validate the results presented in this study with those produced using another lexicon as part of our continued research.

Comparing the specific points of interest regarding sentiment that we noted earlier when using the NRC-10 lexicon, we observe:[16]

- With most of the questions where we detected more negative sentiment than positive using the NRC lexicon, we also see the highest levels of negative sentiment

[16] When making comparisons of results using these two lexica, it is important to note that the actual percentage values will be much lower for the NRC-10 as the polarity represents only two out of ten affect characteristics. In the previous analysis, values for all ten attributes sum to 100%, rather than just two as with Liu's sentiment lexicon.

with Liu's lexicon.

 - T7 ("Issue Involvement") at 65.0% negative
 - T10 ("Impact of the Economic Downturn") at 73.9% negative
 - T28 ("Trust in Information Sources") at 75.0% negative

- Question T12 ("Conservation Actions") is the exception. Although it shows more negative sentiment than positive using NRC, with Liu's lexicon it scores 55.4% negative, which is below the mean of 60.2%.
- Question T6 ("Emotions") scores the third highest level of negative sentiment (71.0%) but was not reported as expressing exceptionally high levels of negative affect when we used the NRC lexicon (13.8% against a mean of 16.0%).[17]
- Question T17 ("Perceived Importance of Conservation Behaviors") at 56.4% positive is the only question that shows a greater expression of positive sentiment than negative at a level of more than 1%. This same question also had the highest score for positive affect per the NRC lexicon.

In general, when comparing the two lexica with respect to a preliminary analysis of the tweets linked to the individual survey questions, we note a number of important points. Certain sentiment polarity trends we discovered using the NRC-10 can indeed also be demonstrated using Liu's Opinion Lexicon. However, this may not be apparent at first glance since the raw levels of expressed affect can generally be quite different due to factors such as the number of terms in the lexica and the proportions of the various affect characteristics represented by those terms. Also, likely not every trend we find using the NRC lexicon will be supported by a parallel observation using Liu's lexicon. Of course, this will always be the case for trends related to emotion (as opposed to sentiment polarity). We should also remember that unlike the NRC-AIL (see Section 3.2.1), we are limited to a coarse "on or off" distinction with respect to word affect characteristics

[17]The title of the question leads us to consider whether its textual content invokes the emotional attributes of the affect lexicon at the expense of the attributes representing polarity. This need not be the case as the NRC lexicon allows a given word to represent emotion and sentiment polarity at the same time. An analysis of the dynamics at play here could prove interesting for continued research.

for both of the lexica used in this chapter. This may make any subtle trends in the data all the more difficult to discover (more on this particular limitation in Section 6.4).

On a more positive note, we can expect that the use of a second lexicon may help to reveal additional trends in the tweet data that, although present, were not immediately apparent when looking at results generated with the first lexicon. We shall endeavour to keep these points in mind as we continue our analysis.

6.2 Green and Denier Categories

In the previous section we used Lucene to find the question hits for all the tweets in the unlabelled dataset, and we examined levels of affect for the tweets linked to each question in the Six Americas study. Now we would like to consider a subset of tweets whose stance has already been determined a priori by analyzing which green or denier leader accounts a user is following on Twitter as described in Section 5.1.1.

6.2.1 Minimum Online Activity

In addition to limiting the data considered to that of labelled users, we have chosen to focus on the subset of these users having a minimum tweet count of 17 for the remaining experiments in this chapter. This choice may strike the reader as rather arbitrary. It is not. Throughout the course of this research, we have consistently focused on an analysis of the bigger players, the high-activity online users. We maintain the working hypothesis that they are to an extent representative of their larger community on social media. Our work up to this point has provided some evidence to support this hypothesis. In Section 3.4 we found correlation in levels of fear and anger between high-activity users and the larger *#globalwarming* community on Twitter. Of course, the experiments in the present chapter do not represent exactly the same scenario. Here we are in essence just focusing on generally bigger players from our 2019 dataset, not only the "Top N" most active users. Additionally, the 17-tweet minimum constraint comes in light of the findings described in Section 5.3. The F1 scores for the denier-oriented survey concept rules (SCR) peak at or around a value of 17, and for all the SCR pairs this value falls within the range that produces the best results. Although our use of the constraint is the same in that phase of the research, the analytical goal in this chapter is quite different. Now we are endeavouring to use IR techniques to analyze sentiment and emotion in texts relating to different questions from the Six Americas, rather than using an ontological model based on that study to identify the users' stances on climate change.

Ideally, we would continue running experiments for a range of minimum activity levels. Such a plan could indeed potentially provide increased confidence in our findings, assuming the results are similar across a range of minimum-tweet thresholds. It could also serve as additional evidence that high-activity users can be representative to an extent of the general online community. Finally, it could further demonstrate that there is a more-or-less optimal activity level for this type of experiment and that moving too far away from this value will decrease the performance of the experimental model. Going much lower may increase the level of noise in the form of the social media "chatter" that we discussed in Section 1.3. Going higher may at best lead to a situation of diminishing returns or at worst cause our model to drift further from the "true" state of affairs on Twitter as we use and potentially overfit our models to an ever smaller subset of the original data. Unfortunately, we are limited here by the scope of the present doctoral research program. Therefore, given our findings from previous phases of this research, we hypothesize that a minimum activity level of 17 tweets will likely provide us with better results to analyze with regard to our goal for this phase. Confirming this choice will be a task for a future research effort specifically focused on better grounding for these results as described in Section 6.4.2.

Although we are now working with a significantly smaller number of microblogs (12,241 tweets), the labelled dataset allows us to identify topic trends with respect to the survey questions that distinguish users in the green category from those in the denier category as previously determined using the follow-the-leader method. Table 6.4 presents the tweet counts by question for users labelled as green or denier along with the total tweets by question for both groups together. Again, all the users considered have at least 17 tweets in the dataset. Given that users on social media will often form virtual "echo chambers," interacting mainly with others who already share their opinions on a matter such as climate change (Williams et al., 2015), we may look at users in these groups as high-activity members of two distinct communities on Twitter. Furthermore, as research indicates that people tend to follow leader figures with respect to their perceived reality and political stance on climate change (Bohr, 2014; Krosnick, 2000), it is possible that these users have some level of influence on the larger community or that to an extent they may serve to represent their respective communities.

Note that even as we found that users in the dataset are about two thirds green and one third denier (see Section 5.2), the categories are even more skewed when we consider the individual tweets. Approximately 82% of the tweets are published by green users, and only 18% come from users labelled as denier. We must keep this imbalance in mind when analyzing the data.

Figure 6.8, graphs the tweet counts for users publishing 17 or more tweets in the dataset. The data for the chart comes from the "Total" column of Table 6.4, and the bars are ordered by descending hit counts for each survey question. The ranked order of questions is markedly similar to what we found when analyzing the full dataset (Figure 6.1). T22 ("Outcome Expectations") is still the question to which most tweets are linked by far. T22 asks whether and to what extent humans will be able to mitigate the effects of global warming (or if this even necessary) as well as whether a "single individual" can make a difference (Leiserowitz et al., 2010). We look into this question in depth in Section 6.3.1. As for the other questions, a number have shifted up or down a few places in the ranking (compare to Figure 6.1), especially over a given range where hit counts are similar to begin with. The following are questions that have increased in popularity by at least four in the ranking:

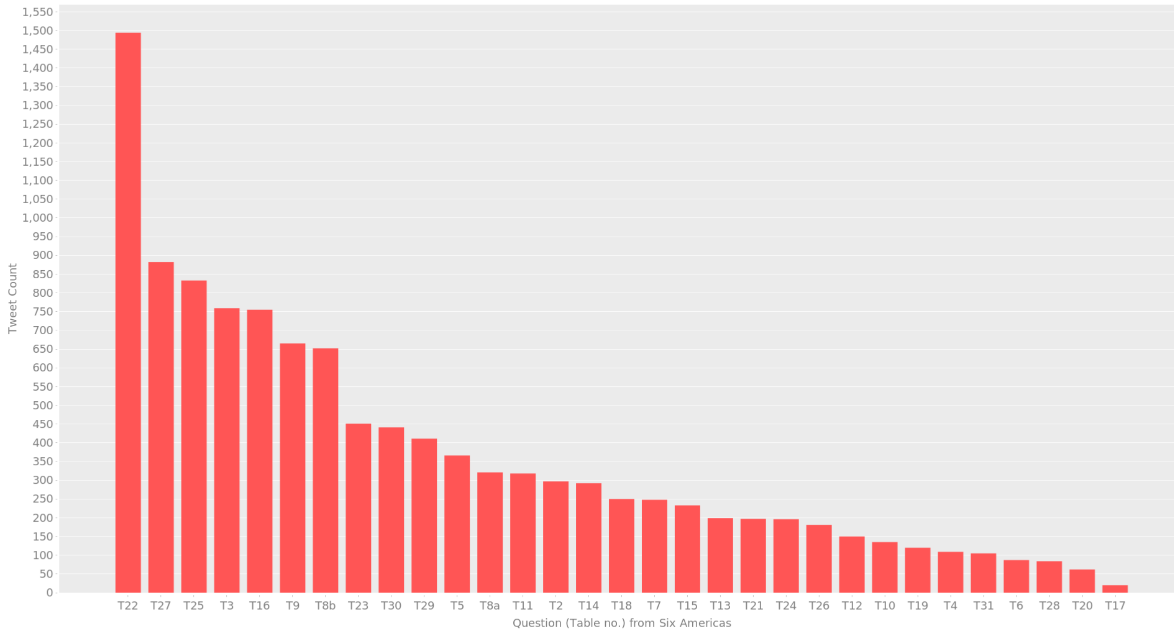
Table 6.4 Survey question hits for labelled users at a minimum activity level of 17.

Question	Green	Denier	Total	Six Americas Title
T2	244	53	297	Attitudinal Certainty and Need for Additional Information to Decide
T3	638	121	759	Questions About Global Warming
T4	84	25	109	Information-Seeking and Attention
T5	302	64	366	Beliefs
T6	74	13	87	Emotions
T7	220	28	248	Issue Involvement
T8a	286	35	321	Risk Perceptions: Who Is at Risk
T8b	586	66	652	Risk Perceptions: When Harm Will Occur
T9	519	146	665	Perceptions of Weather and Climate
T10	115	20	135	Impact of the Economic Downturn
T11	293	25	318	Environmental Protection and Economic Growth
T12	137	13	150	Conservation Actions
T13	147	52	199	Conservation Intentions
T14	251	41	292	Perceived Impact of Own Actions
T15	184	49	233	Consumer Activism
T16	692	63	755	Political Activism
T17	14	6	20	Perceived Importance of Conservation Behaviors
T18	212	38	250	Perceptions of Social Norms
T19	102	18	120	Interpersonal Communication
T20	48	14	62	Family Communication
T21	160	37	197	Opinion Leadership
T22	1358	136	1494	Outcome Expectations
T23	353	98	451	Support for a National Response: Action Desired (Conditions/Magnitude)
T24	83	113	196	Issue Priority
T25	709	124	833	Support for National Response: Specific Climate and Energy Policies
T26	133	48	181	Attention and Response to Climategate
T27	774	108	882	Attention and Response to IPCC Errors
T28	77	7	84	Trust in Information Sources
T29	346	65	411	Media Preferences
T30	391	50	441	Attention to Specific Programs and Media Sources
T31	75	30	105	Party Identification, Political Ideology, and Voter Registration
	483	445	928	*Tweets not referring to a question*
	10090	2151	12241	**Total Tweets**

- T16 (“Political Activism”) has jumped to being the fifth most popular question from 12$^{\text{th}}$ in the full 97,666-tweet unlabelled dataset.
- T8a (“Risk Perceptions: Who Is at Risk”) has jumped to #12 from #16.
- T12 (“Conservation Actions”) has jumped to #23 from #27.

Note here that we consider T12 to represent a cut-off in terms of tweet hit counts. Looking again at Figure 6.8, we see a drop in the hit count on the right side of the chart at T12. This drop is the beginning of a tail of less popular questions at the end of the curve. Referring back to Table 6.4, we see that only 150 of the 12,241 tweets in the 17-tweet minimum dataset are associated with T12. Of these, only 13 are labelled as denier tweets. The low coverage in terms of associated tweets for the less popular questions makes it difficult to have any level of confidence in trends we may observe for these

Figure 6.8 Tweet counts for survey questions at a minimum activity level of 17.



questions. Therefore, as we proceed with our analysis using the smaller, high-activity dataset, we will omit the questions on the right side of the graph.

Continuing with this limitation in mind, we see that there are a number of questions that have decreased in popularity by at least four spots:

- T2 ("Attitudinal Certainty and Need for Additional Information to Decide") has dropped to 14$^{\text{th}}$ place in popularity, down from ninth in the full unlabelled dataset.
- T5 ("Beliefs") has dropped to #11. It was #7 in the unlabelled dataset.
- T13 ("Conservation Intentions") has dropped to #19 from #15.
- T26 ("Attention and Response to Climategate") has dropped to #22 from #14.

Now taking a look at the two categories in this dataset, Figures 6.9 and 6.10 present the tweet counts for the users with a minimum of 17 tweets who have been labelled green or denier respectively. These charts graphically represent the "Green" and "Denier" columns in Table 6.4. Immediately we see the shape of the graph of green users is visually similar to the graph with both categories combined (Figure 6.8). This should not be overly surprising since, as mentioned above, most of the tweets in the combined graph (green + denier) are published from the green category. However, the graph of denier tweets has a shape that is markedly different, and this is a trend we will continue to see in each phase of the analysis.

Not surprisingly, as we look closer at the chart for green tweets (Figure 6.9) and the chart for tweets from *both* groups (Figure 6.8), we see that not only the shape, but also the hit count ranking is notably similar. The three most-referenced questions are the same; the two least-referenced questions are the same, and questions in between follow nearly the same order. A few of these swap one or occasionally two places on the x-axis between the two charts, but the ranking of survey questions is quite comparable. The exception is question T24 ("Issue Priority"), which is referred to less often in green tweets than in the tweets from both groups combined.

References to T24 are much more frequent among denier tweets. This question ranks #5 for users in the denier category, but only #26 for users labelled as green. In the survey, the question asks how high of a priority the president and the United States Congress should consider (1) the issue of global warming and (2) the development of clean energy (Leiserowitz et al., 2010). We will consider question T24 individually in Section 6.3.9.

The rank ordering of hits to other survey questions is noticeably different for denier tweets as well (see Figure 6.10). T9 ("Perceptions of Weather and Climate") is the most popular question, while it ranks only at #7 for tweets from users in the green category. More generally, the very top questions from the green tweets also tend to be top questions for the denier group; however, the order of these questions varies noticeably. This similarity in ranking between the two groups ceases after the first six questions with T9 jumping to the top of the denier list as mentioned above and then T30 ("Attention to Specific Programs and Media Sources") falling from #8 for the green group to #14 for the denier group. After this there is little marked similarity between the two groups until the very end of the ranking. Notably, in both groups T17 ("Perceived Importance

Figure 6.9 Green tweet counts for survey questions at a minimum activity level of 17.

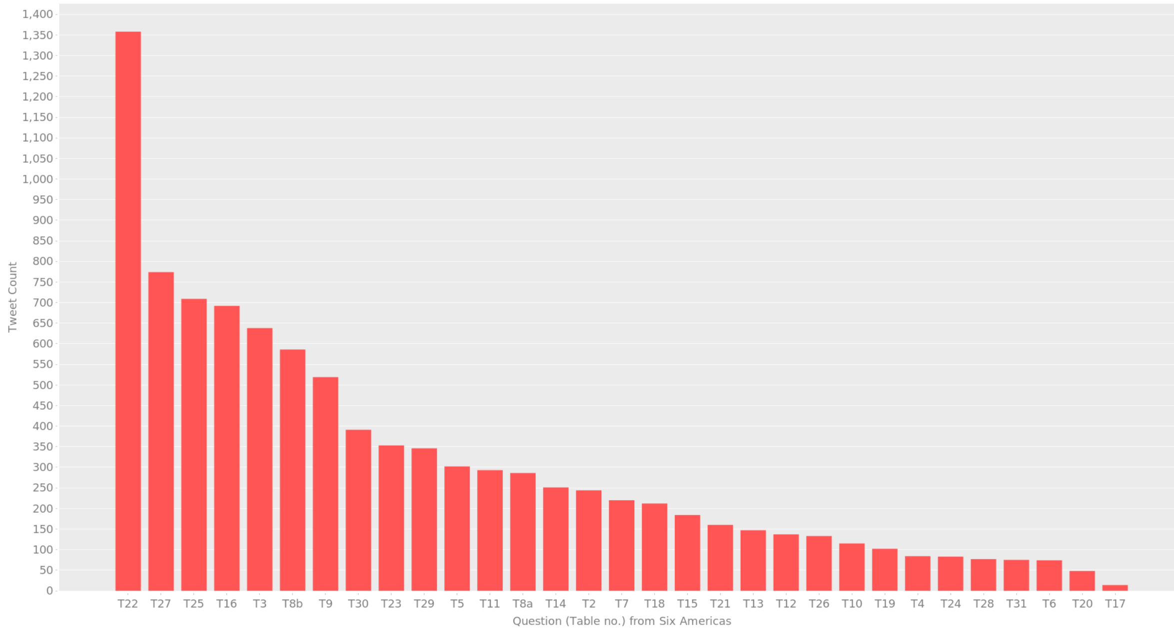
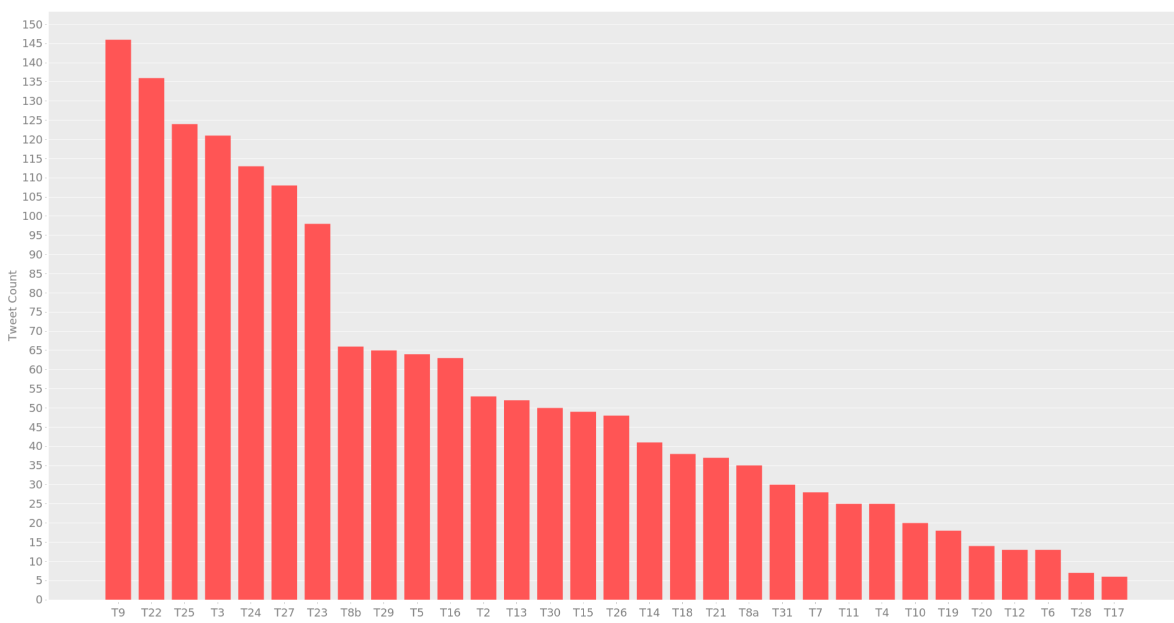


Figure 6.10 Denier tweet counts for survey questions at a minimum activity level of 17.



of Conservation Behaviors") is the least referenced question, marking the end of the tail of questions with low coverage that we mentioned above. We will consider question popularity among users in each category in the sections that follow as we take a closer look at a number of individual questions from the Six Americas.

6.2.2 Affect at 17 or More Tweets

Just as we did in Section 6.1.2, we now use the NRC Word-Emotion Association Lexicon (NRC-10) (Mohammad & Turney, 2013) to analyze the affect expressed in this subset of microblogs marked with the hashtag *#globalwarming* at a minimum participation level of 17 tweets. As these tweets are labelled, we can see how the emotion and sentiment expressed in texts linked to the various survey questions from the Six Americas differ with respect to the user category: green or denier.

Figure 6.11 presents the counts of affective words (y-axis) in the tweets linked to each survey question (x-axis). Once again, we see that the peaks in the bars showing affect are jagged when compared to the bars for tweet hit counts by question (Figure 6.8) as more affect is expressed in relation to some questions over others. This lack of accordance between hit count and the affect expressed in the tweet content actually appears more pronounced for this data subset with the 17-tweet minimum. Most notably, question T16 ("Political Activism"), which jumped seven spots in the popularity ranking to #5, is now the question with the second highest level of expressed affect, where in the full dataset it was the fifth highest (see Figure 6.2). Question T16 asks how active one is in helping to limit global warming by way of organized volunteer work, monetary donations, or communications with media and governmental officials (Leiserowitz et al., 2010).

Table 6.5 displays the underlying data for Figure 6.11 (left side for each affect characteristic) as well as Figure 6.12 (right side of the same). The latter figure presents the affect signature for each survey question. Again, the subdivisions of the bars on the chart give the percentage of each affect characteristic with respect to the overall affect expressed

Figure 6.11 Affect at a 17 tweet minimum by Six Americas survey question.

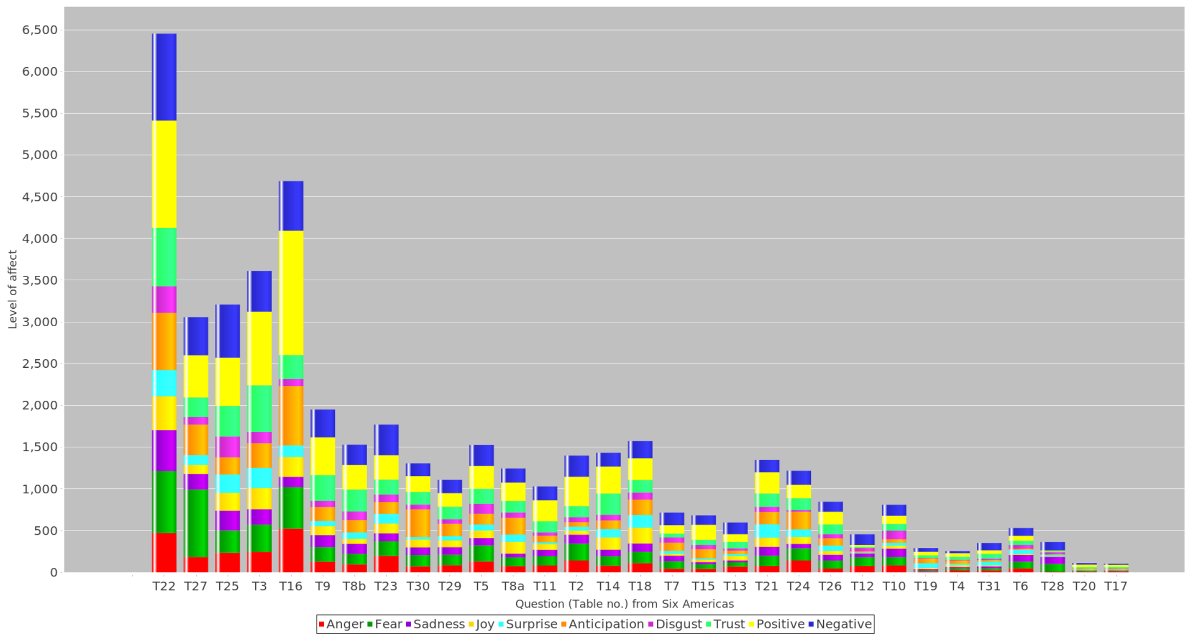


Figure 6.12 Affect signature as percentages at a 17 tweet minimum for survey questions.

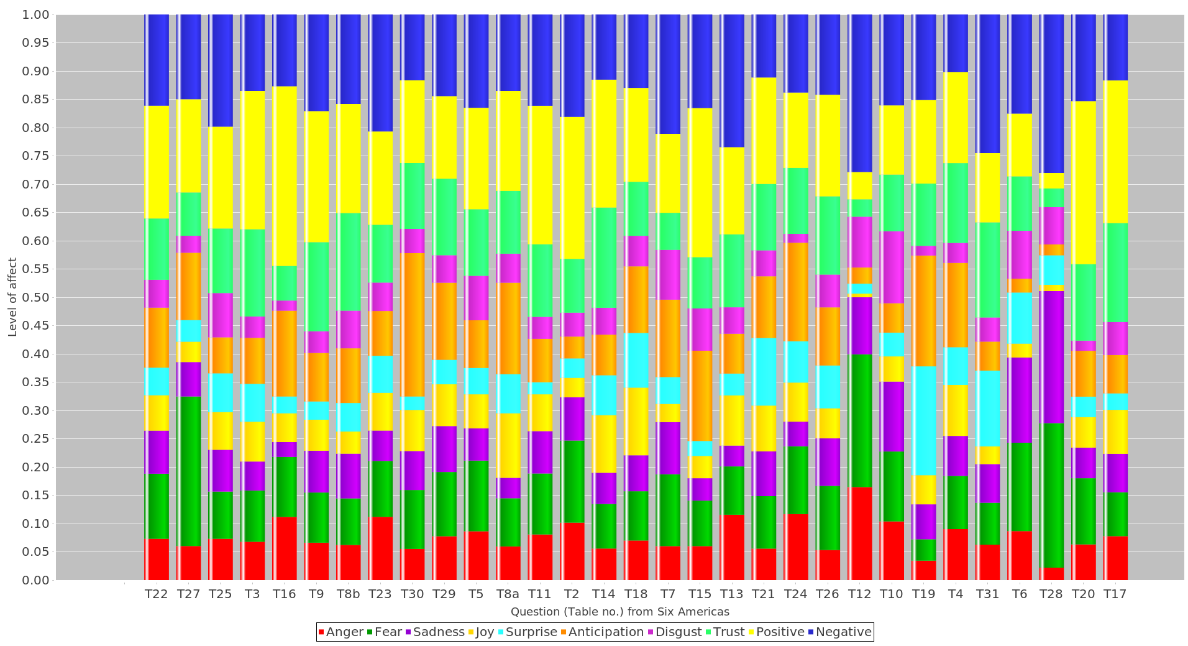


Table 6.5 Affect word counts and percentages at a 17 tweet minimum.

Question	Anger		Fear		Sadness		Joy		Surprise		Anticipation		Disgust		Trust		Positive		Negative	
T2	142	10.2%	203	14.5%	107	7.7%	48	3.4%	48	3.4%	54	3.9%	59	4.2%	133	9.5%	351	25.1%	253	18.1%
T3	245	6.8%	327	9.1%	185	5.1%	254	7.0%	242	6.7%	294	8.1%	136	3.8%	557	15.4%	883	24.5%	488	13.5%
T4	23	9.0%	24	9.4%	18	7.1%	23	9.0%	17	6.7%	38	14.9%	9	3.5%	36	14.1%	41	16.1%	26	10.2%
T5	132	8.6%	191	12.5%	87	5.7%	92	6.0%	71	4.6%	129	8.4%	120	7.9%	180	11.8%	274	17.9%	252	16.5%
T6	46	8.7%	83	15.6%	80	15.1%	13	2.4%	48	9.0%	13	2.4%	45	8.5%	51	9.6%	59	11.1%	93	17.5%
T7	43	6.0%	91	12.7%	66	9.2%	23	3.2%	34	4.7%	98	13.7%	63	8.8%	47	6.6%	100	14.0%	151	21.1%
T8a	74	5.9%	106	8.5%	45	3.6%	142	11.4%	86	6.9%	201	16.2%	64	5.1%	138	11.1%	220	17.7%	168	13.5%
T8b	95	6.2%	126	8.2%	121	7.9%	60	3.9%	77	5.0%	148	9.7%	102	6.7%	264	17.3%	295	19.3%	242	15.8%
T9	129	6.6%	173	8.9%	144	7.4%	107	5.5%	63	3.2%	167	8.6%	75	3.8%	307	15.7%	452	23.2%	333	17.1%
T10	84	10.4%	100	12.4%	100	12.4%	36	4.4%	34	4.2%	42	5.2%	103	12.7%	81	10.0%	99	12.2%	130	16.1%
T11	83	8.1%	111	10.8%	77	7.5%	67	6.5%	22	2.1%	79	7.7%	40	3.9%	132	12.8%	252	24.5%	166	16.1%
T12	75	16.4%	107	23.5%	46	10.1%	3	0.7%	8	1.8%	13	2.9%	41	9.0%	14	3.1%	22	4.8%	127	27.9%
T13	69	11.6%	51	8.5%	22	3.7%	53	8.9%	23	3.9%	42	7.0%	28	4.7%	77	12.9%	92	15.4%	140	23.5%
T14	80	5.6%	113	7.9%	79	5.5%	146	10.2%	101	7.0%	103	7.2%	68	4.7%	254	17.7%	324	22.6%	165	11.5%
T15	41	6.0%	55	8.1%	27	4.0%	27	4.0%	18	2.6%	109	16.0%	51	7.5%	62	9.1%	180	26.4%	113	16.5%
T16	524	11.2%	497	10.6%	123	2.6%	239	5.1%	139	3.0%	711	15.2%	84	1.8%	287	6.1%	1489	31.8%	595	12.7%
T17	8	7.8%	8	7.8%	7	6.8%	8	7.8%	3	2.9%	7	6.8%	6	5.8%	18	17.5%	26	25.2%	12	11.7%
T18	110	7.0%	137	8.7%	100	6.4%	188	12.0%	152	9.7%	185	11.8%	85	5.4%	150	9.5%	261	16.6%	204	13.0%
T19	10	3.4%	11	3.8%	18	6.2%	15	5.2%	56	19.2%	57	19.6%	5	1.7%	32	11.0%	43	14.8%	44	15.1%
T20	7	6.3%	13	11.7%	6	5.4%	6	5.4%	4	3.6%	9	8.1%	2	1.8%	15	13.5%	32	28.8%	17	15.3%
T21	75	5.6%	125	9.3%	107	7.9%	109	8.1%	161	11.9%	147	10.9%	62	4.6%	158	11.7%	254	18.8%	150	11.1%
T22	471	7.3%	743	11.5%	491	7.6%	404	6.3%	315	4.9%	684	10.6%	318	4.9%	700	10.8%	1287	19.9%	1040	16.1%
T23	198	11.2%	175	9.9%	95	5.4%	118	6.7%	116	6.6%	140	7.9%	89	5.0%	181	10.2%	292	16.5%	366	20.7%
T24	142	11.7%	146	12.0%	53	4.4%	84	6.9%	89	7.3%	212	17.4%	19	1.6%	142	11.7%	162	13.3%	168	13.8%
T25	234	7.3%	268	8.4%	237	7.4%	214	6.7%	220	6.9%	204	6.4%	251	7.8%	366	11.4%	578	18.0%	636	19.8%
T26	45	5.3%	96	11.3%	71	8.4%	45	5.3%	64	7.6%	87	10.3%	49	5.8%	117	13.8%	152	18.0%	120	14.2%
T27	184	6.0%	809	26.5%	186	6.1%	110	3.6%	117	3.8%	364	11.9%	92	3.0%	234	7.7%	504	16.5%	458	15.0%
T28	8	2.2%	93	25.5%	85	23.4%	4	1.1%	19	5.2%	7	1.9%	24	6.6%	12	3.3%	10	2.7%	102	28.0%
T29	86	7.8%	126	11.4%	90	8.1%	82	7.4%	48	4.3%	151	13.6%	54	4.9%	150	13.5%	162	14.6%	160	14.4%
T30	72	5.5%	136	10.4%	90	6.9%	95	7.3%	31	2.4%	331	25.3%	56	4.3%	152	11.6%	191	14.6%	152	11.6%
T31	22	6.3%	26	7.4%	24	6.8%	11	3.1%	47	13.4%	18	5.1%	15	4.3%	59	16.8%	43	12.3%	86	24.5%
MEAN		7.7%		11.5%		7.5%		5.9%		6.0%		10.1%		5.3%		11.5%		18.0%		16.5%
STD-DEV		2.81		5.11		3.86		2.71		3.71		5.35		2.48		3.75		6.39		4.61

in the tweets for a given question. Comparing these affect signatures to those from the full dataset (Figure 6.3), we observe the following:

- In the full dataset we saw a general tendency towards positive sentiment over negative. Although this tendency continues in the 17-tweet subset, it is now notably weaker, both in terms of the number of exceptions and how strongly those exceptions contradict the trend:
  - T7 ("Issue Involvement") remains a (slightly stronger) exception at 21.1% negative vs. 14.0% positive.
  - T12 ("Conservation Actions") remains a (now much stronger) exception at 27.9% negative vs. 4.8% positive.
  - T13 ("Conservation Intentions") is now an exception at 23.5% negative vs. 15.4% positive.
  - T25 ("Support for National Response: Specific Climate and Energy Policies") is now an exception at 19.8% negative vs. 18.0% positive.

  Question T24 also becomes an exception, but at less than 1% polarity difference. There are also a few exceptions among the less popular questions on the right side of the chart. These are not listed due to low tweet coverage.
- In addition to T27 ("Attention and Response to IPCC Errors") with expressed fear at 26.5%, question T12 ("Conservation Actions" at 23.5%) is also now associated with increased levels of fear with respect to a mean of 11.5%.
- Question T30 ("Attention to Specific Programs and Media Sources") still shows a high level of anticipation (25.3% against a mean of 10.1%).
- Question T12 ("Conservation Actions") still demonstrates a much higher level of anger (16.4% vs. a mean of 7.7%). Question T10 ("Impact of the Economic Downturn") remains above the mean for anger, but it is no longer notably higher than the other questions.

Affect Signatures for Green and Denier Tweets

The trends visible in the affect signatures at a minimum activity level of 17 tweets do not appear extremely different from those we saw in the signatures for the full dataset. However, when we compare the affect expressed in green tweets and denier tweets as they relate to the Six Americas survey questions, we do indeed observe a number of marked differences between the two groups. Note that for the chart in Figure 6.13 and for all the ones that follow in this section, the order of the survey questions (along the x-axis) is always the same as in Figure 6.8, which shows the tweet hit counts for all the users in the labelled dataset (green and denier categories combined). We varied the order in Figures 6.9 and 6.10 because the purpose of those charts was specifically to show tweet activity by question for users in the green and denier categories respectively. Here we want to keep the question order fixed to facilitate the comparison of the affect expressed in the tweets across the user groups.

Figure 6.13 presents the raw levels of affect (affective word count) expressed in the tweets linked to each survey question *only* for the users labelled as green. Figure 6.14 shows affect signatures as percentages by question for these same tweets. The underlying data for both of these charts is shown in Table 6.6. The affect levels and percentages for the green category seen in these graphs are quite similar to the results for the full 17-tweet minimum dataset (see Figures 6.11 and 6.12). This is not at all surprising since, as we mentioned previously, over 80% of the tweets in this data subset are green.

Users in the denier category, however, are expressing themselves rather differently. Figure 6.15 shows the affect levels for tweets from users labelled as denier. Clearly the biggest disparity is the amount of emotion and sentiment expressed in question T24 ("Issue Priority"), which asks how seriously the U.S. government should be considering the matters of global warming and clean energy (Leiserowitz et al., 2010). Although we are not preserving the order of question popularity in this chart, it is worth mentioning again that T24 ranks fifth for users in the denier group but only 26$^{\text{th}}$ in the green group.

Figure 6.13 Affect for the green category at a 17 tweet minimum.

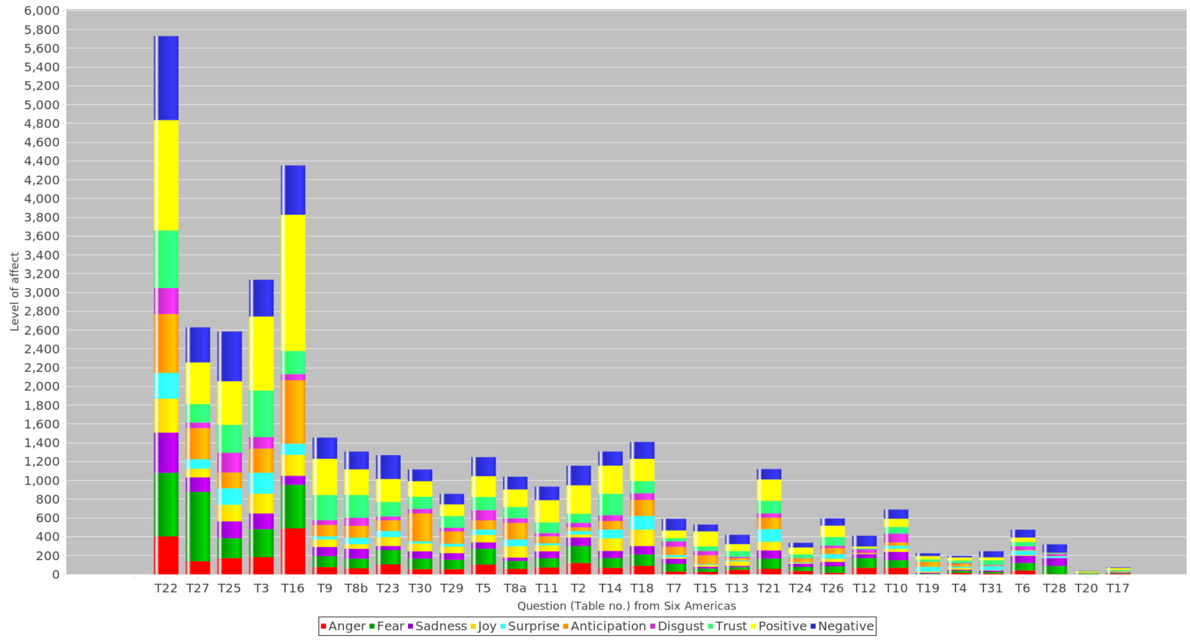


Figure 6.14 Affect signature as percentages for green category at a 17 tweet minimum.

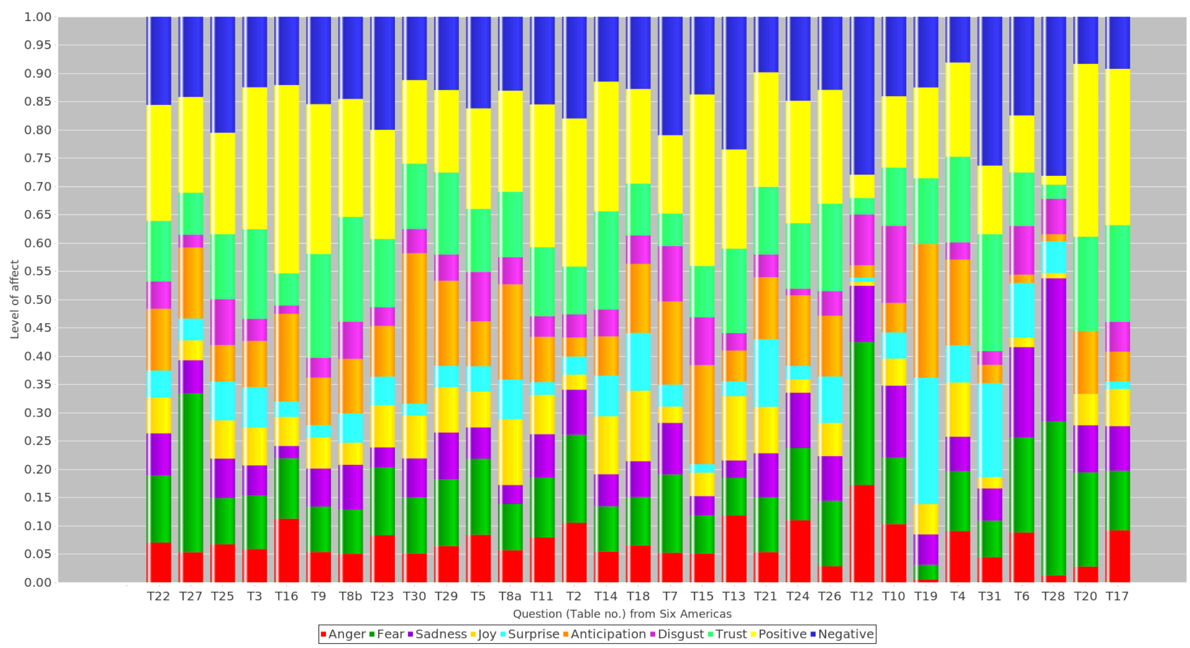


Table 6.6 Affect word counts and percentages: green group at a 17 tweet minimum.

Question	Anger		Fear		Sadness		Joy		Surprise		Anticipation		Disgust		Trust		Positive		Negative	
T2	122	10.5%	180	15.6%	92	8.0%	31	2.7%	37	3.2%	39	3.4%	47	4.1%	98	8.5%	303	26.2%	208	18.0%
T3	185	5.9%	296	9.4%	168	5.4%	210	6.7%	224	7.1%	256	8.2%	122	3.9%	497	15.8%	787	25.1%	391	12.5%
T4	18	9.1%	21	10.6%	12	6.1%	19	9.6%	13	6.6%	30	15.2%	6	3.0%	30	15.2%	33	16.7%	16	8.1%
T5	105	8.4%	168	13.5%	69	5.5%	79	6.3%	56	4.5%	99	7.9%	109	8.7%	139	11.1%	222	17.8%	202	16.2%
T6	42	8.8%	80	16.8%	76	16.0%	8	1.7%	46	9.7%	7	1.5%	41	8.6%	45	9.5%	48	10.1%	83	17.4%
T7	31	5.2%	82	13.9%	54	9.1%	17	2.9%	23	3.9%	87	14.7%	58	9.8%	34	5.7%	82	13.9%	124	20.9%
T8a	59	5.7%	85	8.2%	35	3.4%	121	11.6%	73	7.0%	175	16.8%	50	4.8%	120	11.5%	186	17.9%	136	13.1%
T8b	66	5.0%	102	7.8%	104	8.0%	51	3.9%	68	5.2%	126	9.6%	86	6.6%	242	18.5%	273	20.9%	190	14.5%
T9	78	5.4%	117	8.0%	98	6.7%	80	5.5%	32	2.2%	122	8.4%	51	3.5%	267	18.3%	386	26.5%	225	15.5%
T10	71	10.3%	81	11.7%	88	12.8%	33	4.8%	32	4.6%	36	5.2%	94	13.6%	71	10.3%	87	12.6%	97	14.1%
T11	74	7.9%	99	10.6%	72	7.7%	65	7.0%	21	2.2%	75	8.0%	34	3.6%	114	12.2%	236	25.2%	145	15.5%
T12	71	17.2%	104	25.2%	41	10.0%	3	0.7%	3	0.7%	9	2.2%	37	9.0%	12	2.9%	17	4.1%	115	27.9%
T13	50	11.8%	28	6.6%	13	3.1%	48	11.4%	11	2.6%	23	5.5%	13	3.1%	63	14.9%	74	17.5%	99	23.5%
T14	71	5.4%	105	8.0%	74	5.7%	134	10.2%	94	7.2%	91	7.0%	62	4.7%	227	17.4%	300	22.9%	150	11.5%
T15	27	5.1%	36	6.8%	18	3.4%	22	4.1%	8	1.5%	93	17.5%	45	8.5%	48	9.0%	161	30.3%	73	13.7%
T16	490	11.3%	465	10.7%	95	2.2%	224	5.1%	118	2.7%	675	15.5%	65	1.5%	246	5.7%	1450	33.3%	525	12.1%
T17	7	9.2%	8	10.5%	6	7.9%	5	6.6%	1	1.3%	4	5.3%	4	5.3%	13	17.1%	21	27.6%	7	9.2%
T18	92	6.5%	121	8.6%	89	6.3%	176	12.5%	144	10.2%	172	12.2%	71	5.0%	129	9.1%	236	16.7%	180	12.8%
T19	1	0.4%	6	2.7%	12	5.4%	12	5.4%	50	22.3%	53	23.7%	0	0.0%	26	11.6%	36	16.1%	28	12.5%
T20	1	2.8%	6	16.7%	3	8.3%	2	5.6%	0	0.0%	4	11.1%	0	0.0%	6	16.7%	11	30.6%	3	8.3%
T21	60	5.4%	109	9.7%	87	7.8%	92	8.2%	134	12.0%	123	11.0%	45	4.0%	134	12.0%	227	20.2%	110	9.8%
T22	404	7.1%	678	11.8%	428	7.5%	360	6.3%	276	4.8%	626	10.9%	276	4.8%	613	10.7%	1175	20.5%	893	15.6%
T23	106	8.3%	152	12.0%	45	3.5%	95	7.5%	64	5.0%	114	9.0%	42	3.3%	153	12.0%	245	19.3%	254	20.0%
T24	37	11.0%	43	12.8%	33	9.8%	8	2.4%	8	2.4%	42	12.5%	4	1.2%	39	11.6%	73	21.7%	50	14.8%
T25	174	6.7%	212	8.2%	180	7.0%	175	6.8%	177	6.8%	167	6.5%	210	8.1%	297	11.5%	464	17.9%	530	20.5%
T26	17	2.9%	69	11.6%	47	7.9%	35	5.9%	49	8.2%	64	10.7%	26	4.4%	92	15.4%	120	20.1%	77	12.9%
T27	140	5.3%	739	28.1%	154	5.9%	93	3.5%	101	3.8%	330	12.5%	60	2.3%	195	7.4%	445	16.9%	373	14.2%
T28	4	1.3%	87	27.2%	81	25.3%	3	0.9%	18	5.6%	4	1.3%	20	6.3%	8	2.5%	5	1.6%	90	28.1%
T29	55	6.4%	101	11.8%	71	8.3%	69	8.1%	32	3.7%	129	15.1%	40	4.7%	124	14.5%	125	14.6%	111	13.0%
T30	57	5.1%	111	9.9%	77	6.9%	84	7.5%	24	2.1%	297	26.6%	48	4.3%	129	11.5%	165	14.8%	125	11.2%
T31	11	4.5%	16	6.5%	14	5.7%	5	2.0%	41	16.6%	8	3.2%	6	2.4%	51	20.6%	30	12.1%	65	26.3%
MEAN		7.0%		12.0%		7.6%		5.9%		5.7%		10.3%		4.9%		12.0%		19.1%		15.6%
STD-DEV		3.41		5.82		4.31		3.10		4.71		6.03		3.03		4.47		7.17		5.34

Once again, the affect signatures in Figure 6.16 may prove more useful to visualize the affect that users in the denier group are expressing for the various questions. As usual, we provide the data used to generate the charts for both figures in Table 6.7.

