

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

HASARD ET INDÉTERMINATION DANS LE
DOCUMENTAIRE INTERACTIF

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DOCTORAT CONJOINT EN COMMUNICATION

PAR
JULIA CHAIM SALLES DUCAMIN

MAI 2021

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

À mes parents, Sandra et Paulo

REMERCIEMENTS

Une des leçons les plus importantes que j'ai apprises dans le doctorat est que la recherche est un sport d'équipe (et d'endurance !), et j'ai eu la chance et le bonheur de compter sur une « équipe de rêve » à mes côtés tout au long de mon parcours doctoral. À commencer par mon directeur de thèse, Louis-Claude Paquin, qui m'a indiqué le chemin et m'a soutenu depuis notre première rencontre. Lorsque mes idées étaient encore très floues, il m'a fait confiance et m'a orienté pendant ce long et ardu parcours doctoral, un énorme merci, Louis-Claude !

Je remercie également les membres du jury qui ont généreusement accepté de participer à l'évaluation de cette thèse, ainsi qu'aux examens qui intègrent le parcours doctoral. Un énorme merci à Brian Massumi qui, au-delà de sa contribution lors de l'examen de projet, m'a fait renouer avec la philosophie lors des séminaires que j'ai eu l'honneur et le plaisir de suivre dans le cadre du programme de doctorat conjoint. Je remercie Charles Perraton pour sa participation dans le jury de mes examens de synthèse et de projet. Ses interventions combinaient toujours une rigueur intellectuelle profonde avec une grande douceur. Merci beaucoup à Thierry Bardini, qui m'a fait l'honneur de participer à tous mes examens doctoraux : je vous remercie grandement pour votre contribution depuis mes premiers pas dans le doctorat. Je remercie également Sofian Audry pour sa participation dans l'évaluation de cette thèse. Un grand merci aussi à Sandra Gaudenzi : c'est en lisant ton blogue sur ta recherche doctorale que j'ai trouvé le nom de ce que je voulais faire, le documentaire interactif. Merci, Sandra, de m'avoir donné un champ et d'être une référence centrale à la fois dans la création de médias interactifs et dans l'art de travailler ensemble.

Je remercie aussi toutes et tous mes collègues du programme de doctorat conjoint : la diversité de perspectives et d'approches de la recherche et de la recherche-création en communication que mes collègues ont apportée à chaque séminaire et à chaque rencontre a enrichi de manière inestimable, non seulement mon travail, mais aussi ma vision du monde. Je n'ai pas de mots pour remercier mon collègue et ami, Luciano Frizzera. Sans la perspective critique de Luciano, et son talent pour penser et créer les algorithmes, la création de *On Chance* n'aurait pas été possible. Muito obrigada, Luciano! Estamos juntos, parceiro! Un énorme merci également à ma collègue et amie Myriam Rafla, qui a grandement contribué à la création du récit de *On Chance*. Je me souviens encore de notre première rencontre dans le cadre d'un séminaire et d'avoir la sensation que c'était le début, pas seulement d'une collaboration à venir, mais aussi d'une belle amitié. Je remercie également mes collègues du Forum doctoral, notamment celles et ceux du groupe d'écriture, Momoko Allard, Caroline Martel, Celia Vara, Rad Gagnon, Martin Bonnard et Felix Faucher : nos rencontres étaient pour moi une source incroyable de force et d'inspiration, merci beaucoup!

Je remercie également tous les professeurs et professeuses qui m'ont enseigné pendant mon parcours doctoral : Andra McCartney, Kim Sawchuk, Krista Lynes, Jean Décarie, Éric George, Gilles Coutlé, Jeremy Stolow, et à nouveau, Louis-Claude Paquin et Brian Massumi : quelle chance et quel honneur d'avoir été en classe avec chacune et chacun de vous ! Un grand merci à Yannick Richer pour son soutien et sa grande empathie envers les étudiant.e.s étranger.e.s de l'UQAM.

Pendant la réalisation de mon doctorat, j'ai eu le plaisir de collaborer dans différents projets de recherche et de recherche-création, et chacun de ces projets a été important dans le développement de mon parcours. Je remercie Erin Manning pour son ouverture, son accueil toujours chaleureux et généreux lors des activités du SenseLab. Un grand merci également à Ana Ramos ainsi qu'à toutes et tous les participants du Brazil Hub : certains passages de cette thèse ont été imaginés avec vous, lors de nos discussions, muito obrigada a tod@s vocês ! Ana, querida, muito obrigada por sua luz! Obrigada

também à querida Céline Pereira, toujours solidaire ! Merci à Maude Bonenfant et Gabrielle Trépanier-Jobin qui m'ont accueillie à bras ouvert dans le premier groupe de recherche que j'ai intégré, Homo Ludens. Merci également à Bart Simon, Harry Smoak, Marc Beaulieu et toute l'équipe de Milieux Institute : vous avez réussi à créer un espace où l'on est heureux de faire de la recherche, un grand merci ! Dans Milieux j'ai eu la chance de collaborer avec *so many brilliant minds!* *Huge thanks to* Marco Luna, Daniel Cross, Olivia McGilchrist, Seyed Tabatabaei, Dougy Hérard, Gada Jane, Jessie Marchessault (*we have to finish that paper!* :), Myriam Durocher, Samuel Thulin, a querida Sâmia Pedraça, *and all the others I had the pleasure to collaborate with.* Je remercie spécialement Julia Zamboni, ma très chère collaboratrice et amie : a gente ja fez muita coisa legal juntas, mas tenho certeza que tem mais por vir (inclusive ler o Bateson!), obrigada por sua criatividade e carinho, querida! Je remercie énormément André Paz, grande parceiro, avec qui je collabore depuis de nombreuses années et avec qui j'ai le bonheur d'être amie depuis encore plus d'années, valeu demais, André ! Merci aussi *ao companheiro e amigo querido* Cristiano Therrien qui m'a fait confiance et m'a fait venir dans un coin de l'université où je n'aurais jamais imaginé travailler, la Faculté de Droit :) obrigada pela força, compa!

Je remercie également les professeurs et collègues avec qui j'ai le plaisir de collaborer dans mes projets de recherche actuels, Simon Harel, Normand Roy, Jean-Simon DesRochers, Léonore Brassard, et toute l'équipe du projet « Montréal, territoire et terrain ». Muito obrigada também a toda equipe brasileira do projet, em especial à querida Kelly Russo. Je remercie aussi tous les collègues de projet « Chatbot et récit », notamment Maria Grullon. Je remercie tout particulièrement le professeur Francisco Loiola, grande amigo, companheiro de luta e parceiro de trabalho : muito obrigada por tudo, grande Loiola! Seu apoio e sua confiança foram fundamentais para que eu tivesse condições de terminer esse doutorado! Je remercie également mes collègues du Laboratoire multimédia de la Faculté des arts et des sciences de l'Université de Montréal qui m'ont toujours soutenu avec beaucoup d'efficacité et de bonne humeur

lors de mes charges de cours, merci à Julie Bélanger, David Cherniak, Francis St-Louis, Mario Jacob et toute l'équipe. Un grand merci à l'équipe du Centre Sportif de l'UQAM, spécialement Andrée Dionne, avec qui j'ai eu le plaisir de collaborer sur de nombreux projets. Je remercie également François Lanctôt : merci beaucoup pour ta confiance et pour les échanges fascinants sur la technologie, les sciences, la musique, et tant d'autres intérêts partagés !

Je souhaite remercier également les différents collectifs dans lesquels je me suis engagée pendant le doctorat. Face au trouble politique qui vit le Brésil depuis quelques années, il a été très important pour moi de m'engager dans la communauté brésilienne à l'étranger. Muito obrigada às companheiras e companheiros do Coletivo Brasil-Montréal! Je remercie également mes collègues du collectif Brain (Brasileiros pelo Avanço e Internacionalização do Conhecimento), spécialement Deborah Maia. Je remercie en particulier mes amis qui ont participé des tests et entraînements réalisés dans les différents processus de création du doctorat, un grand merci à mes très chers amis Rosa Peralta, Vanessa Conti, Dalila Vasconcelos, Felipe Pio, Priscylla Joca, Graziela Tanaka, Abel de Castro, Ana Luiza Cascão, Felipe Carrelli, e minha irmã querida Beatriz Salles.

Je remercie mes très chers parrains, Sônia Proto et Raimundo Marcone Proto, pour leur soutien chaleureux et présent, toujours! Muito obrigada à querida Vilma Arêas, pelo carinho e pelo apoio! Merci à toute ma famille, toujours présente même de loin ! Je remercie tout particulièrement mes parents, Sandra Abrahão Chaim et Paulo Fernando Guimarães Salles, pour leur amour et soutien inconditionnels. Je remercie également, du fond du cœur, Simon Ducamin et nos enfants Louise et Thomas : vous m'avez tellement soutenue, supportée, écouté, aidé, et encore pendant ces presque dix ans de doctorat, que les mots me manquent pour exprimer à quel point je vous suis reconnaissante! Vous avez fait de ces années de doctorat un des plus beaux moments de ma vie.

Enfin, je souhaite remercier l'agence *Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior)* du Ministère de l'Éducation du gouvernement fédéral du Brésil, pour le soutien financier qui m'a permis de réaliser les études doctorales. La subvention que j'ai obtenue fait partie d'un moment de l'histoire du Brésil où le gouvernement fédéral avait des politiques publiques d'investissement dans l'éducation, la science et l'innovation. À nos jours, malheureusement les étudiants et chercheurs brésiliens doivent faire face, au-delà des difficultés inhérentes à la recherche et au doctorat, à un gouvernement qui méprise la science, l'éducation et la pensée critique. J'exprime ici ma gratitude au programme *Ciência sem Fronteiras (2011-2017)*, sans lequel je n'aurais pas pu réaliser la présente thèse doctorale.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
LISTE DES FIGURES.....	xi
RÉSUMÉ	xiii
ABSTRACT	xv
CRÉATION : ON CHANCE	xvii
INTRODUCTION	1
1. CHAPITRE I.....	9
Introduction au problème du hasard et de l'indétermination dans le documentaire interactif	9
1.1. Questions de recherche-crédation	9
1.2. Cadre conceptuel	16
1.2.1. Hasard et indétermination.....	16
1.2.2. Documentaire interactif	28
1.3. Posture : Pragmatisme spéculatif	44
1.4. Bricolage méthodologique : La recherche-crédation et le cadrage de l'inopiné	48
Première partie : Discussion théorique	53
2. CHAPITRE II	53
L'analogique : Image-interaction	53
2.1. L'expérience du documentaire interactif.....	56
2.2. Centre d'indétermination et image-interaction.....	64
2.3. Le documentaire interactif comme agencement.....	73

3. CHAPITRE III	80
Le numérique : Hasard et information	80
3.1. Interprétations de la théorie des probabilités.....	83
3.2. Théorie de l’information : hasard et entropie	102
3.3. Théorie algorithmique de l’information : hasard et complexité.....	117
4. CHAPITRE IV	126
Machine de Turing probabiliste, algorithmes randomisés et <i>machine learning</i>	126
4.1. Machines de Turing non déterministes, Machines de Turing probabilistes et algorithmes randomisés	127
4.2. Machine Learning (apprentissage automatique) : Algorithmes de Recast.AI (clustering), réseau neuronal artificiel et <i>probabilistic programming</i>	134
4.3. Machine Learning (apprentissage automatique) : Interprétation des probabilités dans les modèles cognitifs et algorithmes d’apprentissage	146
Deuxième partie : Discussion autour de la création.....	161
5. CHAPITRE V	161
On Chance : un documentaire interactif avec et sur le hasard	161
5.1. Présentation de On Chance.....	161
5.2. Chatbot et récit	163
5.3. La perspective de la vulgarisation scientifique	184
6. CHAPITRE VI.....	192
Récit de pratique	192
6.1. Premières années du doctorat	192
6.2. Stage de création	196
6.3. Forum doctoral	200