Comparing green (Figure 6.14) and denier (Figure 6.16) percentage affect signatures, we can readily see a number of interesting points:

- The tendency for a greater level of positive sentiment (green mean: 19.1%) is largely reversed in tweets from users in the denier category (mean: 14.7%). Most notably, question T23 ("Support for a National Response: Conditions for & Magnitude of Action Desired") shows a very high level of negative sentiment (22.4%) compared to the positive sentiment (9.4%) expressed in these same denier tweets. Question T25 ("Support for National Response: Specific Climate and Energy Policies") is an exception in that the associated denier tweets express a lower level of negative sentiment (17.0%) than the green tweets (20.5%).
- Question T23 ("Support for a National Response: Conditions for & Magnitude of Action Desired") has the most expressed anger for tweets from users in the denier category (18.4% vs. a mean of 9.9%) but only a slightly elevated level of anger for green tweets (8.3% against a mean of 7.0%).
- Questions T27 ("Attention and Response to IPCC Errors") and T12 ("Conservation Actions") show high levels of fear for users in the green category (28.1% and 25.2% respectively against a mean of 12.0%), but relatively less fear from users in the denier category (16.4% and 6.8% against a mean of 9.5%).
- Question T30 ("Attention to Specific Programs and Media Sources") has the highest expressed anticipation for users in both the denier (18.0% vs. a mean of 9.4%) and green (26.6% vs. a mean of 10.3%) groups.

Notable trends are not immediately obvious for the emotions sadness, joy, disgust, and trust in this analysis.[18] It seems interesting that in our study when one emotion in one

[18]Questions T28 and T6 show high levels of sadness in the 17-tweet minimum dataset; however, these

Figure 6.15 Affect for denier category at a 17 tweet minimum.

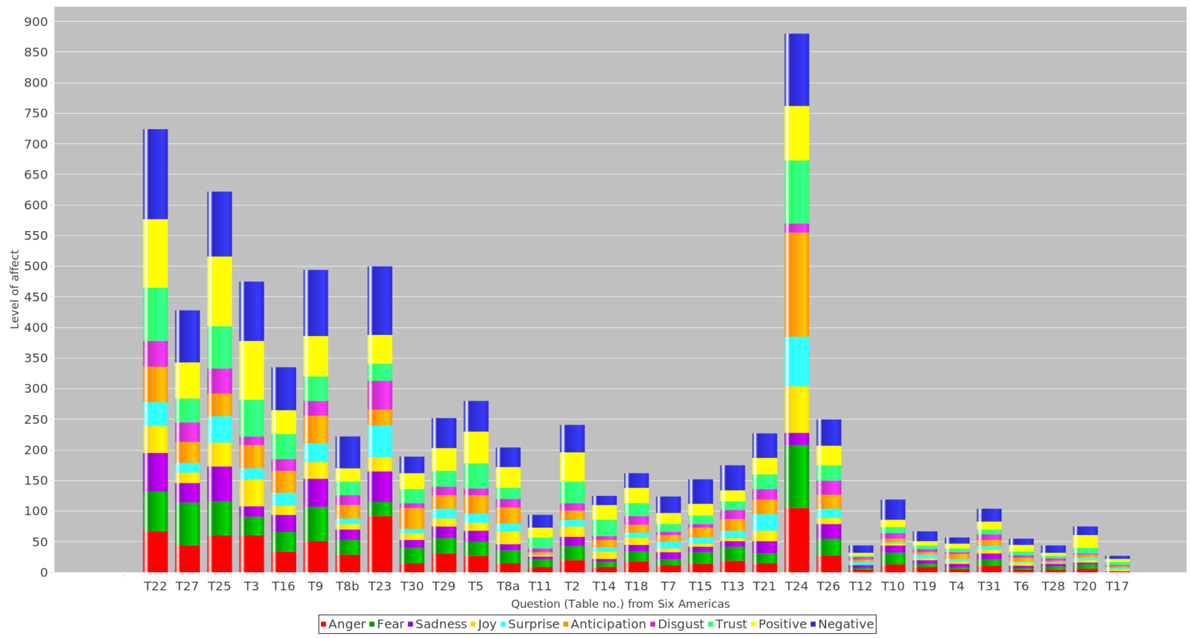


Figure 6.16 Affect signature as percentages for denier category at a 17 tweet minimum.

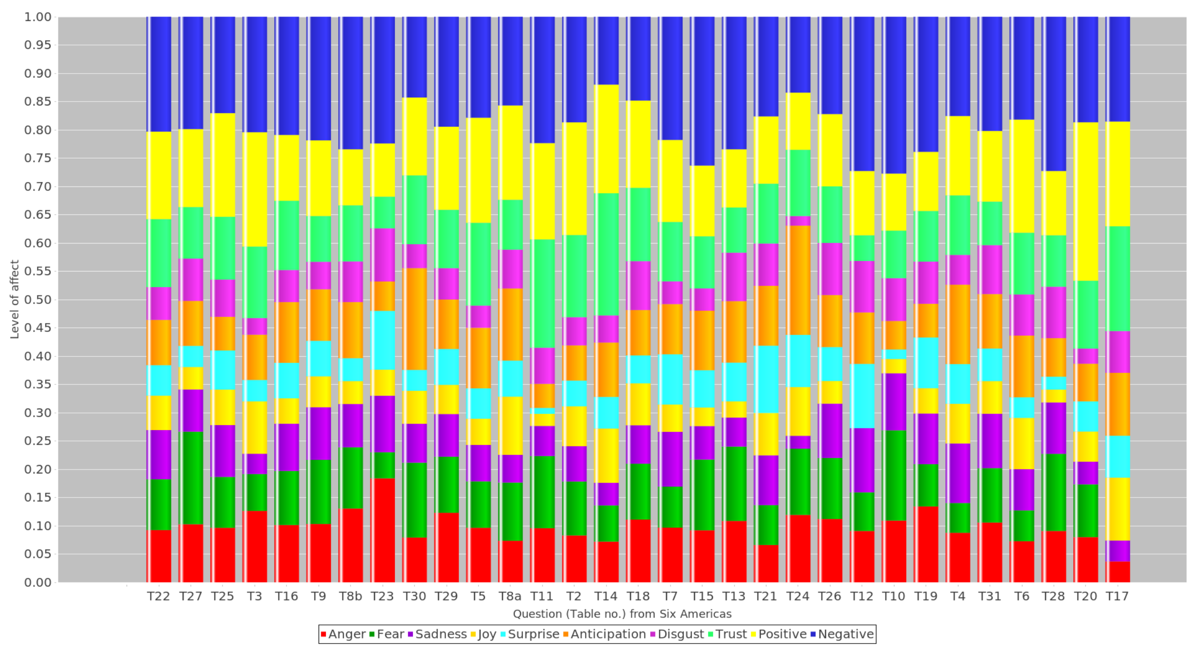


Table 6.7 Affect word counts and percentages: denier group at a 17 tweet minimum.

Question	Anger		Fear		Sadness		Joy		Surprise		Anticipation		Disgust		Trust		Positive		Negative	
T2	20	8.3%	23	9.5%	15	6.2%	17	7.1%	11	4.6%	15	6.2%	12	5.0%	35	14.5%	48	19.9%	45	18.7%
T3	60	12.6%	31	6.5%	17	3.6%	44	9.3%	18	3.8%	38	8.0%	14	2.9%	60	12.6%	96	20.2%	97	20.4%
T4	5	8.8%	3	5.3%	6	10.5%	4	7.0%	4	7.0%	8	14.0%	3	5.3%	6	10.5%	8	14.0%	10	17.5%
T5	27	9.6%	23	8.2%	18	6.4%	13	4.6%	15	5.4%	30	10.7%	11	3.9%	41	14.6%	52	18.6%	50	17.9%
T6	4	7.3%	3	5.5%	4	7.3%	5	9.1%	2	3.6%	6	10.9%	4	7.3%	6	10.9%	11	20.0%	10	18.2%
T7	12	9.7%	9	7.3%	12	9.7%	6	4.8%	11	8.9%	11	8.9%	5	4.0%	13	10.5%	18	14.5%	27	21.8%
T8a	15	7.4%	21	10.3%	10	4.9%	21	10.3%	13	6.4%	26	12.7%	14	6.9%	18	8.8%	34	16.7%	32	15.7%
T8b	29	13.1%	24	10.8%	17	7.7%	9	4.1%	9	4.1%	22	9.9%	16	7.2%	22	9.9%	22	9.9%	52	23.4%
T9	51	10.3%	56	11.3%	46	9.3%	27	5.5%	31	6.3%	45	9.1%	24	4.9%	40	8.1%	66	13.4%	108	21.9%
T10	13	10.9%	19	16.0%	12	10.1%	3	2.5%	2	1.7%	6	5.0%	9	7.6%	10	8.4%	12	10.1%	33	27.7%
T11	9	9.6%	12	12.8%	5	5.3%	2	2.1%	1	1.1%	4	4.3%	6	6.4%	18	19.1%	16	17.0%	21	22.3%
T12	4	9.1%	3	6.8%	5	11.4%	0	0.0%	5	11.4%	4	9.1%	4	9.1%	2	4.5%	5	11.4%	12	27.3%
T13	19	10.9%	23	13.1%	9	5.1%	5	2.9%	12	6.9%	19	10.9%	15	8.6%	14	8.0%	18	10.3%	41	23.4%
T14	9	7.2%	8	6.4%	5	4.0%	12	9.6%	7	5.6%	12	9.6%	6	4.8%	27	21.6%	24	19.2%	15	12.0%
T15	14	9.2%	19	12.5%	9	5.9%	5	3.3%	10	6.6%	16	10.5%	6	3.9%	14	9.2%	19	12.5%	40	26.3%
T16	34	10.1%	32	9.6%	28	8.4%	15	4.5%	21	6.3%	36	10.7%	19	5.7%	41	12.2%	39	11.6%	70	20.9%
T17	1	3.7%	0	0.0%	1	3.7%	3	11.1%	2	7.4%	3	11.1%	2	7.4%	5	18.5%	5	18.5%	5	18.5%
T18	18	11.1%	16	9.9%	11	6.8%	12	7.4%	8	4.9%	13	8.0%	14	8.6%	21	13.0%	25	15.4%	24	14.8%
T19	9	13.4%	5	7.5%	6	9.0%	3	4.5%	6	9.0%	4	6.0%	5	7.5%	6	9.0%	7	10.4%	16	23.9%
T20	6	8.0%	7	9.3%	3	4.0%	4	5.3%	4	5.3%	5	6.7%	2	2.7%	9	12.0%	21	28.0%	14	18.7%
T21	15	6.6%	16	7.0%	20	8.8%	17	7.5%	27	11.9%	24	10.6%	17	7.5%	24	10.6%	27	11.9%	40	17.6%
T22	67	9.3%	65	9.0%	63	8.7%	44	6.1%	39	5.4%	58	8.0%	42	5.8%	87	12.0%	112	15.5%	147	20.3%
T23	92	18.4%	23	4.6%	50	10.0%	23	4.6%	52	10.4%	26	5.2%	47	9.4%	28	5.6%	47	9.4%	112	22.4%
T24	105	11.9%	103	11.7%	20	2.3%	76	8.6%	81	9.2%	170	19.3%	15	1.7%	103	11.7%	89	10.1%	118	13.4%
T25	60	9.6%	56	9.0%	57	9.2%	39	6.3%	43	6.9%	37	5.9%	41	6.6%	69	11.1%	114	18.3%	106	17.0%
T26	28	11.2%	27	10.8%	24	9.6%	10	4.0%	15	6.0%	23	9.2%	23	9.2%	25	10.0%	32	12.8%	43	17.2%
T27	44	10.3%	70	16.4%	32	7.5%	17	4.0%	16	3.7%	34	7.9%	32	7.5%	39	9.1%	59	13.8%	85	19.9%
T28	4	9.1%	6	13.6%	4	9.1%	1	2.3%	1	2.3%	3	6.8%	4	9.1%	4	9.1%	5	11.4%	12	27.3%
T29	31	12.3%	25	9.9%	19	7.5%	13	5.2%	16	6.3%	22	8.7%	14	5.6%	26	10.3%	37	14.7%	49	19.4%
T30	15	7.9%	25	13.2%	13	6.9%	11	5.8%	7	3.7%	34	18.0%	8	4.2%	23	12.2%	26	13.8%	27	14.3%
T31	11	10.6%	10	9.6%	10	9.6%	6	5.8%	6	5.8%	10	9.6%	9	8.7%	8	7.7%	13	12.5%	21	20.2%
MEAN		9.9%		9.5%		7.4%		5.6%		6.1%		9.4%		6.3%		11.1%		14.7%		20.0%
STD-DEV		2.59		3.45		2.38		2.61		2.59		3.36		2.11		3.66		4.18		4.04

of Plutchik's pairs (Plutchik, 2001) shows a trend (or fails to do so), the other emotion in the pair generally follows suit. This was true for our analysis of the big players as well (see Chapter 3) and could be a topic for future research.

Polarity Signatures

As we did for the full dataset, we repeat the analysis using Bing Liu's Opinion Lexicon (Hu & Liu, 2004), which gives us sentiment polarity, only now for texts from users publishing the minimum of 17 tweets. We remind the reader that as we have found a notable difference in the intensity of positive and negative sentiment by survey question when comparing the two lexica, we must use caution when identifying trends in our observations. Here, we have two objectives. First, we wish to see how an analysis with Liu's lexicon on the smaller dataset of high-activity users compares with our full dataset. Second, we continue a comparison of the two lexica, focusing on commonalities and differences as they relate to the 17-tweet minimum dataset.

Figure 6.17 gives sentiment polarity by word count for all users (both the green and denier categories) in the 17-tweet minimum data subset. Again, the order of the questions corresponds to the tweet hit count for the various survey questions (see Figure 6.8). The underlying data for the chart is shown in Table 6.8. When we compare the 17-tweet minimum chart with Figure 6.4, which gives sentiment intensity for the full dataset, we essentially find that not too much has changed now that we are focusing on the high-activity users. Taking into account the slight adjustment in the order of the questions at the 17-tweet minimum, we note the following:

- Question T22 ("Outcome Expectations") has more than double the level of expressed sentiment of any other question in both datasets.
- In contrast, question T30 ("Attention to Specific Programs and Media Sources")

---

are among the questions excluded from the analysis due to low tweet hit count.

Figure 6.17 Sentiment at a 17 tweet minimum.

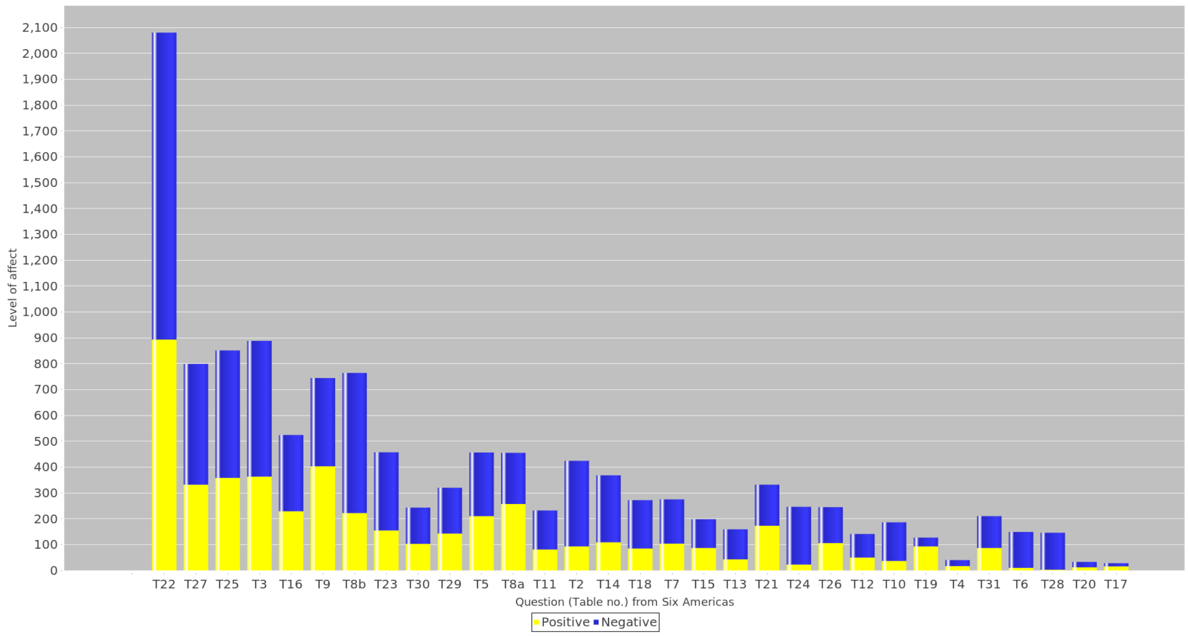


Figure 6.18 Sentiment signature as percentages at a 17 tweet minimum.

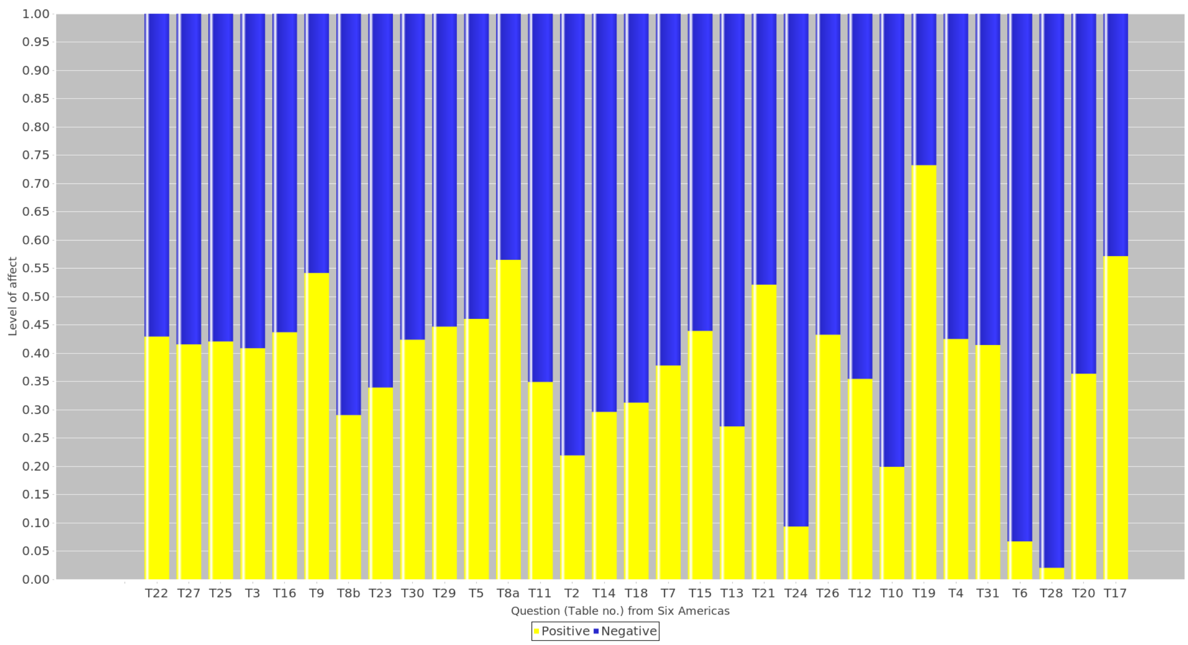


Table 6.8 Sentiment word counts and percentages at a 17 tweet minimum.

Question	Positive		Negative	
T2	93	21.9%	331	78.1%
T3	363	40.9%	525	59.1%
T4	17	42.5%	23	57.5%
T5	210	46.1%	246	53.9%
T6	10	6.7%	139	93.3%
T7	104	37.8%	171	62.2%
T8a	257	56.5%	198	43.5%
T8b	222	29.1%	542	70.9%
T9	403	54.2%	341	45.8%
T10	37	19.9%	149	80.1%
T11	81	34.9%	151	65.1%
T12	50	35.5%	91	64.5%
T13	43	27.0%	116	73.0%
T14	109	29.6%	259	70.4%
T15	87	43.9%	111	56.1%
T16	229	43.7%	295	56.3%
T17	16	57.1%	12	42.9%
T18	85	31.3%	187	68.8%
T19	93	73.2%	34	26.8%
T20	12	36.4%	21	63.6%
T21	173	52.1%	159	47.9%
T22	893	42.9%	1187	57.1%
T23	155	33.9%	302	66.1%
T24	23	9.3%	223	90.7%
T25	358	42.1%	493	57.9%
T26	106	43.3%	139	56.7%
T27	332	41.6%	467	58.4%
T28	3	2.1%	143	97.9%
T29	143	44.7%	177	55.3%
T30	103	42.4%	140	57.6%
T31	87	41.4%	123	58.6%
MEAN		37.5%		62.5%
STD-DEV		15.00		15.00

has a very low level of expressed sentiment despite the fact that it is a relatively popular question. We see a similar drop for questions T11 (“Political Activism”), T15 (“Consumer Activism”), and T13 (“Conservation Intentions”), though these questions are less popular.

- Levels in other questions vary slightly between the two datasets, but generally the high-activity users seem to reflect fairly well the sentiment proportionality seen when considering all the users. There are exceptions for some questions with low hit counts, but we hesitate to make note of possible trends for these due to the relatively small number of associated tweets.

As explained above, we also compare the chart in Figure 6.17 with the corresponding one for NRC-10 at a 17-tweet minimum (Figure 6.11). There are a couple of noteworthy observations:

- T22 (“Outcome Expectations”) is by far the most charged question, both in terms of general affect (per the NRC-10) and sentiment polarity (per Liu’s lexicon).
- Question T16 (“Political Activism”), which NRC-10 reported as having the second highest level of affect intensity in spite of being only the fifth most popular question, shows a relatively low level of intensity in terms of sentiment polarity per Liu’s lexicon.

The associated signatures as percentages of expressed sentiment are displayed in Figure 6.18, which also takes its data from Table 6.8. As we found with the raw intensity levels, for the most popular questions, there is not too much change between these signatures and those for the full dataset (Figure 6.5). The questions showing the most notable changes are: Question T24 (“Issue Priority”) at 90.7% negative sentiment at the 17-tweet minimum vs. 66.5% for the full dataset; T2 (“Attitudinal Certainty and Need for Additional Information to Decide”) at 78.1% vs. 59% for the same; and Question 13 (“Conservation Intentions”) at 73% vs. 58%. Again, there is less regularity in signatures for questions with low tweet counts. A number of these show much higher levels of negativity, but we are focusing the analysis on the more popular questions with a greater number of associated tweets. Notably, the mean percentage of negative sentiment is just over 2% greater in the high-activity dataset.

Finally, we compare the signatures for Liu’s Opinion Lexicon with those for the NRC-10 (Figure 6.12). As we saw with the full dataset, the general trend of increased positive sentiment across the majority of questions is simply not present when we perform the analysis with Liu’s lexicon. Most notably:

- Question T16 (“Political Activism”), which shows the highest level of positive sentiment per NRC-10, scores only 43.7% positive per Liu’s lexicon, which is less than one standard deviation above the mean (see Table 6.8).

- Similarly, popular questions that Liu's lexicon scores relatively high for negative sentiment are generally more positive than negative per the NRC-10 analysis. The most notable examples are: T8b ("Risk Perceptions: When Harm Will Occur" at 70.9% negative) and T2 ("Attitudinal Certainty and Need for Additional Information to Decide" at 78.1% negative).
- An exception is question T23 ("Support for a National Response: Conditions for and Magnitude of Action Desired" at 66.1%), which is more negative than positive according to both lexica.

Looking at the less popular questions on the right side of the chart in Figure 6.18, we see a number of questions with exceedingly high levels of negativity. Although it is tempting to begin speculating on what all the negative sentiment could potentially represent, we must leave this matter for future research efforts on social media which incorporate more data associated with these questions.

Polarity Signatures for Green and Denier Tweets

Once again, as the 17-tweet minimum dataset is labelled, we are able to extend our analysis to look at users found to be in the green category separately from those in the denier category. Figure 6.19 presents the sentiment intensity levels (y-axis) per Liu's Opinion Lexicon for tweets from the green group linked to each of the questions (x-axis) from the Six Americas survey. As before, the questions are listed in order of descending tweet hit count for all users in this dataset (see Figure 6.8) so that we may more easily compare results between related charts. The data used to generate this chart is displayed in Table 6.9. When we compare this chart to its counterpart in Figure 6.17, which shows all users (green and denier) for the 17-tweet minimum, we see very little difference. A few of the more popular questions show some increased negative sentiment, but the changes are not particularly striking. This is unsurprising as we already know that most of the tweets from the 17-tweet minimum dataset are green.

Figure 6.19 Sentiment for green category at a 17 tweet minimum.

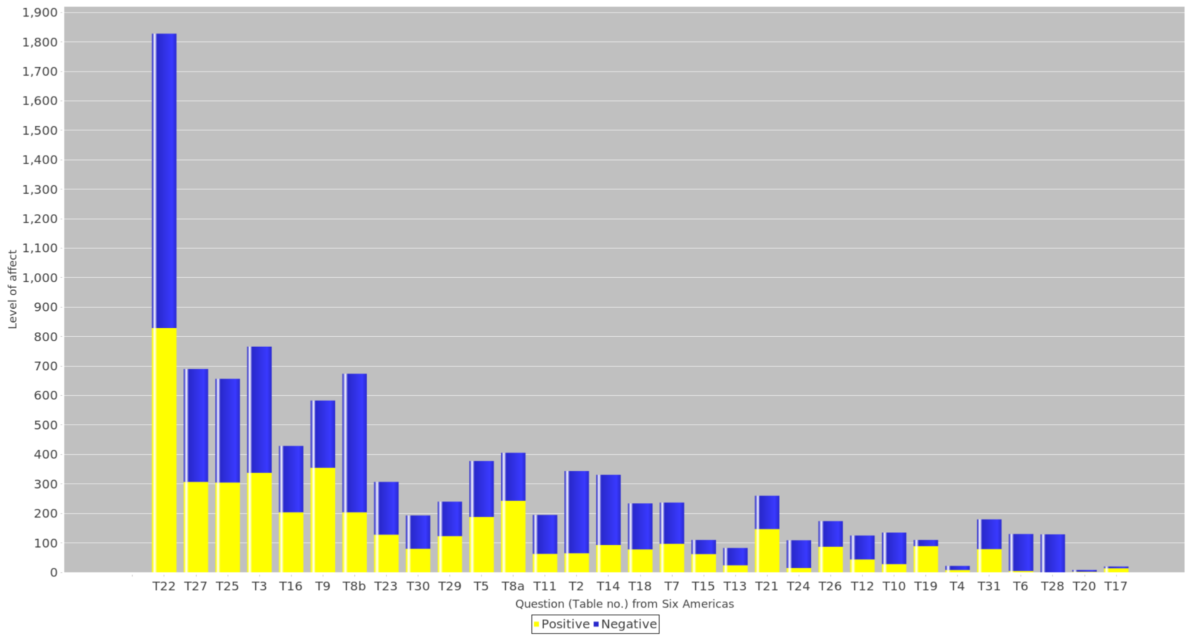
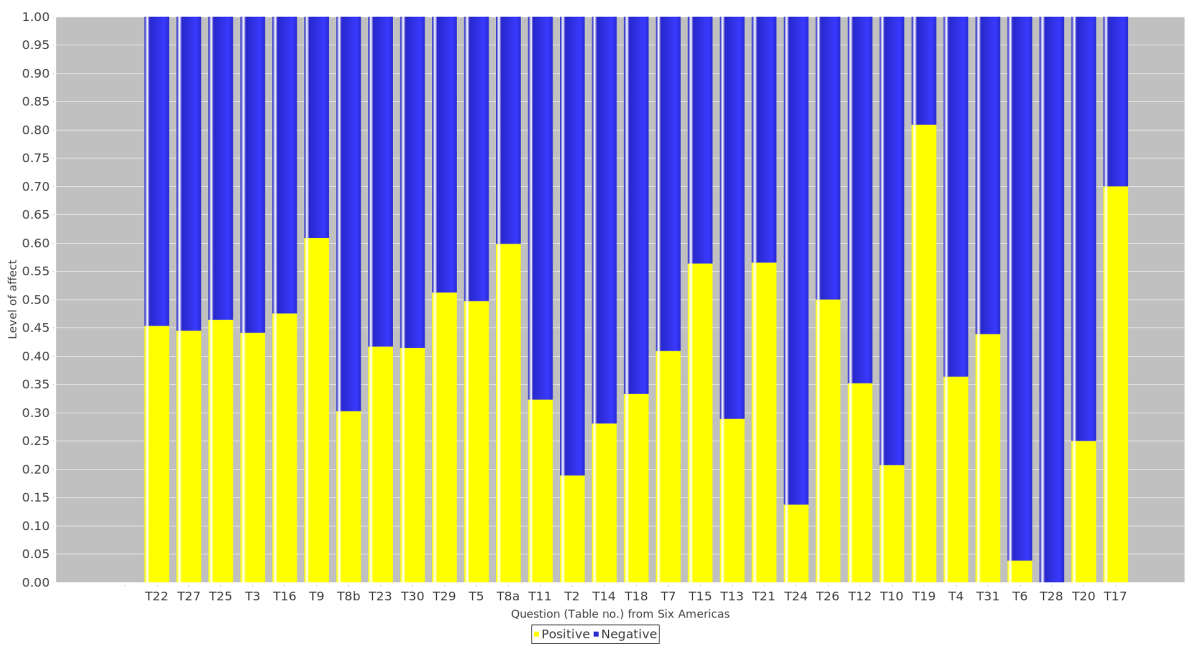


Figure 6.20 Sentiment signature as percentages for green group at a 17 tweet minimum.



Comparing the Liu sentiment levels by question in Figure 6.19 with the corresponding NRC-10 chart for green-category tweets (Figure 6.13), we note the same observations we mentioned above when looking at tweets from all 17-tweet users. Namely, the two lexica seem to agree that tweets linked to T22 ("Outcome Expectations") collectively express much more affect than sets of tweets linked to the other questions (well over twice the level of expressed sentiment per Liu's Opinion Lexicon). However, the lexica disagree as to whether tweets linked to the fifth most popular question, T16 ("Political Activism"), express a relatively high level of affect. The NRC-10 lexicon indicates that they do, while Liu's lexicon scores the associated affect lower than other questions around that ranking.

Table 6.9 Sentiment word counts & percentages: green group at a 17 tweet minimum.

Question	**Positive**		**Negative**	
T2	65	18.9%	279	81.1%
T3	338	44.1%	428	55.9%
T4	8	36.4%	14	63.6%
T5	188	49.7%	190	50.3%
T6	5	3.8%	125	96.2%
T7	97	40.9%	140	59.1%
T8a	243	59.9%	163	40.1%
T8b	204	30.3%	470	69.7%
T9	355	60.9%	228	39.1%
T10	28	20.7%	107	79.3%
T11	63	32.3%	132	67.7%
T12	44	35.2%	81	64.8%
T13	24	28.9%	59	71.1%
T14	93	28.1%	238	71.9%
T15	62	56.4%	48	43.6%
T16	204	47.6%	225	52.4%
T17	14	70.0%	6	30.0%
T18	78	33.3%	156	66.7%
T19	89	80.9%	21	19.1%
T20	2	25.0%	6	75.0%
T21	147	56.5%	113	43.5%
T22	829	45.4%	999	54.6%
T23	128	41.7%	179	58.3%
T24	15	13.8%	94	86.2%
T25	305	46.4%	352	53.6%
T26	87	50.0%	87	50.0%
T27	307	44.5%	383	55.5%
T28	0	0.0%	129	100.0%
T29	123	51.3%	117	48.8%
T30	80	41.5%	113	58.5%
T31	79	43.9%	101	56.1%
MEAN		39.9%		60.1%
STD-DEV		17.88		17.88

Figure 6.20 again stretches these bars to show them as percentages of expressed affect representing the sentiment signatures for each survey question. As with the previous

graph, the underlying data for this chart comes from Table 6.9. Comparing these signatures to those for all the 17-tweet minimum users (Figure 6.18) shows only minor changes in sentiment proportions for any given question. These results are more or less as expected since the green tweets makes up the larger part of the dataset.

Since T22 ("Outcome Expectations") and T16 ("Political Activism") have shown themselves to be questions of interest for the comparison of Liu's Opinion Lexicon with NRC-10, we are particularly interested to see how their Liu sentiment signatures compare to their NRC-10 affect signatures (Figure 6.14) for the green tweets. T22 presents no surprises as we again see the relatively higher level of expressed positive sentiment with NRC-10 but a lower positive sentiment with Liu's lexicon. The positive/negative proportions are not overwhelmingly skewed with either lexicon for T22. With question T16, although the general positive/negative discordance between the lexica exists, it appears rather slight. The NRC-10 scores it as the most positive question for the green group, while the Liu signature gives a positive score of 47.6% against a mean of 39.9%.

To end our comparison analysis using Liu's Opinion Lexicon, we look at tweets from the denier group for the 17-tweet minimum activity level. Figure 6.21 shows the sentiment levels by word count (y-axis) against the Six Americas survey questions (still following the order established in Figure 6.8). The chart takes its data from Table 6.10. We note once again that expressed sentiment looks quite different when comparing denier tweets with tweets from the green group (Figure 6.19) or both groups (Figure 6.17). We also see that the trend for increased negative sentiment when using Liu's lexicon is remarkably stronger for denier tweets than for green or both groups with a mean negative sentiment level of 72.0% vs. 60.1% for the green category and 62.5% for both combined (compare against Tables 6.9 and 6.8 respectively). We also note that although tweets associated with question T22 ("Outcome Expectations") still express the most affect, the difference compared with other high-sentiment questions is much less pronounced. The increase in expressed sentiment for some questions is explained by the fact that the order of questions along the x-axis is fixed for the complete 17-tweet minimum dataset (both green and denier groups as shown in Figure 6.8), but tweet hit counts for the denier group fall somewhat differently. Yet, referring back to Figure 6.10 for the question order for the denier groups further marks the importance of question T22. It is the second most popular question in the denier category, but it still receives the highest expressed sentiment. By contrast, question T9 ("Perceptions of Weather and Climate") is the most popular question for the denier group but has a relatively moderate level of associated sentiment.

Table 6.10 Sentiment word counts & percentages: denier group at a 17 tweet minimum.

Question	Positive		Negative	
T2	28	35.0%	52	65.0%
T3	25	20.5%	97	79.5%
T4	9	50.0%	9	50.0%
T5	22	28.2%	56	71.8%
T6	5	26.3%	14	73.7%
T7	7	18.4%	31	81.6%
T8a	14	28.6%	35	71.4%
T8b	18	20.0%	72	80.0%
T9	48	29.8%	113	70.2%
T10	9	17.6%	42	82.4%
T11	18	48.6%	19	51.4%
T12	6	37.5%	10	62.5%
T13	19	25.0%	57	75.0%
T14	16	43.2%	21	56.8%
T15	25	28.4%	63	71.6%
T16	25	26.3%	70	73.7%
T17	2	25.0%	6	75.0%
T18	7	18.4%	31	81.6%
T19	4	23.5%	13	76.5%
T20	10	40.0%	15	60.0%
T21	26	36.1%	46	63.9%
T22	64	25.4%	188	74.6%
T23	27	18.0%	123	82.0%
T24	8	5.8%	129	94.2%
T25	53	27.3%	141	72.7%
T26	19	26.8%	52	73.2%
T27	25	22.9%	84	77.1%
T28	3	17.6%	14	82.4%
T29	20	25.0%	60	75.0%
T30	23	46.0%	27	54.0%
T31	8	26.7%	22	73.3%
MEAN		28.0%		72.0%
STD-DEV		10.03		10.03

When we compare these sentiment levels against the previously presented affect counts for the denier group according to NRC-10 (Figure 6.15), we now observe a similar chart shape from both lexica, not only for question T22, but indeed for most of the questions in the survey. The raw word counts will be different, of course. This is expected due to the different sizes of the lexica as well as the fact that NRC-10 is covering ten attributes

Figure 6.21 Sentiment for denier category at a 17 tweet minimum.

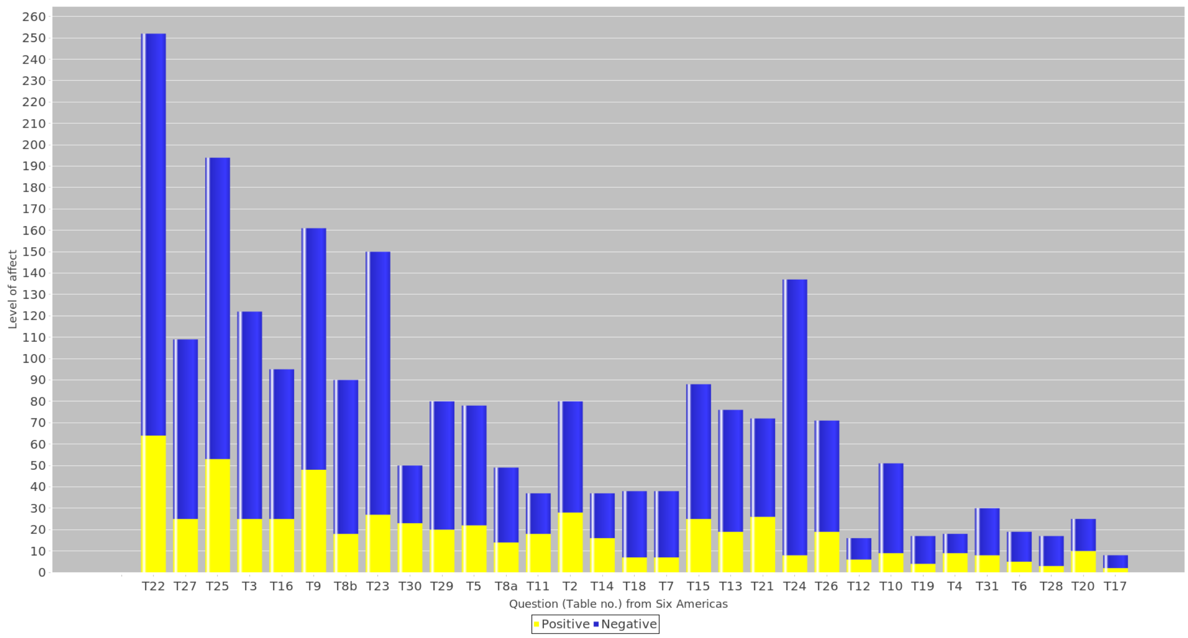
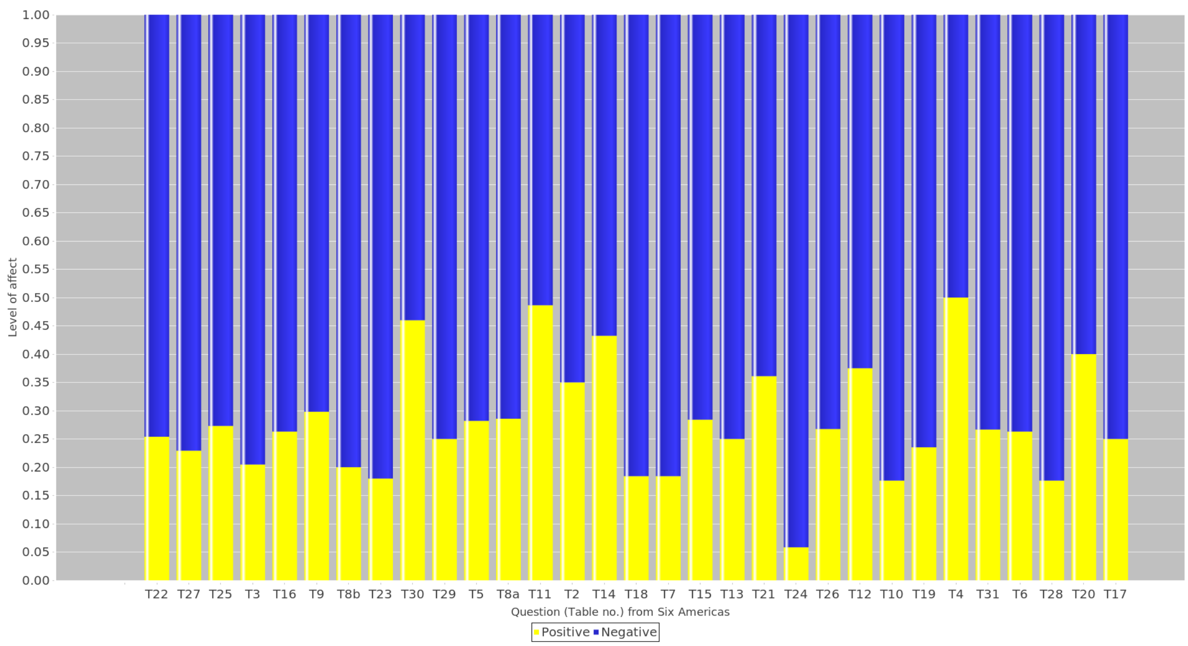


Figure 6.22 Sentiment signature as percentages for denier group at a 17 tweet minimum.



rather than two. The proportion of positive to negative sentiment is also better aligned between the two lexica, but this may simply be a result of the fact that denier tweets are expressing more negative sentiment in general. Given that we did not see this alignment when analyzing the green tweets or both groups together, without further study we cannot simply interpret it as an indicator that one (or potentially both) of these lexica may be better suited for use with our target data than our earlier findings have suggested. Even so, the striking similarity of the affect levels when we link denier tweets to the survey questions clearly calls for continued research. Lastly, we continue here to be interested in question T24 ("Issue Priority"), which both lexica clearly show to be an issue remarkably charged with sentiment and emotion in the denier category.

In Figure 6.22 we move to an analysis using sentiment percentages as survey question signatures. The underlying data for the chart is in Table 6.10. The trend for negative sentiment in denier tweets is clear. No question shows more positive sentiment than negative.[19] Most questions show increased negative sentiment for the deniers as compared to the green group (Figure 6.20). Among the more popular questions for which we have more data, the exceptions are:

- T30 ("Attention to Specific Programs and Media Sources") scores less negative sentiment for denier tweets (54.0%) than for green (58.5% from Table 6.9).
- T11 ("Environmental Protection and Economic Growth") scores 51.4% negative sentiment for denier tweets vs. 67.7% for green.
- T12 ("Conservation Actions") scores 62.5% negative sentiment for deniers vs. 64.8% for greens.
- T14 ("Perceived Impact of Own Actions") scores 56.8% negative sentiment for denier tweets vs. 71.9% for green.

A few of the less popular questions also show less negative sentiment for the denier tweets, but again we are concentrating on questions for which we have more data.

[19] Interestingly, T4 is split 50%–50%, but this question has few linked tweets.

Finally, although T24 ("Issue Priority") is close to the borderline with respect to our cutoff for questions with insufficient tweet counts for a proper analysis, its signature is nevertheless striking. Furthermore, it is still fifth in popularity for denier tweets (see Figure 6.10). It is also clearly the question generating the highest ratio of negative-to-positive sentiment per Liu's lexicon, but not for NRC-10 (see Figure 6.16). Rather, NRC-10 reports that denier tweets are expressing a very high level of anticipation compared to other questions as well as a relatively high level of anger.

Considering users in the denier category and the tweets they are publishing, we see that back in Section 5.3, we had problems identifying these users based on the denier-oriented SCR pairs we modelled in our ontology. It is interesting that now, performing a different type of analysis on the same data, we have a similar problem, but reversed. Here, it is characteristics of users from the green group that are difficult to distinguish out of the mass of tweets, while the denier group sets itself apart relatively clearly. Even the lexica we have considered appear to be in better accord when we are targeting denier tweets. It is not simply that in the present chapter we have turned our attention back to emotion and sentiment. We did model these elements in the ontology. We did not, however, include them as part of the results presented in Chapter 5 because we did not succeed in using description logic to find clear "green vs. denier" patterns that served to identify users in a category based on the affect expressed in their tweets. As we have continuously followed an iterative approach in our research, we now pose the hypothesis that rather than simply modelling affective elements where they occur in the tweets, incorporating the abstraction of an affect signature into the ontology may significantly improve our results from the analysis using description logic. Certainly, we would first need to address the pending issue of determining which lexicon is most applicable to microblogs about climate change. These endeavours represent promising paths for future research.

For the present research, we see a few interesting trends to note. When working with the NRC 10, we see a clear tendency for positive sentiment over negative for both the full and the 17-tweet minimum datasets. This tendency does not appear with Liu's

Opinion Lexicon, but we must consider that the higher number of negative terms in this lexicon is likely a factor leading to this difference in observation. Both lexica indicate increased negative sentiment in tweets from users in the denier category. We see anger and fear, polar opposites in Plutchik's system of basic emotions (Plutchik, 2001), as being integrally tied to the conversation on a number of questions. Twitter users also appear to express increased levels of anticipation for a small number of questions, while other basic emotions do not generally show a trend. In the following section we will take another look at these trends as they relate to a number of selected questions. We should keep in mind, however, that we are seeing exceptions with respect to specific questions and that our initial methodology for linking tweets to survey questions will need to be refined over the course of our continued research (see Section 6.4.2).

6.3 Selected Survey Questions

For the purpose of deciding which Six Americas survey questions we would like to present in depth, we consider the three charts giving tweet hit counts by question for users participating at an activity level of at least 17 tweets. These are Figures 6.8, 6.9, and 6.10 which respectively represent tweets from (1) all users publishing at least this many tweets, (2) those users who have been identified a priori as being in the green category, and (3) those users identified as being in the denier category. Examining these charts, we see a natural division in question popularity, which we have chosen as our selection criterion. On all three charts, after seven questions there is a visible drop in the tweet hit count. In the chart of tweets from all the high-activity users and the chart for those in the green group, the seven questions are the same (though the order varies slightly). Regarding the chart for tweets from users in the denier group, two additional questions are popular, replacing two slots in the lineup of top questions. Combining these rankings from all the charts gives us nine questions which we analyze further in depth in this section.

With the individual analysis of each these questions, we present two sets of affect signatures for the green group, the denier group, and both groups combined. The first set of signatures represent the tweets linked to the question, just as we presented in the previous section. The second set of affect signatures is more encompassing. It represents all the tweets published by users who have at least one tweet linked to the question. In essence we are considering a subcommunity of users who have touched the question and are creating a signature representing the affect in all the microblogs with the hashtag *#globalwarming* from that subcommunity (whether the texts refer to the question or not). Note that we only create affect signatures using the NRC Word-Emotion Association Lexicon (NRC-10) in this section.[20]

6.3.1 Question T22: "Outcome Expectations"

Question T22 is by far the most popular question among users in the green category (Figure 6.9) and the second most popular for those in the denier category (Figure 6.10). In the Six Americas survey T22 has two parts. The first asks the subject whether or not she believes that humans will be sufficiently motivated and capable to do what is necessary to mitigate global warming. She may also answer that there is no global warming (thereby implying there is nothing to be done). The second part asks how strongly she agrees or disagrees with the statement, "The actions of a single individual won't make any difference in global warming" (Leiserowitz et al., 2010).

---

[20]As previously mentioned, we recognize that the lack of close accordance between NRC-10 and Liu's Opinion Lexicon indicate that further work is necessary to determine just how appropriate the NRC-10 is for analyzing microblogs on climate change. Continued research towards this end will serve to either validate or improve upon the results we present in the present document.

Affect Signatures

Table 6.11 gives the affective word count and percentages of expressed affect for question T22 for tweets from users in the green and denier groups as well as both groups combined ("All") for reference. Figure 6.23 presents the associated affect signatures using the data from the table. The numbers inside the bars give the raw word count for the associated affective characteristic. Note that we cannot compare the word counts directly as the sizes of the groups differ significantly. Instead we are using the size of the bar subdivisions, which represent the percentages of affective characteristics with respect to the total expressed affect.

Table 6.11 Affect per NRC-10 lexicon for tweets linked to question T22.

	Greens		**Deniers**		**All**	
Users	75		42		117	
Anger	404	7.1%	67	9.3%	471	7.3%
Fear	678	11.8%	65	9.0%	743	11.5%
Sadness	428	7.5%	63	8.7%	491	7.6%
Joy	360	6.3%	44	6.1%	404	6.3%
Surprise	276	4.8%	39	5.4%	315	4.9%
Anticipation	626	10.9%	58	8.0%	684	10.6%
Disgust	276	4.8%	42	5.8%	318	4.9%
Trust	613	10.7%	87	12.0%	700	10.8%
Positive	1,175	20.5%	112	15.5%	1,287	19.9%
Negative	893	15.6%	147	20.3%	1,040	16.1%

This question does not invoke extreme differences in the signatures between the two groups. Regarding the sentiment polarity, we see the general trend observed previously where denier tweets tend to express more negative sentiment than green tweets. We do find that the denier tweets are expressing a bit more anger, compared to the green ones, and green tweets are showing more fear. Back in Section 3.4 we discussed how the emotions anger and fear are important research topics to the extent that affect is studied in recent research relating to climate change. Of course, we should also recall here that our results reported in Section 3.4.2 suggest that the emotions anger and fear expressed in tweets of high-activity users may to some extent be representative of the larger community. It could be worthwhile to repeat those experiments using the 17-tweet minimum dataset (perhaps focusing on specific questions from the Six Americas) as a future research endeavour.

We also find users in the green category expressing more anticipation for question T22. This emotion seems even less studied than anger and fear in relation to human beings and climate change. Research out of Greenland, where people are seeing drastic loss of both inland ice and sea ice, argues that anticipation should be a major focus for climate change research (Nuttall, 2010). Nuttall proposes this definition for anticipation: "the ways of making choices and decisions based on predictions, expectations or beliefs about the future." Although this definition may take anticipation beyond what we are calling an emotion in this work, it nevertheless underlines why anticipation is an important consideration for climate change research.

Table 6.12 Affect per NRC-10 lexicon for all tweets from users linked to question T22.

	Greens		**Deniers**		**All**	
Users	75		42		117	
Anger	2,683	7.3%	746	10.3%	3,429	7.8%
Fear	4,413	12.0%	755	10.4%	5,168	11.8%
Sadness	2,442	6.7%	523	7.2%	2,965	6.7%
Joy	2,283	6.2%	427	5.9%	2,710	6.2%
Surprise	1,965	5.4%	475	6.5%	2,440	5.6%
Anticipation	3,742	10.2%	734	10.1%	4,476	10.2%
Disgust	1,798	4.9%	421	5.8%	2,219	5.1%
Trust	4,067	11.1%	794	10.9%	4,861	11.1%
Positive	7,671	20.9%	1,019	14.0%	8,690	19.8%
Negative	5,594	15.3%	1,377	18.9%	6,971	15.9%

The second set of affect signatures we consider for question T22 is presented in Figure 6.24. The percentages shown in the signatures are listed in Table 6.12 along with the affective word counts. These signatures are created not only from the tweets which were linked to T22, but all of the *#globalwarming* tweets from the subcommunity of high-activity users on Twitter (17 tweets or more) who had one or more tweets associated with T22. This set of affect signatures gives us a view of the affect expressed in the general communications on global warming of people in the green and denier categories who may likely be concerned with whether or not humans will be able to mitigate the crisis. This includes people from the denier group who may potentially be arguing that

Figure 6.23 Affect per the NRC-10 lexicon for tweets on Question T22.

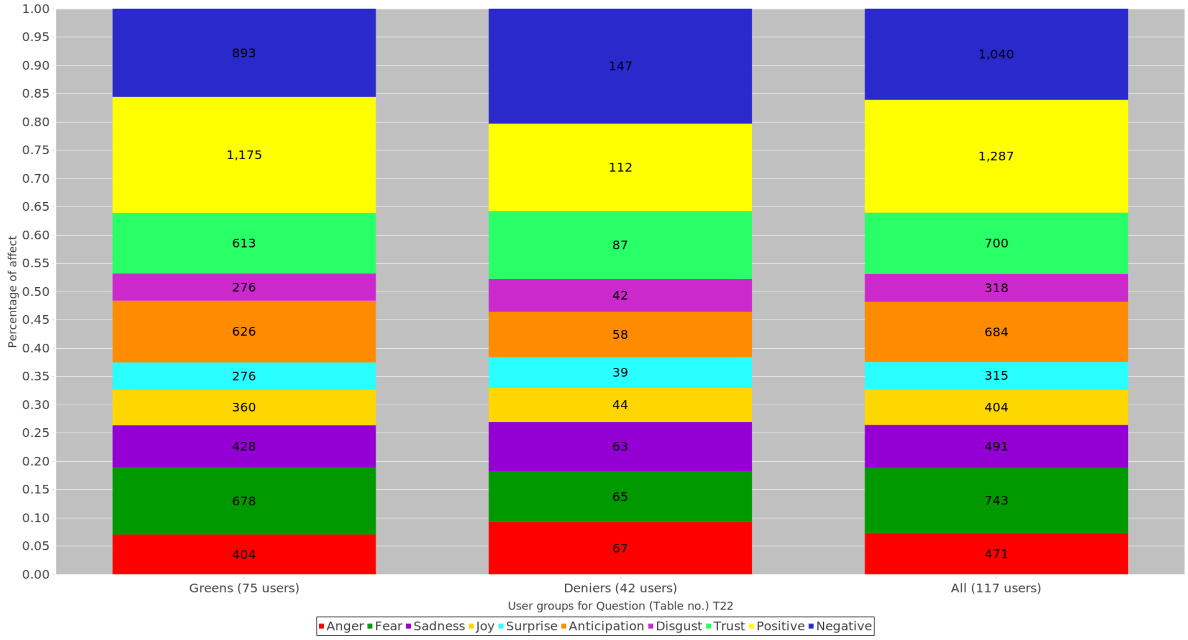
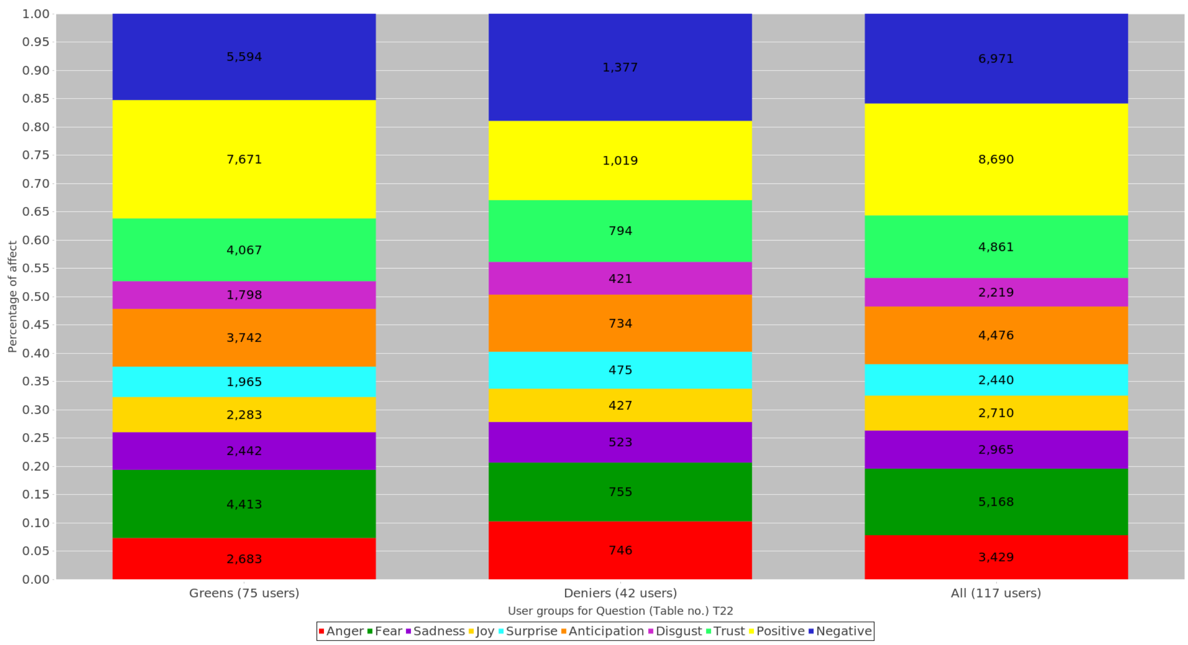


Figure 6.24 Affect per the NRC-10 lexicon for users tweeting about Question T22.



there is no climate crisis to begin with as well as those who are uncertain. These microblogs may also include a debate on whether the actions of one person can even make a difference.

In these affect signatures for the T22 subcommunity we see a similar result where the denier group is expressing more anger while the green group is expressing more fear. However, we see less of a distinction now for anticipation, which users in the denier group are expressing more in their general tweets on *#globalwarming*.

6.3.2 Question T27: "Attention and Response to IPCC Errors"

Question T27 is the second most popular question for users in the green group (Figure 6.9) but only the sixth most popular for users in the denier group (Figure 6.10). The question asks the subject if he is aware of any reports in the media concerning errors on the part of the IPCC. If so, T27 continues, asking the subject how much attention he is giving these reports and to what extent they have affected (1) how certain he is that global warming is happening [or not] and (2) the level of trust he places in climate scientists.

Affect Signatures

Figure 6.25 presents the affect signatures for question T27. The charted percentages are given in Table 6.13 for tweets from users in the green and denier groups as well as both groups combined. Again, the percentage of expressed negative sentiment and anger is greater in denier tweets than in green tweets. We also see a good deal of fear from users in the green category again. Words expressing fear, per the NRC-10, represent well over a quarter (28.1%) of the affective words in their tweets associated with T27.

The affect signatures representing *all* the tweets from users linked to question T27 are shown in Figure 6.26. The underlying data for these signatures is given in Table 6.14.

Figure 6.25 Affect per the NRC-10 lexicon for tweets on Question T27.

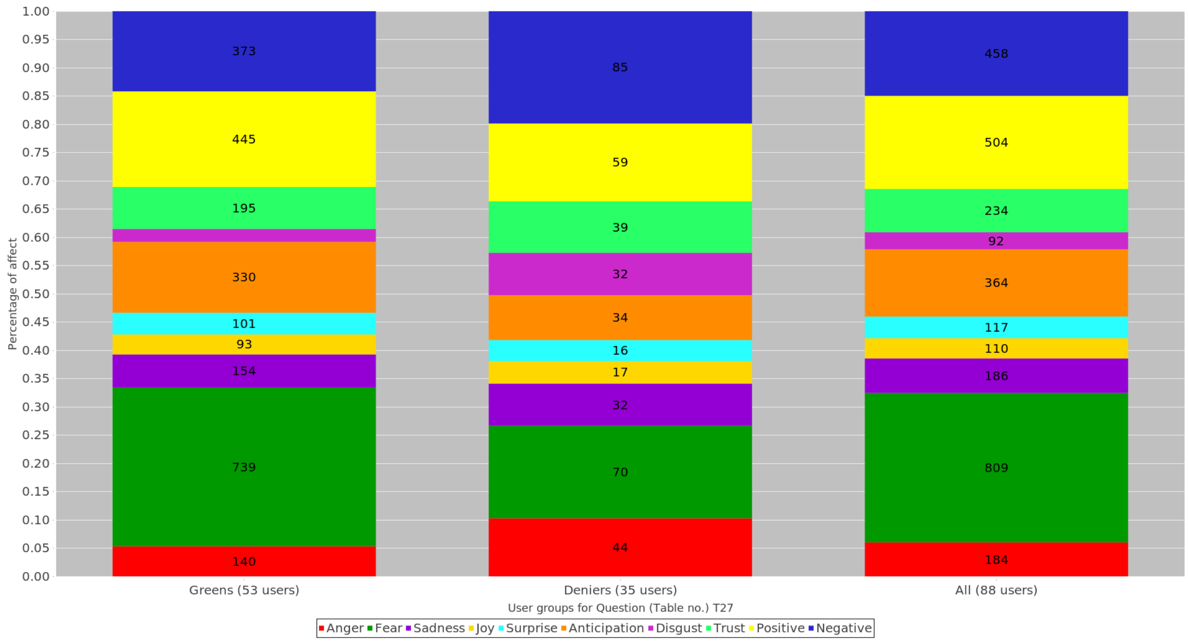


Figure 6.26 Affect per the NRC-10 lexicon for users tweeting about Question T27.

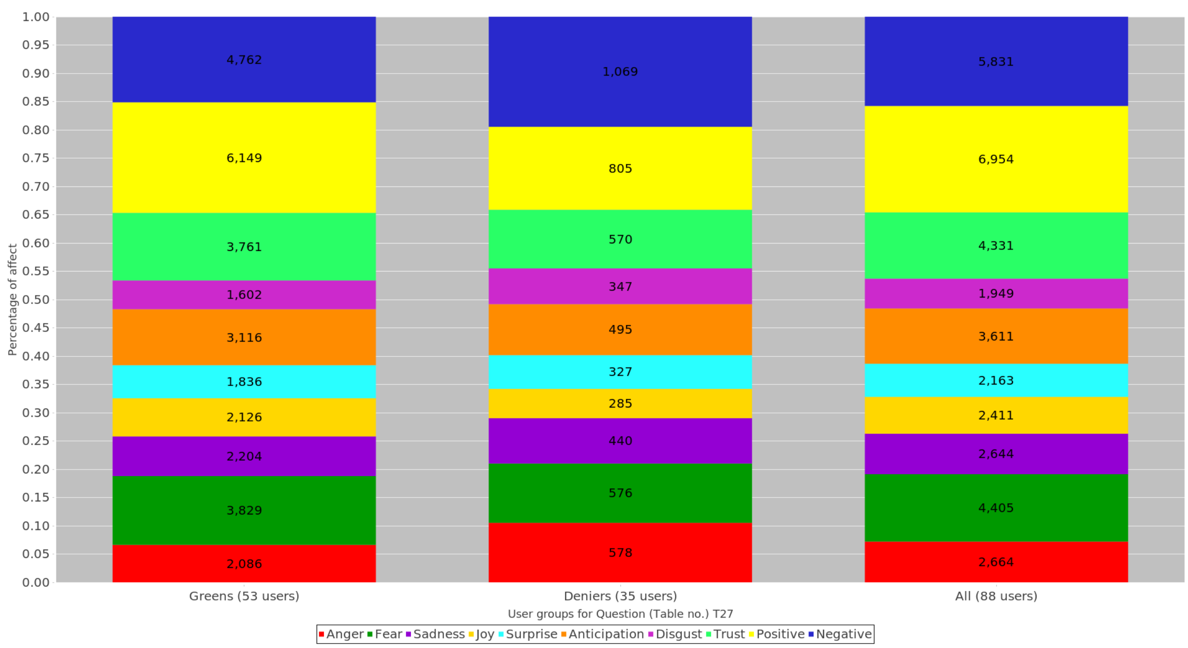


Table 6.13 Affect per NRC-10 lexicon for tweets linked to question T27.

	Greens		Deniers		All	
Users	53		35		88	
Anger	140	5.3%	44	10.3%	184	6.0%
Fear	739	28.1%	70	16.4%	809	26.5%
Sadness	154	5.9%	32	7.5%	186	6.1%
Joy	93	3.5%	17	4.0%	110	3.6%
Surprise	101	3.8%	16	3.7%	117	3.8%
Anticipation	330	12.5%	34	7.9%	364	11.9%
Disgust	60	2.3%	32	7.5%	92	3.0%
Trust	195	7.4%	39	9.1%	234	7.7%
Positive	445	16.9%	59	13.8%	504	16.5%
Negative	373	14.2%	85	19.9%	458	15.0%

Table 6.14 Affect per NRC-10 lexicon for all tweets from users linked to question T27.

	Greens		Deniers		All	
Users	53		35		88	
Anger	2,086	6.6%	578	10.5%	2,664	7.2%
Fear	3,829	12.2%	576	10.5%	4,405	11.9%
Sadness	2,204	7.0%	440	8.0%	2,644	7.2%
Joy	2,126	6.8%	285	5.2%	2,411	6.5%
Surprise	1,836	5.8%	327	6.0%	2,163	5.9%
Anticipation	3,116	9.9%	495	9.0%	3,611	9.8%
Disgust	1,602	5.1%	347	6.3%	1,949	5.3%
Trust	3,761	12.0%	570	10.4%	4,331	11.7%
Positive	6,149	19.5%	805	14.7%	6,954	18.8%
Negative	4,762	15.1%	1,069	19.5%	5,831	15.8%

We note immediately that the high percentage of fear in T27's green tweets drops to 12.2% when we consider all the tweets for these users. Other affective characteristics appear relatively in line between the chart considering tweets linked to T27 (Figure 6.25) and the general *#globalwarming* communications from these same users.

Reflections on T27

In addition to concerns about errors from the IPCC and perhaps other organizations reporting climate science, there is ongoing consideration regarding the statistical language used by the IPCC and the scientific community. Scientific findings are qualified in terms of *confidence* and *likelihood* in order to be explicit about the level of certainty scientists place in an analysis produced by a given probabilistic climate model. (Herrando-Pérez et al., 2019) argue that this type of language can be counterproductive and that public understanding of the level of scientific consensus surrounding climate change is undermined due to a combination of the complexity of the science, the frequency of findings reported with medium or low confidence, and deliberate misinformation campaigns.

Regarding the idea of errors, however, in the present analysis in which we are simply linking tweets to the various Six Americas survey questions, T27 presents an opportunity for a better understanding not only of the limitations imposed by our methodology using Lucene, but more generally of one of the ever-present challenges in the domain of NLP. When a human subject reads the four-part question, which covers a full page in (Leiserowitz et al., 2010) with its multiple choice answers for each part, he automatically understands the sense as we described it above: *do errors from the experts* ***influence*** *his belief in global warming and his trust in climate scientists?* The concept of an error (perhaps extended to uncertainty) is what is key here. With this in mind, consider this microblog from the 17-tweet minimum dataset, which Lucene links to this question:

`Influencers influencing. [web-link] #GlobalWarming`

Lucene, functioning according to an algorithm like Okapi BM25 or TF*IDF, is essentially trying to find a hit which maximizes the number of matches of uncommon words between a tweet and a question while discounting matches of words that are common across the full corpus of questions. While this methodology has proven quite effective in research (Robertson & Zaragoza, 2009) and in practice (Inkpen, 2006), we can see it working as intended, yet still somewhat missing the mark. Only one other question (T26) includes a form of the word "influence."[21] Should this tweet have been linked to T27? We believe there is an argument for both sides. Even if it misses the key point of the question, the tweet still hits part of the intent of the question. What is needed in terms of the language

[21]Using the Porter stemmer (Porter, 2006), the system considers any word which reduces to the basic stem *influenc* as a match.

understanding required for a “true hit” goes beyond the level of IR sophistication we are working with here. Certainly we could try to improve our accuracy by extending the methodology Lucene is employing to give more weight to specific concepts that are important for our particular domain–the various questions in the Six Americas survey. Of course, accomplishing this feat for generic processing in a variety of applications requires a strategy for general, common sense knowledge of agency, psychology, and the physical workings of the natural world. It remains one of the core challenges in AI (Lake et al., 2017).

6.3.3 Question T25: “Support for National Response: Specific Climate and Energy Policies Priority”

The third most popular question for both the green and the denier groups is T25. The question asks the subject how much she “supports or opposes” a number of energy policies ranging from funding for research into renewables, to international treaties intended to cut emissions, to a gasoline tax offset by reductions in basic income tax (Leiserowitz et al., 2010). The policies are numerous, covering three pages of the document. This is likely a contributing factor towards this question’s popularity since the policy list contains many concept-oriented words that would not appear in many other questions. For example, this is the only question in the survey with the words “renewable” and “solar.” A question such as T25 is therefore ideal for the present analysis in that standard IR techniques are likely to correctly identify texts pertaining to the climate-based policies listed in T25. By contrast, the words “support” and “oppose” are used throughout the Six Americas questionnaire. We should therefore expect our analysis to target communications about the policies themselves, rather than indications of a level or support or opposition.

Affect Signatures

Figure 6.27 presents the affect signatures for tweets linked to question T25. The percentages for the affect characteristics which form these signatures are shown in Table 6.15 along with the raw word counts for reference. Interestingly, question T25 shows a reversal of the common trend that denier tweets tend to express more negativity than green tweets in the 17-tweet minimum dataset (compare Figures 6.14 and 6.16 as well as Figures 6.20 and 6.22). For this question it is the green users who are tweeting with increased negative sentiment. Recall that even though Liu's Opinion Lexicon did not show the same tendency for green tweets to be generally more positive than negative as we saw with NRC-10, both lexica show denier tweets to be generally more negative than green tweets. For question T25, however, our results indicate the opposite. We also note that once again we have a higher level of anger in tweets from users in the denier group; however, the difference is not as pronounced as it was in question T27.

Table 6.15 Affect per NRC-10 lexicon for tweets linked to question T25.

	Greens		**Deniers**		**All**	
Users	67		32		99	
Anger	174	6.7%	60	9.6%	234	7.3%
Fear	212	8.2%	56	9.0%	268	8.4%
Sadness	180	7.0%	57	9.2%	237	7.4%
Joy	175	6.8%	39	6.3%	214	6.7%
Surprise	177	6.8%	43	6.9%	220	6.9%
Anticipation	167	6.5%	37	5.9%	204	6.4%
Disgust	210	8.1%	41	6.6%	251	7.8%
Trust	297	11.5%	69	11.1%	366	11.4%
Positive	464	17.9%	114	18.3%	578	18.0%
Negative	530	20.5%	106	17.0%	636	19.8%

The affect signatures for all the tweets from users linked to question T25 are shown in Figure 6.28, whose underlying data is in Table 6.16. In the general *#globalwarming* communications for these users, we see a return to the trend of increased negative sentiment in denier tweets as compared to green tweets. Determining the reason for this shift in sentiment polarity would be an interesting direction for continued research. Less surprisingly, anger continues to be higher for denier tweets, and we return to an

Figure 6.27 Affect per the NRC-10 lexicon for tweets on Question T25.

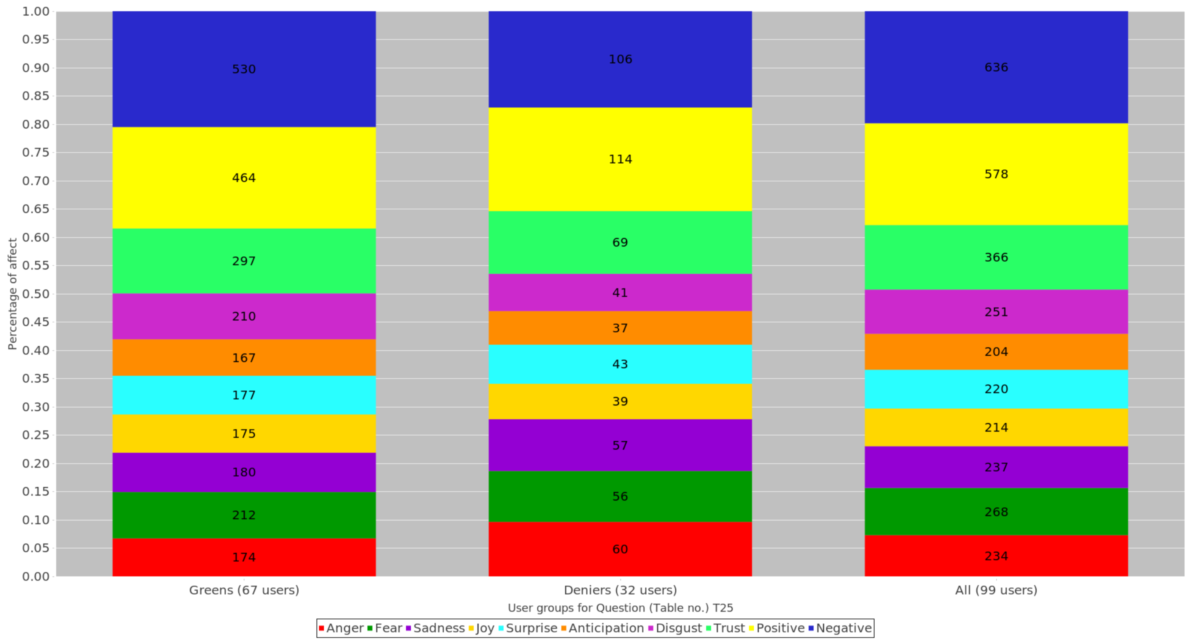


Figure 6.28 Affect per the NRC-10 lexicon for users tweeting about Question T25.

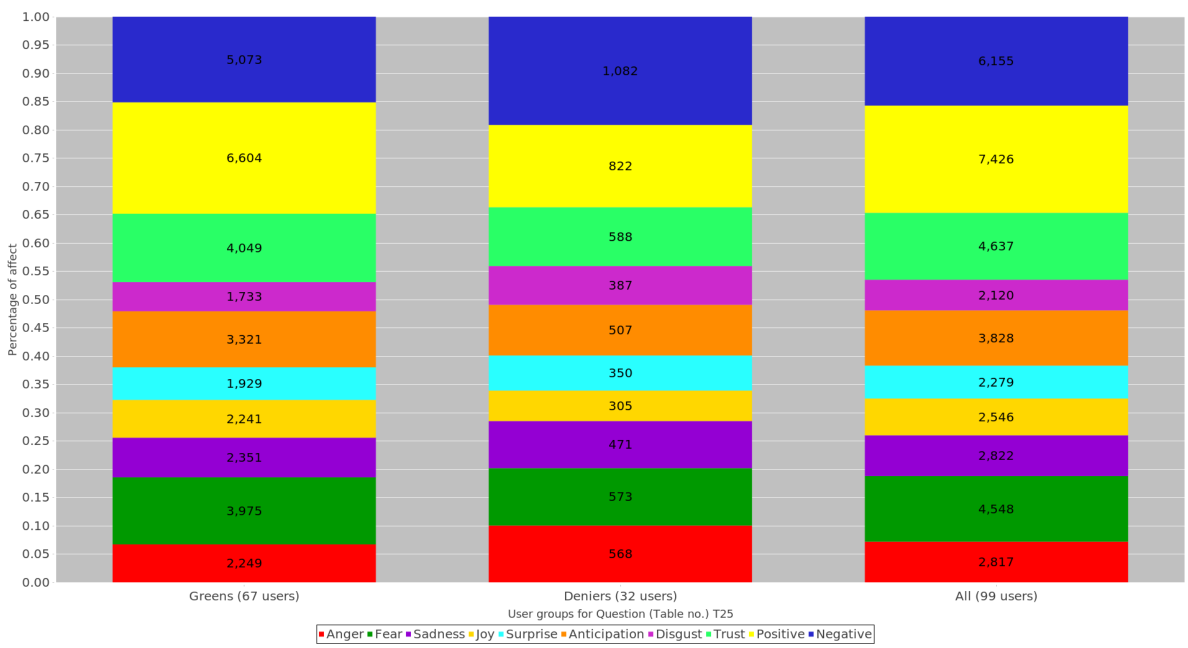


Table 6.16 Affect per NRC-10 lexicon for all tweets from users linked to question T25.

	Greens		**Deniers**		**All**	
Users	67		32		99	
Anger	2,249	6.7%	568	10.0%	2,817	7.2%
Fear	3,975	11.9%	573	10.1%	4,548	11.6%
Sadness	2,351	7.0%	471	8.3%	2,822	7.2%
Joy	2,241	6.7%	305	5.4%	2,546	6.5%
Surprise	1,929	5.8%	350	6.2%	2,279	5.8%
Anticipation	3,321	9.9%	507	9.0%	3,828	9.8%
Disgust	1,733	5.2%	387	6.8%	2,120	5.4%
Trust	4,049	12.1%	588	10.4%	4,637	11.8%
Positive	6,604	19.7%	822	14.5%	7,426	19.0%
Negative	5,073	15.1%	1,082	19.1%	6,155	15.7%

increased level of fear for tweets from users in the green category.

Reflections on T25

Studying social media communications on climate policy from people who tend to be contrarian towards climate science can potentially reap benefits with respect to considerations on how best to frame information about climate change for certain audiences. For example, (Maibach et al., 2009) found that people in the Six Americas' *dismissive* segment (which we map to our denier category along with the *doubtful* segment) would respond to appeals for efficient use of energy even as they were likely to reject findings on climate change from the scientific community. Furthermore, studies based on the Six Americas are finding that that people in the U.S. tend to be more supportive of research in renewables than they are of policies regulating emissions such as a carbon tax (Ballew et al., 2019).

6.3.4 Question T3: "Questions About Global Warming"

Question T3 has two parts. In the first part it has the subject suppose that he is able to ask an expert about global warming and inquires as to what he would ask. The possible choices start with "Is global warming really happening?" and continue with questions such as how does the expert know; what are the causes; what harm will come; what can be done in the U.S.; and what can the subject do personally? There is also a choice for "other." He is to check any questions he would like to ask; however, in the second part of T3, he must indicate which of these same questions he would ask the expert if he is limited to only one (Leiserowitz et al., 2010).

Affect Signatures

Figure 6.29 presents the affect signatures for tweets linked to question T3 for users in the green and denier categories as well as both groups combined. The percentages of expressed affect and the raw counts of affective words are given in Table 6.17. These affect signatures demonstrate the same trends we have seen in previous questions, namely increased anger and negative sentiment in the denier tweets.

Table 6.17 Affect per NRC-10 lexicon for tweets linked to question T3.

	Greens		**Deniers**		**All**	
Users	61		33		94	
Anger	185	5.9%	60	12.6%	245	6.8%
Fear	296	9.4%	31	6.5%	327	9.1%
Sadness	168	5.4%	17	3.6%	185	5.1%
Joy	210	6.7%	44	9.3%	254	7.0%
Surprise	224	7.1%	18	3.8%	242	6.7%
Anticipation	256	8.2%	38	8.0%	294	8.1%
Disgust	122	3.9%	14	2.9%	136	3.8%
Trust	497	15.8%	60	12.6%	557	15.4%
Positive	787	25.1%	96	20.2%	883	24.5%
Negative	391	12.5%	97	20.4%	488	13.5%

Figure 6.30 has the corresponding affect signatures for *all* the tweets from users with at least one tweet linked with question T3. The underlying data for these signatures is presented in Table 6.18. The trends for increased expressed anger and negative sentiment in denier tweets continue here in the general *#globalwarming* communications for these users, suggesting that opinions associated specifically with T3 are not likely a direct cause for the increases. Comparing Figures 6.29 and 6.30, however, we do see that the level of disgust drops when considering only the tweets linked with T3. It could be

Figure 6.29 Affect per the NRC-10 lexicon for tweets on Question T3.

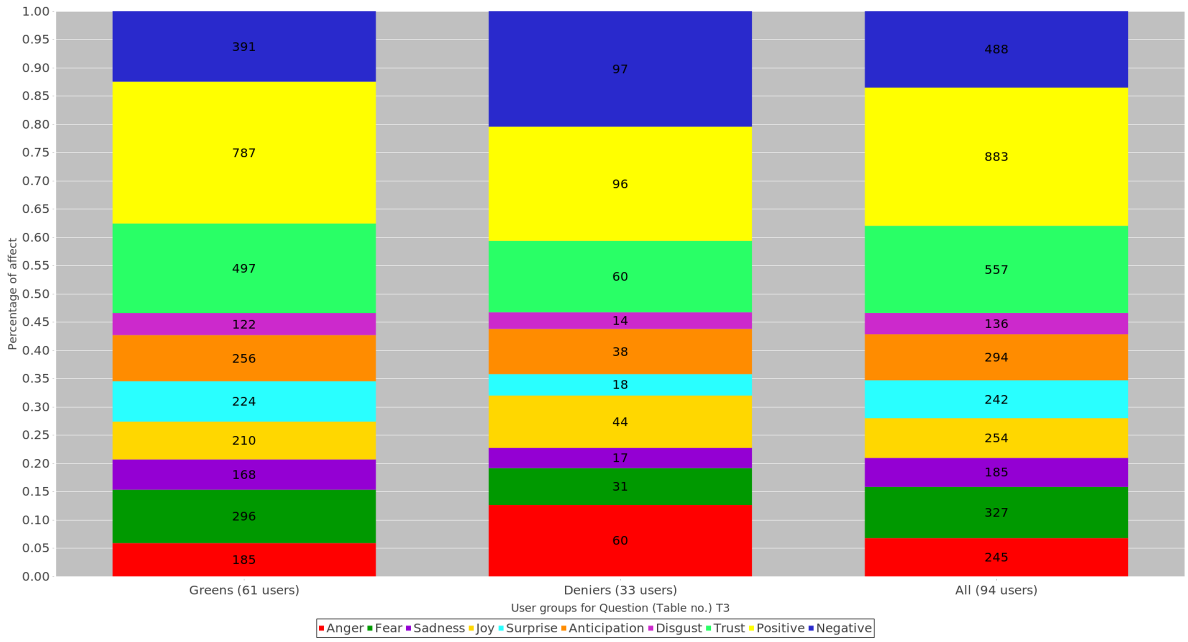


Figure 6.30 Affect per the NRC-10 lexicon for users tweeting about Question T3.

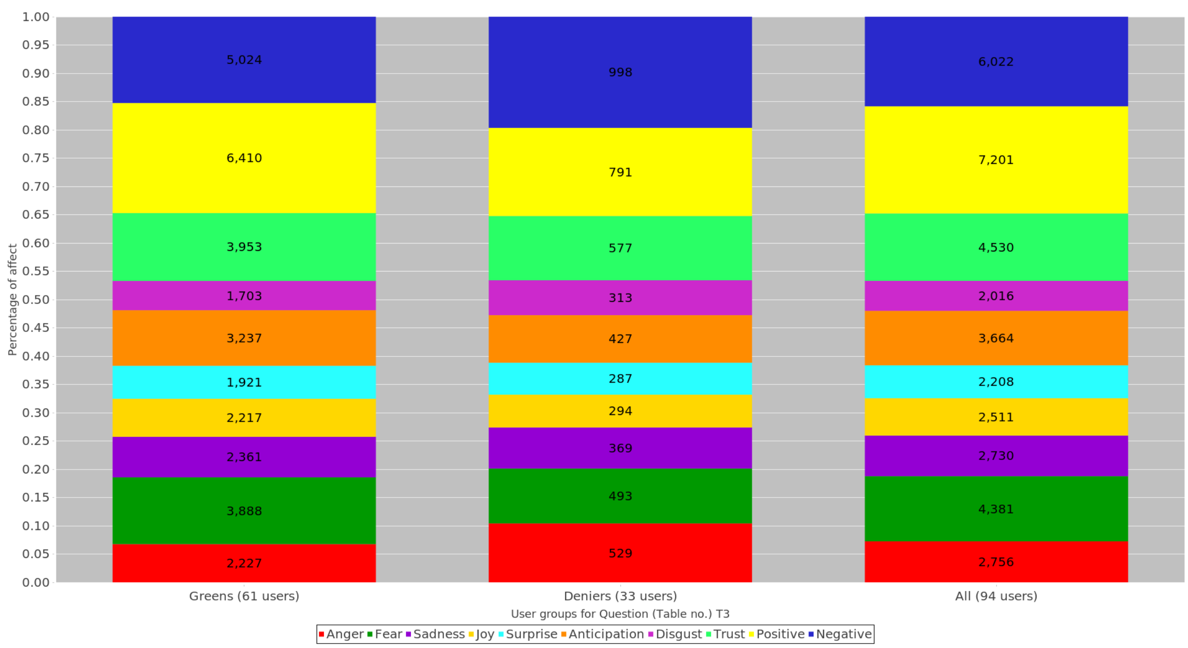


Table 6.18 Affect per NRC-10 lexicon for all tweets from users linked to question T3.

	Greens		**Deniers**		**All**	
Users	61		33		94	
Anger	2,227	6.8%	529	10.4%	2,756	7.2%
Fear	3,888	11.8%	493	9.7%	4,381	11.5%
Sadness	2,361	7.2%	369	7.3%	2,730	7.2%
Joy	2,217	6.7%	294	5.8%	2,511	6.6%
Surprise	1,921	5.8%	287	5.7%	2,208	5.8%
Anticipation	3,237	9.8%	427	8.4%	3,664	9.6%
Disgust	1,703	5.2%	313	6.2%	2,016	5.3%
Trust	3,953	12.0%	577	11.4%	4,530	11.9%
Positive	6,410	19.5%	791	15.6%	7,201	18.9%
Negative	5,024	15.3%	998	19.7%	6,022	15.8%

interesting to explore this finding further, given that we do not see many large changes in the level of disgust expressed among the more popular questions.

Reflections on T3

The Six Americas is an ongoing project, and so research teams have been able to track how people in the United States relate to issues concerning climate change from the first experimental runs in 2008 to present day. Interestingly, the percentage of people in the U.S. who believe global warming is happening dropped from 71% in 2008 to 59% in 2010 and since then has been rising again, back to 71% in more recent studies (Ballew et al., 2019; Leiserowitz et al., 2016). In the 17-tweet minimum dataset we identified approximately two thirds of the users as being in the green category (mapping to the *alarmed* and *concerned* segments of the Six Americas). Though this ratio is not too far from the 71% found in these recent studies, we must still assume that the demographics of the community on Twitter are skewed to some extent with respect to the subjects involved with experiments in the Six Americas project. We discuss this limitation to our methodology further in Section 6.4.1, but we should keep in mind here that even though our model is based on the Six Americas, the analysis pertains to an online community rather than the population of the United States.

Just as the knowledge in the U.S. that climate change is occurring dropped after 2008, only to rise back up to about the same level according to more recent Six Americas studies, so did two related issues, both of which are covered by the set of questions presented to the subject in T3. One is the understanding that humans are causing the earth to warm (58% in 2008, down to 48% in 2010, and gradually back up to 56% in 2017). The other is that there is a scientific consensus on these facts (46% in 2008, down to 33% in 2010, and gradually up again to 53% in 2017). A better understanding of these types of trends can serve to refine efforts to educate the public, especially in the younger generation who generally talk less about climate change and hear about it less often in the media (Ballew et al., 2019). Though it may be more difficult to create a representative sample when collecting online data than it is with a formal survey, the demographics of the online community (e.g., younger users) may be favourable when targeting certain groups with outreach efforts.

6.3.5 Question T16: "Political Activism"

Question T16 is the fifth most popular question overall in the 17-tweet minimum dataset. It is the fourth most popular question among users in the green category but only the eleventh most popular among users in the denier category. This question also has two parts. The first part asks the subject how many times in the last 12 months she has (1) donated her time or money to an organization working to mitigate global warming, (2) posted comments to an online news or a blog post related to global warming, or (3) contacted a government official to push for [or against] action on global warming. In the second part, question T16 asks if she is likely to take these same actions more often, less often, or with about the same frequency over the coming 12 months (Leiserowitz et al., 2010).

Affect Signatures

The affect signatures for T16 are presented in Figure 6.31 for the green and denier groups and for both groups combined. The data for the chart is given in Table 6.19. The usual trend of increased negative sentiment is again present for denier tweets, but not the trend of an increase in expressed anger. For question T16 the associated green tweets have a slightly higher level of anger.

Table 6.19 Affect per NRC-10 lexicon for tweets linked to question T16.

	Greens		**Deniers**		**All**	
Users	42		28		70	
Anger	490	11.3%	34	10.1%	524	11.2%
Fear	465	10.7%	32	9.6%	497	10.6%
Sadness	95	2.2%	28	8.4%	123	2.6%
Joy	224	5.1%	15	4.5%	239	5.1%
Surprise	118	2.7%	21	6.3%	139	3.0%
Anticipation	675	15.5%	36	10.7%	711	15.2%
Disgust	65	1.5%	19	5.7%	84	1.8%
Trust	246	5.7%	41	12.2%	287	6.1%
Positive	1,450	33.3%	39	11.6%	1,489	31.8%
Negative	525	12.1%	70	20.9%	595	12.7%

Table 6.20 Affect per NRC-10 lexicon for all tweets from users linked to question T16.

	Greens		**Deniers**		**All**	
Users	42		28		70	
Anger	2,377	7.1%	576	10.0%	2,953	7.5%
Fear	3,950	11.8%	587	10.2%	4,537	11.5%
Sadness	2,191	6.5%	398	6.9%	2,589	6.6%
Joy	2,099	6.3%	351	6.1%	2,450	6.2%
Surprise	1,836	5.5%	375	6.5%	2,211	5.6%
Anticipation	3,489	10.4%	583	10.1%	4,072	10.4%
Disgust	1,597	4.8%	339	5.9%	1,936	4.9%
Trust	3,800	11.3%	654	11.4%	4,454	11.3%
Positive	7,206	21.5%	816	14.2%	8,022	20.4%
Negative	5,033	15.0%	1,076	18.7%	6,109	15.5%

When considering *all* the tweets from users linked to T16, however, we see that denier tweets are expressing more anger, once again following the general trend as shown in Figure 6.32. The data for these affect signatures are listed in Table 6.20. Additionally, we see the level of disgust drop for green tweets linked to question T16 with respect to the

Figure 6.31 Affect per the NRC-10 lexicon for tweets on Question T16.