6.4. Réalisation	203
6.4.1. Chatbot.....	203
6.4.2. Design de conversation.....	207
6.5. Discussion	213
CONCLUSION.....	218
ANNEXE A	226
Documentation video On Chance.....	226
ANNEXE B.....	227
Scénario de On Chance (version du 2 mai 2021).....	227
BIBLIOGRAPHIE	292

LISTE DES FIGURES

Figure 1 : Capture d'écran de Twitter	1
Figure 2: Aspen Movie Map. Source : Inventing Interactive	29
Figure 3 : Interface Rio de Janeiro — Autorretrato. Source : Capture d'écran	34
Figure 4 : Interface Le Défi des Bâisseurs. Source : Capture d'écran	37
Figure 5 : Interface web Walking the Edit. Source : Capture d'écran	41
Figure 6 : Classes de complexité algorithmique	129
Figure 7 : Processus cognitif. Source : K. Friston et al., 2016.....	151
Figure 8 : Clustering de formes complexes. Source : Ayasdi (https://www.ayasdi.com)	159
Figure 9 : Interface de On Chance	163
Figure 10 : Capture d'écran de discussion sur le réseau social Twitter	165
Figure 11 : Page titre A Tendency to Spill (Grian, 2017a)	166
Figure 12: Page “call to action” A Tendency to Spill (Grian, 2017a)	166
Figure 13 : Interface 2 de Enterre-moi, mon Amour (2019)	170
Figure 14 : Interface 1 de Enterre-moi, Mon Amour.....	170
Figure 15 : Scénario interactif (Fonseca, 2019)	171
Figure 16 : Interfaces de réglages avec les options de « vitesse du jeu » et « vitesse d'apparition des messages »	172
Figure 17 : Interface Nour « occupée »	172
Figure 18 : Interfaces de Fabi (2018).....	173
Figure 19 : Interfaces de Yeshi	175
Figure 20 : Interfaces de HappyBot	176
Figure 21 : Illustration de l'article de Harry Stevens « Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve” (Stevens, 2020)	185

Figure 22 : Modèles de vulgarisation scientifique (Brossard & Lewenstein, 2009, p. 17)	190
.....	
Figure 26 : A Partida de Xadrez, Maria Helena Vieira da Silva, 1943.....	193
Figure 27 : Premiers essais de texture 3D pour l'esthétique du labyrinthe.....	194
Figure 28 : Illustration proposée au début du stage doctoral	198
Figure 29 : Carte de Montréal avec une grille [49x49], utilisée pour le repérage des lieux de tournage	199
Figure 30 : Proposition de structure thématique	202
Figure 31 : Ressources pour chaque thématique.....	203
Figure 32 : Interface de Recast.AI	206
Figure 33 : Interface de Dialogflow	207
Figure 34 : Mindmap du scénario	210
Figure 35 : Scénario	210
Figure 36 : Interface de base de données	211
Figure 37 : Page d'accueil de On Chance	212
Figure 38 : Interface principale de On Chance	212

RÉSUMÉ

Dans cette recherche-cr ation, il sera question de mettre en  vidence les pr esuppos s   la base de l'utilisation des mod les pr edictifs dans les technologies num riques de communication, plus sp cifiquement, dans l'utilisation des algorithmes d'apprentissage automatique comme outil d'interactivit  dans un *chatbot* documentaire. Il s'agira d'explorer les mani res dont les pr esuppos s li s au hasard et   l'ind termination interviennent dans les pratiques de cr ation et d'exp rience du documentaire interactif, d'exp rimer avec des modes de cr ation qui embrassent l'incertitude et l'impr visible et aussi d'approcher le hasard et les algorithmes pr edictifs comme th matique dans la cr ation d'un documentaire interactif. Cette recherche doctorale s'initie par la question de l'interpr tation du r le du hasard et de l'ind termination dans les m dias interactifs, plus sp cifiquement le documentaire interactif. Il est question de comprendre comment le hasard agit dans les processus analogiques et num riques mis en place dans un documentaire interactif. Or, si les documentaires interactifs sont des programmes num riques, des algorithmes  crits dans un code binaire pour effectuer des calculs selon des r gles pr d termin es, quelles sont alors les conditions pour la cr ation d'ind termination dans la dimension num rique du documentaire interactif? Comment comprenons-nous les notions de hasard, ind termination, incertitude, entropie (et ses relations avec l'information et le bruit) dans le contexte des technologies num riques de communication utilis es dans les documentaires interactifs? Comme nous d montrons, ces questions sont importantes non seulement par ses aspects th oriques, mais  galement par les cons quences pratiques de l'utilisation de plus en plus fr quente de mod les pr edictifs bas s sur le traitement statistique de donn es num riques. Ce probl me est abord  par la perspective du pragmatisme sp culatif, ce qui permet de reconna tre les aspects processuels et relationnels du hasard et de l'ind termination, et de proposer une

compréhension non dualiste de l'expérience du documentaire interactif, ainsi que des notions d'information et de probabilité. Le problème est cadré par la création du documentaire interactif conversationnel *On Chance*.

Mots-clés : documentaire interactif ; médias numériques ; *chatbot* ; agent conversationnel ; probabilité ; intelligence artificielle.