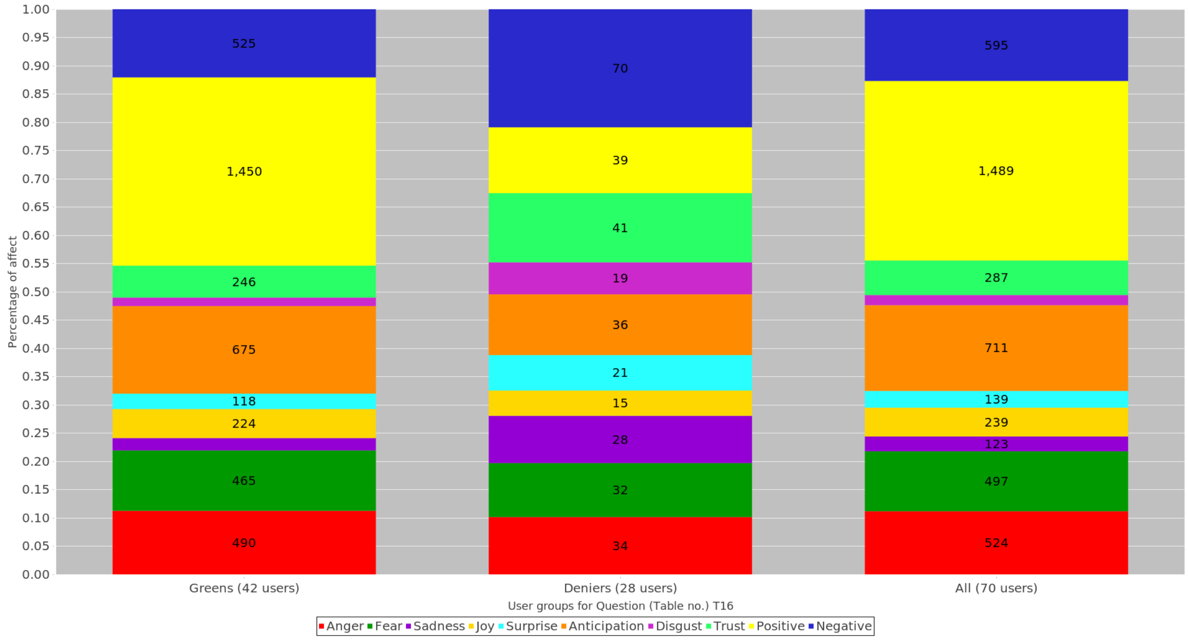
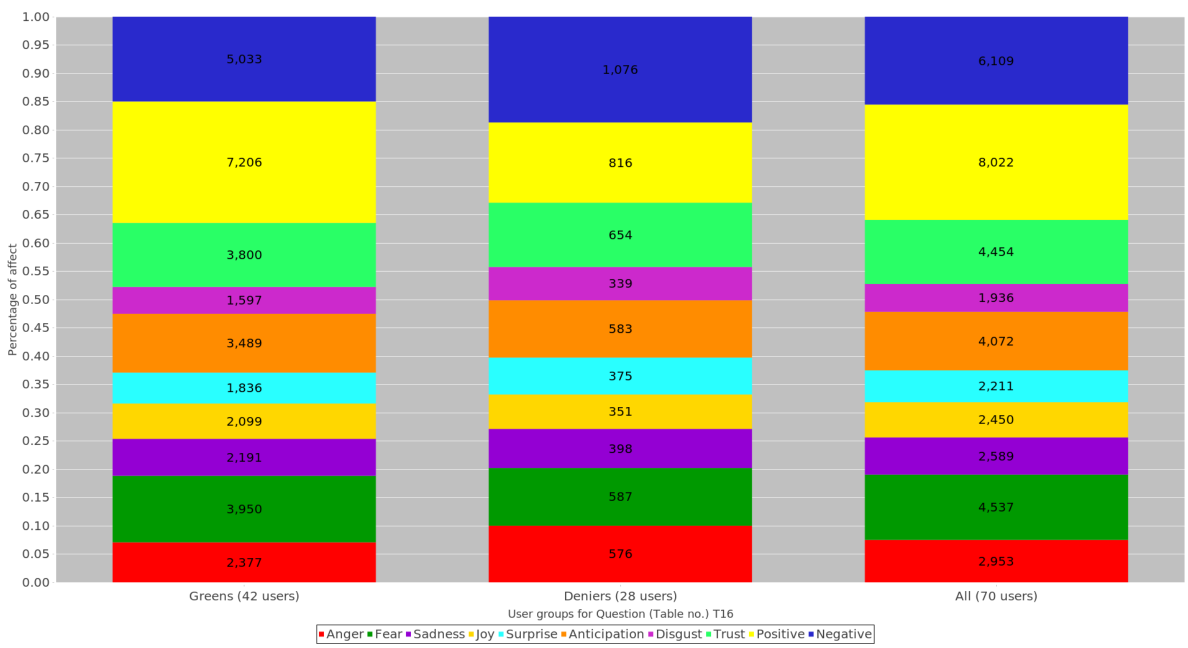


Figure 6.32 Affect per the NRC-10 lexicon for users tweeting about Question T16.



general *#globalwarming* communications for these users. We noted a similar result for question T3 back in Section 6.3.4. Furthermore, users in the green group are expressing more anticipation but less sadness and trust per the NRC-10 lexicon for tweets linked to T16 as compared to the general *#globalwarming* discourse of the users who published those tweets.

Reflections on T16

The link between political affiliation and beliefs concerning climate change in the United States is clear. While Democrats, the party leaning left in the U.S., often consider climate change to be the most important environmental problem facing the country (Bohr, 2014; McCright & Dunlap, 2010), Republicans, the party leaning right, largely maintain a point of view of denial. Moreover, (McCright et al., 2016) argue that Republicans promote denial of anthropogenic climate change more than any other political party in the world. However, users on Twitter tend to be younger with respect to the general population (Mellon & Prosser, 2017; Wojcik & Hughes, 2019), and studies have shown that younger Republicans are more likely to respond positively concerning climate-related issues than older Republicans. Efforts to model and analyze public communications about political engagement on climate issues for online communities such as those linked to question T16 could prove invaluable, especially considering that younger people in general are having relatively fewer conversations about climate change (Ballew et al., 2019).

6.3.6 Question T9: "Perceptions of Weather and Climate"

T9 is an interesting question in terms of popularity. It ranks sixth overall, seventh for users in the green category, but it is #1 among users in the denier category. It first asks the subject if he has personally felt any of the impacts of a changing climate and then continues, asking specifically if recent, severe snowstorms cause him to doubt the existence of global warming and how the winter in his area compared to previous years with respect to temperature, snowfall, and rainfall (Leiserowitz et al., 2010).

Affect Signatures

Figure 6.33 shows the affect signatures for question T9 for associated tweets from the green and denier groups as well as both groups together. Table 6.21 has the underlying percentages and raw affective word counts for these signatures. Again, we note increased negative sentiment and anger in the denier tweets; however, we also see increased fear on the denier side. For the majority of questions it is users in the green category who are publishing tweets with higher levels of fear.

Table 6.21 Affect per NRC-10 lexicon for tweets linked to question T9.

	Greens		**Deniers**		**All**	
Users	47		38		85	
Anger	78	5.4%	51	10.3%	129	6.6%
Fear	117	8.0%	56	11.3%	173	8.9%
Sadness	98	6.7%	46	9.3%	144	7.4%
Joy	80	5.5%	27	5.5%	107	5.5%
Surprise	32	2.2%	31	6.3%	63	3.2%
Anticipation	122	8.4%	45	9.1%	167	8.6%
Disgust	51	3.5%	24	4.9%	75	3.8%
Trust	267	18.3%	40	8.1%	307	15.7%
Positive	386	26.5%	66	13.4%	452	23.2%
Negative	225	15.5%	108	21.9%	333	17.1%

Table 6.22 Affect per NRC-10 lexicon for all tweets from users linked to question T9.

	Greens		**Deniers**		**All**	
Users	47		38		85	
Anger	2,094	6.7%	743	10.6%	2,837	7.4%
Fear	3,747	12.0%	730	10.4%	4,477	11.7%
Sadness	2,202	7.1%	524	7.5%	2,726	7.1%
Joy	2,077	6.7%	398	5.7%	2,475	6.5%
Surprise	1,821	5.9%	460	6.6%	2,281	6.0%
Anticipation	3,066	9.9%	738	10.5%	3,804	10.0%
Disgust	1,598	5.1%	423	6.0%	2,021	5.3%
Trust	3,707	11.9%	728	10.4%	4,435	11.6%
Positive	6,096	19.6%	960	13.7%	7,056	18.5%
Negative	4,716	15.2%	1,318	18.8%	6,034	15.8%

Figure 6.33 Affect per the NRC-10 lexicon for tweets on Question T9.

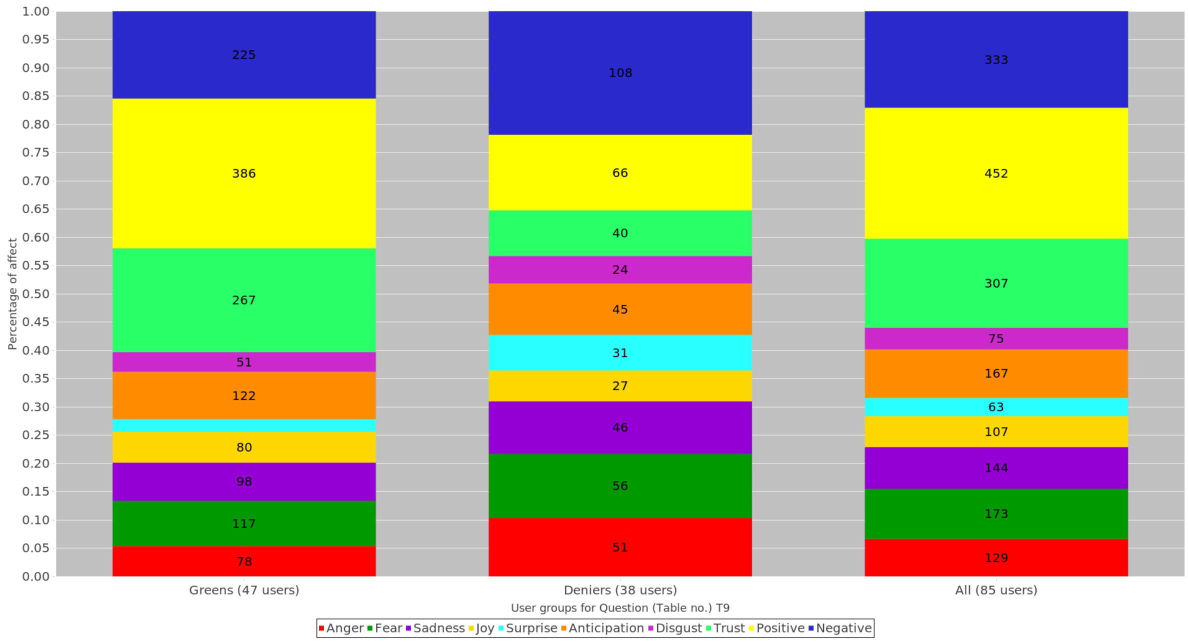
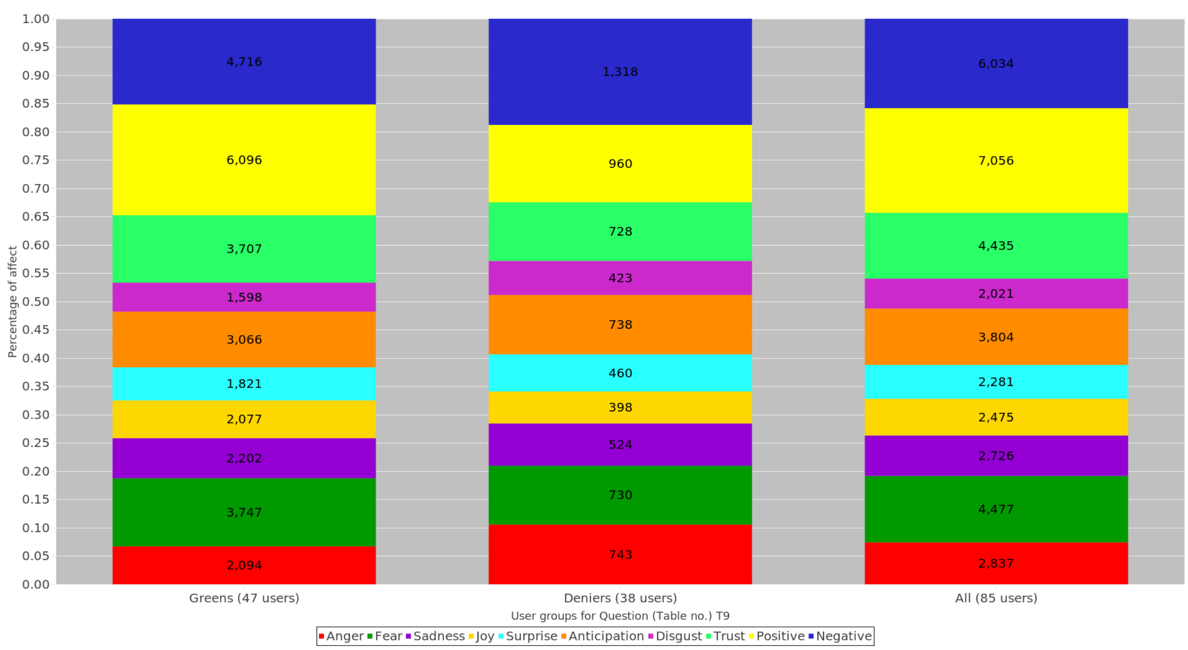


Figure 6.34 Affect per the NRC-10 lexicon for users tweeting about Question T9.



The affect signatures for *all* the tweets from users who have a tweet linked to T9 are shown in Figure 6.34. Table 6.22 gives the data for this chart. Interestingly, the green tweets are now expressing a higher level of fear, more along the lines we have come to expect. This result indicates a path for further research on why green tweets associated with climate perceptions, such as those linked to T9, may be less fearful or possibly less fear-invoking, while denier tweets may show higher levels of expressed fear. Comparing Figures 6.33 and 6.34, we also note a lower level of surprise for green tweets linked to T9 than for the full set of green tweets comprising the general *#globalwarming* communications for users with a tweet linked to this question. We also see that the level of disgust drops for the tweets linked to T9, but the difference is less remarkable than we have seen for other questions.

Reflections on T9

If we consider the relationship between the perception of a changing climate and the perception of the risk it imposes, we see that we are touching a key issue with question T9. Recent studies indicate that approximately half of the people in the United States do not see climate change as a risk to themselves personally (Ballew et al., 2019). Recalling our previous analysis on members of the Republican party in the U.S. (see Section 6.3.5), (Botzen et al., 2016) report that members of this party readily reject probable links to climate when hearing about severe floods and other ecological disasters.

More generally, however, psychological studies reveal that human beings are not particularly well equipped to fully comprehend the type of emergency that climate change represents. According to the dual process theory of cognition (Epstein, 1994; Evans & Stanovich, 2013), our affective system (S1 processes) responds to a threat with an emotion like fear, serving to motivate a person to get out of a dangerous situation. Yet, complications arise because S1 processes generally function in response to a known and immediate danger. The less a threat is inherently understood and the further out in time it may occur, the more the brain relies on its analytical system (S2 processes). When the threat is climate change, the less a person has had personal and recent experience with droughts, hurricanes, wild fires, or other severe effects of global warming, the more his S2 processes will be assessing the level of personal risk it poses. The problem with a severe but abstract and long-term danger like climate change is that when the two cognitive systems are in conflict, it is usually the S1 processes that dominate the decisions the person makes and the actions he will take. Even so, the emotion associated with the affective system (S1 processes) can work to guide the analytical reasoning in the S2 processes (Weber, 2006). One might say that the affective interactions between the two cognitive systems effectively represent a rather tricky cognitive puzzle, and research is needed to solve it. As we noted above, our results for question T9 show an increased level of expressed fear from users in the denier group in their tweets linked to this question. We propose that continuing this research and using this type of affective analysis on social media may help lead to a better understanding of this puzzle and generate insight into how emotion may play an important part in clear and effective reasoning about the perception of climate change as a serious and personal threat.

6.3.7 Question T8b: "Risk Perceptions: When Harm Will Occur"

Question T8b is the seventh most popular question overall. If we consider the green and denier categories together, then this is the last of the popular questions before there is a dip in associated tweet counts from the 17-tweet minimum dataset (see Figure 6.8). For users in the green category T8b is the sixth most popular question. In the denier category it is the eighth most popular. Question T8b has two parts. In the first part T8b asks the subject when she thinks global warming is going to begin harming people in the U.S. The multiple choice answers are now, in 10, 25, 50, or 100 years, or never. In the second part she is asked the same question, but for people across the world instead of just in the U.S.

Affect Signatures

Figure 6.35 presents the affect signatures for question T8b for tweets from users in the green group, the denier group, and both groups combined. Table 6.23 displays the percentages for the chart as well as the corresponding affective word counts. The usual trends for increased levels of expressed anger and negative sentiment in denier tweets are particularly strong for this question. The somewhat less common trend for increased fear in green tweets is not present for T8b. There is, however, a relatively high level of trust in the green tweets.

Table 6.23 Affect per NRC-10 lexicon for tweets linked to question T8b.

	Greens		**Deniers**		**All**	
Users	55		26		81	
Anger	66	5.0%	29	13.1%	95	6.2%
Fear	102	7.8%	24	10.8%	126	8.2%
Sadness	104	8.0%	17	7.7%	121	7.9%
Joy	51	3.9%	9	4.1%	60	3.9%
Surprise	68	5.2%	9	4.1%	77	5.0%
Anticipation	126	9.6%	22	9.9%	148	9.7%
Disgust	86	6.6%	16	7.2%	102	6.7%
Trust	242	18.5%	22	9.9%	264	17.3%
Positive	273	20.9%	22	9.9%	295	19.3%
Negative	190	14.5%	52	23.4%	242	15.8%

Table 6.24 Affect per NRC-10 lexicon for all tweets from users linked to question T8b.

	Greens		**Deniers**		**All**	
Users	55		26		81	
Anger	2,148	6.6%	565	10.6%	2,713	7.2%
Fear	3,795	11.7%	563	10.6%	4,358	11.6%
Sadness	2,346	7.3%	363	6.8%	2,709	7.2%
Joy	2,175	6.7%	305	5.7%	2,480	6.6%
Surprise	1,914	5.9%	340	6.4%	2,254	6.0%
Anticipation	3,212	9.9%	592	11.1%	3,804	10.1%
Disgust	1,713	5.3%	282	5.3%	1,995	5.3%
Trust	3,855	11.9%	579	10.9%	4,434	11.8%
Positive	6,264	19.4%	716	13.5%	6,980	18.5%
Negative	4,931	15.2%	1,012	19.0%	5,943	15.8%

The affect signatures of *all* the tweets for users with at least one tweet linked to question T8b are shown in Figure 6.36. The underlying data for these signatures are listed in

Figure 6.35 Affect per the NRC-10 lexicon for tweets on Question T8b.

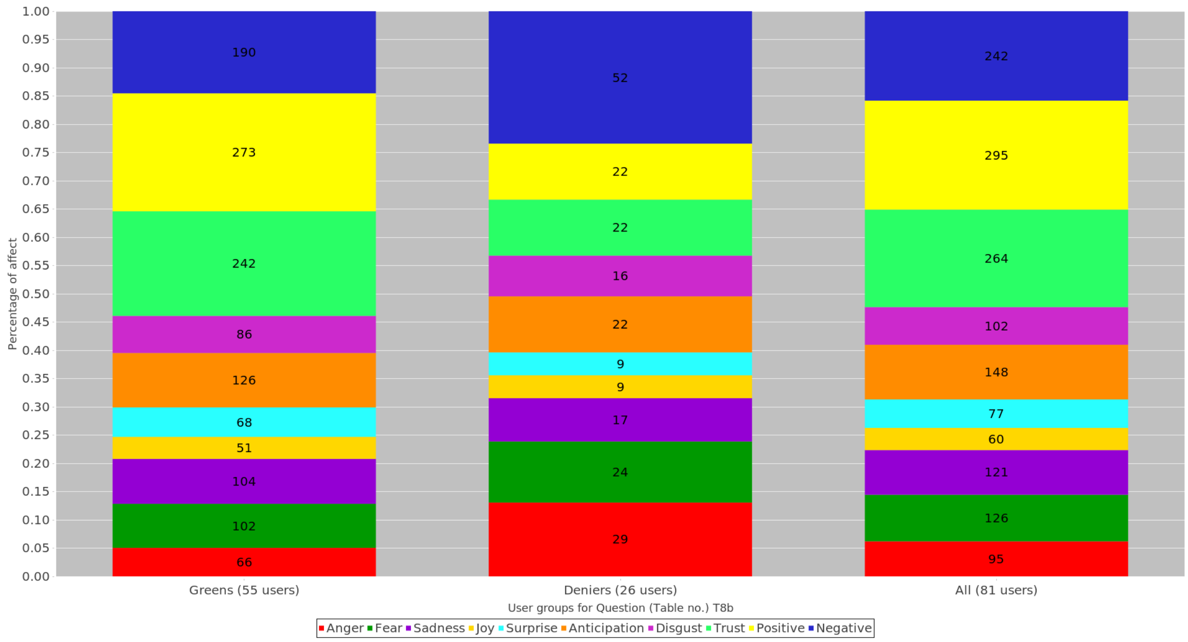


Figure 6.36 Affect per the NRC-10 lexicon for users tweeting about Question T8b.

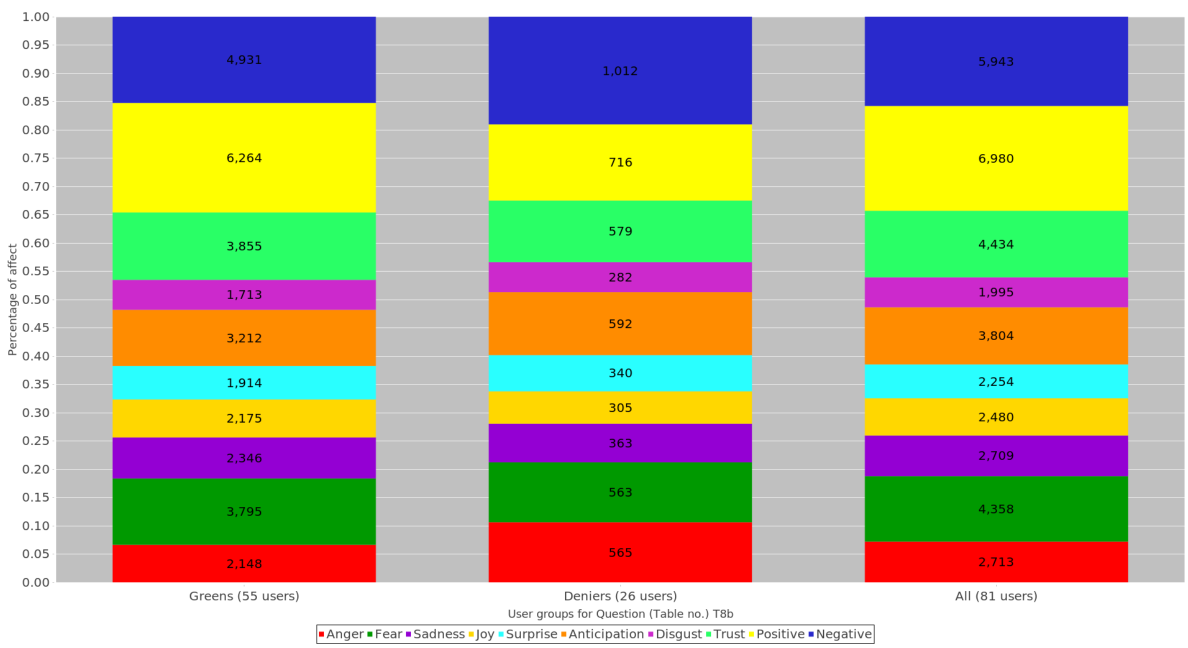


Table 6.24. Comparing these signatures to those for the T8b tweets only, we see that the trends for increased anger and negativity in tweets from the denier group are still present but not as strong. Also, expressed fear is proportionally higher again in the green tweets, which is the more common trend. Finally, the level of trust in green tweets has dropped in the general *#globalwarming* communications for these users from the relatively high level with only their T8b tweets.

Reflections on T8b

Question T8b ties back to the perception of risk. As we discussed in the previous section, a perceived risk that seems far in the future will more likely be handled by the brain's more analytical, less affect-oriented S2 processes. When this is the case, a human being is pushed to action by conscious and more-or-less objective reasoning. However, this push is terribly weak compared to the automatic shove one receives from the emotion and sentiment generated when S1 processes work to get the person out of harm's way in the face of immediate danger (Weber, 2006). With this in mind, we argue that an ongoing look at the affect in conversations on social media concerning the question of when online users think climate change will be affecting them personally could be invaluable. This kind of continued research would not only serve towards a better understanding of people's perceptions of climate risk but could also aid efforts to properly target messages to their intended audiences for information campaigns and online climate education. Conversation and communication are important. Recent results from the ongoing Six Americas survey-based research indicates that people in the U.S. who talk about climate change at least occasionally are much more likely to believe that it is causing harm in the U.S. right now (71%). Of the subjects who do not discuss climate change or speak of it only rarely, just 39% believe that it is happening now (Ballew et al., 2019).

6.3.8 Question T23: "Support for a National Response: Conditions for & Magnitude of Action Desired"

Question T23 only meets our criteria for question selection when considering denier tweets in the 17-tweet minimum dataset. The question ranks eighth overall, only ninth in the green group, but seventh in the denier group. This question also has two parts. The first part asks the subject on what condition he would support efforts by the United States to reduce emissions: (1) absolutely, no matter what the rest of the world does; (2) only if other industrialized nations take action; or (3) only if other industrialized nations as well as developing nations take action. He may also answer that the U.S. should *not* take action at all or that he doesn't know. The second part allows him to say how great that effort to reduce emissions should be (large, medium, or small) when considering the associated cost to the economy. He may also reiterate here that no effort should be made whatsoever.

Affect Signatures

The affect signatures for question T23 are given in Figure 6.37 for the green category, the denier category, and both combined. Table 6.25 lists the percentages of expressed affect making up the subdivisions of the bars in the chart as well as the affective word counts. This question seems exceptionally charged for users in the denier group given the amount of expressed anger in the associated denier tweets. The tweets from this group are also expressing a relatively high level of disgust. These two emotions are not particularly high in green tweets; however, T23 is one of the few questions which shows (albeit only slightly) higher negative sentiment than positive in associated green tweets. We also note the common trend that the green tweets show a higher level of expressed fear.

Affect signatures for *all* tweets whose authors have a tweet linked to question T23 are presented in Figure 6.38. Table 6.26 contains the underlying data for the chart. These

Figure 6.37 Affect per the NRC-10 lexicon for tweets on Question T23.

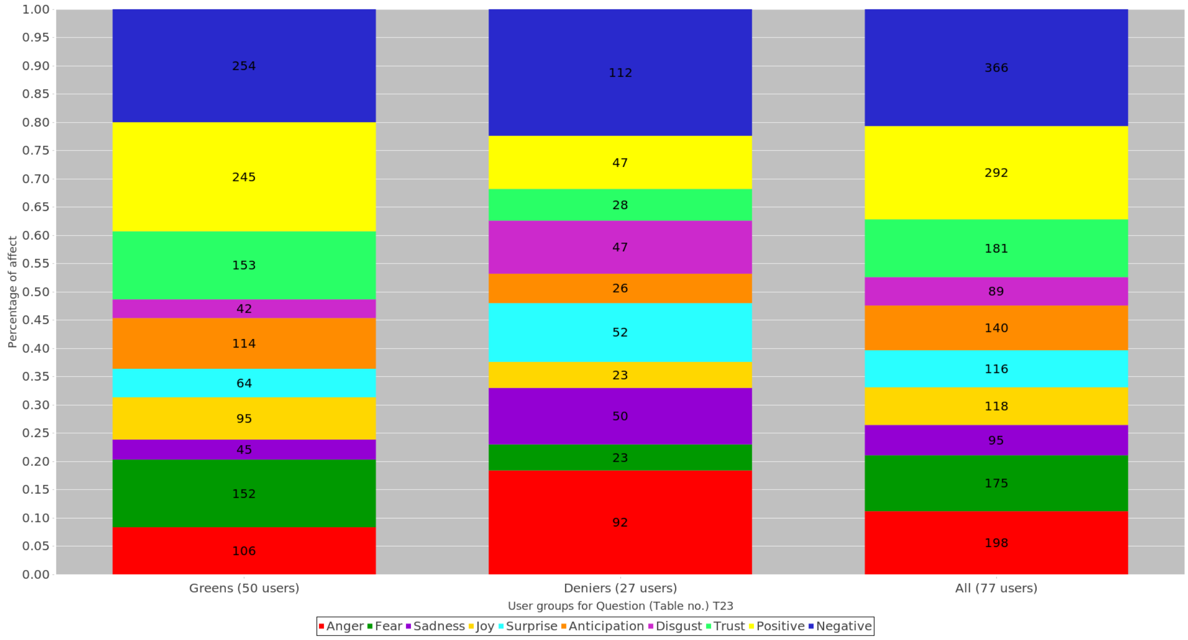


Figure 6.38 Affect per the NRC-10 lexicon for users tweeting about Question T23.

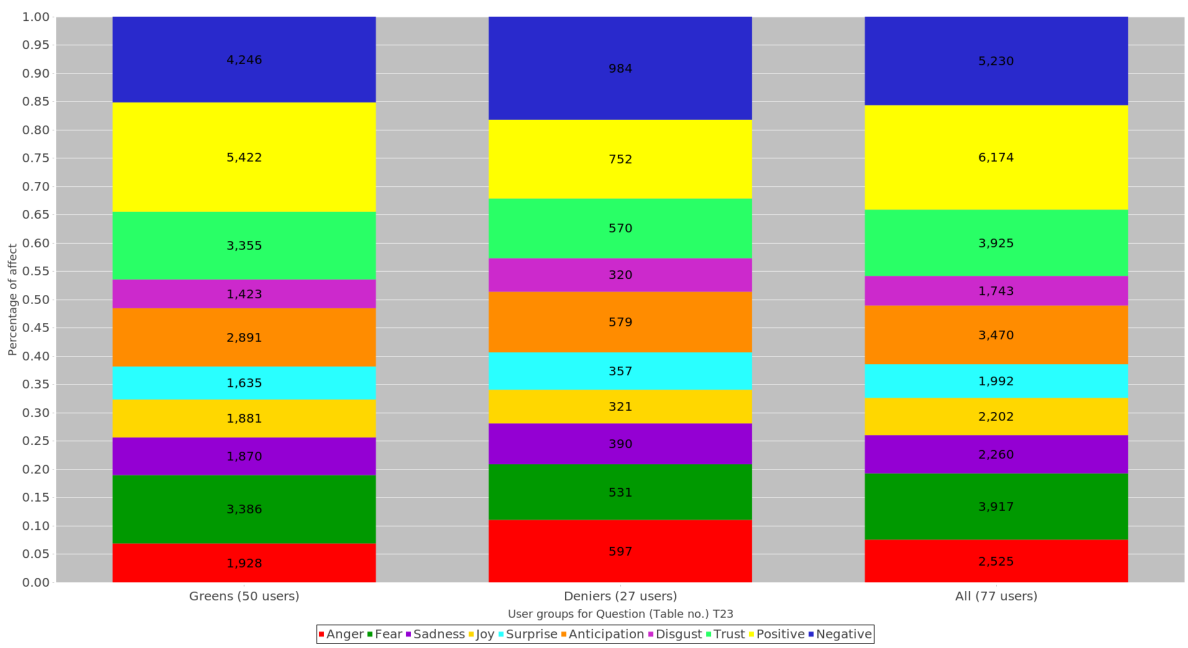


Table 6.25 Affect per NRC-10 lexicon for tweets linked to question T23.

	Greens		**Deniers**		**All**	
Users	50		27		77	
Anger	106	8.3%	92	18.4%	198	11.2%
Fear	152	12.0%	23	4.6%	175	9.9%
Sadness	45	3.5%	50	10.0%	95	5.4%
Joy	95	7.5%	23	4.6%	118	6.7%
Surprise	64	5.0%	52	10.4%	116	6.6%
Anticipation	114	9.0%	26	5.2%	140	7.9%
Disgust	42	3.3%	47	9.4%	89	5.0%
Trust	153	12.0%	28	5.6%	181	10.2%
Positive	245	19.3%	47	9.4%	292	16.5%
Negative	254	20.0%	112	22.4%	366	20.7%

Table 6.26 Affect per NRC-10 lexicon for all tweets from users linked to question T23.

	Greens		**Deniers**		**All**	
Users	50		27		77	
Anger	1,928	6.9%	597	11.1%	2,525	7.6%
Fear	3,386	12.1%	531	9.8%	3,917	11.7%
Sadness	1,870	6.7%	390	7.2%	2,260	6.8%
Joy	1,881	6.7%	321	5.9%	2,202	6.6%
Surprise	1,635	5.8%	357	6.6%	1,992	6.0%
Anticipation	2,891	10.3%	579	10.7%	3,470	10.4%
Disgust	1,423	5.1%	320	5.9%	1,743	5.2%
Trust	3,355	12.0%	570	10.6%	3,925	11.7%
Positive	5,422	19.3%	752	13.9%	6,174	18.5%
Negative	4,246	15.1%	984	18.2%	5,230	15.6%

signatures for the general *#globalwarming* communications for these users seem to reflect a relatively even expression of the affect characteristics from the NRC-10 lexicon compared to the signatures for only those tweets linked to T23.

Reflections on T23

The part of question T23 which asks if the U.S. should make an effort only if other nations do so as well is particularly interesting when we consider that the U.S. is responsible for about 25% of global greenhouse gas emissions despite the fact that the country makes up only 5% of the world's population (Leiserowitz, 2006). As of 2016, research based on the Six Americas surveys indicate that 61% of people in the U.S. believe that the country should cut its emissions without regard to whether other nations are doing the same. As T23 made the list of selected questions only because of its popularity in the denier group, what may be more salient are the 13% who say it does indeed depend on what other nations do and the 6% who respond that the U.S. should not cut emissions in any case (Leiserowitz et al., 2016). The present study shows high levels of expressed anger and disgust in the denier tweets linked to these ideas per an association with question T23. Continued research is merited in order to better explore possible reasons for these emotions.

6.3.9 Question T24: "Issue Priority"

Question T24 is the last of the Six Americas survey questions selected for a specific analysis in the present work. The question only ranks #21 in popularity overall and is close to the right side of the graph in Figure 6.8 where lack of associated tweets makes us hesitate to perform an analysis. Referring back to Table 6.4, we see it is the only question with a higher denier tweet hit count than green hit count. Among green users it ranks #26 out of the 31 questions, but it is fifth in popularity for users in the denier category. As with many questions in the survey, T24 has two parts. In the first part it asks the subject if she thinks the president and the U.S. Congress should consider global warming to be a "low, medium, high, or very high priority." In the second part T24 asks how high of a priority the president and the Congress should consider the development of clean energy sources.

Affect Signatures

Figure 6.39 presents the affect signatures for tweets linked to question T24 for the green and denier groups as well as both groups together. The percentages of expressed affect for these signatures as well as the affective word counts are given in Table 6.27. The signatures for T24 have a number of notable differences from those of other questions we have analyzed. First of all, the common trends of increased anger and negative sentiment in denier tweets, although present, are not particularly strong for this question. The trend of a higher level of fear in the green tweets is also not very prominent here. However, we are seeing notable differences for several other emotions. There is a low level of disgust for both groups. Expressed surprise is low in green tweets, and sadness is low in denier tweets. Finally, the level of anticipation is exceptionally high in associated tweets from the denier group.

Table 6.27 Affect per NRC-10 lexicon for tweets linked to question T24.

	Greens		**Deniers**		**All**	
Users	23		17		40	
Anger	37	11.0%	105	11.9%	142	11.7%
Fear	43	12.8%	103	11.7%	146	12.0%
Sadness	33	9.8%	20	2.3%	53	4.4%
Joy	8	2.4%	76	8.6%	84	6.9%
Surprise	8	2.4%	81	9.2%	89	7.3%
Anticipation	42	12.5%	170	19.3%	212	17.4%
Disgust	4	1.2%	15	1.7%	19	1.6%
Trust	39	11.6%	103	11.7%	142	11.7%
Positive	73	21.7%	89	10.1%	162	13.3%
Negative	50	14.8%	118	13.4%	168	13.8%

Table 6.28 Affect per NRC-10 lexicon for all tweets from users linked to question T24.

	Greens		**Deniers**		**All**	
Users	23		17		40	
Anger	1,314	6.5%	445	11.1%	1,759	7.3%
Fear	2,237	11.1%	429	10.7%	2,666	11.0%
Sadness	1,660	8.2%	290	7.2%	1,950	8.1%
Joy	1,238	6.1%	242	6.0%	1,480	6.1%
Surprise	1,323	6.6%	284	7.1%	1,607	6.6%
Anticipation	1,947	9.6%	448	11.2%	2,395	9.9%
Disgust	1,050	5.2%	214	5.3%	1,264	5.2%
Trust	2,440	12.1%	407	10.1%	2,847	11.8%
Positive	3,820	18.9%	536	13.4%	4,356	18.0%
Negative	3,154	15.6%	715	17.8%	3,869	16.0%

The affect signatures for *all* tweets from users with at least one tweet associated with T24 are shown in Figure 6.40. Table 6.28 shows the corresponding percentages of expressed affect and the word counts. Again, in these signatures representing the general *#globalwarming* communications for these users we see the levels of expressed affect

Figure 6.39 Affect per the NRC-10 lexicon for tweets on Question T24.

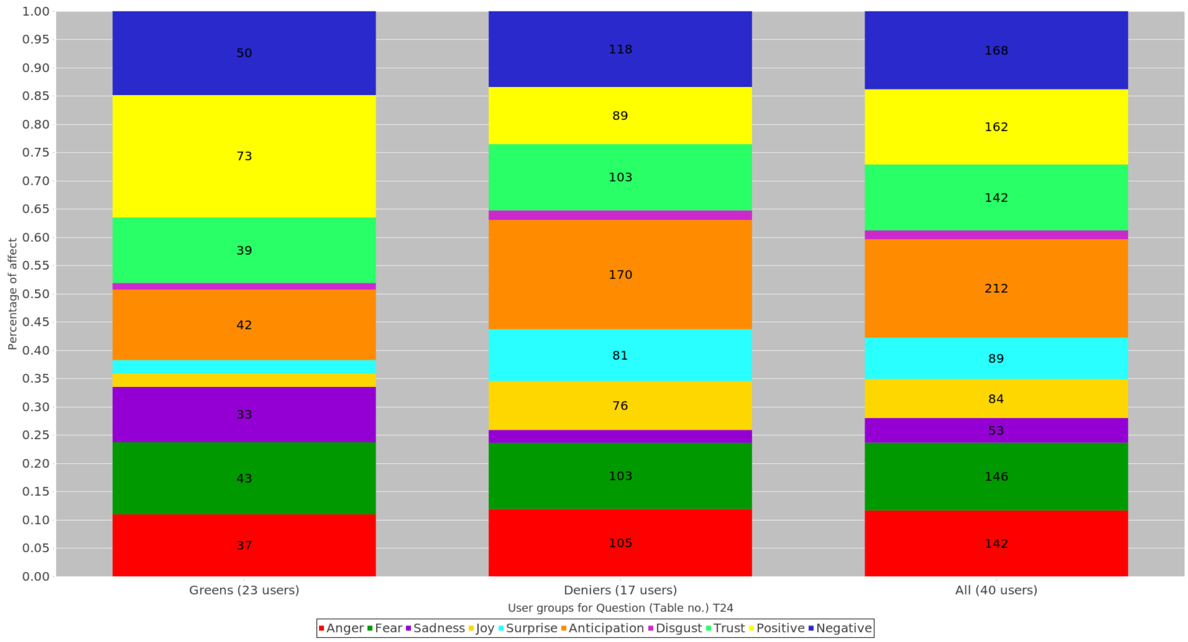
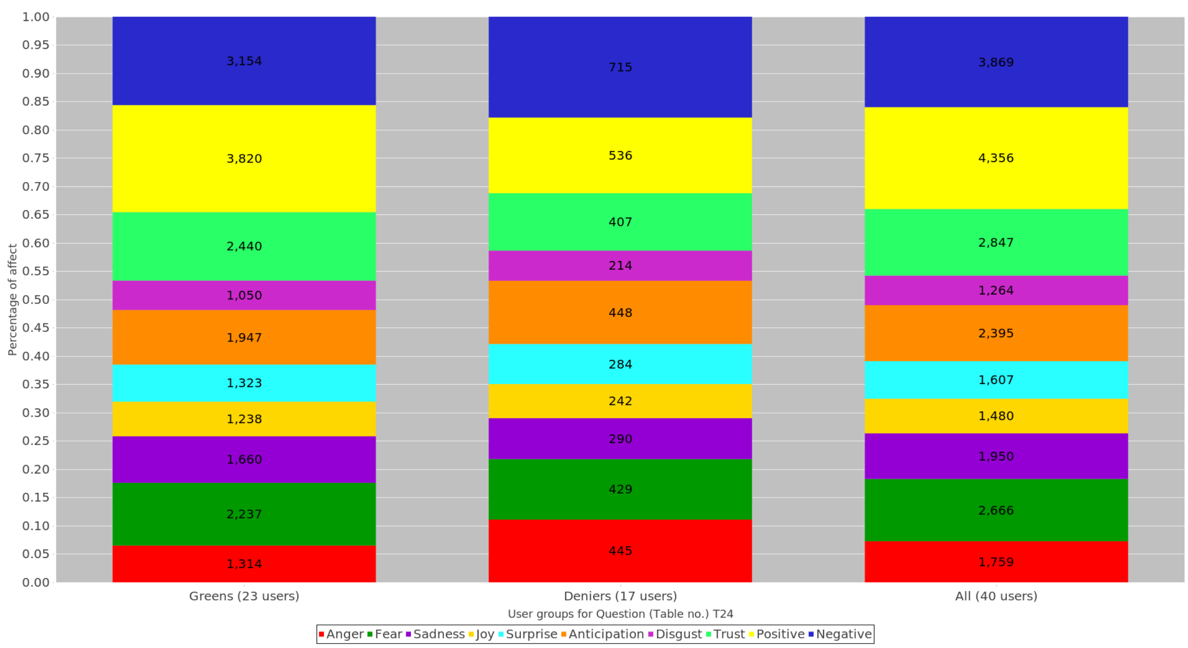


Figure 6.40 Affect per the NRC-10 lexicon for users tweeting about Question T24.



return to something near what appears to be the common configuration. The trends for increased anger and negative sentiment in denier tweets are present. The trend for higher expressed fear in green tweets is also present, though not especially pronounced. The differences noted above are therefore likely tied to the discourse on question T24.

Reflections on T24

In our discussion of question T16 above (Section 6.3.5) we noted the relationship between politics and beliefs about climate change and a person's support or opposition to climate policy. In a meta-analysis studying predictors of people's stances related to climate change in the United States, (McCright et al., 2016) found political orientation to be the second strongest indicator (after pro-environmentalism). This can have ramifications beyond climate policy matters within the U.S. as the country has often played "an obstructionist role" in negotiations concerning climate change on the international stage. At the personal level, political affiliation tends to serve as a "filter" for a person's understanding of climate-based issues in the United States. Republicans and Democrats get different viewpoints from their news sources and different cues on what to believe from their political leaders (McCright, 2011).

Question T24 truly represents an opportunity for continued research in an effort to mine social media for a better understanding not only of what online users think the U.S. president and Congress should be tackling when it comes to climate change, but also how these opinions shed light on the influence of political orientation with regard to their fundamental beliefs and attitudes on the matter. McCright makes this point clear: "At least into the foreseeable future, the Internet is likely to remain a world where climate change denialism thrives and where falsities and half-truths endure" (McCright, 2011).

6.3.10 Overall Observations

As we look at specific questions, it is most interesting to note which ones break the general trends. For example, we usually see more negative sentiment and anger in tweets from users in the denier category, but in one question, T25 ("Support for National Response: Specific Climate and Energy Policies Priority"), it is the users in the green category who are expressing more negative sentiment in their tweets. In one other, T16 ('Political Activism"), those green users are expressing more anger. Similarly, it is generally users in the green category who express more fear; yet for one question, T9 ("Perceptions of Weather and Climate"), it is the denier tweets that register as more fearful. We see increased anticipation in two questions, one on the green side with T22 ("Outcome Expectations") and one on the denier side with T24 ("Issue Priority"). Finally, the dynamic involved in expressions of disgust appear different in that for three of the four questions showing increased levels of this emotion, the increase was present in both the green and the denier tweets. These three questions were T3 ("Questions About Global Warming"), T16 ("Political Activism"), and T24 ("Issue Priority"). In the fourth question showing increased levels of disgust, T23 ("Support for a National Response: Conditions for & Magnitude of Action Desired"), the increase was only on the denier side.

These breaks from general trends seem particularly salient when we consider that once we incorporate all the tweets from authors linked to a specific question, the general trend seems to reestablish itself. The results presented in this chapter point to multiple opportunities for continued research aimed at gaining a better understanding of why specific topics on the subject of climate change may be linked to an increase or a decrease in a given emotion with respect to the general conversation online. However, before looking much further along this line of research, we should address the limitations of our current methodology, most specifically the points addressed in Section 6.4.2. Doing so will increase our confidence that the architecture is, in most cases, categorizing tweets according to the most relevant survey questions.

6.4 Limitations

Although we have endeavoured to be thorough in our analysis, the breadth of the research effort presented in this chapter has left a significant amount of work to be done. Of course, there are inherent limitations in any research project, but the scope of this doctoral program does not allow us to cover all aspects of the study as we would like. In this section we list the major limitations associated with the work presented in this chapter and indicate, where possible, how we may address these limitations through continued research.

6.4.1 Online Users versus the General Population

In this chapter we have made numerous comparisons of the results from an analysis of affect in microblogs published by a group of users on Twitter to an analysis of survey results from the series of experiments representing the Six Americas project. However, while the survey efforts take care to choose a sample of subjects representative of the population of the United States based on census data (Maibach et al., 2009), the present study does not perform this step as all the data necessary for a proper demographic analysis is not readily available for users on Twitter. Although a number of users disclose some of this data in their personal online profiles, it is presented in a free-form manner. Collecting it presents an additional research challenge as does ascertaining the veracity of the data obtained. Furthermore, studies have shown that online users tend to be younger and lean liberal more often than in the general population (Mellon & Prosser, 2017; Wojcik & Hughes, 2019). We are, for example, likely seeing the effects of this type of discrepancy in our results for question T3 as we noted in Section 6.3.4.