ABSTRACT

In this research-creation, my aim is to examine the presuppositions underlying the use of predictive models in digital communication technologies, more specifically, in the use of machine learning algorithms as a tool for interactivity in a documentary chatbot. I will explore the ways in which the presuppositions about randomness and indeterminacy intervene in the practices of the interactive documentary, to experiment with modes of creation that embrace uncertainty and unpredictable events and also to approach chance and predictive algorithms as themes for an interactive documentary. This doctoral research is initiated by the question of the role of chance and indeterminacy in interactive media, more specifically interactive documentaries. The research addresses the problem of chance in the analog and digital processes acting in interactive documentary. If interactive documentaries are digital programs, algorithms written in binary code to perform calculations according to predetermined rules, then what are the conditions for the creation of indeterminacy in the digital dimension of the interactive documentary? How do we understand the notions of chance, indeterminacy, uncertainty, entropy (and its relationship to information and noise) in the context of digital communication technologies used in interactive documentaries? As I demonstrate, these questions are important not only for their theoretical aspects, but also for the practical consequences of the increasingly frequent use of predictive models based on the statistical processing of numerical data. This problem is approached from the perspective of speculative pragmatism, which makes it possible to recognize the processual and relational aspects of chance and indeterminacy, and to offer a non-dualistic understanding of the interactive documentary experience, as well as notions of information and probability. The problem is framed by the creation of the conversational interactive documentary *On Chance*.

Keywords: interactive documentary; digital media; chatbot; probability; artificial intelligence.

CRÉATION : ON CHANCE

La thèse est formée du présent essai doctoral et du travail de création *On Chance*, un documentaire interactif disponible sur la page suivante : <https://onchance.net/>

INTRODUCTION

« AI is made of three parts: a data set, a learning algorithm, and a prediction »

(Hao, 2020)

L'idée que l'intelligence artificielle est un ensemble de technologies capables de faire (ou du moins d'aider à faire) des prévisions est très présente dans les communications de science et technologie contemporaines. Parfois cette idée est poussée à des titres éloquentes notamment dans les publications de vulgarisation scientifique, comme celui utilisé par la revue *New Scientist* pour publiciser un article (Lawton, 2019) sur le réseau social *Twitter* :



Figure 1 : Capture d'écran de Twitter

La confiance dans la capacité des modèles et systèmes informatiques à faire des prévisions justes a des conséquences pratiques importantes si l'on considère, par exemple, que de plus en plus d'opérations courantes de la vie quotidienne (tels que les décisions sur les demandes de crédit, les processus d'embauche, entre autres) sont basées sur des prévisions d'algorithmes d'intelligence artificielle. Dans l'article mentionné ci-dessus Lawton décrit la *multi-agent artificial intelligence* (MAAI), une technologie que "(...) allows predictions to be made with extraordinary accuracy by testing them in highly detailed simulations that amount to entire artificial societies." (Lawton, 2019)

Les applications contemporaines des systèmes prédictifs du type de la MAAI sont nombreuses et peuvent aller des algorithmes de recommandation de contenu sur internet jusqu'à des systèmes plus complexes, comme celui développé dans le domaine des soins à domicile qui utilise l'analyse des données biométriques pour réduire le nombre d'hospitalisations des seniors¹, ou encore les systèmes de « police prédictive » (*predictive policing*, en anglais), comme *PredPol*, développé aux États-Unis et utilisé aujourd'hui par des bureaux de police dans plusieurs pays (Maguire, Rao, & Zurawski, 2018). Cette vocation des algorithmes d'intelligence artificielle, et plus particulièrement des technologies dites « d'apprentissage automatique » (ou *machine learning*, en anglais), à faire des prévisions est fortement liée à l'utilisation de modèles statistiques et de probabilités pour l'analyse de larges bases de données. La probabilité et la statistique sont les outils mathématiques pour introduire dans ces algorithmes les aspects incertains et aléatoires présents dans les réalités complexes qu'ils essaient de modéliser.

¹ Système développé par l'entreprise canadienne AlayaCare.

Derrière l'idée que les algorithmes d'apprentissage automatique sont des outils efficaces dans la création des modèles prédictifs, il existe des présupposés (souvent non dits) concernant la vision du monde à la base des prévisions, l'interprétation du rôle de la probabilité dans les modèles prédictifs et les conceptions de hasard et d'indétermination. La posture souvent implicite est qu'il suffit d'avoir plus de données et des modèles statistiques plus raffinés pour être capable de prédire, ou encore de contrôler, les événements à venir. Mais comme Maguire suggère dans le contexte des tactiques prédictives de la police, il est important de mettre en évidence et discuter les présupposés des systèmes prédictifs :

(...) if the critical social sciences simply engage with new policing and security technologies in terms of their possible nefarious uses, we will lose the possibility of genuine critique, by which I mean an understanding of the core assumptions from which those technologies emerged and the possible alternatives available at root. (Maguire et al., 2018, p. 154)

Malgré les tendances déterministes qui peuvent être associées à cet idéal de prévision des algorithmes d'apprentissage automatique, les concepts de hasard, d'entropie et d'indétermination sont au cœur de nombreuses théories à la base des champs de développement de ces mêmes technologies, comme la théorie de l'information (Atlan, 2006 [1972]; Shannon, 1948) et la théorie de l'information algorithmique (Chaitin, 1975). Comme on le verra plus en détail, les études en communication et technologies de l'information s'insèrent dans un tournant non déterministe qui commence à la fin du 19^e siècle marqué notamment par l'utilisation de modèles statistiques et de la probabilité dans les sciences (Morin, 1982, 1990; Prigogine, 2001; Prigogine & Stengers, 1979; Shannon, 1948). Selon cette perspective, l'incertitude et l'indétermination sont au cœur des systèmes d'information. Comment aborder alors cette apparente contradiction entre la vocation prédictive des algorithmes d'apprentissage automatique et la centralité des notions de hasard, d'indétermination et d'entropie dans les systèmes d'information ?

Il est important de réaffirmer que ces questions concernant les présupposés liés à l'utilisation des modèles prédictifs peuvent engendrer des effets pratiques importants, considérant les nombreuses applications utilitaires de ces modèles. Autrement dit, lorsque les algorithmes et les modèles prédictifs basés sur l'apprentissage automatique interviennent dans un large éventail de pratiques sociales, il est important de mettre en évidence les présupposés et les *a priori* à la base du développement et de l'usage de ces systèmes informatiques.

Cette problématique sera abordée dans cette recherche-création dans le contexte des médias interactifs, plus précisément dans le documentaire interactif. En effet, dans cette thèse il sera question de mettre en évidence les présupposés à la base de l'utilisation des modèles prédictifs dans les technologies numériques de communication, plus spécifiquement, dans l'utilisation des algorithmes d'apprentissage automatique comme outil d'interactivité dans un *chatbot* documentaire. Il s'agira d'explorer les manières dont les présupposés liés au hasard et à l'indétermination interviennent dans les pratiques de création et d'expérience du documentaire interactif, d'expérimenter avec des modes de création qui embrassent l'incertitude et l'imprévisible et aussi d'approcher le hasard comme thématique dans la création d'un documentaire interactif.

L'utilisation croissante de dispositifs numériques interactifs dans les récits documentaires audiovisuels soulève d'autres aspects relatifs à l'intersection entre le numérique et l'indétermination. Dans le cas des pratiques documentaires audiovisuelles, le rôle du hasard et de l'inopiné a déjà été démontré par plusieurs auteurs (Burch, 1986 [1967]; Lioult, 2008b; Philibert, 2009). Comme il a été montré, le hasard est partie inhérente du documentaire audiovisuel dans le sens où il émerge de pratiques de cadrage de l'inopiné. Néanmoins, dans le cas du documentaire interactif le rôle joué par le hasard et l'indétermination reste à être précisé, notamment dans l'aspect numérique de ce média. Certains auteurs (Gaudenzi, 2013b; Gifreu, 2013) remarquent que l'utilisation de l'interactivité dans le récit documentaire ouvre l'expérience à une indétermination supplémentaire dans le sens où les gestes posés par

l'interacteur rendent imprévisible (au moins partiellement) le format du contenu tel qu'il sera expérimenté par le public. La mobilité des documentaires géolocalisés contribue également à augmenter l'incertitude dans le documentaire interactif notamment dans les aléas liés au mouvement dans l'espace public (la météo, la circulation et toutes les rencontres inattendues sont des exemples des facteurs que la mobilité apporte comme incertitude). Mais est-ce que l'incertitude du documentaire interactif est aussi dans son aspect numérique, ou l'indétermination est uniquement dans la rencontre avec l'interacteur ? Est-ce que la préprogrammation des codes numériques exclut le hasard de cette dimension de l'expérience du documentaire interactif (la dimension numérique) ?

Ces problèmes sont à l'origine de cette thèse doctorale en recherche-crédation, composée du présent essai écrit et d'un travail de création médiatique, le documentaire interactif *On Chance*. *On Chance* est un documentaire interactif sur et avec le hasard. Les concepts de hasard et d'indétermination sont traités de différentes façons dans le documentaire. En premier lieu, dans la thématique abordée par le documentaire, car il s'agit d'un travail de vulgarisation scientifique, composée d'une variété de matériel d'archive concernant l'intersection entre les technologies numériques prédictives et la philosophie et les sciences du hasard et de l'indétermination. Dans une exploration des possibilités qu'offrent les médias interactifs au champ de la vulgarisation scientifique, ce documentaire utilise la notion de hasard pour aborder le tournant non déterministe des théories scientifiques du 20^e siècle. Par ailleurs, le hasard est intégré au processus de création du documentaire selon la méthodologie théorisée par Jean-Luc Lioult du « cadrage de l'inopiné » (Lioult, 2008a). Enfin, l'interactivité du documentaire est basée sur un dispositif qui fait appel à des algorithmes d'apprentissage automatique, un agent conversationnel (ou *chatbot*) et, en conséquence, utilise de modèles probabilistes de traitement d'information numérique. La deuxième partie de cet essai doctoral abordera différents aspects du processus de création de *On Chance*.

La première partie du présent essai apporte une perspective théorique sur les problématiques concernant le hasard et l'indétermination dans le champ du documentaire interactif, dans ses aspects à la fois analogiques et numériques. Il s'agira d'explorer les possibilités qu'offre la posture pragmatiste spéculative (Debaise & Stengers, 2016; James, 2016; Manning & Massumi, 2014) pour l'étude des technologies numériques, notamment dans l'intersection entre les concepts de hasard et d'indétermination et les algorithmes d'apprentissage automatique. Le pragmatisme spéculatif est à la base du programme de recherche proposé par Isabelle Stengers et Didier Debaise (Debaise & Stengers, 2016) en dialogue avec les travaux de nombreux philosophes, comme Alfred N. Whitehead, Charles S. Peirce, William James, Gilles Deleuze, Erin Manning et Brian Massumi. Cette posture se propose de défaire les dualismes fréquents dans les théories modernes et de considérer la multiplicité de l'expérience. Comme affirment Debaise et Stengers, le pragmatisme spéculatif consiste à :

(...) ne pas admettre de principe de jugement hors situation, qui domestiquerait cette multiplicité en termes de catégories ou d'exigences qui lui sont étrangères. Toute pensée est de ce point de vue absolument située, engagée dans la situation dont elle émerge et qui lui donne sens (Debaise & Stengers, 2016, p. 84).

Dans cette recherche-crédation, il s'agira notamment de s'attaquer au dualisme sujet/objet dans différents contextes : l'expérience des médias interactifs, les interprétations de la théorie des probabilités, les théories de l'information et du numérique. Comme on le verra, l'approche pragmatiste spéculative permet d'élaborer une conception relationnelle et processuelle du hasard, de l'indétermination et de l'information.