Yet even if our sample of users on social media is not truly representative of the people in the U.S. in the way that the Six Americas project on which we base our model strives to maintain, studies of online communities are important for research on climate change in their own right. Simply reporting "the science" in microblogs and online posts is not the effective solution many would like it to be (Auer et al., 2014). Furthermore, there is a standing interdisciplinary call to research towards understanding the dynamics of online communications and debates on climate change (Schäfer, 2012).

An interesting alternative to research efforts aiming to bring the demographics of online communities closer in line to those of the original study in the human sciences would be to perform a series of case studies which target different groups of online users, subsets of the larger community on social media. In the context of the Six Americas, the obvious extension to the present research would be to take a closer look at the particularities of online activity distinct to the different segments of the Six Americas (ideally, after we have enhanced our work to the point that it is identifying all six). A case study would allow for a more qualitative analysis than we have performed in the present work, An in-depth look at the tweets being published in each segment should enable us to begin to describe various aspects common to the online users in each group that could then be compared with the results obtained from the original study.

However we proceed, given the massive online presence and the influence of social media in today's world, the importance of research on social media seems clear, both as a ready source of public opinion and as a tool for public outreach and education.

6.4.2 Grounding Results

One challenge in analyzing the affective signatures we have created in this chapter is that we do not have known examples from which to ground our results. If we had a set of microblogs whose affective characteristics were known a priori, we could then use the known values as a base from which to quantitatively compare our results. If the system generates signatures with high accuracy for texts in this test set, we may have increased confidence for the affective signatures generated for tweets whose affective characteristics we do not know. We have a similar limitation with respect to linking tweets to Six Americas survey questions. Although Lucene using BM25 for similarity scoring represents a state-of-the-art approach for identifying relevant documents from a corpus (Robertson & Zaragoza, 2009; Trotman et al., 2014), we are using this IR technique in a nonstandard manner, treating microblogs as queries to find find a connection to one of a small number of survey questions. This makes the lack of a test set of microblogs with known links to questions even more critical with respect to the goal of evaluating our methodology.

To allow for a quick idea of how well our IR strategy is performing, we randomly sampled 200 tweets from the 2019 *#globalwarming* dataset and had a single human evaluator classify them as one of the 31 survey questions from the Six Americas (or as none if unable to link a tweet to a question). The human and the system were in agreement with only 28 (14%) of these tweets.[22] In addition to being a clear indicator that further work is required before we might expect the architecture to be able to make efficient use of the techniques presented in this chapter, the manual labelling process sheds some light on the challenges inherent in automating that process.

1. The subject of a tweet is often only indirectly linked to a specific question, and a fair amount of background knowledge may be needed to understand these links. For example, our human interpreted "put your heating on in June" as an unusual weather condition, and linked the tweet to question T9 "Perceptions of Weather and Climate." The automated system, however, chose T13 "Conservation Intentions," a question which mentions setting "the thermostat to 76 degrees or warmer [to] use less air conditioning" during Summer (Leiserowitz et al., 2010).
2. The human was often torn between two (or very occasionally three) survey ques-

[22] It is interesting to note that the most popular survey concept pair, human–cause, is indeed attached to the most popular question, T22 ("Outcome Expectations"), but it is in the sense of "humans can make [cause] a difference," rather than "humans are a primary cause of climate change." T22 also represents a hit for half of the CO_2–cut concept pair, asking if people can "reduce" [cut] global warming (rather than reducing carbon).

tions which seemed readily applicable to the subject of a tweet. In Section 6.1.1 we discussed using only the best hit per Lucene for a match to a question. Our intention was to begin our research efforts with a relatively simple configuration to establish a baseline and then see how we could improve upon the initial results by experimenting with these types of parameters. Our experience with this short test indicates that allowing matches to a pair of questions may improve the capabilities of the automated system.

3. With several tweets, the human made a decision based on background knowledge about named politicians (e.g., *@realdonaldtrump*), activists (e.g., *@GretaThunberg*), places (e.g., "Arctic"), and other proper nouns (e.g. "Green Deal"). Incorporating a component for named-entity recognition (NER) would likely improve performance as we continue our work with this methodology.[23]
4. The human needed to split a number of word clusters in the tweets (e.g., "AsTheAmountof #IceWorldWideDecreases, TheSpeedof #GlobalWarming Increases"). *Say Sīla* already incorporates this functionality for basic clusters (those able to be separated based on capitalization, numerals, and punctuation),[24] but it is not currently used with in the IR component of the architecture.

Clearly, linking tweets to a survey question is not always a trivial process for a human either. We can attempt to address some of the issues with the human-labelled sample as we continue with research aiming to refine our methodology. However, we will still have the basic problem that we are working with large unlabelled datasets. When we designed the experiments for analyzing emotion in tweets using the "Big Players" as described in Section 3.2, we faced a similar challenge. We had no known data on which to ground our results. Our solution there was to perform the same experiment using multiple time blocks, essentially sweeping a twelve-month window across 20 months of tweets. We

[23] A recent project for NER in the domain of climate change research may provide a good starting point (Holloway, 2015).

[24] https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/social.clj#L28

could then identify the correlations showing across the full set of experimental runs. The next phase of research for the affective signatures presented in this chapter may benefit from a similar evaluation methodology. In addition to an extended analysis over several time spans, we will need to compare the analysis presented here with experimental runs using a range of minimum activity levels. Our choice of a 17-tweet minimum was based on a hypothesis that this would likely be near an ideal level. Although in the present research we have compared our observations for the 17-tweet minimum with those of the full dataset, we need to conduct a complete evaluation as part of our continued research efforts in order to determine the validity of this hypothesis.

A complementary strategy to that of performing extended experimental runs across a range of parameter values would be to extend our small random sample of tweets to a more significant size, annotating them by means of a crowdsourcing platform such as Amazon's Mechanical Turk (Sheehan, 2018) with respect to both their link to a question from the Six Americas and their affective characteristics.

6.4.3 Choice of Lexica

In Section 6.1.2 we mentioned the importance of evaluating an affect lexicon with respect to how effectively it covers the emotion and sentiment expressed in texts for a given domain. In our case this domain is climate change as discussed in online microblogs. Further efforts in this direction are a must for our continued research. In the present work we used Bing Liu's Opinion Lexicon (Hu & Liu, 2004) for a comparison against our results obtained using the NRC Word-Emotion Association Lexicon (NRC-10) (Mohammad & Turney, 2013) for positive and negative sentiment polarity. We did observe a number of similar patterns between the two lexica; however, the results are not exceedingly well aligned with respect to these two affective characteristics. This indicates that further work is warranted to evaluate which lexicon (if either) is a good choice for the textual data we are analyzing. Furthermore, we have not performed this level of evaluation for the other eight emotional attributes covered by the NRC-10.

We hypothesized that the reason for the disagreement between these two lexica is likely at least partially due to the difference in the ratio of negative to positive words as well as the difference in size between the two lexica. The lexica we have used here are hand-tagged. Automated and semi-automated methods allow for the creation of much larger lexica, which may have better coverage of the words used in the microblogs we are analyzing. Yet, we must take care as these methods are known to introduce a certain level of noise, often due to the inclusion of neutral words. Another refining strategy may be to employ a strength-based lexicon such as the new NRC Emotion Intensity Lexicon (NRC-EIL), mentioned in Section 3.2.1. In all cases we must perform a preliminary analysis to evaluate how well a given lexicon conforms to our target domain. We may also try to determine if combining a number of approaches, such as handling n-grams or pairing the affect associated with a word to its part of speech, will significantly improve results (Bravo-Marquez et al., 2014).

6.4.4 Information Retrieval Enhancements

Even as Lucene is a powerful tool for Information Retrieval (IR), our usage of this tool for the experiments in this chapter arguably represents a baseline. Since we are seeking only a single Six Americas survey question to link to a given tweet, we performed no exploration of query expansion techniques that could still possibly improve results. As mentioned in Section 6.1, such techniques most often work to enhance recall in a search (Voorhees, 1994). Essentially, this would mean that we could more readily link additional questions to a tweet. This is not our goal; however, we should not take it as a foregone conclusion that query enhancements such as the inclusion of synonyms, hypernyms, or hyponyms would yield no improvement at all on search precision.

A somewhat less ambitious strategy for potentially improving accuracy when linking questions has to do with preparing the available information in a way which maximizes the effectiveness of Lucene. For example, when we indexed the corpus of questions for the Six Americas, we put the full text of each question including the question title into a Lucene document field called `TEXT`. This method provides us with a baseline for results. However, it is common for document titles to go into a separate `TITLE` field, and search queries can then indicate that certain search terms should match a document's title, while other terms should match text in the body of the document. Perhaps this configuration would be of limited use in our study as we are simply using the full text from a tweet as our query. However, we can create custom fields as appropriate for our model, such as a field for key words related to each question. We may also boost important search terms in our query when these terms are present in the tweet from which it is formed (McCandless et al., 2010). Experimentation with these approaches could yield significant improvements for our model, especially if it includes a test set of known microblog–question associations as discussed in Section 6.4.2.

6.5 Contributions and Continued Research

The work presented in this chapter, perhaps best visualized as the numerous affect signatures we have produced, serves to present a novel view of the Six Americas as embodied by online communities tweeting about global warming. We have utilized the state of the art in information retrieval (Trotman et al., 2014) to link the conversation on Twitter to specific questions from the Six Americas survey. This process has allowed us to perform an analysis of the emotion and sentiment expressed online with regard to each question. We have examined a number of tendencies which we have observed for users in the green and denier categories, and we have compared the results using two lexica designed specifically for work with online communications. For the most popular questions, we performed a more in-depth analysis, comparing the affect expressed for a given question to the affect observed in the general online conversation on global warming by the subcommunity demonstrating a connection to that question. Finally, we have extended our observations from the automated analysis for these popular questions by comparing our findings to existing research in the human sciences.

From the perspective of our ongoing research, we are imagining a system that can generate affect signatures like the ones presented in this chapter and use results from the survey-based experiments of the Six Americas project as an anchor to model communities of online users. These models could then be directly compared to the segments of the U.S. population as reported by the survey-based research, especially on matters concerning the emotional state of the online communities. In the scope of this doctoral program, we have essentially produced a series of snapshots of these emotional states for *#globalwarming* communications on Twitter in 2019. Certainly, these snapshots require further research efforts with respect to the limitations discussed in the previous section, but what we have presented here is an important first step and will serve to direct future research efforts. We realize that we have perhaps raised more questions than we have answered. Yet, a number of these questions seem of particular interest with respect to the efforts of the scientific community to understand people's attitudes towards climate change. For example, an extremely concise version of the Six Americas survey called the Six Americas Short SurveY (SASSY) has been found to be 70–87% accurate at identifying the classification of subjects in terms of the six segments[25] (Chryst et al., 2018). The four questions in SASSY are from questions T7 and T8a as presented here.[26] We would expect these questions, which come closest to being as effective as the full questionnaire, to represent common themes in the *#globalwarming* conversation on Twitter. Yet, in our study T7 and T8a were not found to be popular questions for the high activity users tweeting about global warming. Determining why may help to shed light on essential differences between online communities and people in everyday life—if not generally, then at least as these differences relate to the issue of global warming. Recent research connected to the Six Americas project includes affective elements (worry)

[25] Alarmed, concerned, cautious, disengaged, doubtful, or dismissive.

[26] Each Six Americas "question" (T2–T31) in the present document represents one of the tables describing the survey in (Leiserowitz et al., 2010). Each table generally covers a number of individual questions from the questionnaire.

in its modelling of how people in the U.S. perceive the scientific consensus on climate change (van der Linden et al., 2019). Parallel research efforts aimed at monitoring and evaluating the sentiment and emotion expressed in the online discourse on social media sites could prove invaluable.

Our ultimate goal is to create an automated system with the level of understanding necessary to read hundreds of thousands of tweets and link them to survey questions or other key concepts for a completely clear view of what folks are discussing on social media. We are certainly not there yet; however, researchers in psychology and sociology can still benefit from an approximate view of how these communications relate to important questions about climate change or indeed any field of study.

CHAPTER VII

MINING USER STANCE ON GLOBAL WARMING

At various points in the present document we mention that we have followed an iterative approach for our research methodology. While this approach is crucial for the creation of our ontological model (see Section 5.1), the methodology is also beneficial in a more general way. Findings in earlier stages of the research help both to create a foundation for the later stages and to improve the associated research effort. One notable example is our initial work regarding the big players (Chapter 3). Even though that began as a stand-alone project primarily meant to get us started analyzing tweets on the *Say Sila* architecture, once we started developing and experimenting with the ontological model, we incorporated the idea of using high-activity users as a technique to stratify our data, leading to an analysis at multiple levels of minimum user participation in the conversation on Twitter. Requiring a minimum level of participation for inclusion in the dataset also reduced the amount of preprocessing necessary before a series of experiments since we did not need to parse dependencies for a massive amount of online posts from one-tweet users, nor did we need to query the Twitter Developer API to determine their followers.

The big players project was based on machine learning, but to an extent it served to enhance our work with description logic and the say-sila ontology. In a similar manner, the results from our work with the ontological model can be brought back into the domain of machine learning. This includes not only the ontological elements ultimately used for classification (the survey concept pair indicators), but also the emotion and sentiment expressed in the tweets and the parts of speech associated with the words, emojis, etc. In this chapter we use these elements from the model for a series of classification experiments using machine learning algorithms and thereby build on what we have created in order to take another step towards our current goal of identifying the stance on global warming for users on Twitter.

Note, however, that this step is not intended to be our final one. The experiments described in this chapter represent a preliminary look at how to utilize information as it flows from the top stage of *Say Sĩla*, the description logic level, back down to the bottom stage, the machine learning level. The results here improve upon those produced by a single pass through the architecture, but our ultimate goal goes beyond simply using an ontology to organize the incoming data and transform it into attributes which machine learning algorithms can leverage for a classification problem. The architecture we are creating is recursive. Results from a second pass through the machine learning level are intended to serve as a signal for error correction in the description logic level so that the architecture may refine its ontological model. This dynamic is inspired by theories in biological cognition which demonstrate that neural plasticity in one part of the brain, specializing in a given function, can be influenced by error information coming from other areas of the brain, responsible for other functions (Clark, 2013; Eliasmith, 2015).

Although pure machine learning solutions are likely to produce superior results for a classification problem at the present stage of our ongoing research, the architecture focuses on the ontology, considering it to be the final output model of the system. In this way, the ontological model can be used independently after *Say Sĩla* has created it (or created a "snapshot" for a model that evolves as the architecture continuously analyzes incoming posts from social media). A domain expert can then analyze or even edit the model using a tool like Protégé (Musen, 2015).[1] Researchers can also query the ontology as they might a database using a language like SPARQL (SPARQL Protocol

[1] https://protege.stanford.edu/

and RDF Query Language) (Sahoo et al., 2009)[2] or import part or all of the ontology into their own systems. Any of a vast array of Semantic Web technologies are at their disposal to access the final output model. Of course, we are not there yet. We are conducting these experiments in order to evaluate a prototype, knowing that a final version of the architecture extends well beyond the scope of this doctoral program. This chapter opens a window into the machine learning level of *Say Sīla* so that we may assess its ability at the current stage of our research to make use of information gained from the ontological model in the architecture's description logic level.

7.1 Machine Learning Models

We have selected two standard machine learning algorithms for classification: logistic regression and decision trees. The classification task is the same as it was in Chapter 5, namely to determine if users tweeting about global warming are in the green category or the denier category with respect to their position on climate change. We have chosen these two learners as they are standard algorithms which generate "clear box" models, allowing the researcher to understand which characteristics of the data have led to a given classification result.

Understand that there is a great deal of potential for improvement on the results we present in this chapter. Experiments aiming to tune the hyperparameters for the algorithms we have chosen may yield a significant increase in classification performance. Likewise, more sophisticated learners such as deep artificial neural networks or support vector machines may yield substantial improvements with regard to these initial results. Finally, preprocessing methods such as principal component analysis can help determine an ideal selection of (possibly transformed) attributes to create an enhanced classification model. Yet, all of these paths represent continued research endeavours. Our immediate purpose is to effectuate machine learning models that can build on knowledge about

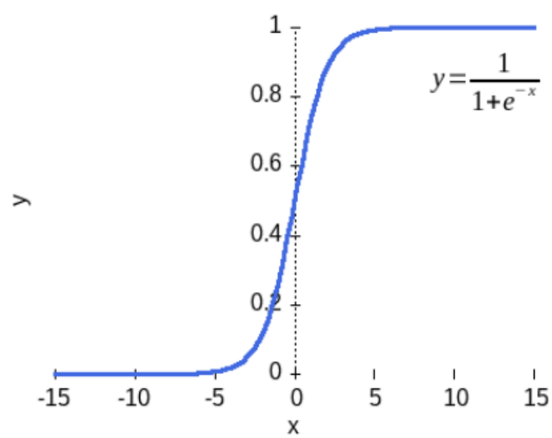
[2] https://www.w3.org/TR/sparql11-query/

Twitter users and their tweets as modelled in the say-sila ontology. In taking this step, we are essentially providing a baseline for future work aimed at coupling machine learning with our ontological model.

7.1.1 Logistic Regression

In Section 3.3.3 we describe how we use the machine learning algorithm linear regression to predict levels of expressed emotion in the tweets of the *#globalwarming* community on Twitter. For the research phase covered in the present chapter we use a related algorithm called logistic regression. At a practical level, the difference between linear regression and logistic regression is that linear regression handles regression problems, where we are predicting a numeric value, and logistic regression (despite its counterintuitive name) handles classification problems, where we are predicting one of a fixed number of classes (e.g., green or denier) (Goodfellow et al., 2016).

Figure 7.1 Logistic sigmoid function.



These two types of problems are quite distinct. The odd name stems from the fact that the logistic regression algorithm essentially builds on a regression step in order to perform classification. Using our own classification problem as an example, the algorithm determines the probability that an instance belongs to the green category based on the weights computed for a linear model. Of course, the probability that the instance belongs to the denier category is then $1 - P[green]$. Ideally, we would want to determine:

$$P[green|a_1, a_2, ..., a_k], \tag{7.1}$$

where a_j is the value of a given attribute for a data instance across k attributes in a given dataset. Unfortunately, however, a linear function cannot accurately model this target random variable. Therefore, the algorithm seeks instead to maximize the log likelihood:

$$log\left(\frac{Pr[green|a_1, a_2, ..., a_k]}{1 - Pr[green|a_1, a_2, ..., a_k]}\right) \tag{7.2}$$

Rather than ranging between 0 and 1 as is needed for a probability, values for the log likelihood may vary from negative to positive infinity. Therefore, the algorithm squashes this value using a logistic sigmoid, a non-linear "S-shaped" curve. This function is illustrated in Figure 7.1. The equation describing the standard logistic sigmoid is:

$$y = \frac{1}{1 + e^{-x}} \tag{7.3}$$

Thus, the equation expressing the transformed models is:

$$P[green|a_1, a_2, ..., a_k] = \frac{1}{1 + exp(-w_0 - w_1a_1 - w_2a_2 - ...w_ka_k)}, \tag{7.4}$$

where w_j is the weight assigned to a given attribute a_j. As mentioned above, the logistic regression algorithm determines these weights by maximizing the log likelihood of the model, rather than by minimizing the mean squared error as is the case with linear regression (Goodfellow et al., 2016; Witten & Frank, 2005).

Although the *Say Sila* application creates the datasets used for testing with the machine learning models (see Section 7.2), we actually use the *Logistic* learner (le Cessie & van Houwelingen, 1992)[3] in conjunction with the Weka machine learning platform's Experimenter tool (Frank et al., 2016) to run the series of experiments for the logistic

[3]The fully qualified class name is *weka.classifiers.functions.Logistic*.

regression algorithm.

As we are interested in establishing a baseline for this component of the architecture, we use Weka's default hyperparameters for the algorithm. Notable among these is a ridge regularization parameter.[4] The algorithm uses this parameter as a regularizer. It includes a penalty term as it calculates the the log likelihood, essentially working to maximize the value:[5]

$$\sum_{i=1}^{n} \left(y_i A_i^\top W - log(1 + e^{A_i^\top W}) \right) - \lambda W^\top W \tag{7.5}$$

where λ is the ridge parameter, y_i is the class value, and W is the vector of weights (excluding the intercept) to be applied to the vector of attributes $A_i = \{a_{i,1}, a_{i,2}, \ldots, a_{i,k}\}$ for each data instance i in the set of n instances. The ridge penalty term works to keep the attribute weights from growing overly large and provides a trade off, increasing the bias of predicted values by a small amount, while decreasing prediction variance by a large amount. Thus, small changes to attribute values in the training data do not result in as large of a change to the weights. It generally serves to limit a model's complexity, which helps to reduce overfitting that model to the training data (Goodfellow et al., 2016; Pereira et al., 2016).

We should note that although they are not directly available for Weka's *Logistic* class, other regularizers exist as well. Lasso (least absolute shrinkage and selection operator) is a well known example which also adds a penalty, not only to limit the size of weight coefficients but also to reduce them to zero if possible, further simplifying the model by eliminating an attribute completely (Tibshirani, 1996). Another regularization technique is least absolute deviation (LAD), which works to minimize invalid weight estimates caused by outliers (Bloomfield & Steiger, 1983; Li & Arce, 2004).

---

[4]Weka uses a value of 10^{-8} for the ridge (λ) in the log likelihood.

[5]Weka actually performs the equivalent operation of minimizing the negative log likelihood (and so the ridge term is added onto the value, rather than subtracted from it). For details, consult the Weka documentation: https://weka.sourceforge.io/doc.stable/index.html?weka/classifiers/trees/J48.html

As a remaining default hyperparameter, Weka has the logistic regression algorithm run until convergence as opposed to stopping it after a fixed number of iterations.

7.1.2 Decision Trees

The second machine learning algorithm we use for the experiments in this chapter is the decision tree. This is another learner created for classification problems. The decision tree algorithm leverages a "divide and conquer" approach to generate a tree which a modeller may then use to map data instances to class values. Beginning with the root, each internal node of the tree is associated with a specific attribute from the dataset. Branches off these nodes each correspond to a range of possible values for the associated attribute. When traversing the tree with respect to a given data instance, the modeller follows the branch that corresponds with that instance's value for the attribute. Doing so will lead to a subsequent branch, corresponding to a different attribute or else to a leaf node which indicates the class to be assigned to the instance (Quinlan, 1993; Witten & Frank, 2005).

Figure 7.2 Example decision tree for emotion attributes in a tweet.

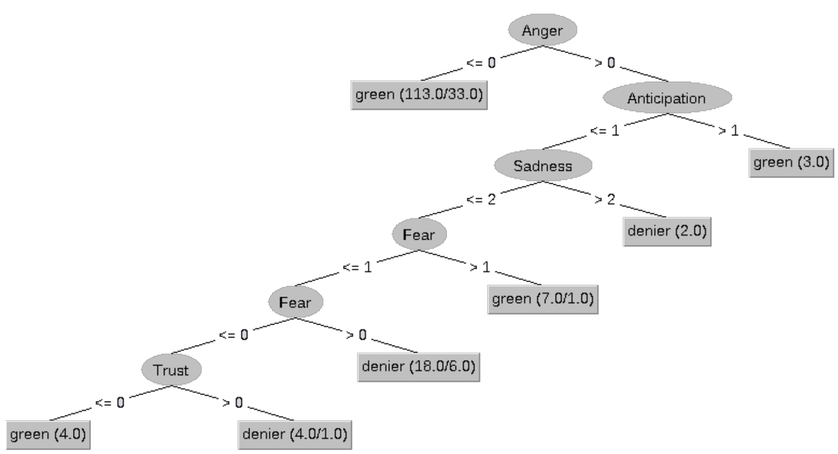


Figure 7.2 presents an example of a decision tree from this work created using the Weka machine learning platform (Frank et al., 2016). The model is used to classify a Twitter user as being in the green or the denier category based on her tweets from the 2019

#globalwarming dataset (see next section). If we traverse the tree, starting at the root, we consider how many tokens expressing anger are in the user's tweets. If there are none,[6] then we move left to a leaf node, and the instance representing this user is classified as green. Otherwise there are tokens expressing anger, and we move right to a new interior node. This new node is associated with the count of tokens expressing anticipation. If the user's tweets include more than one, then we move right to a leaf node that again classifies the user as green. However, if there is only one anticipatory token or none, we move left. In this case we must continue the traversal of the tree, checking the count of tokens expressing sadness and then possibly other attributes until we finally arrive at a leaf node which will classify the instance either as green or as denier. Finally, the reader may have noticed that the leaf nodes in the figure all have a number in parentheses after the name of the class or alternatively two numbers separated by a slash. These numbers indicate how well the model fits the training data used to create it. The first is the number of instances from the training dataset that were correctly classified for the label represented by the leaf node. The second number, if it is present, indicates how many training instances were labelled as belonging to the other (incorrect) class (Frank et al., 2016).

As is often the case with tree-based structures in computer science, the algorithm to create a decision tree is a recursive one. At a high level it works as follows (Witten & Frank, 2005):

1. Select a data attribute to represent the root node.
2. Create branches off this node with each branch corresponding to certain values in the training dataset for the attribute representing the node. If the attribute is nominal, create one branch per possible value. If it is a numeric attribute, determine a threshold and create a branch for values below (or equal) and another

[6]The check is actually for an attribute value less than or equal to zero. We know that there will never be a negative count of angry tokens, but the algorithm is generalized to work for any range of numeric data values.

for values above that threshold.

3. Split the training data into subsets with respect to the selected attribute in accordance with these branching conditions.
4. Repeat these steps recursively for each newly created branch with its associated data subset.

The key question is how to select which attribute should represent the node. Every time the decision tree algorithm creates a node, its goal is to maximize the purity of the training data subsets in the new leaf nodes connected by branches created for the node currently being processed. Note that these new "leaf nodes" may not be leaves in the final tree. Rather, they may become interior nodes themselves once the algorithm processes them as it recursively generates the tree. Ideally, a new leaf node will cover training instances of only one class (i.e., either all green or all denier). In this case, the new leaf node completes the branch, and the algorithm terminates the recursive path which generated it (continuing then, of course, with the other paths associated with other branches) (Witten & Frank, 2005).

The algorithm attempts to generate smaller trees,[7] and so it always generates pure nodes first when it is able to. Purity is given by the *information value* of a node, or its *entropy*. Entropy is measured in bits,[8] and it is calculated as follows for classification problems with two classes, e.g., green (G) and denier (D):

$$entropy(p_G, p_D) = -p_G \cdot log_2(p_G) - p_D \cdot log_2(p_D) \tag{7.6}$$

where p_G and p_D denote the percentages of coverage in the data subset for the current branch with respect to the green and denier instances. Note that calculating entropy for

---

[7]Larger and more complex decision trees are generally more prone to overfitting, essentially memorizing the training data rather than modelling the general patterns it contains.

[8]Though it is now used quite regularly in computer science, a *bit* as a unit of measure originally came from communication theory. It is an abbreviated form for *binary digit* associated with Claude Shannon's proposal to use base-2 logarithms as a measure of information (Shannon, 1948).

a pure leaf node would mean zero percent coverage for one of the classes, resulting in a term of $log_2(0)$, which is undefined. The algorithm simply ignores this term. Thus, pure leaf nodes always have an entropy of 0.

When processing a node, the information associated with the node is the average of the entropy values of its child nodes. To select an attribute for the node, the algorithm iterates through all possible attributes, calculating information values based on the average entropy of the children that would result from a split on that attribute. The decision to assign a given attribute to the node is based on the gain for that attribute:

$$gain(attribute) = info(node) - info(attribute) \tag{7.7}$$

where the $info(node)$ is the entropy for the node being processed (before the split) as given by Equation 7.6, and $info(attribute)$ is the average entropy of the child nodes if the node were to branch based on that attribute. The algorithm then selects the attribute that provides the greatest gain. A larger gain entails finding the attribute with the lowest possible information value, and here zero represents the ideal–a pure node (Witten & Frank, 2005).

Of course, the above description applies to nominal attributes, and for our classification problem all the attributes are numeric. The process is essentially the same, except that now the algorithm must perform an additional step. Rather than splitting data according to the class values for an attribute, it needs to determine a threshold for the split. To do so, it sorts the values for the attribute in the training data[9] and calculates the information gain at each possible threshold per Equation 7.7. These thresholds are the midway points between two attribute values when the corresponding training instances are not of the same target class (i.e., one is green and the other denier). The algorithm selects the threshold providing the largest gain. This is also the gain used to

[9]To increase performance, rather than repeatedly performing sorts at every node, an implementation will typically sort values for all attributes before it begins the process of tree generation.

compare the attribute to the others when determining which attribute will be used for branching on the node being processed as described above (Witten & Frank, 2005).

Like logistic regression, a decision tree is a clear box algorithm in that the learner provides a complete description of how it arrives at a given classification result. Looking at Figure 7.2, one can easily follow the sequence of tests made on a data instance in order to determine if a user is in the green or the denier category. Decision trees are a popular and relatively effective algorithm for machine learning. However, they are also criticized since their method of constantly splitting data against a single attribute means that the cuts which partition a given problem's search space are always aligned with the axes of that space (since an axis represents the range of values for an attribute) (Goodfellow et al., 2016). So unlike logistic regression, in the classification problem we are tackling here, a decision tree would not be able to model the type of pattern in the data where, for example, users are publishing tweets with more tokens expressing sadness than tokens expressing anger.

As mentioned in the previous section, we perform the experiments covered in this chapter using Weka's Experimenter tool (Frank et al., 2016). We use Weka's *J48* decision tree implementation, which corresponds to revision 8 of the well-known C4.5 decision tree algorithm by J. Ross Quinlan.[10] This revision is the last for which Quinlan published the source code as his C5.0 version is a commercial product (Quinlan, 1993; Witten & Frank, 2005).

For purposes of establishing a baseline with the decision tree algorithm, we use Weka's default hyperparameters. These include ensuring a minimum of 2 instances per leaf node and a 0.25 pruning confidence threshold. Pruning involves simplifying the decision tree with the goal of making it more general, reducing the chance of overfitting the model to the training data. One type of pruning is called *subtree replacement*, whereby the algorithm replaces an interior decision node (and the subtree under it) with a leaf node.

[10]The fully qualified class name is *weka.classifiers.trees.J48*.

A decision tree may also be pruned via *subtree raising*, which is when the decision node is replaced by a second decision node found in the subtree under the first node. To determine whether to leave a decision node as is, or to put a leaf or a smaller subtree in its place, the algorithm compares the error rate for both versions against unseen (future) data using the confidence threshold parameter to calculate an upper confidence limit on the standard deviation from the mean for that error rate. Of course, the algorithm cannot calculate the real error rate of any unseen data, and so it estimates the future error rate assuming a worst case based on the training data. (Quinlan, 1993; Witten & Frank, 2005).

7.2 The 2019 Global Warming Tweeter Dataset

The data we use for the experiments in this chapter is effectively the same data compiled for the 2019 *#globalwarming* tweet dataset as described in Section 5.2; however, we have transformed it. Each instance in the original dataset represents a single tweet, while here we need data instances that represent the users we are modelling. Essentially, we use the 2019 *#globalwarming* dataset to create our ontological model, and then we use the model to generate the datasets needed for the machine learning experiments covered in this chapter.

Like our analysis using the say-sila ontology in Chapter 5, the experiments using machine learning are conducted repeatedly for a range of minimum levels of user participation in the form of tweets posted during 2019 with the hashtag *#globalwarming*. The minimum-tweet levels still range from 2 up to 20. Accordingly, the first step for dataset generation is to create an ontological model for each minimum-tweet level and then to extract the required datasets in ARFF format for each of these models.[11] Each instance in these datasets corresponds to a modelled Twitter user, and the attributes are the counts of

[11] The `world->arff` function generates these datasets. It may be referenced here: https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/sila.clj#L4160

individuals in the ontology for the specific elements we wish to model here. For example, the number of tweets published by the user is one such attribute.

The other attributes may be separated into groups according to their type. We create a separate ARFF dataset for each group. These groups are:

- **POS** (parts of speech): These attributes represent token counts in the published tweets with respect to the parts of speech identified by the TweeboParser (Kong et al., 2014) as listed in Table 4.1. There is one attribute for each part of speech for a total of 25 attributes.
- **EMO** (emotions): The ten attributes in this group represent the counts of tokens expressing affect across all of a user's tweets as per the NRC Word-Emotion Association Lexicon (NRC-10) (Mohammad & Turney, 2013) as described in Section 6.1.2. There is one attribute for each of the emotions anticipation, anger, disgust, fear, joy, sadness, surprise, and trust as well as for positive and negative sentiment polarity.
- **ONT** (ontology): This group represents text-level weak and strong concept pair indicators from the say-sila ontology as described in Section 5.1.3. There are eight attributes corresponding to the number of a user's tweets which are members of the following classes: *WeakHumanCauseText*, *WeakNatureCauseText*, *WeakCO2CutText*, *WeakEconomicGrowthText*, *StrongHumanCauseText*, *StrongNatureCauseText*, *StrongCO2CutText*, or *StrongEconomicGrowthText*.
- **QST** (questions): This group covers 31 attributes, one for each of the questions of the Six Americas survey as listed in Table 6.1.1. Each attribute gives the number of the user's tweets that have been linked to the associated question using the analysis with Lucene (McCandless et al., 2010) as described in Section 6.1.
- **ALL** This final group is simply the aggregate of all the attribute groups described above.

For each minimum participation level we create five datasets, each corresponding to one of these sets of elements modelled in the say-sila ontology. The group code, shown in bold face in the list above, acts as a tag for the associated datasets and also identifies the group in the tables and charts presented with our results in Section 7.3. Note that we generate 19 datasets for each group (for minimum participation levels of 2 to 20) for a total of 95 ARFF datasets for our analysis with each of the two learners.[12] Additionally, all the datasets include the base tweet count attribute mentioned above. This count may be interpreted as a measure of how big a player each user is.

7.3 Experimentation, Results, and Analysis

This section presents the results we obtained using logistic regression and decision trees with the Weka Experimenter tool (Frank et al., 2016), running the datasets with different data elements from previous stages of the present research as described in the section above. For all experiments we use the F1 measure to score the model, which we evaluate using 10-fold cross-validation. However, we run the series of experiments twice, looking at the F1 scores from a different perspective each time.

When running the series the first time, we measure F1 with respect to the green category. This method parallels our F1 measurements when analyzing the same 2019 *#globalwarming* data using the ontological model as described in Section 5.3. When working with the ontology as it currently stands, this perspective makes sense because we have designed the model to infer which users are in the green category.[13] For this series of experiments with machine learning, however, we take a second perspective and also report the F1 scores as a weighted average of the F1 values as measured for each class: green and denier. Note that neither the models nor the datasets change for this second run, only

---

[12] In the code the `make-gw2019` function creates these datasets for the attribute groups: https://github.com/dendrown/say_sila/blob/uqam-dic/apps/say_sila/priv/fnode/say/src/say/sila.clj#L4202

[13] To be clear, our intention was to also have ontological classes capable of identifying users in the denier category, but we found that the methodology for inference explored in this research does not function for the denier category the way it does for the green.

the calculation of the F1 scores. Referring back to Equations 5.31 and 5.32 for precision and recall and Equation 5.33 for the F1 score, the positive category is green for the first experimental run. For the second run, Weka effectively calculates two scores, $F1_G$ and $F1_D$, where the positive category is considered to be green and denier respectively.[14] A given F1 score for this run is then simply the weighted average of $F1_G$ and $F1_D$. This approach represents a more comprehensive evaluation of a classifier model as both classes are taken into account in a single reported metric.

7.3.1 Green-based Scores

We begin with the experiments using logistic regression. Table 7.1 lists the F1 scores with respect to the green class for this learner for datasets representing minimum-tweet user participation levels from 2 to 20. The columns in the table correspond to the different groups of modelled data elements as described in Section 7.2, with the "ALL" column representing datasets with attributes that are a combination of all these elements. Figure 7.3 displays these same results graphically with a different coloured line representing each group of data elements.

The F1 scores in Table 7.1 marked with a dagger (†) indicate that the logistic regression models for a dataset performed significantly better per a corrected, resampled T-Test (Nadeau & Bengio, 2003)[15] at a threshold of 0.05 as compared to models generated using a weak reference algorithm on that same dataset. Likewise, an asterisk (*) indicates that the algorithm performed worse than the reference at this same threshold. We have chosen the machine learning algorithm 1R (One Rule) for this reference.[16] The 1R

---

[14] $F1_G$ represents the same F1 score as was obtained in the first experimental run.

[15] As implemented via the Weka class *weka.experiment.PairedCorrectedTTester*, this adaptation of a statistical T-Test takes into account the variability of data in the training dataset, rather than simply calculating based on the variability in the test dataset. This adjustment avoids an overestimation of the performance of a model when using cross-validation to evaluate that model.

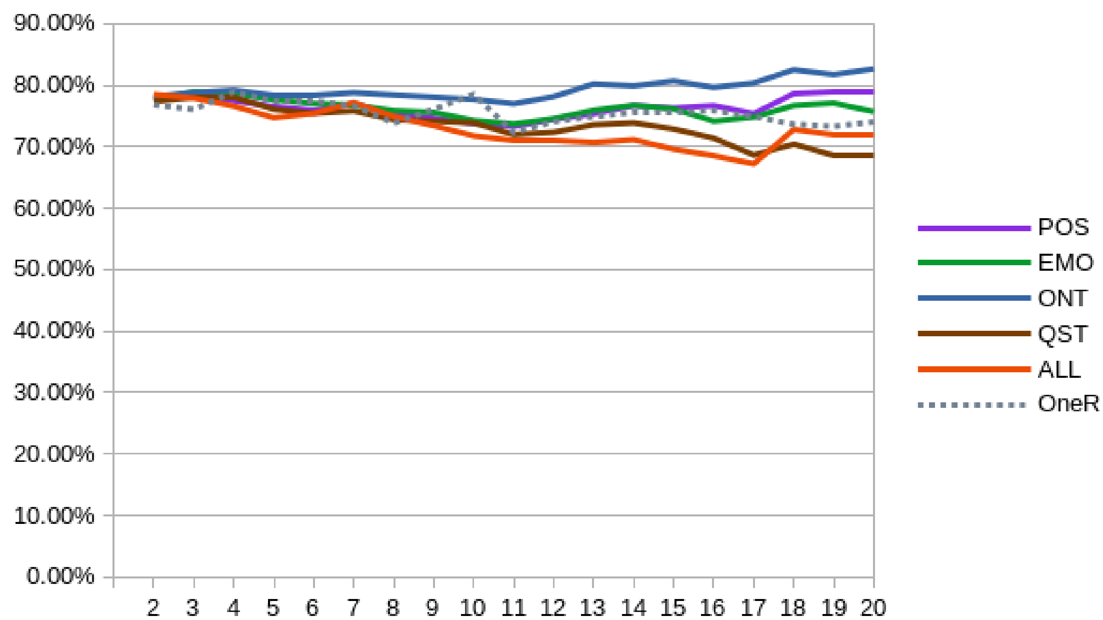
[16] We are using the Weka class *weka.classifiers.rules.OneR* with default parameters.

Table 7.1 Green-based F1 scores for models using logistic regression.

Min. Tweets	Logistic Regression Models POS	EMO	ONT	QST	ALL
2	77.83%	77.68%	78.16%	77.37%	78.49%†
3	78.07%†	78.95%	78.80%	77.97%	77.91%
4	77.49%	78.67%	79.22%	77.90%	76.61%
5	76.52%	77.66%	78.44%	76.11%	74.71%
6	75.93%	77.03%	78.40%	75.45%	75.43%
7	77.01%	76.79%	78.79%	75.82%	77.18%
8	75.33%	75.86%	78.40%	74.39%	74.93%
9	74.64%	75.58%	78.05%	74.18%	73.42%
10	73.78%*	74.27%	77.70%	73.87%	71.77%*
11	73.29%	73.69%	76.99%	71.97%	71.03%
12	74.45%	74.57%	78.15%†	72.35%	70.90%
13	75.38%	75.93%	80.22%†	73.57%	70.69%
14	76.69%	76.75%	79.89%	73.87%	71.14%
15	76.31%	76.16%	80.69%†	72.88%	69.59%
16	76.69%	74.15%	79.65%	71.39%	68.54%
17	75.38%	74.80%	80.36%†	68.65%	67.21%
18	78.64%	76.70%	82.51%†	70.42%	72.82%
19	78.91%	77.10%	81.71%†	68.46%	71.80%
20	78.85%	75.71%	82.67%†	68.62%	71.87%

† significantly better than 1R at a level of 0.05
∗ significantly worse than 1R at a level of 0.05

Figure 7.3 Charted green-based F1 scores for logistic regression models vs. 1R(ALL).



algorithm simply creates a rule based on each attribute in the dataset and then discards all of these rules except the one which gives a minimum error rate for classification. When considering 1R for a reference, we should note that it is not a completely trivial algorithm as it can map multiple ranges of the selected attribute and essentially performs as an extremely pruned decision tree (Holte, 1993). We might prefer the 0-R (Zero Rule) algorithm instead.[17] For classification problems, it simply always picks the majority class. However, 0-R will be problematic when we average green-based and denier-based F1 scores in Section 7.3.2 as we need a reference algorithm that is capable of choosing the denier category.[18]

There are a number of interesting observations regarding the results for logistic regression. First of all, we see that for lower levels of minimum participation, the various groups of data elements demonstrate a similar predictive capability with respect to the green category. However, as the minimum level of participation increases (i.e., users must have published more tweets to be included in the datasets), the predictive capability of these groups begins to vary markedly. The datasets representing counts of tweets with weak and strong indicators from the Six Americas (ONT) demonstrate the highest F1 scores for minimum participation levels of four tweets or greater. This result is important as it represents a confirmation by means of an independent process of the say-sila ontology's pertinence with respect to its predictive capability. Somewhat surprisingly, the ONT datasets outperform even those which contain all the data attributes (including the attributes from ONT). This finding also demonstrates that in this machine learning context we are seeing results that reflect our observations with the ontological model. During our experimentation with the ontology we were unable to find combinations of concept pairs, emotions, and parts of speech that led to increased predictive capability. Additionally, the relatively poor scores for the ALL datasets point to the importance

[17] *weka.classifiers.rules.ZeroR*

[18] We can report, however, that experiments using 0-R as the reference do not generally improve the green-based results presented here in Tables 7.1 and 7.2.

of incorporating a strategy for attribute selection when working with machine learning models.

It is also interesting that the datasets incorporating counts of the parts of speech (POS) for tokens in the tweets score relatively well, ending up as the second best performer for minimum participation levels of 15 tweets or greater. As mentioned above, the ontological model ignores the POS elements for logical inference of users in the green category (see Chapter 5). This is not to say, however, that the exploratory phase of our research did not seek to incorporate them. On the contrary, we spent an appreciable amount of time working with parts of speech in the ontological model, but our results with the survey concept indicators were ultimately more successful. Now that we have extracted these elements from the ontology for use in this machine learning phase of the research, we see a strong indication that they may indeed prove useful for determining stance on climate change. When a given user is one of the bigger players, factors such as writing style as reflected in the predominate parts of speech used for the words in his tweets may indeed be important elements in the model. This finding reinforces the value of the iterative methodology we have adopted throughout the course of this research.

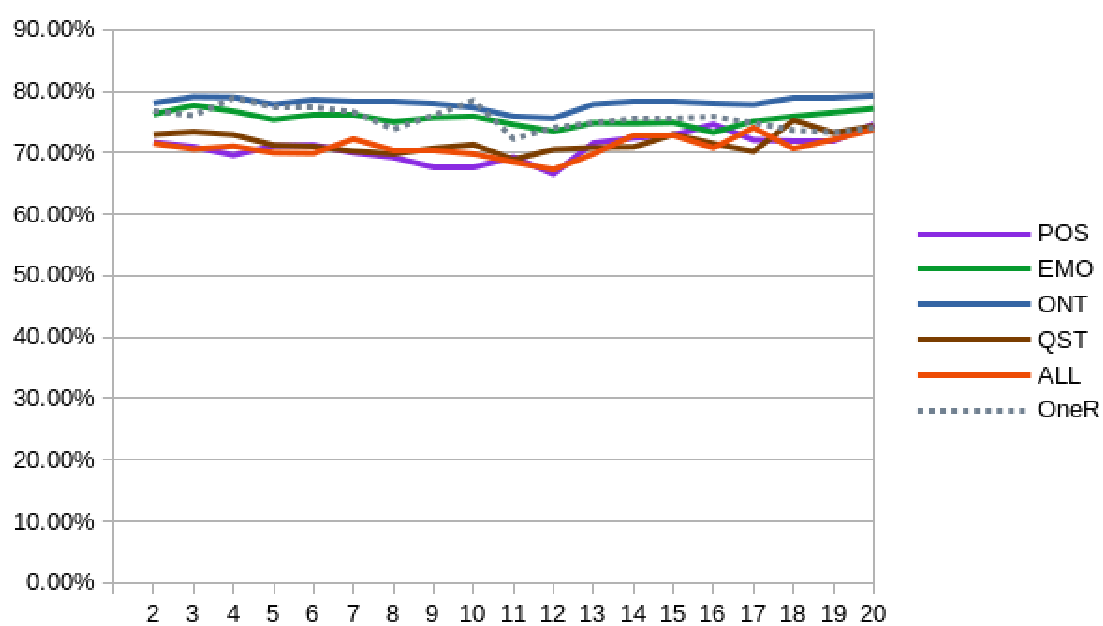
We must, however, be cautious when considering results at the highest levels of minimum participation as the low number of instances in the corresponding datasets works to increase a learner's tendency to overfit the data. The datasets for a minimum participation level of 20 tweets contain only 119 instances, and all the datasets for levels 14 and higher have less than 200 (see Table 5.5). Additionally, since we are using 10-fold cross-validation, 10% of the training instances are removed from the dataset before constructing the learner. The problem is particularly difficult in the datasets with more attributes (QST and ALL) as the small number of instances can do little to cover the high-dimensional feature space (Domingos, 2012). The logistic regression algorithm is using ridge regularization, which works to avoid overfitting, but we still see the performance falling off at $N = 15$ for these datasets with a high attribute count.