Dans la première étape de la recherche théorique seront abordés les aspects analogiques de l'expérience interactive et ses relations avec l'indétermination et le hasard (chapitre 2). Dans une première application des principes du pragmatisme spéculatif, il s'agira d'étudier l'expérience du documentaire interactif dans une perspective qui dépasse le

dualisme sujet/objet. En effet, il sera question de proposer une idée de l'expérience interactive qui ne se définit pas par un sujet (interacteur) séparé de l'objet (documentaire) qu'il perçoit, mais plutôt par les relations émergentes de l'événement qui constitue l'expérience (Manning & Massumi, 2014; Massumi, 2011). Par la suite (chapitres 3 et 4), l'intersection entre hasard, indétermination et documentaire interactif sera abordée dans son aspect numérique. Il s'agira à nouveau de partir des principes du pragmatisme spéculatif, mais cette fois-ci pour étudier les processus numériques qui intègrent l'expérience du documentaire interactif et ses relations avec les notions de hasard et d'indétermination. Comme on le verra plus en détail, le système numérique d'interactivité utilisé dans la création de *On Chance* est basé sur le *chatbot*, un dispositif d'interface conversationnelle qui utilise souvent des algorithmes d'apprentissage automatique. Or, comme il a déjà été mentionné, ces types d'algorithmes font ample utilisation d'analyses statistiques et des probabilités pour produire des modèles de prévision basés sur des données historiques. L'interprétation de la probabilité dans les modèles d'apprentissage automatique est souvent divisée selon le dualisme subjectif/objectif en deux types : la fréquentiste (objective) et l'interprétation bayésienne (subjective). Ainsi, dans un premier moment, il sera important d'étudier les différentes interprétations philosophiques du rôle du hasard dans la théorie des probabilités, pour ensuite proposer une approche pragmatiste spéculative de la question (Deleuze, 1993 [1968]; Gillies, 2000; Charles S. Peirce, 1957) et, encore une fois, défaire le dualisme sujet/objet.

Dans une perspective pratique du problème du hasard dans le documentaire interactif, la deuxième partie du présent essai doctoral aborde des aspects du processus de création du webdocumentaire *On Chance*. Dans un premier moment (chapitre 5) il s'agira de présenter le webdocumentaire et réviser la littérature et un corpus d'œuvres à l'intersection du récit numérique et des agents conversationnels (*chatbots*). Enfin, le dernier chapitre se consacre au récit de pratique de *On Chance*.

Comme prône la recommandation habituelle, cette introduction a été écrite à la fin du processus de rédaction du présent essai doctoral. Il importe de le mentionner, car dans une recherche-crédation doctorale d'une durée de presque dix ans, les inquiétudes, les motivations et les technologies à la base de ce travail ont évolué pendant toutes ces années d'étude et de création. Il est intéressant de remarquer que les technologies utilisées dans la partie de la création de cette thèse, notamment dans le *chatbot*, n'étaient pas envisageables, ou n'existaient pas tout simplement au début de mon parcours doctoral, en 2011. Néanmoins, il existe un fil d'Ariane qui traverse toutes ces années de recherche-crédation : la notion de hasard et ses relations avec les médias numériques. Le premier chapitre se consacre à introduire cette ligne conductrice, le problème central de cette recherche-crédation, le hasard et l'indétermination dans les médias interactifs, notamment le documentaire interactif.

1. CHAPITRE I

INTRODUCTION AU PROBLÈME DU HASARD ET DE L'INDÉTERMINATION DANS LE DOCUMENTAIRE INTERACTIF

Dans ce premier chapitre, il s'agira de formuler les questions que pose la présente thèse de recherche-crédation et d'établir le cadre conceptuel mobilisé, la posture adoptée pour approcher ces questions, ainsi que les méthodologies utilisées. La suite de la thèse se divise en deux parties. Dans la première, il sera question d'une approche théorique de la question du hasard et de l'indétermination dans l'expérience du documentaire interactif (dans ses aspects analogique et numérique). Dans la deuxième partie, ces questions sont abordées par la perspective de la discussion autour de la création d'un webdocumentaire, *On Chance*.

1.1. Questions de recherche-crédation

Dans les dernières années, l'interactivité s'est infiltrée dans les pratiques du documentaire audiovisuel et a fait éclore de nouvelles formes de narration sur une grande variété de supports, tels que le web, des applications pour *smartphone*, des installations ou encore les réalités augmentée et virtuelle. Le développement des technologies numériques a joué un rôle fondamental dans l'émergence des narrations interactives et des nouvelles écritures documentaires. À la suite de cet éclatement vers une plus grande diversité de formats et d'écritures, divers auteurs ont étudié les impacts de l'utilisation de l'interactivité dans l'expérience du documentaire (Aston, Gaudenzi, & Rose, 2017; Gaudenzi, 2013b; Gifreu, 2011; Levin, 2013; O'Flynn, 2012). Le champ du documentaire interactif se situe dans un champ plus large, celui des médias interactifs, avec lequel il partage de nombreuses questions.

Ces nouvelles écritures non linéaires, comme cela a été observé, transforment le statut du spectateur, qui devient un interacteur (Paquin, 2006, p. 6). Grégory Chatonsky a remarqué que les médias interactifs ont le potentiel de rajouter des éléments d'indétermination à l'expérience : « La faculté d'agir de l'interacteur n'est pas réductible à une volonté de contrôle ou de maîtrise, mais consiste en la possibilité d'entrouvrir l'indétermination, le possible devant alors être conçu comme une variabilité dont on ne peut anticiper le résultat. » (Chatonsky, 2004, p. 90). Chatonsky remarque que dans une expérience interactive, certains aspects de l'expérience émergent des actions posées par l'interacteur, ce qui la rend imprévisible sous certains égards pour le créateur. La notion d'indétermination est liée ainsi à celle d'un déplacement de la prise de décision sur certains aspects de l'expérience, de l'auteur vers l'interacteur. Ce passage est souvent considéré comme une perte de contrôle, même que partielle, de la part de l'auteur (Gifreu, 2013). À l'inverse, il y aurait un gain de contrôle de la part de l'interacteur, dans le sens où il décide certains aspects de son expérience dans les limites des possibilités prévues par l'auteur. Il s'agit ainsi d'un transfert de la puissance d'agir (*agency*), ou de l'agentivité, du créateur vers l'interacteur.

Dans le champ du documentaire interactif, il est aussi question du parallèle entre interactivité et contrôle (ou perte de) : « The biggest difference (and conflict) between the objectives of the traditional documentary and the goals of the interactive documentary are in the area of authorship and control over the narrative discourse. » (Gifreu, 2011, p. 14 , chap. 5) Du fait de ce passage partiel de la puissance d'agir du créateur vers l'interacteur, il a souvent été argumenté que le documentaire interactif aurait tendance à être plus imprévisible que le documentaire linéaire (Galloway, McAlpine, & Harris, 2007; Gaudenzi, 2013b; Gifreu, 2011), et ainsi plus indéterminé.

Comme Gaudenzi a constaté dans le cas des documentaires géolocalisés², l'interactivité a aussi le potentiel d'ouvrir le documentaire interactif aux rencontres imprévisibles avec l'espace :

If interactivity is going to happen in a physical and open environment, and maybe with other people, the interaction will depend on multiple variables (people's reactions, weather, slippery shoes, traffic etc...) most of which are not predictable. (Gaudenzi, 2013b, p. 62)

Dans le cas d'une expérience interactive géolocalisée, l'indétermination et l'ouverture viennent aussi de la possibilité d'inscrire cette expérience dans un espace physique particulier et d'y intégrer des aspects propres à cet espace.

Néanmoins, comme le démontre Nathaniel Stern (Stern, 2013), plusieurs auteurs remettent en question l'idée que l'interactivité implique nécessairement une plus grande indétermination. Il mentionne le travail de Brian Massumi : « There can be 'a kind of tyranny to interaction' (Massumi, 2011: 47), that forces particular and thus predictive movements, which then may as well be static because they are conceptually, and perhaps sensibly, pre-formed. » (Stern, 2013, p. 63) Or, si les documentaires interactifs sont des programmes numériques, des algorithmes écrits dans un code binaire pour effectuer des calculs selon des règles prédéterminées, quelles sont alors les conditions pour la création d'indétermination dans la dimension numérique du documentaire interactif ? Le travail de Chatonsky cité ci-dessus est traversé par une question similaire, il écrit : « comment les programmes peuvent-ils engendrer de l'indétermination, de l'improbable et de l'improgrammable ? » (Chatonsky, 2004, p. 80) Ainsi, il est légitime de poser la question de savoir si l'ouverture dont il est question dans l'expérience interactive, plus précisément l'expérience du documentaire

² Il est important de remarquer qu'il existe une multiplicité de modes d'interactivité qui n'a pas encore été traitée. Les *locative media*, ou médias géolocalisés, sont un de ces modes. Dans la prochaine section, les différents modes d'interactivité seront abordés.

interactif, est nécessairement prédéterminée, autrement dit, une indétermination préprogrammée, ou une pseudo-indétermination.

Il est important de noter que le dispositif d'interactivité choisi dans la création de *On Chance*, le *chatbot* (*agent conversationnel*, en français), permet à la fois d'ouvrir et de cadrer la discussion sur l'indétermination et le hasard dans le cadre du documentaire interactif. Comme nous verrons plus en détail, l'interactivité de *On Chance* est de mode conversationnel et basé sur un *chatbot*, ou un agent automatisé qui répond aux entrées de l'interacteur en simulant une conversation. Habituellement un *chatbot* utilise des techniques de Machine learning (ML, apprentissage automatique, en français) dans le traitement des entrées de l'interacteur. Or, un des principaux modes d'ouverture au hasard dans le champ du numérique est par la vaste utilisation de modèles statiques et probabilistes dans les algorithmes, notamment dans le traitement massif de données (Big Data) et dans les algorithmes de Machine Learning. Ces modèles numériques ravivent l'ambition prévisionniste du déterminisme sur la base d'analyses de grandes collections de données. Certains développeurs de ces modèles affirment être en mesure de réduire l'incertitude et d'augmenter le contrôle des interactions avec le monde :

New research suggests that by cross-referencing your geographical data with that of your friends, it's possible to predict your future whereabouts with a much higher degree of accuracy. This is what's known as predictive modeling, and it requires nothing more than your cellphone data. (Maass & Rajagopala, 2012, p. "s.p.").

Je défends que cette vaste utilisation des probabilités et de modèles prédictifs dans les technologies numériques de communication et d'information demande une réflexion sur les interprétations du rôle de la probabilité dans ces modèles. Autrement dit, j'estime qu'il est important de mettre en évidence quels sont les présupposés conceptuels et en termes de vision de monde (*worldview*) sur lesquels se bâtissent ces algorithmes probabilistes et modèles prédictifs. Quelle interprétation fait-on des

statistiques, des probabilités et de l'information dans le cadre des médias interactifs ? Comment comprenons-nous les notions de hasard, indétermination, incertitude, entropie, information et bruit dans le contexte des technologies numériques de communication utilisées dans les documentaires interactifs ? Ces questions seront posées dans le contexte des algorithmes utilisés dans le développement de *On Chance*.

Entre les nombreuses implications de l'intégration de l'interactivité, et donc du numérique, aux écritures documentaires audiovisuelles, je m'intéresse ainsi à la question du hasard et de l'indétermination, autrement dit : comment créer de l'indétermination dans le numérique, ou plus précisément dans l'expérience interactive ? Quelles sont les conséquences d'intégrer le hasard dans le documentaire interactif ? Quels sont les présupposés liés à l'utilisation des modèles probabilistes dans le numérique ? Il s'agira de comprendre (1) comment agit le hasard dans le processus de création et dans l'expérience d'un documentaire interactif et (2) comment le hasard et l'indétermination agissent dans la dimension numérique de l'interactivité. Il sera question d'étudier le rôle du hasard dans les méthodes de création et dans l'expérience des nouvelles écritures documentaires, et aussi d'explorer le concept du hasard dans sa dimension numérique. Ainsi, ce travail de recherche-crédation se propose d'étudier et d'expérimenter à plusieurs niveaux les relations entre le hasard, l'indétermination et le documentaire interactif.