Table 7.2 Green-based F1 scores for models using decision trees.

Min. Tweets	Decision Tree Models POS	EMO	ONT	QST	ALL
2	71.73%*	76.26%*	78.09%	72.98%*	71.55%*
3	70.99%*	77.73%	79.11%	73.41%*	70.60%*
4	69.61%*	76.77%	79.04%	72.95%*	71.08%*
5	71.10%*	75.41%	77.88%	71.25%*	70.01%*
6	71.26%*	76.27%	78.62%	71.02%*	69.91%*
7	70.03%*	76.10%	78.21%	70.27%*	72.25%
8	69.24%*	75.03%	78.40%	69.82%	70.48%
9	67.56%*	75.72%	78.02%	70.75%	70.40%
10	67.65%*	75.92%	77.37%	71.35%	69.81%*
11	69.29%*	74.60%	75.92%	68.90%	68.50%
12	66.58%*	73.48%	75.63%	70.55%	67.30%
13	71.61%	74.83%	77.90%	70.85%	69.82%
14	72.43%	74.77%	78.31%	70.94%	72.96%
15	72.92%	74.91%	78.21%	73.01%	72.73%
16	74.64%	73.32%	78.03%	71.51%	70.80%
17	72.11%	75.14%	77.85%	70.23%	74.09%
18	71.81%	75.93%	79.06%	75.30%	70.70%
19	71.96%	76.57%	78.83%	73.25%	72.18%
20	74.64%	77.22%	79.28%	74.34%	73.78%

∗ significantly worse than 1R at a level of 0.05

Figure 7.4 Charted green-based F1 scores for decision tree models vs. 1R(ALL).



Moving to our second machine learning algorithm, we repeat this series of experiments, this time using decision trees. The datasets and the experimental method remain the same, the only difference is the learner. Table 7.2 lists the F1 scores with respect to the green category for the different groups of data elements. Models which performed significantly worse than 1R at a threshold of 0.05 per a corrected, resampled T-Test are marked with an asterisk. For this learner, none performed significantly better. Figure 7.4 displays these scores graphically. Although we are not improving on 1R, we might still remark that the overall results using decision trees are not too dissimilar to those for logistic regression. There are some notable differences, however. The F1 scores tend to be somewhat lower in general, but the most notable difference is that the minimum level of participation appears to be less of a defining factor when using decision trees. The F1 scores for a given group of data elements are more consistent as the minimum tweet count increases, and so the lines on the chart do not "fan out" as we move towards the right on the graph. The algorithm does not arrive at nearly the same level of predictive capability for all the data groups when testing at low minimum levels of participation. Rather, it appears to be utilizing the various data elements with differing degrees of success.

The ontological elements (ONT) continue to be the best performing data group with decision trees, just as they were for logistic regression. We see another similarity in that the algorithm does not show any marked improvement with the datasets that include all the attributes. Indeed, the ALL datasets show similar F1 scores to those containing parts of speech (POS) and those representing topic links to questions from the Six Americas (QST). We also note that the datasets with attributes for emotion and sentiment (EMO) allow this learner to achieve its second best scores in this round of experiments.

Finally, we do not see the same decrease in F1 scores at higher levels of minimum participation as occurred when using logistic regression. As explained above in Section 7.1.2, rather than using a regularization penalty term, the decision tree algorithm (by default) attempts to prune the tree in order to avoid overfitting the training dataset. To the extent that the decision trees are successful at this classification problem (given that they are not performing significantly better than 1R), the algorithm's default pruning parameters generally appear to be working to avoid overfitting.

7.3.2 Green-Denier Average Scores

For the second series of experiments we use the averaged F1 measure. To generate these scores, the Weka Experimenter evaluates F1 with respect to the green category as we reported above, then evaluates F1 again with respect to the denier category, and finally calculates the mean of these two values as the final F1 score for the model. Note that the experiments presented here are otherwise identical to those of the previous section. The datasets are the same, as are the two learners. The only difference is how the F1 metric is calculated,

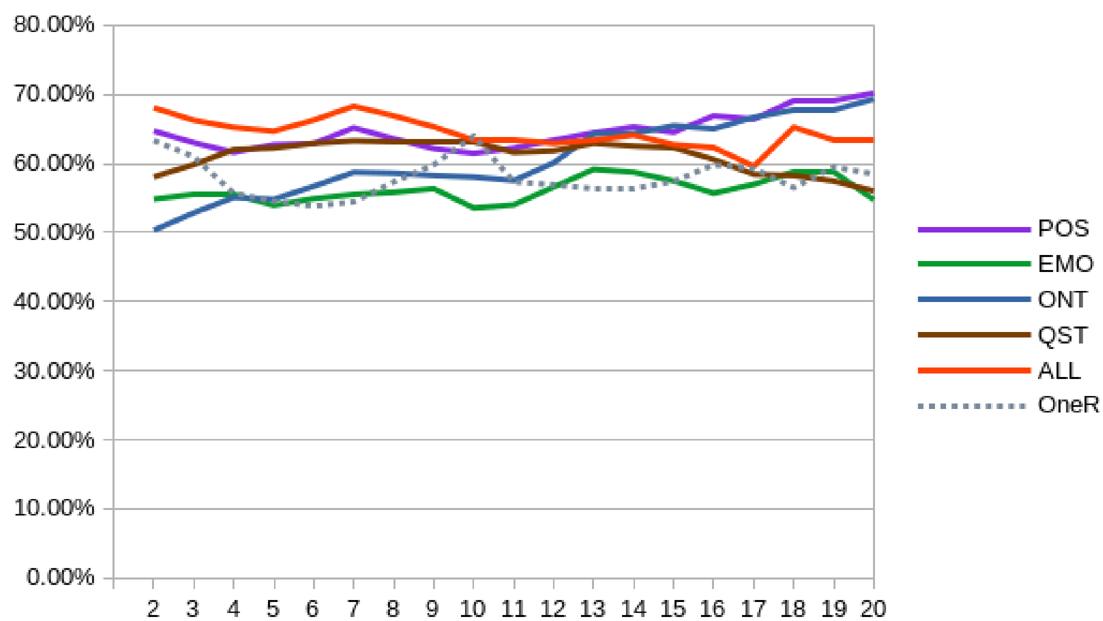
Table 7.3 reports the averaged F1 scores for the green and denier classes for logistic regression. As noted earlier, models marked with a dagger perform significantly better than the 1R reference algorithm according to a corrected, resampled T-Test(Nadeau & Bengio, 2003) at a threshold of 0.05. Of course, we see immediately that while more models are improving on 1R, the scores themselves are lower than the green-based F1 scores (see Table 7.1) for all datasets across all levels of minimum participation. This is not surprising for a number of reasons. As the F1 measure now represents a weighted average across the two classes, any lack in the model's ability to demonstrate fair levels of precision and recall for *either* of the two classes will be reflected in the averaged F1 score. Machine learning classification models typically perform better with respect to predictions for the majority class (Monard & Batista, 2002). In this case the denier category is the minority class. Therefore, we expect the model to demonstrate lower predictive capability for this class, and indeed we observe that performance for the denier class is effectively pulling the final F1 scores down as compared to what we obtained on the previous experimental run. Additionally, as we had difficulty using the ontological

Table 7.3 Averaged F1 scores for models using logistic regression.

Min. Tweets	Logistic Regression Models				
	POS	EMO	ONT	QST	ALL
2	64.70%	54.86%†	50.32%	58.06%†	68.06%†
3	63.01%	55.55%†	52.86%	59.84%†	66.25%†
4	61.61%†	55.62%†	55.08%	62.01%†	65.21%†
5	62.72%†	53.93%	54.83%	62.24%†	64.68%†
6	62.87%†	54.92%	56.71%	62.91%†	66.27%†
7	65.16%†	55.52%	58.72%†	63.26%†	68.28%†
8	63.57%	55.83%	58.58%†	63.21%	66.89%†
9	62.15%	56.37%	58.27%	63.11%†	65.31%
10	61.47%	53.56%	58.04%	63.16%†	63.53%
11	62.21%	54.02%	57.53%	61.60%	63.40%
12	63.39%	56.60%	60.10%	61.83%†	62.92%
13	64.43%	59.15%	64.31%†	62.91%	63.37%
14	65.29%	58.73%	64.35%	62.52%	64.10%
15	64.53%	57.48%	65.45%†	62.29%	62.73%
16	66.90%	55.70%	65.01%†	60.57%	62.30%
17	66.42%	56.98%	66.67%†	58.43%	59.59%
18	69.20%†	58.74%	67.87%†	58.26%	65.23%
19	69.03%†	58.84%	67.80%†	57.45%	63.40%
20	70.18%†	54.77%	69.32%†	56.03%	63.50%

† significantly better than 1R at a level of 0.05

Figure 7.5 Charted averaged F1 scores for logistic regression models vs. 1R(ALL).



model to infer users in the denier category (see Chapter 5), we should not be surprised now when the machine learning models also perform worse when predicting this class value. If nothing else, this series of experiments serves to further the idea that there may be some particular quality inherent in the users who publish denier tweets that merits exploration through continued research.

In Figure 7.5 we have the graph of the averaged F1 scores for logistic regression. We note right away that the lines do not fan out in the way they did with logistic regression in the previous section. Now that the predictive capability for the denier category is represented in the score, we no longer see all groups of data elements as being roughly equivalent at low levels of minimum participation with regard to the information the learner is able to utilize from them. This finding suggests that while there may be distinct characteristics associated with tweets from higher activity users in the green category, the same is likely not the case for those in the denier category. This hypothesis becomes altogether more interesting as we realize that while the learner was not able to leverage the ALL datasets particularly well for the green-based measure, now for minimum activity levels of ten tweets or less, logistic regression performs best when it can make use of all the available data elements. Further investigation is certainly warranted for a better understanding of how these data elements may be used to quantify certain aspects of style and expression in online microblogs and to link these aspects to an author's stance on climate change.

Looking at the data elements individually, we see that while the learner made best use of the ontological indicators (ONT) for green-based scoring, now the ONT datasets only show improved predictive capability for minimum-tweet activity levels of around 13 and higher. Furthermore, below the 13-tweet minimum, the ONT data elements generate the second worst scores until a participation level of four tweets or less where they generate the worst. The logistic regression algorithm continues to do relatively well with the parts of speech (POS) datasets, performing comparably to the ONT models and often surpassing them, even at higher levels of minimum activity where the learner is using the ONT data effectively. The sentiment and emotion (EMO) data elements provide comparatively poor results when both the green and denier categories are considered. This finding seems somewhat in contrast to the previous set of experiments where this group of data elements generated mid-range results for logistic regression.

In a similar manner as we saw for the green-based F1 scores for logistic regression (see Table 7.1 and Figure 7.3), we note some drop in the performance of this learner for the high-dimensional QST datasets with greater minimum levels of activity ($N \geq 15$), likely due to overfitting.[19] For the ALL datasets, the F1 score seems to fall and recover, and so it is not as clear if we are seeing the same degradation in the learner's ability to classify unseen instances when targeting both the green and denier categories as we did with the green-based scores.

Our last set of experiments involves the averaged F1 scoring for the decision tree algorithm. Table 7.4 lists the classification performance for these models. We see again that decision trees did not perform quite as well as logistic regression. We also note once again that the scores are lower for all datasets and all minimum-tweet activity levels as compared to the green-based measures for this learner (see Table 7.2). This is to be expected for the reasons we have stated above. The graph for this data is shown in Figure 7.6. As was the case for logistic regression, the decision tree algorithm is able to leverage the ALL datasets more effectively now that we are taking both the green and the denier categories into account. Indeed this data group generates the highest results fairly often over the entire range of minimum activity levels. Also, as we observed with logistic regression for the averaged F1 measure, we see that the scores for the datasets with ontological indicators (ONT) generally increase as the minimum-tweet participation level goes up. Here, however, the learner does not perform exceptionally well with

---

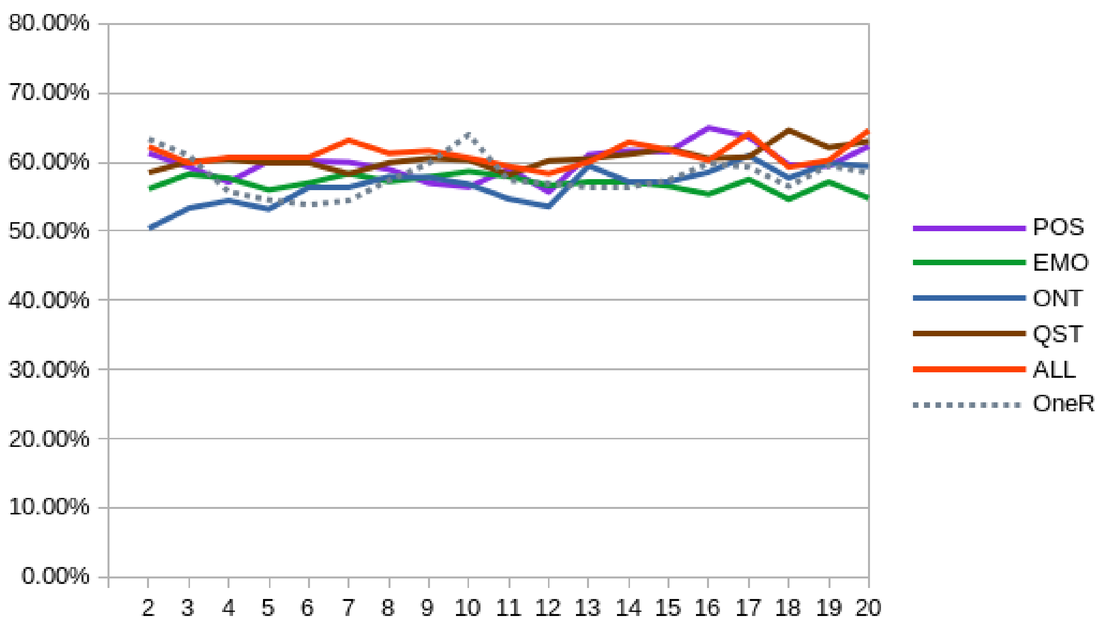
[19]We could likely improve our results by optimizing the ridge parameter (e.g., using Weka's *weka.classifiers.meta.GridSearch*); however, we have chosen to use default values for all hyperparameters in the present work in order to establish a baseline for these learners with respect to their use as modules in the machine learning level of *Say Sīla*. Conducting follow-up experiments with more tweets from more users would arguably be a better strategy for limiting overfitting in our baseline models at high levels of minimum participation. We could then proceed to studies that involve both the optimization of hyperparameters for these algorithms and the use of more sophisticated learners.

Table 7.4 Averaged F1 scores for models using decision trees.

Min. Tweets	Decision Tree Models POS	EMO	ONT	QST	ALL
2	61.30%	56.13%†	50.40%	58.48%†	62.23%
3	59.27%*	58.27%†	53.33%	60.08%†	59.86%
4	57.12%	57.66%†	54.44%	60.38%†	60.60%†
5	60.10%†	55.99%†	53.18%	59.93%†	60.66%†
6	60.27%†	56.96%†	56.37%	59.82%	60.57%†
7	59.97%†	58.35%†	56.28%	58.31%	63.15%†
8	58.96%	57.21%	57.88%	59.90%	61.28%
9	56.98%	57.83%	57.80%	60.55%	61.65%
10	56.39%*	58.64%	56.80%	60.30%	60.62%
11	59.01%	57.93%	54.67%	58.17%	59.40%
12	55.68%	56.60%	53.57%	60.20%	58.34%
13	61.06%	57.20%	59.47%	60.50%	60.01%
14	61.65%	57.14%	57.13%	61.13%	62.88%
15	61.51%	56.52%	57.05%	61.94%	61.75%
16	64.96%	55.35%	58.55%	60.52%	60.28%
17	63.62%	57.50%	60.94%	60.75%	64.13%
18	59.61%	54.63%	57.71%	64.59%	59.28%
19	59.45%	57.13%	59.82%	62.11%	60.25%
20	62.30%	54.81%	59.41%	63.00%	64.58%

† significantly better than 1R at a level of 0.05
∗ significantly worse than 1R at a level of 0.05

Figure 7.6 Charted averaged F1 scores for decision tree models vs. 1R(ALL).



the ontological indicators as compared to the other data groups. In general, the predictive capability of the learner varies less with respect to the different data elements than we have seen for the previous experiments. The algorithm appears to make better use of datasets with links to questions from the Six Americas (QST), but this may simply be the result of the general "evening out" of its performance across the various data elements.

Regarding the QST datasets, one might wonder how the learners performed significantly better than 1R for several minimum activity levels (see Tables 7.3 and 7.4), given that we did not have exceedingly good results with the initial check for grounding our method for linking tweets to survey questions as described in Section 6.4.2. This will be an interesting question to revisit regularly as we work to improve the methodology for identifying these links. A reasonable hypothesis is that although a link might not represent the question a human annotator would have chosen, it nevertheless represents represents a latent feature, which the machine learning algorithm can utilize for its prediction.[20] The idea might be compared to how untrained artificial neural networks, such as those used in extreme learning machines (Huang, 2015) or reservoir computing (Maass, 2002) utilize the "features" identified by their randomized connection weights. Of course, even if the algorithm can leverage inexact links between tweets and survey questions, we must understand that we ultimately loose the benefit of a clear box model for this part of the analysis.

The 1R Reference

As mentioned above, the datasets for the 1R algorithm (Holte, 1993), used as our reference in the significance test in Tables 7.1, 7.2, 7.3, and 7.4 match the algorithm we are testing (1R with ONT vs. logistic regression with ONT, etc.), However, the dotted line representing 1R (OneR) in the accompanying graphs in Figures 7.3, 7.4, 7.5, and

[20]This includes the reference algorithm 1R as can be seen in Table 7.5.

7.6 match the results for 1R with the ALL dataset for each level of minimum activity. Since 1R always selects the attribute which allows for a rule that minimizes classification error, using the ALL dataset shows the best results possible using this reference algorithm for the attributes under consideration. Due to its simplicity, we have chosen 1R as a reference for our two baseline algorithms (with the intention that they, in turn, serve as a reference for more sophisticated algorithms in future research relating to the *Say Sīla* architecture). Hence, the 1R results for the ALL dataset provide the best reference performance possible for a clear visual comparison in the graphs.

Table 7.5 Chosen attribute and averaged F1 scores for 1R(ALL) reference.

Min. Tweets	Attribute	Dataset	Reference F1 Score
2	Address	POS	63.32%
3	Interjection	POS	60.98%
4	T9	QST	55.71%
5	T9	QST	54.53%
6	T9	QST	53.84%
7	T9	QST	54.44%
8	T9	QST	57.34%
9	AtMention	POS	59.81%
10	AtMention	POS	63.97%
11	AtMention	POS	57.35%
12	AtMention	POS	56.92%
13	AtMention	POS	56.29%
14	T9	QST	56.40%
15	AtMention	POS	57.41%
16	T9	QST	59.81%
17	T9	QST	59.27%
18	T9	QST	56.50%
19	T9	QST	59.49%
20	T9	QST	58.47%

Of course, in addition to serving as a reference, 1R allows for a quick check to determine which attribute best merits an analysis if we could pick only one. Accordingly, Table 7.5 reports the averaged F1 measure for 1R as displayed in Figures 7.5, and 7.6 along with the attribute that 1R has determined produces a minimum error for classification.[21]

[21] As our principal focus in this chapter is to analyze the baseline performance given by logistic regression and decision trees, rather than analyzing the reference algorithm used to evaluate that baseline, we are not including F1 scores for the individual POS, EMO, ONT, and QST datasets; nor are we including the scores for 1R for the green-based models from Section 7.3.1.

When consulting the table, the reader should recall that *Address* represents URLs and email addresses, *AtMention* is when the author indicates another user in the tweet (see Table 4.1), and T9 is the question about "Perceptions of Weather and Climate" from the Six Americas survey (see Table 6.1).

Selected Models

As mentioned above, in addition to serving as a baseline when studying more sophisticated algorithms in future research efforts, logistic regression and decision trees provide a clear box model, allowing us to understand why the model has chosen a specific category for a given data instance. Of course, our goal in this chapter is not to attempt a detailed interpretation of 190 machine learning models (2 algorithms $\times$ 5 dataset categories $\times$ 19 minimum levels of participation). Our work here, rather, is about feeding results from the description logic level of the $\mathcal{S}ay\ \mathcal{S}\vec{\imath}La$ architecture back into the machine learning level and improving upon it. Due to the scope of this doctoral program, we are presenting the results from an initial round of experiments using a prototype of this part of the architecture, aided by the Weka Experimenter tool. As part of our ongoing research, we will continue working to integrate the methodology presented here into the architecture itself, endeavouring to have it learn incrementally as information flows through the various levels of the system.

When considering explainable models as they relate to subsequent versions of $\mathcal{S}ay\ \mathcal{S}\vec{\imath}La$, one might question how valuable they really will be in an architecture which simply uses them as an intermediate node, responsible for detecting a specific feature that is to be incorporated in an ontological model. Even when serving as a modular component, however, the value of an explainable model remains. There is an obvious reason in that for the entire architecture to be clear box, each of its components must constitute a visible model. There is also the potential benefit that as we incrementally work on different components of the architecture, having clarity in modules which proceed a given component will aid with the design both of the new component and of the experiments it allows us to perform. Finally, once the architecture has obtained a given goal state with respect to a specific research project, the modular aspect of our design should allow us to substitute more sophisticated algorithms whenever classification performance becomes more important than explainability.

Table 7.6 Logistic regression model for the POS dataset at a 5-tweet minimum.

POS Attr. Weight	**Green**
ExistentialPlusVerbal	45.805
NominalPossessive	0.256
Preposition	0.182
Continuation	0.174
Adjective	0.120
Determiner	0.114
Address	0.082
CoordinatingConjunction	0.055
Hashtag	0.029
Verb	0.022
Count	0.007
Numeral	-0.004
AtMention	-0.017
Punctuation	-0.053
CommonNoun	-0.065
Emoticon	-0.070
Other	-0.096
VerbParticle	-0.098
NominalVerbal	-0.110
ProperNounPlusPossessive	-0.117
Adverb	-0.134
ProperNoun	-0.144
Pronoun	-0.153
ExistentialThere	-0.457
Interjection	-0.744
Intercept	0.853

Hence, we have selected a number of models generated in the experiments in this section as an example of how one may peek in on specific components within the machine learning level of *Say Sila*. As the most interesting models will be those shown to perform better than our reference, we have selected ones for each of the four data categories where the lowest level of minimum participation produces a logistic regression model and a decision tree model that both demonstrate significant improvement over 1R.[22] For

[22]This is the lowest minimum tweet level where the F1 scores for both models are tagged with a

the POS (part of speech) data, Table 7.6 shows the weight coefficients for the logistic regression model at a minimum participation level of five tweets. Figure 7.7 shows the J48 decision tree for the same dataset.

Table 7.7 Logistic regression model for the EMO dataset at a 2-tweet minimum.

EMO Attr. Weight	**Green**
Fear	0.362
Positive	0.255
Anticipation	0.061
Count	0.003
Sadness	-0.030
Joy	-0.095
Trust	-0.148
Anger	-0.153
Disgust	-0.194
Negative	-0.204
Surprise	-0.245
Intercept	0.573

Table 7.7 shows the coefficients of the logistic regression model for the EMO (emotion) dataset at a minimum participation level of two tweets. Figure 7.8 displays the corresponding decision tree model.

Table 7.8 Logistic regression model for the ONT dataset at a 7-tweet minimum.

EMO Attr. Weight	**Green**
WeakEconomicGrowthText	78.277
StrongNatureCauseText	1.348
StrongCO2CutText	0.989
Count	0.004
WeakNatureCauseText	-0.187
WeakHumanCauseText	-0.243
WeakCO2CutText	-0.919
StrongHumanCauseText	-1.107
StrongEconomicGrowthText	-62.037
Intercept	0.633

For the ONT (ontology) data category, which contains the count of tweets that users published with weak and strong indicators (see Section 5.1.3), Table 7.8 contains the weighted coefficients for the logistic regression model at a minimum participation level

dagger (†) in Tables 7.3 and 7.4.

of seven tweets. Unfortunately, we cannot select an accompanying decision tree model as the decision tree algorithm did not perform significantly better than the 1R reference for any minimum tweet level for this group of datasets.

Table 7.9 Logistic regression model for the QST dataset at a 2-tweet minimum.

QST Attr. Weight	Green
T11	0.718
T19	0.534
T14	0.233
T30	0.195
T8a	0.186
T27	0.171
T5	0.125
T29	0.121
T22	0.111
T23	0.088
T25	0.081
T6	0.056
T18	0.043
T8b	0.024
T16	0.019
T15	0.016
T7	-0.007
Count	-0.015
T24	-0.054
T17	-0.061
T2	-0.069
T4	-0.079
T3	-0.096
T26	-0.144
T10	-0.175
T28	-0.181
T13	-0.195
T21	-0.245
T9	-0.275
T12	-0.299
T20	-0.448
T31	-0.662
Intercept	0.524

Our last data grouping is QST (questions), which contains counts of users' tweets linked to survey questions from the Six Americas using the information retrieval techniques discussed in Section 6.1. Table 7.9 describes the logistic regression model for this data group at a minimum participation level of two tweets, while Figure 7.9 displays the decision tree for this same dataset. Again, we are not attempting to perform a detailed

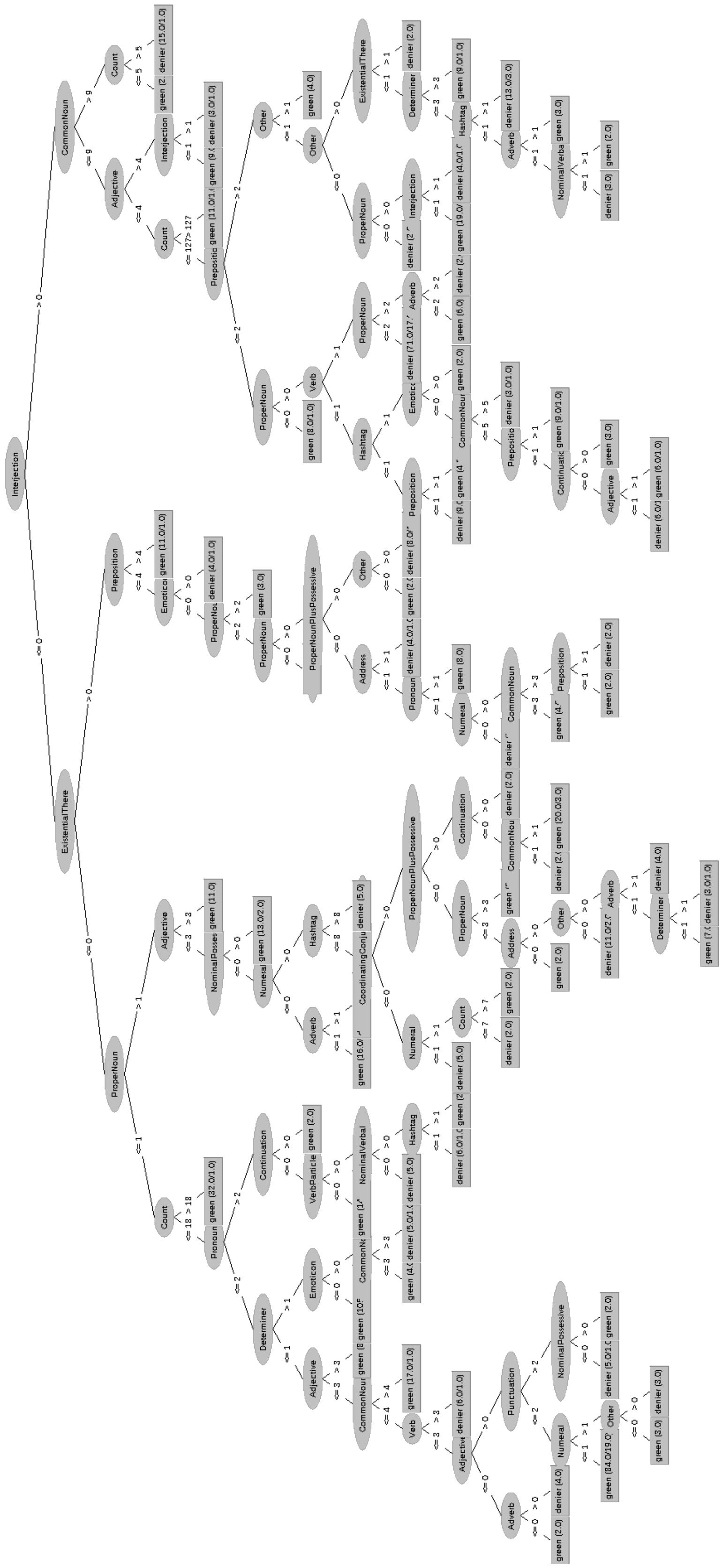


Figure 7.7 Decision tree model for the POS dataset at a 5-tweet minimum.

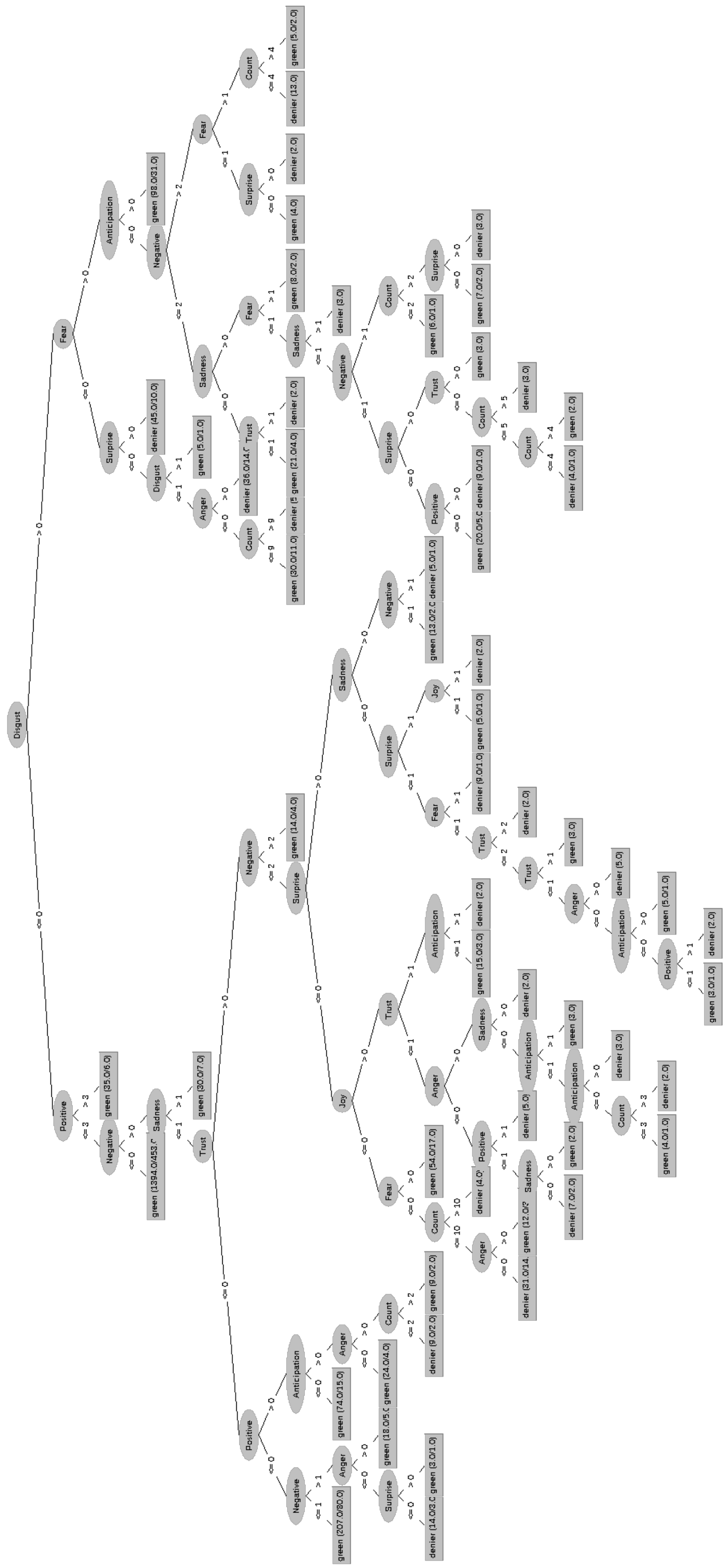


Figure 7.8 Decision tree model for the EMO dataset at a 2-tweet minimum.

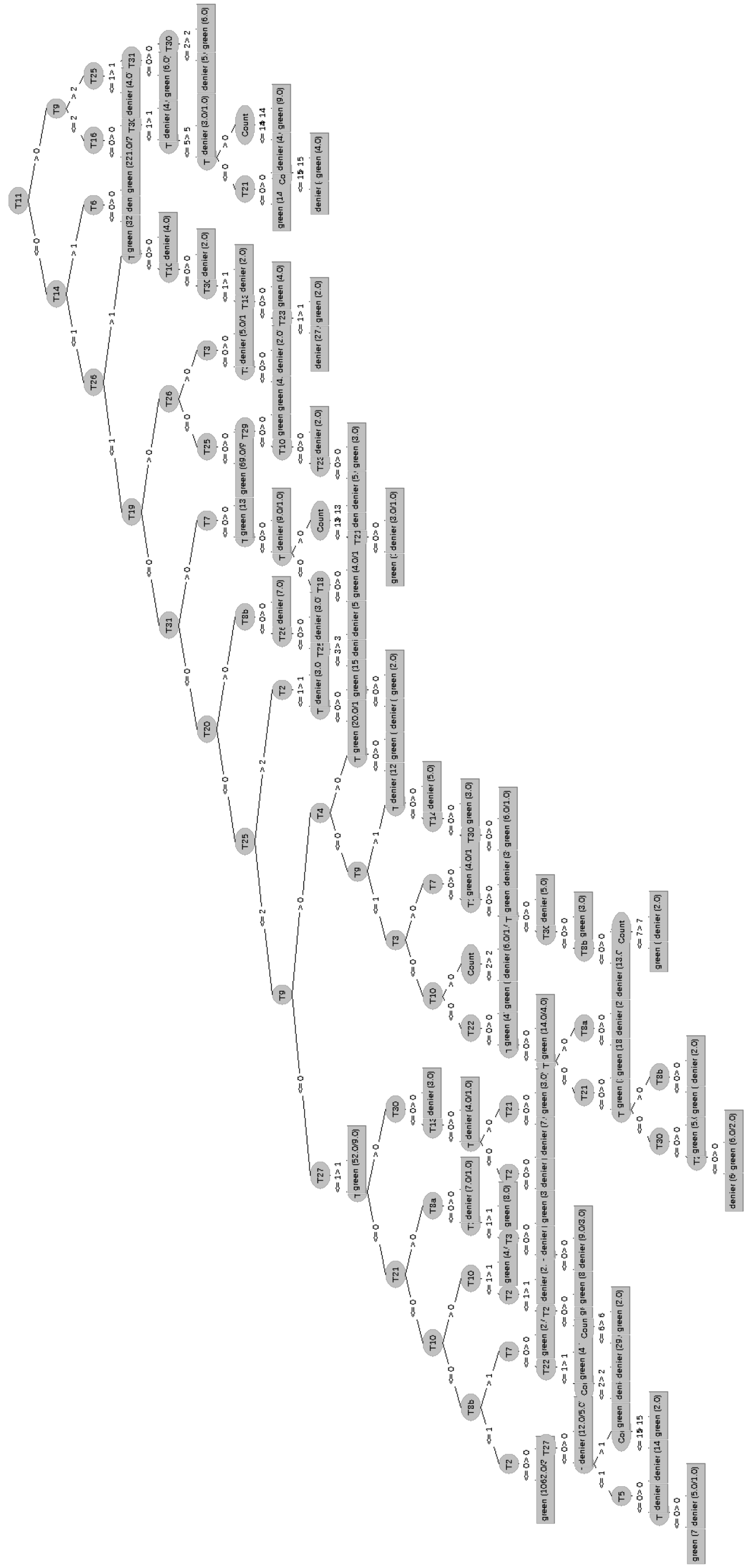


Figure 7.9 Decision tree model for the QST dataset at a 2-tweet minimum.

interpretation of these models, and we have no reason to expect that this sampling of models is indicative of general trends across different levels of minimum participation. Our primary concern here is how machine learning models can receive and process information from the description logic level in the context of the $\mathcal{S}ay\ \mathcal{S}\vec{\imath}la$ architecture.

Collinearity

One important factor to consider when creating these types of machine learning models is the possibility of collinearity between individual attributes in the datasets. Ideally, all attributes would be independent. Absent a significant level of collinearity in the data, a learner will tend to generate a model that is stable and relatively easy to interpret (assuming a clear-box algorithm). However, when strong correlations exist between data attributes, this may no longer be the case. Small changes in the value of one data input may potentially cause wild and unexpected changes in the model's output.

To test for collinearity in the datasets we are using in this chapter, we performed a test using a technique known as Principal Components Analysis (PCA).[23] PCA is a commonly used technique in machine learning to reduce dimensionality by linearly transforming the original dataset attributes to an new set of attributes. The original data values are essentially projected onto a new set of axes which have the property of being orthogonal to each other. The original dataset is represented using a correlation or a covariance matrix, and the eigen vectors of this matrix correspond to the new axes representing the transformed attributes. The linear combination of largest eigenvalue and its associated eigenvector is the first principal component, and it covers the largest percentage of variance in the data. Subsequent principal components will cover increasingly smaller percentages of variance (Jolliffe & Cadima, 2016; Goodfellow et al., 2016).

To have a clearer picture of collinearity between attributes in our datasets, we can consult the correlation matrix used by the PCA algorithm. Table 7.10 shows the matrix

[23] We used the Weka filter *weka.attributeSelection.PrincipalComponents* to conduct the analysis.

Table 7.10 Analysis of collinearity for the POS dataset at a 2-tweet minimum

Com. Noun	Pron.	Nom. Poss.	Prop. Noun	Prop. Noun Plus Poss.	Nom. Verb.	Prop. Noun Plus Verb.	Verb	Adj.	Adv.	Int.	Det.	Crd. Conj.	Verb Part.	Exist. There	Exist. Plus Verb.	Prep.	Hash.	Addr.	Emo.	At Ment.	Cont.	Num.	Pnct.	Cnt.	
1.00	0.20	0.02	0.20	0.11	0.09	**0.56**	0.43	0.24	0.04	**0.56**	0.36	0.12	0.09	0.03	**0.65**	-0.15	0.10	-0.03	0.03	0.07	0.27	**0.51**	0.05	-0.02	CommonNoun
	1.00	0.04	0.04	0.03	0.22	0.50	0.18	0.33	0.14	0.29	0.31	0.14	0.12	0.03	0.30	-0.08	-0.11	0.06	0.00	0.05	0.11	0.28	0.04	0.04	Pronoun
		1.00	0.04	0.12	0.09	0.03	0.02	0.05	0.06	0.06	0.05	0.13	0.09	-0.01	0.03	-0.02	0.09	0.02	0.06	0.07	0.07	0.04	0.08	0.14	NominalPossessive
			1.00	0.13	0.08	0.17	0.11	0.10	0.04	0.14	0.20	0.12	0.08	0.00	0.20	-0.03	0.08	0.02	0.02	0.04	0.25	0.24	0.07	0.05	ProperNoun
				1.00	0.13	0.05	0.10	0.08	0.05	0.09	0.11	0.16	0.14	-0.01	0.11	-0.03	0.11	0.03	0.00	0.08	0.20	0.08	0.09	0.15	ProperNounPlusPossessive
					1.00	0.21	0.14	0.26	0.22	0.14	0.21	0.19	0.15	0.00	0.14	-0.02	-0.05	0.05	0.03	0.05	0.11	0.14	0.07	0.11	NominalVerbal
						1.00	0.38	0.43	0.10	**0.53**	0.39	0.20	0.15	0.03	**0.61**	-0.15	-0.02	0.03	0.04	0.05	0.20	0.48	0.04	0.00	ProperNounPlusVerbal
							1.00	0.22	0.06	0.35	0.29	0.12	0.08	-0.01	0.43	-0.13	0.03	-0.01	0.01	0.02	0.18	0.37	0.00	-0.02	Verb
								1.00	0.16	0.27	0.30	0.19	0.14	0.02	0.30	-0.09	-0.05	0.04	0.00	0.07	0.18	0.29	0.07	0.06	Adjective
									1.00	0.09	0.12	0.17	0.21	-0.01	0.03	-0.05	-0.06	0.02	0.03	0.09	0.05	0.14	0.10	0.11	Adverb
										1.00	0.32	0.17	0.15	0.03	**0.55**	-0.15	-0.01	0.00	0.07	0.05	0.14	0.36	0.03	-0.01	Interjection
											1.00	0.19	0.14	0.01	0.34	-0.08	-0.02	0.02	0.03	0.06	0.20	0.31	0.08	0.03	Determiner
												1.00	0.16	-0.01	0.14	-0.05	0.04	0.05	0.04	0.07	0.14	0.15	0.10	0.09	CoordinatingConjunction
													1.00	-0.01	0.09	-0.02	0.02	0.07	0.03	0.06	0.11	0.12	0.06	0.16	VerbParticle
														1.00	0.03	0.00	-0.02	-0.01	0.01	-0.01	0.01	0.01	0.01	0.00	ExistentialThere
															1.00	-0.16	0.05	-0.02	0.04	0.02	0.26	0.39	0.01	-0.03	ExistentialPlusVerbal
																1.00	0.12	0.02	-0.03	0.04	-0.07	-0.05	0.08	0.00	Preposition
																	1.00	-0.03	0.00	0.07	0.12	0.04	0.07	0.04	Hashtag
																		1.00	0.03	0.08	0.02	0.02	0.07	0.02	Address
																			1.00	0.06	0.02	0.05	0.05	-0.01	Emoticon
																				1.00	0.05	0.09	0.19	0.08	AtMention
																					1.00	0.26	0.08	0.06	Continuation
																						1.00	0.09	-0.01	Numeral
																							1.00	0.05	Punctuation
																								1.00	Count

for the part of speech (POS) dataset. The eleven attributes in the dataset run along the top of the table and again down the right side. A given value in the table represents the correlation in the original data between the attribute heading that column and the attribute named at the end of the row. Ideally, we would like these values to be as close to zero as possible.[24] Values in bold font in the table indicate a particularly strong correlation of greater than 0.5 (or a strong inverse correlation of below -0.5). We see just a few pairs of attributes with a notably high correlation in the POS dataset These likely represent aspects of common linguistic usage such as a common nouns joined with a numeral or an interjection.

In Table 7.11 we have the correlation matrix for the sentiment and emotion (EMO) dataset. We note that there are several more strongly correlated attributes than we had with the POS dataset. While this is not optimal, it is not necessarily surprising when we note which attributes show a relatively high correlation. Anger, as one example, is strongly correlated with negative sentiment and with fear and disgust. Joy, as another, shows a high level of correlation with positive sentiment and with trust. We can expect that these sorts of combinations of affect will occur frequently in social media posts for a given theme in a message.

Table 7.12 presents the PCA correlation matrix for the ontological indicators (ONT) dataset. The few pairs of highly correlated attributes are also not surprising here as they represent the weak and strong text indicators for a given survey concept pair. Since the strong indicators each form a subclass of their corresponding weak indicator, we can be certain these attributes will demonstrate a level of collinearity. We should note that although the weak and strong nature–cause indicators are not in bold, the 0.49 correlation score just misses our threshold.

The datasets up to now have each shown a relatively small number of strongly correlated

[24] The values along the diagonal will always be 1.0 as an attribute will always correlate perfectly with itself.

Table 7.11 Analysis of collinearity for the EMO dataset at a 2-tweet minimum

Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Negative	Positive	Count	
1.00	0.19	**0.63**	**0.51**	0.16	0.50	0.32	0.15	**0.64**	0.15	-0.01	**Anger**
	1.00	0.12	0.18	0.46	0.14	0.35	0.37	0.19	0.38	0.03	**Anticipation**
		1.00	0.38	0.12	**0.51**	0.26	0.12	**0.57**	0.09	-0.02	**Disgust**
			1.00	0.09	0.50	0.25	0.11	**0.60**	0.12	0.02	**Fear**
				1.00	0.11	0.36	**0.60**	0.14	**0.59**	0.02	**Joy**
					1.00	0.25	0.12	**0.67**	0.10	0.02	**Sadness**
						1.00	0.26	0.25	0.23	0.03	**Surprise**
							1.00	0.19	**0.66**	0.01	**Trust**
								1.00	0.17	0.00	**Negative**
									1.00	-0.01	**Positive**
										1.00	**Count**

Table 7.12 Analysis of collinearity for the ONT dataset at a 2-tweet minimum

Weak Human Cause Text	Weak CO2 Cut Text	Weak Nature Cause Text	Weak Economic Growth Text	Strong Human Cause Text	Strong CO2 Cut Text	Strong Nature Cause Text	Strong Economic Growth Text	Count	
1.00	0.11	0.23	0.04	**0.60**	0.12	0.09	0.03	0.13	**WeakHumanCauseText**
	1.00	0.01	0.01	0.07	**0.73**	0.02	0.02	0.07	**WeakCO2CutText**
		1.00	0.02	0.18	0.03	0.49	0.03	0.10	**WeakNatureCauseText**
			1.00	0.05	0.03	-0.01	**0.84**	0.04	**WeakEconomicGrowthText**
				1.00	0.07	0.03	0.03	0.08	**StrongHumanCauseText**
					1.00	0.04	0.04	0.09	**StrongCO2CutText**
						1.00	0.00	0.03	**StrongNatureCauseText**
							1.00	0.04	**StrongEconomicGrowthText**
								1.00	**Count**

Table 7.13 Analysis of collinearity for the QST dataset at a 2-tweet minimum

T2	T3	T4	T5	T6	T7	T8a	T8b	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30	T31	Cnt	
1.00	**0.56**	0.04	**0.77**	0.11	**0.51**	**0.83**	0.44	**0.79**	0.14	0.06	**0.82**	**0.73**	**0.73**	0.39	0.32	0.08	**0.61**	0.28	**0.77**	**0.72**	**0.84**	0.48	0.05	**0.56**	0.08	**0.81**	0.30	0.41	0.21	**0.72**	**0.79**	**T2**
	1.00	0.17	**0.73**	0.45	**0.70**	**0.66**	**0.78**	**0.74**	0.46	0.34	0.47	**0.66**	**0.70**	0.34	0.25	0.22	**0.67**	**0.66**	0.38	**0.76**	**0.72**	**0.68**	0.19	**0.79**	0.22	**0.55**	**0.74**	**0.62**	0.25	0.41	**0.83**	**T3**
		1.00	0.21	0.02	0.04	0.08	0.04	0.14	0.05	0.12	0.01	0.09	0.12	0.16	0.03	0.03	0.12	0.02	0.01	0.21	0.07	0.13	0.10	0.19	0.07	0.06	0.01	0.05	0.01	0.05	0.14	**T4**
			1.00	0.09	**0.76**	**0.92**	**0.52**	**0.93**	0.20	0.11	**0.79**	**0.83**	**0.71**	0.42	0.31	0.06	**0.59**	**0.54**	**0.71**	**0.78**	**0.89**	**0.71**	0.11	**0.69**	0.11	**0.82**	**0.56**	**0.52**	0.27	**0.66**	**0.89**	**T5**
				1.00	0.02	0.04	**0.57**	0.10	0.37	0.23	0.00	0.08	0.40	0.07	0.09	0.25	0.47	0.18	0.01	0.34	0.18	0.04	0.18	0.30	0.18	0.03	0.27	0.26	0.03	0.06	0.27	**T6**
					1.00	**0.70**	0.44	**0.77**	0.24	0.09	0.41	**0.64**	0.36	0.25	0.16	0.05	0.24	**0.79**	0.29	**0.51**	**0.61**	**0.85**	0.07	**0.58**	0.10	**0.51**	**0.84**	**0.53**	0.28	0.30	**0.68**	**T7**
						1.00	0.47	**0.94**	0.14	0.05	**0.90**	**0.85**	**0.76**	0.43	0.34	0.04	**0.62**	0.42	**0.83**	**0.78**	**0.95**	**0.65**	0.05	**0.66**	0.11	**0.89**	0.44	0.47	0.27	**0.76**	**0.90**	**T8a**
							1.00	**0.54**	0.45	0.25	0.35	0.47	**0.67**	0.25	0.22	0.25	**0.66**	0.48	0.30	**0.66**	**0.58**	0.43	0.20	**0.66**	0.21	0.40	**0.58**	**0.51**	0.18	0.33	**0.68**	**T8b**
								1.00	0.22	0.10	**0.82**	**0.87**	**0.75**	0.44	0.32	0.08	**0.61**	**0.53**	**0.73**	**0.79**	**0.93**	**0.73**	0.13	**0.74**	0.13	**0.83**	**0.56**	**0.54**	0.29	**0.68**	**0.92**	**T9**
									1.00	0.22	0.04	0.23	0.29	0.19	0.07	0.19	0.29	0.31	0.03	0.32	0.21	0.30	0.15	0.43	0.22	0.21	0.35	0.32	0.11	0.07	0.34	**T10**
										1.00	0.01	0.15	0.20	0.14	0.04	0.16	0.49	0.12	0.01	0.21	0.14	0.18	0.11	0.37	0.18	0.11	0.12	0.24	0.05	0.08	0.25	**T11**
											1.00	**0.77**	**0.79**	0.43	0.37	0.01	**0.66**	0.12	**0.94**	**0.74**	**0.93**	0.40	0.01	**0.55**	0.02	**0.90**	0.12	0.33	0.20	**0.85**	**0.81**	**T12**
												1.00	**0.69**	0.49	0.30	0.07	**0.58**	0.42	**0.70**	**0.73**	**0.84**	**0.67**	0.18	**0.73**	0.21	**0.80**	0.42	0.49	0.27	**0.68**	**0.85**	**T13**
													1.00	0.42	0.36	0.19	**0.85**	0.20	**0.76**	**0.85**	**0.87**	0.38	0.14	**0.69**	0.18	**0.76**	0.26	0.46	0.19	**0.73**	**0.85**	**T14**
														1.00	0.20	0.09	0.35	0.11	0.41	0.43	0.45	0.36	0.11	**0.57**	0.19	0.44	0.09	0.26	0.14	0.43	0.50	**T15**
															1.00	0.04	0.32	0.07	0.35	0.33	0.38	0.16	0.04	0.26	0.04	0.34	0.08	0.16	0.08	0.32	0.47	**T16**
																1.00	0.22	0.10	0.03	0.22	0.10	0.10	0.10	0.22	0.21	0.07	0.10	0.17	0.03	0.06	0.16	**T17**
																	1.00	0.15	**0.65**	**0.77**	**0.76**	0.27	0.14	**0.65**	0.19	**0.66**	0.21	0.43	0.15	**0.63**	**0.77**	**T18**
																		1.00	0.02	0.40	0.37	**0.70**	0.14	0.45	0.13	0.23	**0.90**	**0.51**	0.22	0.05	0.49	**T19**
																			1.00	**0.69**	**0.86**	0.29	0.01	0.48	0.05	**0.87**	0.00	0.27	0.17	**0.83**	**0.73**	**T20**
																				1.00	**0.86**	**0.52**	0.16	**0.73**	0.18	**0.75**	0.42	**0.51**	0.23	**0.68**	**0.87**	**T21**
																					1.00	**0.58**	0.09	**0.73**	0.11	**0.90**	0.40	**0.50**	0.26	**0.80**	**0.94**	**T22**
																						1.00	0.13	**0.72**	0.29	**0.52**	**0.71**	**0.55**	0.28	0.33	**0.72**	**T23**
																							1.00	0.21	0.12	0.04	0.11	0.13	0.04	0.04	0.18	**T24**
																								1.00	0.36	**0.61**	0.49	**0.58**	0.26	**0.52**	**0.83**	**T25**
																									1.00	0.18	0.11	0.24	0.07	0.12	0.23	**T26**
																										1.00	0.23	0.41	0.23	**0.80**	**0.85**	**T27**
																											1.00	**0.53**	0.23	0.03	**0.53**	**T28**
																												1.00	0.19	0.30	**0.60**	**T29**
																													1.00	0.17	0.33	**T30**
																														1.00	**0.72**	**T31**
																															1.00	**Count**

attributes, but Table 7.13, which reports the correlation matrix for the QST dataset, is potentially a bit more alarming. Here we see many pairs of attributes with correlation scores in bold as they exceed the 0.5 threshold. This may well indicate a significant problem with this dataset, but as we consider the objective associated with the dataset, we should perhaps not be too surprised. This objective was determine which of 31 survey questions on the common theme of global warming to link a tweet. In Section 6.4.2 we discussed a considerable limitation of the current architecture in reference to this goal. Additionally, even as the system improves its ability to link tweets and survey questions, we should expect that the fact that as these are posts on a common theme, and many survey questions are similar in nature, we will likely continue to see a fairly strong degree of collinearity among these attributes.[25]

We also generated a correlation matrix for the ALL dataset which includes all of the attributes from the four datasets we have just examined. The analysis on this dataset allows us to detect collinearity between the groups of attribute types. We have not included this correlation matrix in the present document as the large number of attributes in the dataset makes it difficult to fit the table on a single page. We can report, however, that we did not find any notable correlation between pairs of attributes across the different types. The one exception is the common "Count" attribute, which represents the number of *#globalwarming* tweets published by a user in 2019. This attribute was strongly correlated with several of the QST attributes. Of course, these correlation scores can also be seen in Table 7.13 which covers the QST dataset.

To complete our analysis of collinearity in our datasets, we reran our comparison of the performance of logistic regression and decision trees against the 1R reference algorithm for the ALL dataset. This time, however, we are using the transformed dataset resulting from Weka's *PrincipalComponents* filter.[26] A dataset transformed via PCA has the

[25]In our continued research we will be interested in experimenting with how the predictive feedback signal from the higher level in the hybrid architecture may help to mitigate this type of issue.

[26]We use the default parameters, which standardize the data attributes, perform PCA using a cor-

Table 7.14 Learner comparison between original data and PCA for ALL dataset.

Min Tweets	Original Data 1-Rule	Original Data Logistic Regression	Original Data Decision Trees	PCA Transformed 1-Rule	PCA Transformed Logistic Regression	PCA Transformed Decision Trees
2	63.32%	68.06%†	62.23%	55.15%	65.96%†	60.75%†
3	60.98%	66.25%†	59.86%	56.55%	63.87%†	60.11%†
4	55.71%	65.21%†	60.60%†	54.96%	62.60%†	60.30%†
5	54.53%	64.68%†	60.66%†	55.16%	62.95%†	56.76%
6	53.84%	66.27%†	60.57%†	59.30%	64.70%	57.31%
7	54.44%	68.28%†	63.15%†	56.26%	66.23%†	57.03%
8	57.34%	66.89%†	61.28%	54.27%	66.35%†	55.65%
9	59.81%	65.31%	61.65%	56.11%	63.70%	57.94%
10	63.97%	63.53%	60.62%	56.94%	63.26%	56.23%
11	57.35%	63.40%	59.40%	53.12%	62.50%†	53.21%
12	56.92%	62.92%	58.34%	56.34%	62.90%	54.48%
13	56.29%	63.37%	60.01%	57.48%	63.86%	60.29%
14	56.40%	64.10%	62.88%	54.54%	62.91%	58.64%
15	57.41%	62.73%	61.75%	55.13%	62.24%	58.27%
16	59.81%	62.30%	60.28%	57.35%	62.82%	57.41%
17	59.27%	59.59%	64.13%	50.80%	61.37%	54.96%
18	56.50%	65.23%	59.28%	55.93%	63.15%	58.14%
19	59.49%	63.40%	60.25%	58.51%	62.17%	56.15%
20	58.47%	63.50%	64.58%	63.12%	63.07%	55.01%

† significantly better than 1R at a level of 0.05

advantage of being free from any issues due to collinearity among the new attributes. The major disadvantage is that we no longer have data points that represent a direct interpretation of the measured real-world data. The resulting model loses a significant level of the clear-box explainability it had with the original data inputs. Table 7.14 displays two groups of results. The "Original Data" group on the left repeats the results from our comparison of the learners using the original datasets. These results were reported in the "ALL" column in Figures 7.3 for logistic regression and 7.4 for decision trees. We repeat them here so they be easily compared to the results from the same experiment using the corresponding ALL datasets with principal components. These are shown in the "PCA-Transformed" group on the right of Table 7.14. As with the previous comparisons, the F1 scores marked with a dagger (†) are higher than the 1R reference algorithm (on the same dataset) at a significance threshold of 0.05 percent per

relation (as opposed to covariance) matrix, and return the number of principal component attributes needed to cover 95% of the variance in the data.

a corrected, resampled T-Test (Nadeau & Bengio, 2003).

Figure 7.10 Logistic regression with original data vs. PCA for the ALL dataset.

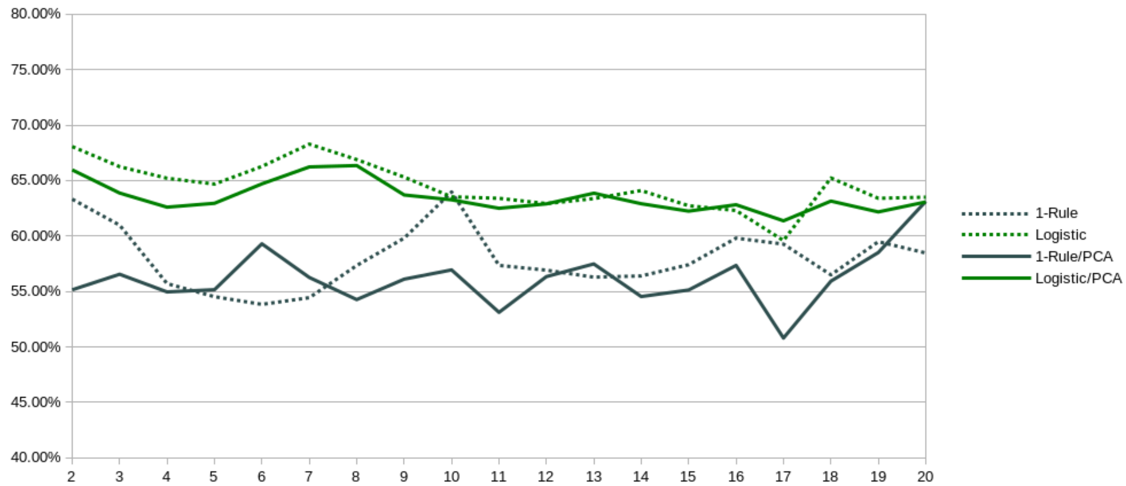
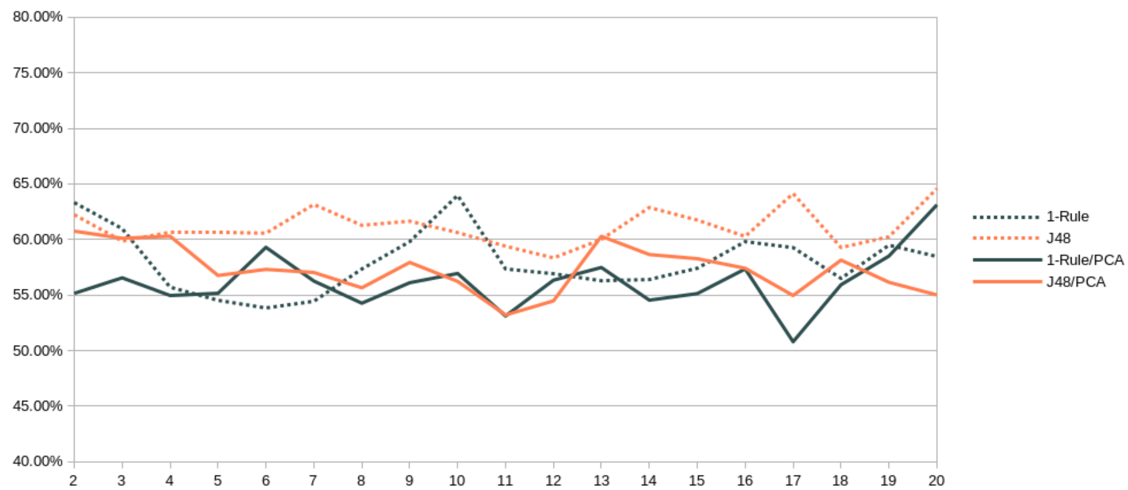


Figure 7.10 graphs these results for logistic regression (Weka's *Logistic* learner) for both the original and the transformed datasets. Figure 7.11 does the same for decision trees (Weka's *J48* learner). In both figures the 1R reference learner is shown in grey. For all learners the dotted lines give the F1 performance (y-axis) when using the algorithm with the original ALL datasets, while the solid lines show the algorithm's performance with the PCA-transformed dataset. For most (though certainly not all) levels of minimum participation (x-axis), both logistic regression and decision trees gives slightly better results when using the original, untransformed dataset.

As we see, using PCA to eliminate collinearity in our dataset does not necessarily mean our results will improve. We have, however, benefited from this analysis as it leads us to a better understanding of the data we are using in our machine learning models. In the POS, EMO, and ONT datasets, we have identified attributes that have a strong correlation and formed hypotheses for why the correlation exists. This can help to limit surprises when we seek to interpret our results. For the QST dataset, the large number of attributes that are highly correlated serves to reinforce the conclusion that there are inherent limitations to the usefulness of this dataset within the scope of the current implementation of the *Say $\mathcal{S}\vec{\iota}La$* architecture as we discussed in Section 6.4.2. We must

Figure 7.11 Decision tree with original data vs. PCA for the ALL dataset.



work to address these limitations in our continued research efforts.

7.3.3 Overall Observations

Looking back over both series of experimental runs, our analysis has revealed a number of similarities and differences with respect to the various types of data elements across the range of minimum levels of participation under consideration in this study. Using two separate machine learning algorithms helps us to discount any artifacts that might be associated with a given learner. As a core component of our research involves description logic and the say-sila ontology, it is especially interesting that this phase of the research serves to confirm that the ontological indicators we developed in Chapter 5 appear generally to be one of the more effective data elements for predicting stance on climate change, most notably for higher levels of minimum participation. We also see that the part-of-speech elements appear to be a noteworthy characteristic in the data. We eventually ended up putting these elements aside when we conducted the experiments with the ontology. However, the analysis in this chapter suggests that our future research efforts should continue looking into what parts of speech people are using in their online posts.

Of course, emotion is also a core focus of the present research. We are left here with the observation that expressed emotion in tweets has not turned out to be a primary characteristic that our machine learning algorithms could leverage. This finding is somewhat surprising, especially given that the big players project (see Chapter 3) indicated that the emotions anger and fear did show a certain level of predictive capability.[27] However, the big players project represents a different problem in a number of ways. First of all, that predictive capability was demonstrated when predicting future levels of emotion expressed within the online community. We did not see the same capability when predicting during the actual time period being analyzed, which is arguably a situation that better parallels the experiments in the present chapter. Secondly, the machine learning models in the big players project were using big player emotion to predict emotion expressed in the regular players' tweets, rather than predicting the stance on climate change for those users. In spite of these differences, the results from the big players project may still serve as an indicator that emotions, particularly anger and fear, merit continued research with regard to modelling online communities and their attitudes and beliefs concerning climate change. We stated back in Section 1.4 that emotion likely plays a rather subtle role in matters concerning how humans relate to climate change. It seems that we must continue our efforts if we wish to effectively model the subtlety of emotion's role.

7.4 Limitations

Our results for classification here certainly improve on those we obtained in Chapter 5. This means that using machine learning models that include output from our ontological analysis allows us to surpass the predictive capability of the ontological model itself in its current implementation. This is an interesting finding in its own right, but some readers may ask why we have stopped here. Why limit ourselves to default parameters

---

[27]Note that these two emotions form an opposing pair in Plutchik's system of basic emotions (Plutchik, 2001).

for two basic machine learning algorithms? Why not use a grid search to optimize hyperparameters? Why not put a little more focus on attribute selection or on analysis and transformation of the input data? Why not use stronger learners? What about deep learning?

These are all excellent questions, and each one suggests an interesting project for future research. When we opened this chapter, we stated that our purpose in conducting the series of experiments presented here is to set a baseline. If our research were specifically following an approach based on machine learning, this baseline would essentially represent an important first step, and we would aim to use it as a guide when working with more sophisticated learners (Domingos, 2012). However, the present research effort is primarily concerned with the ontological model we have developed. The machine learning baseline can thus serve as a tool which aids in evaluating our work with this model, identifying its strengths and weaknesses, and generating ideas for improvements. These findings also serve to direct future research endeavours. What we have discovered in this machine learning phase of the research supports us as we continue with the iterative approach we have chosen for the long-term project.

We may also note that even at this stage of the iterative process we have a similar limitation as we had with the big players project in that we are not taking the machine learning experiments conducted in this chapter to completion. We have produced 190 models, but only presented a few of them to the reader as samples without providing an in-depth interpretation of those few, let alone a synthesized analysis of the full set (or at least perhaps those which performed significantly better than 1R). The reason for this limitation is essentially the same as it was for the big players project. We are primarily concerned with the flow of information collected from our online community as it moves through the *Say $S\vec{\iota}La$* architecture. Referring back to Figure 0.1, we are working here with data which was processed first in the machine learning level, transformed and then processed in the description logic level, and then transformed and processed yet again in the machine learning level.[28] The models we are analyzing in this chapter are intermediate models with respect to their role in the architecture. *Say $\mathcal{S}\vec{\iota}$la* will take information obtained from processing these machine learning models and move it up into the description logic level (potentially with new data coming in from Twitter). The intention of the *Say $\mathcal{S}\vec{\iota}$la* design is that "final results" are reported from the description logic level as shown in the figure. This functionality goes beyond mere reporting, however. Our intention is that virtually any tool based on Semantic Web technologies can access a final descriptive model from *Say $\mathcal{S}\vec{\iota}$la* and use that model as needed for whatever task it is trying to accomplish.

7.5 Contributions and Continued Research

Our goal behind the machine learning models presented in this chapter may have been to establish a baseline for our continued work, but the contribution associated with this part of our research reaches much further. The experimentation in this phase has allowed us to come full circle in the application with respect to two distinct AI technologies which the *Say $\mathcal{S}\vec{\iota}$la* architecture is using to model communities on social media and their online communications about climate change. We mentioned in Chapter 3 that the big players project was something of an independent effort with regard to the main thrust of our research. Yet, central ideas from the big players, such as minimum levels of participation, continue to play an integral role throughout the development and analysis of the say-sila ontology. From a certain point of view, the big players project was not as independent as we ourselves had thought at the outset. The big players experiments represent the initial phase of our research, which involved machine learning models. The next research phase was essentially quite different, involving description logics and ontological modelling; yet, our experience with the big players project helped to create the methodology for

[28] There are two important considerations regarding this information flow: (1) the middle Bayesian logic level is optional and will be incorporated in future studies, and (2) we used the Weka Experimenter to analyze these models because this part of the data flow has not yet been fully integrated into the architecture.

that subsequent phase. Now we are employing the results from the ontological model and returning to where we started, the domain of machine learning.

Additionally, we have extended our iterative strategy, originally inspired by the EU Decarbonet project (Maynard & Bontcheva, 2015). We employed this approach for the development of the say-sila ontology (see Chapter 5), and now we have taken knowledge inferred using the ontology and applied it to a new type of model. Note that we could not have created the machine learning models presented in this chapter from the raw data we have collected from Twitter. Instances from an ARFF dataset created directly from this data would necessarily represent tweets. Here we are representing users, the authors of those tweets. The *Say S$\vec{\imath}$La* application creates an ontological model of these users based on what they publish online and then outputs datasets to encode essential knowledge from that model as ARFF files. In essence, our system generates the means for us to evaluate that same system here using an independent framework. Now we can take our findings from this evaluation and return to the ontological model, endeavouring to improve it further.

Predicting the stance on climate change of users on social media is no small endeavour. We have given ourselves a notable design advantage by implementing an architecture with the ability to shift our model from the domain of machine learning, to description logics and back to machine learning. Our objective to build an ontological model capable of distinguishing users in the green and denier categories will ultimately lead to the even more ambitious goals, the first of which will be modelling online communities based on all six segments of the Six Americas. Our ability to leverage independent machine learning models to gain critical insight into the observations we make using the ontological model will most certainly prove invaluable as we continue our research with *Say S$\vec{\imath}$La*.

CONCLUSION

The course of the research conducted for this doctoral program in cognitive informatics (*informatique cognitive*) at the *Université du Québec à Montréal* has touched on many areas in the domains of both artificial intelligence and cognitive science. We have incorporated Plutchik's theory of basic emotions (Plutchik, 2001) into a number of models based both on machine learning and description logics. We have also used methods from information retrieval to link the emotion expressed in online posts about global warming to questions from the survey for the socio-psychological study known as the Six Americas (Maibach et al., 2011).

The nature of this project has necessitated the creation of our own methodology. Unfortunately, the National Institute of Standards and Technology (NIST) provides no labelled dataset for our chosen research topic that we could use to compare our results with the *Say $S\vec{\imath}La$* application against those from many other researchers using a variety of methods for classification. We collected our own dataset from Twitter and developed a method to establish a ground truth for the examples it contains. Our ontological model is able to use this ground truth to evaluate its own logical inference as to the stance on climate change of online users. We also took knowledge generated by our ontological model and used it as input to a final round of machine learning models for a second evaluation using an independent process. Our iterative methodology has allowed us not only to refine each of the various stages of our research, it has also allowed us to use each phase of our work as a foundation on which to build the next phase.

Our goal for this research has been an ambitious one. We have sought to construct an architecture which is at once inspired by theories of human cognition and based on proven methods in artificial intelligence. We have aimed to use this architecture to model attitudes and beliefs of people on social media based on established research from the human sciences. The benefits of this system may go well beyond the interesting results revealed by the model. Ideally, experiments conducted with our architecture may be used to augment findings from survey methods as employed in sociology, psychology, and philosophy. Thus, it may serve as an additional means to confirm the results obtained through traditional experimental means, perhaps even identifying specific observations not seen in the original study but ultimately meriting continued research.

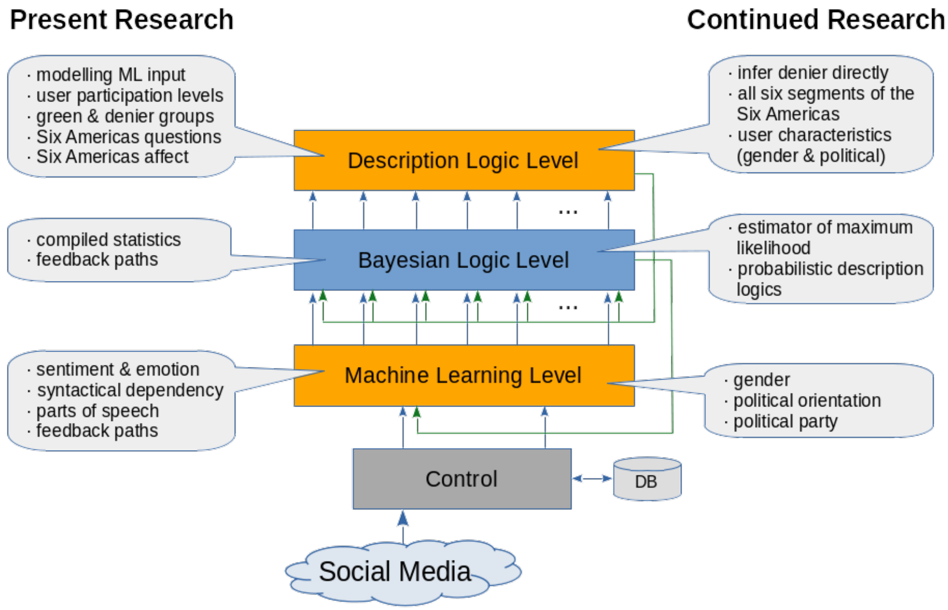
8.1 Continuing with *Say S$\vec{\imath}$La*

The *Say S$\vec{\imath}$La* application[29] has been an integral part of this research project. Throughout the various phases of the study we have designed and developed an infrastructure which not only allows us to perform the experimental work presented in this thesis but also stands as the foundation for a sophisticated research tool providing access to communities on social media for studies in the human sciences. Our future research aims to continue building upon this foundation. As we described back in Section 0.4, the *Say S$\vec{\imath}$La* architecture is inspired by the theory of hierarchical cognition (Clark, 2013). Of course, the hierarchy we are integrating into the architecture is simplified compared to biological cognition, and it is adapted for informational systems. There are three levels in the current design. We have touched all of these to an extent with regard to the present research, but there is still significant work to come for each of them.

Figure 8.1 displays the *Say S$\vec{\imath}$La* architecture, identifying key functionality linked to the present research for each of the three modelling levels as well as functionality which we will be addressing in future research efforts. We describe the specific points corresponding to each layer of the architecture in the sections that follow. The description starts at the lowest modelling level, where the system receives raw input from social media and proceeds towards the highest modelling level, where it produces the output representing

[29] https://github.com/dendrown/say_sila/tree/uqam-dic

Figure 8.1 Present and future research linked to the Say Sĭla architecture.



the modelled online communities.

8.1.1 Machine Learning Level

The machine learning level is responsible for identifying particular traits in the text and the metadata of the online posts flowing into the system. Work at this layer of the architecture within the scope of present research includes: identifying sentiment and emotion in the incoming texts, mapping syntactical dependency between tokens in the texts, and determining grammatical elements such as parts of speech. Additionally, we have taken a solid step towards creating a feedback signal from the highest modelling layer, the description logic level, back into the machine learning level. Our final experiments involving machine learning algorithms, whose input comes from the output of the say-sila ontology, represent an initial proof of concept for this part of the design (see Chapter 7).

A number of modules in the machine learning level of the architecture remain to be researched and developed. For example, the results from the Six Americas studies indicate that gender, political orientation, and party affiliation can be significant indicators of a person's stance on climate change (Maibach et al., 2009). Researchers have demonstrated some success using tweets and their associated metadata to identify gender (Burger et al., 2011), to discriminate between a conservative and a liberal political stance (Preoţiuc-Pietro et al., 2017), and also to classify users as Republican or Democrat (Sylwester & Purver, 2015). We have had some encouraging, albeit preliminary, results using logistic regression and support vector machines to classify users as male or female. We need to incorporate these components into the machine learning layer of the architecture, thereby continuing our efforts with gender and building on current research for political stance and party.

8.1.2 Bayesian Logic Level

Necessarily, the Bayesian logic level in the *Say Sīla* application received less attention than the other two modelling layers over the course of this doctoral program. This is not altogether surprising since the architecture is designed such that this layer is optional, and the machine learning level may instead feed directly into the description logic level. For this reason, we focused our main efforts on the research relating to these other two layers which are essentially responsible for handling the input and the output of the system. The Bayesian logic level is intended to improve upon the results that the machine learning and description logic levels are capable of producing on their own. The idea is that this level should act as a maximum likelihood estimator (Goodfellow et al., 2016), effectively creating a probabilistic filter intended to mitigate the noise due to informal language, topic shift, and other complicating factors inherent to posts on social media.

Our principal contribution towards this layer is the compilation of the statistical data associated with our experiments for the say-sila ontology. The statistical information for the ontological model at the top of the layered architecture forms the feedback signal which guides the filtering process. The data points coming up from the machine learning level correspond to what the system perceives about the outside world by means of the online posts it reads. The feedback coming down from the ontological model represents what the system expects those inputs to be. In other words, the information going up (input towards output) functions as an error correcting signal, while the information going down (output towards input) serves as a prediction signal.

At present, the work for this level is rudimentary. It essentially represents the bottom-up and the top-down paths as data flows through the system. Our continued research will involve the implementation of the Bayesian logic level in code as well as its formal interface with the modelling layers above and below it. Our initial design for this level centred around a classifier based on Naive Bayes which learns incrementally by means of expectation maximization (Dempster et al., 1977). We have considered a critique of this design, which argues that it is simply layering a machine learning model on top of another machine learning model. We do not believe this is entirely accurate, given that the Naive Bayes classifier is using the feedback signal from the description logic layer as an input in addition to the signal it receives from the machine learning layer. The model of the online community as it stands as well as the modelled aspects of the underlying study, i.e., the Six Americas, will represent the a priori probabilities for the classifier. Nevertheless, we recognize that the critique may still have identified a potential problem if the feedback signal from the ontology does not have a significant tempering effect as we are hypothesizing it will have. With the possibility of this problem in mind, we would also like to explore an alternate strategy based on probabilistic description logics (Bellodi et al., 2017; Ochoa-Luna et al., 2011) and compare results for the two approaches.

What is interesting about these two strategies is that the first implies an implementation that stacks machine learning algorithms, while the second implies the stacking of systems of description logics. Hence, one might envision a future version of the architecture as a design which could essentially be described as either a double-layer of machine learning or a double-layer of description logics. Do we dare consider a double-layer of both?

Of course, exploring different potential configurations for the architecture also means looking into effective methods of transmitting information up and down the hierarchy. This effort may well involve delving deeper into the subject of human cognition in order to evaluate how the biological hierarchy may be modelled effectively as an informational one. We have many questions to answer regarding this level of the *Say Sĩla* architecture and many paths to explore as we seek answers to these questions. One can easily imagine that the research involved for a full implementation of the Bayesian logic level will necessarily represent a new thesis in and of itself. We are looking forward to the next steps in our work for this important component of *Say Sĩla*.

8.1.3 Description Logic Level

The description logic level is the highest modelling layer, essentially representing the output model for the architecture. Many of the contributions associated with the present research involve this layer of the system. This level of the hierarchy goes well beyond that of a module for basic knowledge representation. It comprises the ontological model, the implementation in code of the underlying description logic, as well as the functionality needed to populate the model from incoming data and to analyze that populated model in the context of a given research problem. This layer translates the output of the machine learning level (or alternatively the Bayesian logic level) into DL constructs, linking concepts as they come up the hierarchy according to the properties associated with modelled knowledge about Twitter (e.g., *TokenA* follows *TokenB*) and modelled knowledge of the Six Americas (e.g., *TokenA* is a *CauseToken*). Importantly, the description logic level effectuates modelling at the user level, rather than simply at the tweet level, which corresponds to the raw input data. At the user level new knowledge emerges. The system can model relative levels of participation and create a linkage between users and specific questions in the Six Americas survey. It can also perform logical inference to make a judgment as to whether a user is in the green or the denier category. Finally, it can associate expressed sentiment and emotion with specific questions from the Six Americas.

As notable as these accomplishments may be, the direction of continued research associated with the description logic level is clear. We shall initially look to refine the model so that it may infer users in the denier category directly (rather than just including users who are not green). The next major step, of course, is to extend this level of the architecture to model all six of the population segments included in the Six Americas. To do this, we will need to incorporate the additional user characteristics which the updated version of the machine learning level will be identifying in the online posts coming into the system. Finally, although we have created initial paths for feedback signals running back to lower modelling levels of the architecture, we must refine this preliminary implementation and conduct a series of experiments to establish its effectiveness with regard to the overall architecture.

8.1.4 So Many Tweets

We began collecting microblogs with the hashtags *#climatechange* and *#globalwarming* from Twitter in conjunction with the initial phase of our research, the big players project (Chapter 3). To date, we have more than 30 million tweets with these hashtags covering a period of over four years, and *Say Sĩla* is still collecting them. We have used only a small fraction of these for the experiments presented in the present work. *Say Sĩla* takes its input as tweets. Its output, in accordance with a scientific study such as the Six Americas, describes online users as modelled by the tweets they author. There is obvious value in continuing the work we have begun here, developing the architecture for research projects which extend the machine learning and ontological models to incorporate millions of users.

Of course, there are constraints as to the time and computing resources needed to work towards this goal. The current version of *Say Sĩla* already utilizes the concurrency constructs in Clojure to speed up the process of logical inference in the description logic level. Future work leveraging the distributed and concurrent computing paradigm that is part of Erlang/OTP can aid us in handling datasets involving massive numbers of online users by extending the application's computing power over multiple servers and making efficient use of multiple processing cores on those servers. Ultimately, we are aiming to create an automated "expert" with respect to a given study in the human sciences. This is to be an expert capable of reading tens or even hundreds of millions of online posts that may potentially relate to that study and who can then tie what specific communities are publishing on social media to the findings established using traditional survey-based research methods.

8.2 Looking Forward

We come to the end of this doctoral program recognizing that while we have accomplished much throughout the course of our research, there is also much which remains to be done. Emotion is a central theme of this research, yet our focus on emotion has perhaps served to generate at least as many questions as it has answered. We are curious, for example, what new insights we may find as we extend our machine learning models of big player communications from two opposing pairs in Plutchik's system of basic emotions (Plutchik, 2001) to all four pairs for the full set of eight emotions. We would like to explore how we may enhance our models to account for irony and sarcasm expressed in the online posts. Ultimately, we are aiming to extend our contribution not only to psychology but also potentially to domains like political science by examining questions such as how one arrives at the role of "influencer" on social media.

Continued research projects abound for the *Say S$\vec{i}$La* architecture as well. Considering the theme of emotion, we are particularly interested in comparing our observations using the present model based on a classical theory of basic emotion to those seen when incorporating a model based on the theory of constructed emotion (Barrett, 2017). Before moving in this direction, however, we have significant work ahead of us to complete the full research tool as we have presented it in Figure 8.1.

It should be noted that while our present research endeavour has been centred around communications about climate change on social media, the iterative methodology we have demonstrated for creating a model and using it as a means of inferential analysis may be applied to a great number of other research problems in the human sciences. One of the next steps for us specifically is to enhance the model so that it covers all six categories of the Six Americas. Yet, the methodology we are employing to reach this goal is general, as is the underlying design of the architecture for *Say S$\vec{u}$La*. Thus, our core contributions are relevant with regard to research endeavours in sociology, psychology, philosophy, or any of a vast number of other domains.

Certainly, research involving online social media is in no way intended to replace more traditional survey-based methods of gathering information. There are numerous challenges when using this exceedingly informal source of public opinion, not the least of which is the fact that online users will not constitute a representative sample with respect to a target demographic. Nonetheless, online posts are potentially an extremely valuable information resource. They represent an expression of the beliefs and opinions of real people, many of them with some level of capability to influence others. Politicians use social media; so do celebrities and news sources (Fownes et al., 2018; Wojcik & Hughes, 2019). Our vision is that data from social media may be used to expand the scope of traditional experimental methods that rely on formal surveys. The initial survey-based work can provide the foundation for studies which model online communities. The results from these studies may then be used to confirm findings from the original experiments. They may also serve to identify novel details of interest leading to new directions in the research associated with the original work. Thus, the opportunity exists to apply the iterative methodology, which has served us well in the present research, across projects and across teams of researchers.

As we look forward, we should remark that although the *Say S$\vec{u}$La* architecture is intended to be general such that it may be used as a modelling tool for a vast array of research topics, we ourselves intend to continue our efforts studying communications on climate change. This work has been an answer to a call for research on this existential crisis, threatening not only humanity but countless ecosystems and essentially the planet as we know it. Our hope is that our contributions with this work and with what follows will aid in some way towards mitigating the danger we have brought on ourselves and upon our world. We also hope that our efforts will serve to encourage others to apply their talents so that human beings may ultimately avert the planetary catastrophe we seem unwittingly to have caused.

GLOSSARY

ABox (assertion box) The part of an ontology containing declarations of named individuals (concept assertions) and property instantiations (role assertions). 47, 48, 53, 57

affect With respect to the present work *affect* means human emotion or positive/negative sentiment as felt or expressed in words. Outside of this work, the exact definition varies significantly from one expert to another in psychology, philosophy, sociology, and other fields that study this concept. In a general sense, however, the term usually refers to "feelings" that a being experiences at some level, be it emotional or physiological. 26, 31, 33, 34, 37, 38, 68, 96, 112, 113, 191, 199, 210, 219, 280, 281, 297

artificial intelligence (AI) A subfield in the domain of computer science which seeks to incorporate intelligent behaviour into computing systems. Often, the goal is to create programs capable of "thinking" in a manner analogous to human reasoning, but common objectives also include applications able to perform a specific task generally considered to require intelligence (e.g., playing chess or operating a vehicle). 3, 4, 14, 16, 20, 21, 31, 40, 43, 44, 55, 210, 250, 327, 351

Attribute-Relation File Format (ARFF) A text-based data format for the machine learning platform Weka (Frank et al., 2016). Unlike the more common comma-separated values (CSV) format, ARFF allows the data modeller to declare specific types for data attributes, name the data relationship, and add comments in the file where necessary. 71, 296–298, 327, 351

BM25 Okapi "Best Match" implementation #25. A similarity weighting algorithm generally considered to be an improved version of TF*IDF. The algorithm imposes a saturation limit on the weight contribution due to a given term's frequency in a document, and it normalizes document length across the corpus. BM25 is the default similarity weighting algorithm in Lucene. 196, 197, 249, 277

Clojure A functional programming language that is a member of the Lisp family of programming languages. Clojure features built-in constructs for concurrency and runs on the Java Virtual Machine (JVM), giving the developer full access to a vast array of Java-based libraries. 20, 21, 78, 103–105, 132, 335

cross-validation (N-fold cross-validation) A technique which enables the evaluation of a machine learning model without the use of an independent test dataset. Instead, the training dataset is divided randomly into N folds. Quite often $N = 10$, and in this case the method is called *10-fold cross-validation*. To evaluate a given model, the system essentially creates N copies of that model. Each copy has a different data fold set aside for testing, and is trained with the remaining $N - 1$ folds. The system tests each of the copies separately and returns a final evaluation result which is the average of the results from the N copies of the model. 90, 91, 93, 298, 302

decision tree A machine learning algorithm which utilizes a "divide and conquer" approach for classification problems. The algorithm creates a tree for which the root and interior nodes each correspond to a specific data attribute. Based on the value for that attribute in a given data instance, these nodes will branch either to other interior nodes, associated with other attributes, or to leaf nodes which correspond to the different classes associated with the target attribute. 291–293, 295, 298, 304, 305, 308

denier (in the present work) A person who generally does not believe in climate change or that it is caused by human activity and does not support policy to mitigate it. In this study we link the term "denier" to the the doubtful and dismissive categories in the Six Americas survey (Leiserowitz et al., 2010). 11, 12, 44, 114, 120–124, 131, 141, 142, 145–147, 185, 187, 188, 191, 200, 212, 219, 287, 298, 328, 334, 335

description logics (DL) A family of logics for knowledge representation through concepts and the roles these concepts take in relation to each other. 43–52, 54, 56, 57, 59, 61–63, 102, 108, 112, 119, 334, 351

Dolce+D&S Ultralite (DUL) A top-level ontology based on DOLCE Lite (an implementation of DOLCE into OWL) and integrating the Descriptions and Situations, Plan, Information Objects, and Collection foundational ontologies (Presutti & Gangemi, 2016). 106–113, 131, 184, 185, 351

emotion mining A subfield of NLP, closely related to sentiment analysis, which seeks to reveal information about a person's emotional state with respect to a given subject by way of an analysis of texts the person authors (or speech that she utters). 25, 31, 38

Erlang A functional programming language designed specifically for highly-concurrent, distributed server-based systems which require a high level of fault tolerance. 19–21, 336

eXtensible Markup Language (XML) A standardized language for annotated documents intended to be readable by both humans and computers. XML is used in many information processing applications; however, in the context of the present work it serves as the foundation for the Web Ontology Language (OWL). 59, 61, 103, 104, 107, 353

F1 (F-measure or F-score) A measure of classification performance which incorporates precision and recall into a single metric. It is calculated thus: $F1 = \frac{2 \times recall \times precision}{recall + precision}$ (Witten & Frank, 2005). 147, 152, 153, 177, 178, 180, 181, 186, 298, 299, 305, 307

green (in the present work) A person who believes in anthropogenic climate change and generally supports policies to mitigate the threat. Although colloquial use of this term varies widely, in this work we consider a person who is "green" to be in either the alarmed or the concerned category of the Six Americas survey (Leiserowitz et al., 2010). 11, 12, 44, 114, 120–124, 131, 141, 142, 145–147, 185, 187, 188, 191, 200, 212, 219, 287, 298, 328, 334, 335

hashtag A word or a short concatenated phrase prefixed with a hash symbol (#) and used to mark the subject matter of a text on social media. One or more hashtags may be listed at the beginning or the end of a given text. Alternatively, a hashtag may take the place of the words it represents in the text itself (e.g., "progress has been made in mainstreaming *#climatechange* issues"). 70, 86, 110, 111, 119, 141, 142, 191, 194

hypernym A word signifying a concept which encompasses and acts as a superset to other concepts. For example, the meaning of the word *text* encompasses a number of more specialized words, such as *microblog*, *post*, and *article*. *Text* is a hypernym of these words. 120, 193, 281

hyperparameter (machine learning) The options used to configure a machine learning algorithm and direct the actions it takes when modelling a dataset. Using linear regression as an example, the weights associated with each data attribute in the linear equation are the *parameters* of the model. The *hyperparameters* are values provided to the algorithm on invocation, instructing it to attempt to exclude collinear attributes, use a specific value for the ridge penalty to control the size of the weights, etc. 79, 287, 290, 295, 325

hyponym A word signifying a concept which is a specialization and acts as a subset of another concept. For example, the meaning of the word *tweet* is encompassed by the more general word *microblog*. *Tweet* is a hyponym of *microblog*. 120, 193, 281

inclusion (description logics) An axiom defining a concept or role ($Sub \sqsubseteq Super$), stating that the defined element (Sub) is included in or subsumed by the referenced element ($Super$). Hence, all individuals belonging to Sub also belong to $Super$, and Sub may be considered a specialization of $Super$. 49, 50, 53

information retrieval (IR) A subfield of natural language processing which allows one to search a large collection of informational material (sc. documents) to find those entities that are pertinent to a given request. 192–194, 197, 198, 213, 250, 277, 279, 281, 352

Intergovernmental Panel on Climate Change (IPCC) An intergovernmental organization dedicated to providing policymakers with assessments of current scientific knowledge on climate change. It was formed in 1988 by the World Meteorological Organization and the United Nations Environment Programme (IPPC, 2013). 8, 46, 112, 113, 246, 248, 351

International Resource Identifier (IRI) A character string forming a unique name which may include non-Latin characters. An IRI functions in the same way as a URI, standing as a reference for a Web resource. 58, 352

learner (data mining and machine learning) A machine learning algorithm providing a model generally intended to handle tasks of classification or regression. 287, 298

lemmatization A form of NLP processing which reduces a word to its lemma or dictionary-based root. The process is comparable to stemming, but it encompasses a more formal and rigorous procedure closely tied to the grammar of the language being analyzed. For this reason lemmatization typically entails a higher cost in terms of time and computational resources for NLP applications. 131, 195

linear regression A machine learning algorithm which maps a vector of attributes onto a scalar. This scalar represents the dependent (predicted target) variable. 74, 80, 85, 87, 88, 90, 95, 98, 288, 289

Lisp A programming language first created in 1958 for symbolic computation and list processing, largely intended for use in artificial intelligence (McCarthy, 1960). The original language LISP has evolved into a family of Lisp programming languages which include Common Lisp, Scheme, and Clojure among many others. 20, 21, 104

Liu's Opinion Lexicon A sentiment polarity lexicon created manually with 6789 terms, each associated with a binary value to represent either positive or negative sentiment. The lexicon includes slang terms, misspelled words and morphological variations of words to aid in the analysis of informal texts such as those frequently found online (Hu & Liu, 2004). 34, 200, 205, 207, 209, 211, 229, 232, 233, 235, 236, 240, 242, 251, 280

logistic regression A machine learning algorithm for classification which uses a logistic sigmoid function to transform a linear model based on log likelihood to probabilities of membership with respect to the nominal classes associated with the target attribute for a dataset. 288, 289, 295, 298, 299, 301, 305, 307, 308, 310, 332

Lucene An open source Java framework for information retrieval from the Apache Software Foundation. 73, 194–198, 200, 201, 212, 249, 250

machine learning A subfield of artificial intelligence in which statistical algorithms are used to detect patterns in sets of example data. The patterns identified may be used either to better understand the underlying organization of the data or to predict targets characteristics in new data (i.e., data not seen by the algorithm until after it has been trained using the example data). 19, 21, 40, 71, 78–80, 90, 98, 99, 285, 286, 288, 289, 291, 296, 302, 325–328, 331–333

Mechanical Turk A crowdsourcing service offered by Amazon which they call "artificial artificial intelligence." The name comes from "The Turk," an 18$^{\text{th}}$ century chess-playing automaton. It won against many players; however, it was later revealed to be a hoax as inside the machine hid a master chess player who was controlling the game (Sheehan, 2018). 199, 280

microblog A short unit of online content (sc. a tweet), generally posted via an Internet application such as Twitter or Tumblr. Most often they are texts published for a wide variety of reasons: expressing an opinion, transmitting information, supporting a cause, or criticizing one. However, other media such as digital images or video may supplement the textual content. 16, 37, 38, 40, 67, 68, 70, 76, 96, 118, 192, 193, 276, 277, 280, 282, 335

n-gram A sequence of n tokens or words which are interpreted semantically as a single concept, often with a meaning significantly different from the that of the individual components. A unigram is a single word. A bigram (e.g., *Black Friday*) is made of two tokens. A trigram (e.g., *Forth of July*) has three. 281

namespace (XML, RDF, OWL) A collection of related elements identified by a common URI. Often the URI is a network location, which serves as a Web resource for these elements. 58, 59, 61

natural language processing (NLP) A subfield of computational linguistics and artificial intelligence which deals with the analysis of human language, such that an informational system is able to "understand" it to a certain extent for a variety of purposes such as information retrieval, translation, summarization, or communication between human and machine. 28, 31, 33, 34, 39, 40, 67, 68, 70, 73, 96, 110, 125, 126, 182, 183, 192, 196, 249, 352

NRC Affect Intensity Lexicon A lexicon from the National Research Council Canada which gives levels of intensity on a scale of 0 to 1 for the emotions anger, fear, sadness, and joy. Version 0.5 (used in the present research) contains 5814 words. The lexicon was created manually using a process called best-worst scaling (Mohammad & Bravo-Marquez, 2017). 71–73, 80, 199, 200, 211, 352

NRC Word-Emotion Association Lexicon (also known as EmoLex) A lexicon from the National Research Council Canada which provides binary (yes/no) values for 10 affective attributes: Plutchik's eight base emotions plus positive and negative sentiment polarity. The lexicon contains 14,182 terms, manually annotated by crowdsourcing with Amazon's Mechanical Turk (Mohammad & Turney, 2013). 200, 201, 205, 207, 209–211, 219, 231–233, 235–237, 240, 242, 246, 251, 260, 270, 280, 297, 352

ontology (computer science) "[A] formal, explicit specification of a shared conceptualisation" (Borst, 1997; Gruber, 1993; Studer et al., 1998). In the context of the present work an ontology may more precisely be thought of as a formal hierarchical system of knowledge representation consisting of (1) classes specifying a set of concepts and representing a vocabulary, (2) individuals representing particular instantiations of those classes, and (3) the relationships between these elements. 40, 43, 44, 46, 285

Pearson correlation coefficient (PCC) A statistical measure ranging from -1 to 1, indicating the correlation between two series of data points. A strong positive correlation means that when values from one dataset are high (or low), values from the other dataset will also be high (or low). Likewise, for a strong negative correlation, when values from one dataset are high (or low), values from the other dataset will be low (or high). Considering the extremes, a PCC of 1 indicates the data are correlated exactly, whereas a PCC of -1 means there is an exact but negative correlation. A PCC of 0 means that no correlation at all exists between the data sets (Witten & Frank, 2005). 87, 89, 90, 352

polarity (sentiment) The expression of a positive (good) or negative (bad) opinion or point of view in a text, word, or phrase. An analysis of sentiment polarity may also include neutral sentiment, which generally indicates descriptive, fact-based text that does not express any particular opinion on the part of the author. 34, 35, 37, 38, 47, 69, 113, 200, 280

precision (synonym: efficiency) An information metric defined as the ratio of the number of true positives over the number of instances a model reports as true (true positives + false positives) (Witten & Frank, 2005). In the say-sila ontological model this corresponds to $\frac{green\ inferred\ green\ accounts}{inferred\ green\ accounts}$. 145–147, 152, 177, 178, 186, 193, 281, 299, 305

punning (Web Ontology Language) The use of the same name for a class and an individual (or a property), thereby allowing the individual to stand for the class. The type of axiom (class or individual) gives the context necessary to distinguish what the name refers to. 53, 112

RDF Schema A mid-level structural component of Semantic Web architectures. It incorporates classes to represent ontological concepts and subclasses to allow the creation of a class hierarchy. In a similar manner, it uses properties and subproperties to support the modelling of a hierarchical role structure. 60, 61, 63, 104

recall (synonym: effectiveness) An information metric defined as the true positive rate or the ratio of the number of true positives over the number of true instances (Witten & Frank, 2005). In the say-sila ontological model this corresponds to $\frac{green\ inferred\ green\ accounts}{green\ accounts}$. 145–147, 152, 153, 177, 178, 186, 193, 281, 299, 305

Resource Description Framework (RDF) A structural component of the Semantic Web which enables the definition of sets of binary properties of the form $\langle subject, property, object\rangle$, generally used to define relationships between various concepts in an ontology. 60, 61, 352

retweet A tweet republished by another Twitter user with credit given to the original author, usually by means of Twitter's *RT* tag. 50, 52, 75, 76

Semantic Web An extension of the World Wide Web which makes available a global network of machine accessible information in a manner analogous to the Web's vast collection of linked hyperdocuments intended for human readers. 19, 54–56, 58, 60, 61, 64, 102, 103, 106, 115, 124, 184, 185, 287

sentiment analysis (opinion mining) A subfield of NLP which deals with the analysis of text (or speech) in order to determine a person's opinion as well as whether that person feels positively, negatively, or is neutral concerning a given matter. 25, 33, 34, 37, 39, 201

similarity weighting (information retrieval) A computational process whose goal is to determine the level at which a specified term (word) is associated with a given document from a corpus. 195

Six Americas An on-going series of survey-based studies, which began in 2008, to better understand the beliefs, attitudes, politics, and level of engagement of people in the United States with respect to global warming. 9–13, 17, 43, 44, 65, 102, 103, 114, 117–121, 131, 134, 141, 145, 191, 192, 194, 195, 198–200, 212, 219, 241, 277, 280–283, 332–335

stemming A process which converts words to their common stems with respect to other grammatical forms of the word. For example, all of these words, *change*, *changes*, *changing*, and *changed* have the same stem: *chang*. 73, 125, 131, 195, 197

stop list The complete group of stop words used in a given project. 73, 197

stop word A very common word which is generally of no analytical value in information retrieval (e.g., *a* and *of*). 73, 125, 129, 195, 197

subsumption (see inclusion) 45, 49

survey concept rule (SCR; in the present work) A programmatic ontological construct that ensures a word in a microblog refers to an important survey concept. The ontological model considers SCRs in pairs so that the text and its author will be modelled as instances of green or denier indicator classes associated with the concept pair. 131, 132, 134–140, 181, 186–188, 213, 240, 352

synset A collection of words, synonymous in meaning, which together exemplify a concept. Synsets are the elemental structure used to create the hierarchy of concepts in the WordNet database (Fellbaum, 1998). 132, 149

TBox (terminological box) The part of an ontology which contains definitions of concepts and roles. 47, 48, 57

term frequency × inverse document frequency (TF*IDF) A similarity weighting algorithm in which the weight of a term increases with respect to a given document when that term appears often in the document but decreases as the term appears more frequently across various documents in the corpus. 196, 197, 249, 353

tokenization The process of separating a string of text into its component *tokens*, i.e., words, punctuation, emojis, etc., and creating a list or similar container for those tokens for use in subsequent stages when computationally processing natural language. 195

top-level ontology (synonyms: foundational ontology, formal ontology, upper-level ontology) A general ontology intended to be common across many domain, reference, and application ontologies in order to align them and thereby promote and facilitate a shared representational structure. The classes in a top-level ontology are used as parent nodes to the root classes in these lower-level ontologies (Arp et al., 2015, pp. 37-38). 102, 106–108, 115, 184, 185

tweet A microblog posted on the online application Twitter. The term may also be used as a verb to indicate the action of publishing these microblogs online. 16, 37–39, 49, 50, 52, 62, 67, 68, 70, 73–77, 112, 118–122, 125, 129, 141, 142, 147, 182, 191–193, 195, 204, 212, 277, 279–281, 296, 335

Uniform Resource Identifier (URI) A character string which serves as a unique name by which to reference a Web resource. 58–60, 107, 353

Web Ontology Language (OWL) A family of XML-based languages providing implementations of certain description logics for information systems, particularly those associated with the Semantic Web. Official specifications for OWL Lite, OWL DL, OWL Full, and OWL 2 are maintained by the World Wide Web Consortium (W3C). 48, 54–63, 65, 105–108, 112, 352

ACRONYMS

ACT-R Adaptive Control of Thought–Rational 3

AI artificial intelligence 1, 3, 4, 14, 16, 20, 21, 31, 40, 43, 44, 55, 210, 250, 327

API application programming interface 21, 70, 102–104, 119, 121, 123, 142

ARFF Attribute-Relation File Format 71, 296–298, 327

ASCII American Standard Code for Information Interchange 194, 198

CMU Carnegie Mellon University 108, 110

CSV comma separated values 20

DAML DARPA Agent Markup Language 56, 59

DARPA Defense Advanced Research Projects Agency 56

DIC Doctorat en informatique cognitive 22

DL description logics 43–52, 54, 56, 57, 59, 61–63, 102, 108, 112, 119, 334

DOLCE Descriptive Ontology for Linguistic and Cognitive Engineering 106, 107, 184

DUL Dolce+D&S Ultralite 103, 106–113, 131, 184, 185

FOAF Friend of a Friend 184, 185

HTML HyperText Markup Language 59

IPCC Intergovernmental Panel on Climate Change 1, 8, 46, 112, 113, 246, 248

IR information retrieval 192–194, 197, 198, 213, 250, 277, 279, 281

IRI International Resource Identifier 58

JSON JavaScript Object Notation 20

JVM Java Virtual Machine 20, 21

NIST National Institute of Standards and Technology 329

NLP natural language processing 25, 28, 31, 33, 34, 39, 40, 67, 68, 70, 73, 96, 110, 125, 126, 182, 183, 192, 196, 249

NRC National Research Council of Canada 35, 70, 72, 199, 207, 210, 211, 219, 240, 280, 297

NRC-10 NRC Word-Emotion Association Lexicon 199–201, 205, 207, 209–211, 219, 231–233, 235–237, 240, 242, 246, 251, 260, 270, 280, 297

NRC-AIL NRC Affect Intensity Lexicon 70–73, 80, 199, 200, 211

OIL Ontology Inference Layer 56, 59

OTP Open Telecom Platform 20, 336

OWL Web Ontology Language 44, 48, 54–63, 65, 105–108, 112

PCA Principal Components Analysis 319

PCC Pearson correlation coefficient 87, 89, 90

POS part of speech 110, 125, 126

RDF Resource Description Framework 60, 61

SCR survey concept rule 131, 132, 134–140, 181, 186–188, 213, 240

SGML Standard Generalized Markup Language 59

SHOE Simple HTML Ontology Extension 56, 59

SIOC Semantically-Interlinked Online Communities 184, 185

SPARQL SPARQL Protocol and RDF Query Language 286

SPAUN Semantic Pointer Architecture Unified Network 3

TF*IDF term frequency × inverse document frequency 195–197, 249

TREC Text REtrieval Conference 196

UML Unified Modelling Language 108

UQÀM Université du Québec à Montréal 22

URI Uniform Resource Identifier 58–60, 107

URL Uniform Resource Locator 22, 73, 110, 363

W3C World Wide Web Consortium 54, 56, 59–61, 115, 184, 185

XML eXtensible Markup Language 59, 61, 103, 104, 107

BIBLIOGRAPHY

Akerlof, K. & Maibach, E. W. (2011). A rose by any other name ...?: What members of the general public prefer to call “climate change”. *Climatic Change*, *106*(4), 699–710.

Aleahmad, A., Karisani, P., Rahgoza, M. & Oroumchian, F. (2016). OLFinder: Finding opinion leaders in online social networks. *Journal of Information Science*, *42*(5), 659–674.

Alm, E. C. O. (2008). *Affect in Text and Speech.* (PhD thesis). University of Illinois at Urbana-Champaign.

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.

Anderson, J. R. & Lebiere, C. (1998). *The Atomic Components of Thought.* Psychology Press.

Andresen, S. L. (2002). John McCarthy: Father of AI. *IEEE Intelligent Systems*, *17*(5), 84–85.

André, P., Bernstein, M. & Luther, K. (2012). Who gives a tweet?: Evaluating microblog content value. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW'12)*, pp. 471–474.

Antoniou, G., Groth, P., van Harmelen, F. & Hoekstra, R. (2012). *A Semantic Web Primer* (3e éd.). Cambridge: The MIT Press.

Antoniou, G. & Harmelen, F. V. (2009). Web Ontology Language: OWL. In S. Staab & R. Studer (eds.), *Handbook on Ontologies* pp. 91–110. Berlin: Springer.

Armstrong, J. (2013). *Programming Erlang: Software for a Concurrent World* (2e éd.). Pragmatic Bookshelf.

Arnold, M. B. (1960). *Emotion and Personality.* New York: Columbia University Press.

Arp, R., Smith, B. & Spear, A. D. (2015). *Building Ontologies with Basic Formal Ontology.* Cambridge, MA: MIT Press.

Auer, M. R., Zhang, Y. & Lee, P. (2014). The potential of microblogs for the study of public perceptions of climate change. *WIREs Climate Change*, *5*(5), 291–296.

Baader, F., Horrocks, I., Lutz, C. & Sattler, U. (2017). *An Introduction to Description Logic*. Cambridge University Press.

Baader, F., Horrocks, I. & Sattler, U. (2009). Description logics. In S. Staab & R. Studer (eds.), *Handbook on Ontologies* pp. 21–43. Berlin: Springer.

Ballew, M. T., Leiserowitz, A., Roser-Renouf, C., Rosenthal, S. A., Kotcher, J. E., Marlon, J. R., Lyon, E., Goldberg, M. H. & Maibach, E. W. (2019). Climate change in the American mind: Data, tools, and trends. *Environment*, *61*(3), 4–18.

Barrett, L. F. (2017). *How Emotions Are Made: The Secret Life of the Brain*. Boston: Houghton Mifflin Harcourt.

Bellodi, E., Lamma, E., Riguzzi, F., Zese, R. & Cota, G. (2017). A web system for reasoning with probabilistic OWL. *Software: Practice and Experience*, *47*(1), 125–142.

Bigonha, C., Cardoso, T., Moro, M., Gonçalves, M. & Almeida, V. (2012). Sentiment-based influence detection on Twitter. *Journal of the Brazilian Computer Society*, *18*(3), 169–183.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bloomfield, P. & Steiger, W. L. (1983). *Least Absolute Deviations: Theory, Applications, and Algorithms*. Boston, MA: Birkhäuser.

Bohr, J. (2014). Public views on the dangers and importance of climate change: Predicting climate change beliefs in the United States through income moderated by party identification. *Climatic Change*, *126*(1), 217–227.

Borgo, S. & Masolo, C. (2009). Foundational choices in DOLCE. In S. Staab & R. Studer (eds.), *Handbook on Ontologies* pp. 361–381. Berlin: Springer.

Borst, W. N. (1997). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. (PhD thesis). University of Twente, Netherlands.

Botzen, W. J. W., Michel-Kerjan, E., Kunreuther, H., de Moel, H. & Aerts, J. C. J. H. (2016). Political affiliation affects adaptation to climate risks: Evidence from New York City. *Climatic Change*, *138*(1), 353–360.

Boyd, D., Golder, S. & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In *Proceedings of the 43rd Hawaii International Conference on System Sciences*, pp. 1–10.

Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and*

Experimental Psychiatry, *25*(1), 49–59.

Bratko, I. (2012). *Prolog programming for artificial intelligence* (4 éd.). Pearson Education.

Bravo-Marquez, F., Frank, E., Pfahringer, B. & Mohammad, S. M. (2019). AffectiveTweets: a WEKA package for analyzing affect in tweets. *Journal of Machine Learning Research*, *20*(92), 1–6.

Bravo-Marquez, F., Mendoza, M. & Poblete, B. (2014). Meta-level sentiment models for big social data analysis. *Knowledge-Based Systems*, *69*, 86–99.

Breitman, K. K., Casanova, M. A. & Truszkowski, W. (2007). *Semantic Web: Concepts, Technologies and Applications*. London: Springer-Verlag.

Brown, P., Dellapietra, V., Souza, P., Lai, J. & Mercer, R. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, *18*(4), 467–479.

Buchholz, S. & Marsi, E. (2006). CoNLL-X shared task on multilingual dependency parsing. In *Proceedings of the 2006 Conference on Computational Natural Language Learning*, pp. 149–164.

Buechel, S. & Hahn, U. (2017). Readers vs. writers vs. texts: Coping with different perspectives of text understanding in emotion annotation. In *Proceedings of the 11th Linguistic Annotation Workshop*, pp. 1–12.

Bühmann, L., Lehmann, J. & Westphal, P. (2016). DL-Learner–A framework for inductive learning on the Semantic Web. *Semantics: Science, Services and Agents on the World Wide Web*, *39*, 15–24.

Burger, J. D., Henderson, J., Kim, G. & Zarrella, G. (2011). Discriminating gender on Twitter. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, (EMNLP-2011)*, pp. 27–31.

Chaffar, S. & Inkpen, D. (2011). Using a heterogeneous dataset for emotion analysis in text. In C. Butz & P. Lingras (eds.). *Advances in Artificial Intelligence*, pp. 62–67. Springer Berlin Heidelberg.

Chapman, D., Lickel, B. & Markowitz, E. (2017). Reassessing emotion in climate change communication. *Nature Climate Change*, *5*(12), 850–852.

Chen, L., Chun, L., Ziyu, L. & Quan, Z. (2013). Hybrid pseudo-relevance feedback for microblog retrieval. *Information Science*, *39*(6), 773–788.

Chryst, B., Marlon, J., van der Linden, S., Leiserowitz, A., Maibach, E. & Roser-Renouf, C. (2018). Global warming's "Six Americas short survey": Audience segmentation of climate change views using a four question instrument. *Environmental Communication*, *12*(8), 1109–1122.

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S. & Danforth, C. M. (2015). Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLoS: ONE*, *10*(8), e0136092.

Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R. L., Verheggen, B., Maibach, E. W., Stuart, J., Carlton, Lewandowsky, S., Skuce, A. G., Green, S. A., Nuccitelli, D., Jacobs, P., Richardson, M., Winkler, B., Painting, R. & Rice, K. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, *11*(4), 048002.

Corry, O. & Jørgensen, D. (2015). Beyond 'deniers' and 'believers': Towards a map of the politics of climate change. *Global Environmental Change*, *32*, 165–174.

Cotfas, L.-A., Delcea, C., Roxin, I. & Păun, R. (2015). Twitter ontology-driven sentiment analysis. *Studies in Computational Intelligence*, *598*, 131–139.

Cowlishaw, M. (ed.) (1990). *IBM Jargon and General Computing Dictionary* (10e éd.). Winchester: IBM UK Laboratories.

Crona, B., Wutich, A., Brewis, A. & Gartin, M. (2013). Perceptions of climate change: Linking local and global perceptions through a cultural knowledge approach. *Climatic Change*, *119*(2), 519–531.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Deng, L. & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, *7*(3–4), 197–387.

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A. & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS: ONE*, *6*(12), e26752.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87.

Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books.

Domingos, P. & Lowd, D. (2019). Unifying logical and statistical AI with Markov logic. *Communications of the ACM*, *62*(7), 74–83.

Drown, D. J., Villemaire, R. & Robert, S. (2020). Big players: Emotion in Twitter communities tweeting about global warming. In *Proceedings of the 33rd Canadian*

Conference on Artificial Intelligence, pp. 189–200.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3/4), 169–200.

Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129.

Ekman, P., Friesen, W. V. & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman (ed.), *Emotion in the Human Face* pp. 98–110. Cambridge University Press.

Eliacik, A. B. & Erdogan, N. (2018). Influential user weighted sentiment analysis on topic based microblogging community. *Expert Systems With Applications*, *92*, 403–418.

Eliasmith, C. (2015). *How to Build a Brain: A Neural Architecture for Biological Cognition*. Oxford University Press.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*(8), 709–724.

Evans, J. S. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. S. B. T. Evans & K. Frankish (eds.), *In Two Minds: Dual Processes and Beyond* pp. 33–54. Oxford University Press.

Evans, J. S. B. T. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3).

Farbotko, C. & Mcgregor, H. V. (2010). Copenhagen, climate science and the emotional geographies of climate change. *Australian Geographer*, *41*(2), 159–166.

Farzindar, A. & Inkpen, D. (2015). *Natural Language Processing for Social Media*. San Rafael, CA: Morgan & Claypool.

Fellbaum, C. (ed.) (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Fersini, E. (2017). Sentiment analysis in social networks: A machine learning perspective. In F. A. Pozzi, E. Fersini, E. Messina, & B. Liu (eds.), *Sentiment Analysis in Social Networks* pp. 91–111. Morgan Kaufmann.

Fownes, J. R., Yu, C. & Margolin, D. B. (2018). Twitter and climate change. *Wiley: Sociology Compass*, *12*(6), e12587.

Frank, E., Hall, M. A. & Witten, I. H. (2016). *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"* (4e éd.). Morgan Kaufmann.

Füegi, J. & Francis, J. (2003). Lovelace & Babbage and the creation of the 1843 'notes'. *IEEE Annals of the History of Computing*, *25*(4), 16–26.

Gandon, F., Faron-Zucker, C. & Corby, O. (2012). *Le Web Sémantique : Comment lier les données et les schémas sur le Web ?* Paris: Dunod.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. & Schneider, L. (2002). Sweetening ontologies with DOLCE. *Lecture Notes in Computer Science*, *2473*, 166–181.

Gangemi, A. & Peroni, S. (2016). The Information Realization pattern. In A. Gangemi, P. Hitzler, K. Janowicz, A. Krisnadhi, & V. Presutti (eds.), *Ontology Engineering with Ontology Design Patterns: Foundations and Applications*, volume 25 of *Studies on the Semantic Web* pp. 81–103. Berlin: Akademische Verlagsgellschaft.

Gil, R., Virgili-Gomá, J., García, R. & Mason, C. (2015). Emotions ontology for collaborative modelling and learning of emotional responses. *Computers in Human Behavior*, *51*, 610–617.

Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, D., Flanigan, J. & Smith, N. A. (2011). Part-of-speech tagging for Twitter: Annotation, features, and experiments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pp. 42–47.

Glimm, B., Horrocks, I., Motik, B., Stoilos, G. & Wang, Z. (2014). HermiT: An OWL 2 reasoner. *Journal of Automated Reasoning*, *53*(3), 245–269.

Gómez-Pérez, A., Fernández-López, M. & Corcho, O. (2004). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. London: Springer.

Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. The MIT Press.

Gray, J. A. (1982). *Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford Psychology Series. Oxford: Oxford University Press.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Guerra, P. C., Meira, W. & Cardie, C. (2014). Sentiment analysis on evolving social streams: How self-report imbalances can help. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining (WSDM 2014)*, pp. 443–452.

Haltinner, K. & Sarathchandra, D. (2018). Climate change skepticism as a psychological coping strategy. *Sociology Compass*, *12*(6), 84–85.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*, 335–346.

He, K., Zhang, X., Ren, S. & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 1026–1034.

Herrando-Pérez, S., Bradshaw, C. J. A., Lewandowsky, S. & Vieites, D. R. (2019). Statistical language backs conservatism in climate-change assessments. *BioScience*, *69*(3), 209–219.

Hickey, R. (2008). The Clojure programming language. In *Proceedings of the 2008 Symposium on Dynamic Languages*, pp. 1.

Hickey, R. (2020). A history of Clojure. In *Proceedings of the ACM on Programming Languages*, volume 4, issue HOPL, pp. 1–46.

Hitzler, P., Krötzsch, M. & Rudolph, S. (2010). *Foundations of Semantic Web Technologies*. Boca Raton: CRC Press.

Hofmann, T. (1999). Probabilistic Latent Semantic Analysis. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, pp. 289–296.

Holloway, S. W. (2015). *Named Entity Recognition in the Climate Change domain: An examination of NER systems for climatological knowledge discovery*. (Master's thesis). Norwegian University of Science and Technology.

Holte, R. C. (1993). Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, *11*, 63–91.

Honeycutt, C. & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *Proceedings of the 42nd Hawaii International Conference on System Sciences*.

Horridge, M. & Bechhofer, S. (2011). The OWL API: A Java API for OWL ontologies. *Semantic Web*, *2*(1), 11–21.

Horrocks, I. (2012). Semantics $\sqcap$ scalability $\models \bot$? *Journal of Zhejiang University: Science C*, *13*(4), 241–244.

Horrocks, I., Kutz, O. & Sattler, U. (2006). The even more irresistible $\mathcal{SROIQ}$. In *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR)*.

Horrocks, I., Patel-Schneider, P. F. & Harmelen, F. V. (2003). From SHIQ and RDF to OWL: The making of a Web Ontology Language. *Journal of Web Semantics*, *1*(1), 7–26.

Horrocks, I. & Sattler, U. (2004). Decidability of $\mathcal{SHIQ}$ with complex role inclusion axioms. *Artificial Intelligence*, *160*(1), 79–104.

Howarth, C. C. & Sharman, A. G. (2015). Labeling opinions in the climate debate: a critical review. *Wiley Interdisciplinary*, *6*(2), 239–254.

Hu, M. & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177.

Huang, G.-B. (2015). What are extreme learning machines? Filling the gap between Frank Rosenblatt's dream and John von Neumann's puzzle. *Cognitive Computation*, *7*(3), 263–278.

Iglesias, C. A., Sánchez-Rada, J. F., Vulcu, G. & Buitelaar, P. (2017). Linked data models for sentiment and emotion analysis in social networks. In F. A. Pozzi, E. Fersini, E. Messina, & B. Liu (eds.), *Sentiment Analysis in Social Networks* pp. 91–111. Morgan Kaufmann.

Inkpen, D. (2006). *Information Retrieval on the Internet*. Technical report, University of Ottawa.

IPPC (2013). Foreword. In T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. Midgley (eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

IPPC (2021). Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Izard, C. E. (1971). *The Face of Emotion*. Century Psychology Series. Appleton-Century-Crofts.

James, W. (1884). What is an emotion? *Mind*, *9*(34), 188–205.

Jolliffe, I. T. & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosphical Transactions of the Royal Society A*, *374*, 20150202.

Kaplan, A. M. & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, *54*(2), 105–113.

Kimmig, A., Mihalkova, L. & Getoor, L. (2015). Lifted graphical models: a survey. *Machine Learning*, *99*(1), 1–45.

Knill, D. C. & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.

Kong, L., Schneider, N., Swayamdipt, S., Bhatia, A., Dyer, C. & Smith, N. A. (2014). A dependency parser for tweets. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1001–1012.

Koyré, A. (1952). An unpublished letter of Robert Hooke to Isaac Newton. *Isis*, *43*(4), 312–337.

Krosnick, J. A. (2000). The impact of the fall 1997 debate about global warming on American public opinion. *Public Understanding of Science*, *9*(3), 239–260.

Ku, L.-W., Liang, Y.-T. & Chen, H.-H. (2006). Opinion extraction, summarization and tracking in news and blog corpora. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*.

Lafferty, J., McCallum, A. & Pereira, F. C. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the 18th International Conference on Machine Learning*, pp. 282–289.

Laird, J. E. (2012). The Soar cognitive architecture. *AISB Quarterly*, *171*, 224–235.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B. & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253.

le Cessie, S. & van Houwelingen, J. (1992). Ridge estimators in logistic regression. *Applied Statistics*, *41*(1), 191–201.

Leduc, T. B. (2007). Sila dialogues on climate change: Inuit wisdom for a cross-cultural interdisciplinarity. *Climatic Change*, *85*(3).

Leiserowitz, A. (2006). Climate change risk perception and policy preferences: The role of affect, imagery, and values. *Climatic Change*, *77*(1), 45–72.

Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinberg, G. & Rosenthal, S. (2016). *Climate Change in the American Mind: March 2016*. Yale project on climate change, Yale University and George Mason University

Leiserowitz, A., Maibach, E., Roser-Renouf, C. & Smith, N. (2010). *Global Warming's Six Americas, June 2010*. Yale project on climate change, Yale University and George Mason University

Leiserowitz, A., Maibach, E. W., Kotcher, J. E., Wang, X., Carman, J., Goldberg, M., Lacroix, K. & Marlon, J. (2021a). *Climate activism: A Six-Americas analysis – December 2020*. Technical report, Yale University and George Mason University, New Haven, CT: Yale Program on Climate Change Communication.

Leiserowitz, A., Roser-Renouf, C., Marlon, J. R. & Maibach, E. W. (2021b). Global warming's Six Americas: a review and recommendations for climate change communication. *Current Opinion in Behavioral Sciences*, *42*, 97–103.

Lewin, K. (1938). *The Conceptual Representation and the Measurement of Psychological Forces*. Duke University Press.

Li, Y. & Arce, G. R. (2004). A maximum likelihood approach to least absolute deviation regression. *EURASIP Journal on Advances in Signal Processing*, *2004*(12), 1762–1769.

Light, J. S. (1999). When computers were women. *Technology and Culture*, *40*(3), 455–483.

Lin, K. H.-Y., Yang, C. & Chen, H.-H. (2007). What emotions do news articles trigger in their readers? In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 733–734.

Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge: Cambridge University Press.

Lord, P. (2013). The Semantic Web takes wing: Programming ontologies with Tawny-OWL. *CoRR*, *abs/1303.0213*.

Louviere, J. J., Flynn, T. N. & Marley, A. A. J. (2015). *Best-Worst Scaling: Theory, Methods and Applications*. Cambridge University Press.

Lu, H. & Schuldt, J. P. (2015). Exploring the role of incidental emotions in support for climate change policy. *Climatic Change*, *131*(4), 719–726.

Maass, W. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, *14*(11), 2531–2560. URLhttp://dx.doi.org/10.1162/089976602760407955

Maibach, E., Roser-Renouf, C. & Leiserowitz, A. (2009). *Global Warming's Six Americas: An Audience Segmentation Analysis*. George Mason University.

Maibach, E. W., Leiserowitz, A., Roser-Renouf, C. & Mertz, C. K. (2011). Identifying like-minded audiences for global warming public engagement campaigns: An audience segmentation analysis and tool development. *PLoS: ONE*, *6*(3), e17571.

Makki, R., Soto, A. J., Brooks, S. & Milios, E. (2015). Active information retrieval for linking Twitter posts with political debates. In *Proceedings of the IEEE 14th International Conference on Machine Learning and Applications (ICMLA-2015)*, pp. 238–245.

Marcus, M. P., Santorini, B. & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, *19*(2), 313–330.

Martins, A. F. T., Almeida, M. B. & Smith, N. A. (2013). Turning on the turbo: Fast third-order non-projective turbo parsers. In *Proceedings of the Association for Computational Linguistics*.

Martins, A. F. T., Smith, N. A., Aguiar, P. M. Q. & Figueiredo, M. A. T. (2011). Dual decomposition with many overlapping components. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP-2011)*.

Marwick, A. E. & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133.

Maynard, D. & Bontcheva, K. (2015). Understanding climate change tweets: an open source toolkit for social media analysis. In *Proceedings of EnviroInfo and ICT for Sustainability*, pp. 242–250.

McCandless, M., Hatcher, E. & Gospodnetić, O. (2010). *Lucene in Action*. Stamford, CT: Manning.

McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, part I. *Communications of the ACM*, *3*(4), 184–195.

McCright, A. M. (2011). Political orientation moderates Americans' beliefs and concern about climate change. *Climatic Change*, *104*(2), 243–253.

McCright, A. M. & Dunlap, R. E. (2010). Anti-reflexivity: The American conservative movement's success in undermining climate science and policy. *Theory, Culture & Society*, *27*(2-3), 100–133.

McCright, A. M. & Dunlap, R. E. (2011). Cool dudes: The denial of climate change among conservative white males in the United States. *Global Environmental Change*, *21*(4), 1163–1172.

McCright, A. M., Marquart-Pyatt, S. T., Shwom, R. L., Brechin, S. R. & Allen, S. (2016). Ideology, capitalism, and climate: Explaining public views about climate change in the United States. *Energy Research & Social Science*, *21*, 180–189.

McDougall, W. (1926). *An Introduction to Social Psychology* (revised éd.). Boston: John W. Luce.

Mellon, J. & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*, *4*(3), 1–8.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*, *7*(3), 141–144.

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S. & Danforth, C. M. (2013). The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS: ONE*, *8*(5), e64417.

Mohammad, S. M. (2012). #Emotional tweets. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics*, pp. 246–255.

Mohammad, S. M. (2018). Word affect intensities. In *Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*, pp. 174–183.

Mohammad, S. M. & Bravo-Marquez, F. (2017). Emotion intensities in tweets. In *Proceedings of the Joint Conference on Lexical and Computational Linguistics*, pp. 65–77.

Mohammad, S. M. & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465.

Monard, M. C. & Batista, G. E. A. P. A. (2002). Learning with skewed class distributions. In *Advances in Logic, Artificial Intelligence and Robotics (LAPTEC-2002)*, pp. 173–180.

Mowrer, O. H. (1960). *Learning Theory and Behavior*. New York: Wiley.

Mulligan, K. & Scherer, K. R. (2012). Toward a working definition of emotion. *Emotion Review*, *4*(4), 345–357.

Munikar, M., Shakya, S. & Shrestha, A. (2019). Fine-grained sentiment classification using BERT. In *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, pp. 1–5.

Musen, M. A. (2015). The Protégé project: A look back and a look forward. *AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence*, *1*(4), 4–12.

Nabi, R. L., Gustafson, A. & Jensen, R. (2018). Framing climate change: Exploring the role of emotion in generating advocacy behavior. *Science Communication*, *40*(4), 442–468.

Nadeau, C. & Bengio, Y. (2003). Inference for the generalization error. *Machine Learning*, *52*, 239—281.

Negnevitsky, M. (2011). *Artificial Intelligence : A Guide to Intelligent Systems* (3e éd.). Harlow, England: Addison Wesley.

Newell, A. (1990). *Unified Theories of Cognition*. The William James Lectures. Harvard University Press.

Nivre, J. (2010). Dependency parsing. *Language and Linguistics Compass*, *4*(3), 138–152.

Nuttall, M. (2010). Anticipation, climate change, and movement in Greenland. *Études/Inuit/Studies*, *34*(1), 21–37.

Oatley, K. & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, *1*(1), 29–50.

Ochoa-Luna, J. E., Revoredo, K. & Cozman, F. G. (2011). Learning probabilistic description logics: A framework and algorithms. In I. Batyrshin & G. Sidorov (eds.). *Advances in Artificial Intelligence (MICAI 2011)*, volume 7094 of *Lecture Notes in Computer Science*, pp. 28–39. Springer.

Ortony, A. & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315–331.

Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N. & Smith, N. A. (2013). Improved part-of-speech tagging for online conversational text with word clusters. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2013)*, pp. 380–390.

Pan, J. Z. (2009). Resource Description Framework. In S. Staab & R. Studer (eds.), *Handbook on Ontologies* pp. 71–90. Berlin: Springer.

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, *5*(3), 407–422.

Parrott, W. G. (ed.) (2001). *Emotions in Social Psychology: Essential Readings.* Psychology Press.

Pereira, J. M., Basto, M. & da Silva, A. F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance*, *39*, 634–641.

Pinker, S. (1997/2009). *How the Mind Works.* New York: W. W. Norton.

Plutchik, R. (1980). A general psychoevoltionatry theory of emotion. In R. Plutchik & H. Kellerman (eds.), *Emotion: Theory, Research, and Experience*, volume 1 of *Theories of Emotion* pp. 3–33. Academic Press.

Plutchik, R. (2000). *Emotions in the Practice of Psychotherapy: Clinical Implications of Affect Theories.* Washington, D.C: American Psychological Association.

Plutchik, R. (2001). The nature of emotions. *American Scientist*, *89*(4), 344–350.

Porter, M. F. (1980/2006). An algorithm for suffix stripping. *Program*, *40*(3), 211–218.

Pozzi, F. A., Fersini, E., Messina, E. & Liu, B. (2017). Challenges of sentiment analysis in social networks: An overview. In F. A. Pozzi, E. Fersini, E. Messina, & B. Liu (eds.), *Sentiment Analysis in Social Networks* pp. 1–11. Morgan Kaufmann.

Preoţiuc-Pietro, D., Liu, Y., Hopkins, D. J. & Ungar, L. (2017). Beyond binary labels: Political ideology prediction of Twitter users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL-2017)*, pp. 729–740.

Preoţiuc-Pietro, D., Volkova, S., Lampos, V., Bachrach, Y. & Aletras, N. (2015). Studying user income through language, behaviour and affect in social media. *PLoS ONE*, *10*(9), e0138717.

Presutti, V. & Gangemi, A. (2016). Dolce+D&S Ultralite and its main ontology design patterns. In A. Gangemi, P. Hitzler, K. Janowicz, A. Krisnadhi, & V. Presutti (eds.), *Ontology Engineering with Ontology Design Patterns: Foundations and Applications*, volume 25 of *Studies on the Semantic Web* pp. 81–103. Berlin: Akademische Verlagsgellschaft.

Quinlan, J. R. (1992). Learning with continuous classes. In *5th Australian Joint Conference on Artificial Intelligence*, pp. 343–348., Singapore. World Scientific.

Quinlan, J. R. (1993). *C4.5 Programs for Machine Learning*. Morgan Kaufmann.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Ritter, R. S., Preston, J. L. & Hernandez, I. (2013/2014). Happy tweets: Christians are happier, more socially connected, and less analytical than atheists on Twitter. *Social Psychological and Personality Science*, *5*(2), 243–249.

Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M. & Gatford, M. (1995). Okapi at TREC-3. In *Special Publication (NIST SP) - 500-225: Overview of the Third Text REtrieval Conference (TREC-3)*, pp. 109–126.

Robertson, S. E. & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval*, *3*(4), 333–389.

Romera-Paredes, B. & Torr, P. (2015). An embarrassingly simple approach to zero-shot learning. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37, pp. 2152–2161., Lille, France. PMLR.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Sahoo, S. S., Halb, W., Hellmann, S., Idehen, K., Thibodeau Jr., T., Auer, S., Sequeda, J. & Ezzat, A. (2009). A survey of current approaches for mapping of relational databases to RDF. *W3C*.

Salguero, A. & Espinilla, M. (2016). Description logic class expression learning applied to sentiment analysis. In W. Pedrycz & S.-M. Chen (eds.), *Sentiment Analysis and Ontology Engineering*, volume 639 of *Studies in Computational Intelligence* pp. 93–111. Springer.

Salguero, A. & Espinilla, M. (2018). A flexible text analyzer based on ontologies: An application for detecting discriminatory language. *Language Resources and Evaluation*, *52*(1), 185–215.

Samwald, M. & Adlassnig, K.-P. (2008). The bio-zen plus ontology. *Applied Ontology*, *3*(4), 213–217.

Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H. & Jin, C. (2007). Red Opal: Product-feature scoring from reviews. In *Proceedings of the Eighth Annual Conference on Electronic Commerce*, pp. 182–191.

Schäfer, M. S. (2012). Online communication on climate change and climate politics: a literature review. *Wiley Interdisciplinary Reviews: Climate Change*, *3*(6), 527–543.

Schneider, N., O'Connor, B., Saphra, N., Bamman, D., Faruqui, M., Smith, N. A., Dyer, C. & Baldridge, J. (2013). A framework for (under)specifying dependency syntax without overloading annotators. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*.

Searle, J. R. (1995). *The Construction of Social Reality*. New York: Free Press.

Shadbolt, N., Berners-Lee, T. & Hall, W. (2006). The Semantic Web revisited. *IEEE Intelligent Systems*, *21*(3), 96–101.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(4), 623–656.

Sheehan, K. B. (2018). Crowdsourcing research: Data collection with Amazon's Mechanical Turk. *Communication Monographs*, *85*(1), 140–156.

Studer, R., Benjamins, V. & Fensel, D. (1998). Knowledge engineering: Principles and methods. *Knowledge Engineering*, *25*(1), 161–197.

Suler, J. (2005). The online disinhibition effect. *International Journal of Applied Psychoanalytic Studies*, *2*(2), 184–188.

Sulis, E., Irazú Hernández Farías, D., Rosso, P., Patti, V. & Ruffo, G. (2016). Figurative messages and affect in Twitter: Differences between #irony, #sarcasm and #not. *Knowledge-Based Systems*, *108*, 132–143.

Sylwester, K. & Purver, M. (2015). Twitter language use reflects psychological differences between Democrats and Republicans. *PLoS: ONE*, *10*(9), e0137422.

Szeredi, P., Lukácsy, G. & Benkö, T. (2014). *The Semantic Web Explained: The Technology and Mathematics behind Web 3.0*. Cambridge University Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *59*(1), 267–288.

Tomkins, S. (1984). Affect theory. In K. R. Scherer & P. Ekman (eds.), *Approaches to Emotion* pp. 163–195. Hillsdale, NJ: Erlbaum.

Trotman, A., Puurula, A. & Burgess, B. (2014). Improvements to BM25 and language models examined. In *Proceedings of the Australasian Document Computing SymposiumNovember*, pp. 58–65.

van der Linden, S., Leiserowitz, A. & Maibach, E. (2019). The gateway belief model: A large-scale replication. *Journal of Environmental Psychology*, *62*, 49–58.

Vigna, S. (2016). Further scramblings of Marsaglia's xorshift generators. *Journal of Computational and Applied Mathematics*, *315*, 175–181.

Voorhees, E. M. (1994). Query expansion using lexical-semantic relations. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-1994)*, pp. 61–69.

Wang, F. L., Yin, J., Lau, R. Y. K., Xie, H., Rao, Y. & Li, X. (2017). Bootstrapping social emotion classification with semantically rich hybrid neural networks. *IEEE Transactions on Affective Computing*, *8*(4), 428–442.

Wang, Y. & Witten, I. H. (1997). Induction of model trees for predicting continuous classes. In *Poster papers of the 9th European Conference on Machine Learning*. Springer.

Watson, J. B. (1930). *Behaviorism.* Chicago University Press.

Weber, E. (2006). Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climate Change*, *77*(1), 103–120.

Weiner, B. & Graham, S. (1984). Attributional approach to emotional development. In C. E. Izard, J. Kagan, & R. B. Zajonc (eds.), *Emotions, Cognition, and Behavior* pp. 167–191. New York: Cambridge University Press.

Weisberg, S. (2014). *Applied Linear Regression* (4e éd.). Hoboken, NJ: Wiley.

Weiss, G. M. (2004). Mining with rarity: a unifying framework. *ACM SIGKDD Explorations Newsletter*, *6*(1), 7–19.

Williams, H. T., Mcmurray, J. R., Kurz, T. & Lambert, F. H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, *32*, 126–138.

Witten, I. H. & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2e éd.). Morgan Kaufmann.

Wojcik, S. & Hughes, A. (2019). *Sizing Up Twitter Users*. Technical report, Pew Research Center.

Yamada, H. & Matsumoto, Y. (2003). Statistical dependency analysis with support vector machines. In *Proceedings of the 2003 International Conference on Parsing Technologies*, pp. 195–206.

Yang, H., Alsadooa, A., Prasad, P. W. C., Al-Dala'in, T., Rashid, T. A., Maag, A. & Alsadoon, O. H. (2022). Deep learning neural networks for emotion classification from text: enhanced leaky rectified linear unit activation and weighted loss. *Multimedia Tools and Applications: An International Journal*, *8*(11), 15439–15468.

Ye, N., Chai, K. M. A., Lee, W. S. & Chieu, H. L. (2012). Optimizing F-measures: A tale of two approaches. In *Proceedings of the 29th International Conference on Machine Learning (ICML-2012)*, pp. 289–296.