Ces questions seront abordées en trois étapes : les deux premières étapes sont d'ordre plutôt théorique et la dernière propose une discussion autour de la création d'un documentaire interactif, *On Chance*. Les relations entre le hasard et le documentaire interactif seront étudiées théoriquement dans cette thèse sous deux aspects. Le premier, abordé dans le deuxième chapitre, concerne le rôle attribué au hasard et à l'indétermination dans l'expérience du documentaire interactif, ce que j'ai nommé les processus analogiques du documentaire interactif. Le deuxième aspect, à être traité dans les troisième et quatrième chapitres, se concentre sur la présence du hasard dans la dimension informationnelle du documentaire interactif, ce que j'ai nommé les

processus numériques du documentaire interactif. Je reprends la différence établie par Massumi entre les processus analogiques et numériques, pour organiser l'approche de la présente recherche. Il définit un processus analogique comme « the continuous transformation of an impulse from one qualitatively different medium into another. » (Massumi, 1998, p. 307), alors que le processus numérique concerne uniquement le passage par le code binaire. Il est important de remarquer que tout processus numérique est néanmoins entremêlé par des processus analogiques. Massumi donne l'exemple d'un texte numérisé : l'écriture du texte est un processus analogique (la transformation de la pensée en écriture). Si l'écriture se fait sur un ordinateur, il y a alors un processus de numérisation (la transformation en code binaire, par exemple l'ASCII³). Néanmoins, la lecture est un processus analogique, car il y a un nouveau passage entre des moyens qualitativement différents. Ainsi, les processus numériques dont il est question dans cette thèse sont souvent des étapes entre divers processus analogiques, et le numérique doit être pensé dans ce va-et-vient avec l'analogique.

Dans le chapitre sur les processus analogiques du documentaire interactif, il sera question du rôle que peuvent jouer le hasard et l'indétermination dans l'expérience d'un documentaire interactif. Les questions auxquelles il s'agira de répondre sont alors : comment penser l'expérience interactive dans une perspective non déterministe ? Qu'est qu'une image dans le contexte du documentaire interactif et que se passe-t-il quand cette image est rapportée à un centre d'indétermination, l'interacteur ? Cette démarche s'inspire largement du travail de Gilles Deleuze dans les deux volumes de *Cinéma* (Deleuze, 1983, 1985). Ces questions seront abordées dans le contexte d'un corpus divers de documentaires interactifs. Par la suite, il s'agira d'explorer la dimension numérique des médias interactifs.

³ *American Standard Code for Information Interchange* (Code américain normalisé pour l'échange d'information, en français)

Dans les chapitres 3 et 4, l'aspect numérique du documentaire interactif sera également mis en relation avec les notions de hasard et d'indétermination. Or, à première vue cela peut paraître contradictoire de penser à la présence du hasard dans un système d'algorithmes souvent considéré comme déterministe. Néanmoins, comme l'exprime Henri Atlan :

[Ç]a ne choque plus personne de dire qu'il n'est pas nocif d'ajouter un peu de hasard aux programmes d'ordinateur, parce que cela les rendra peut-être capables de faire des choses que, sans cela, ils ne pourraient pas faire. C'est devenu presque une banalité, alors qu'il y a vingt ans, c'était considéré comme une folie. (Atlan, 2006 [1972], p. xi)

Dans un premier moment, il s'agira alors d'aborder les différentes interprétations de la théorie des probabilités, notamment dans la perspective des théories de l'information (troisième chapitre) pour ensuite les aborder dans le contexte des algorithmes d'apprentissage automatique utilisés dans la présente recherche-crédation (quatrième chapitre). Cette étape de la discussion théorique sera cadrée par les outils numériques utilisés dans création proposée par la présente thèse, *On Chance*.

Enfin, dans la partie consacrée à la création de *On Chance*, il s'agira de discuter l'utilisation du hasard et de l'indétermination dans le processus créatif de ce documentaire interactif et de présenter la démarche et l'interface mises en place. Comme on le verra, dans la création je cherche à activer les dimensions affectives autour des notions de hasard et d'indétermination et de les aborder dans leur multiplicité, de revenir à la fois à l'usage courant et à l'usage scientifique de ces termes pour les articuler dans une interface conversationnelle.

Si le rôle du hasard dans le processus créatif du documentaire linéaire et des pratiques artistiques a déjà été démontré ((BPI), 2009; Burch, 1986 [1967]; Philibert, 2009; Pollock, 2012; Saurisse, 2007; Vogel, 2005 [1974]), dans le domaine du documentaire interactif cela reste encore à être étudié plus en profondeur. Dans l'introduction du livre

Pratiques du hasard (Pollock, 2012), Jonathan Pollock fait un appel à ce que les éléments indéterminés des pratiques de création artistique soient davantage étudiés :

Il y a bien longtemps que les sciences de la nature — sous les coups de boutoir de l'évolution des espèces, la turbulence des fluides ou la physique quantique — ont abandonné une conception de la causalité oscillant entre la volonté rationnelle (ou inconsciente) et un strict déterminisme mécanique et légal. (...) La volonté et la nécessité ne recouvrent pas l'intégralité du champ de la création : la part laissée dans l'ombre et que l'on désigne au moyen du vocable *hasard* en appelle à de plus amples investigations. (Pollock, 2012, p. 10, italique original)

C'est dans la lignée de cet appel que cette recherche-crédation se situe.

1.2. Cadre conceptuel

1.2.1. Hasard et indétermination

Cette recherche-crédation se situe dans l'intersection entre le documentaire interactif et les concepts de hasard et d'indétermination. Le hasard et l'indétermination sont des concepts polysémiques qui se déploient autant dans le langage courant que dans différents champs de la recherche académique et qui ont beaucoup changé au long de l'histoire (Hacking, 1990; Lestienne, 1993; Mlodinow, 2009; Noël, 1991; Prigogine, 2001). Je m'intéresse notamment à l'évolution de ces concepts depuis la fin du 19^e siècle en philosophie et philosophie des sciences, pour ensuite les articuler au domaine du documentaire interactif. Le choix de cette période précise, la fin du 19^e siècle, se justifie, car il s'agit d'un moment de transition important dans la philosophie des sciences, le moment où des ruptures avec le déterminisme scientifique commencent à faire surface. Mais avant d'examiner plus en détail ce tournant dans les concepts de hasard et d'indétermination, il est important de revenir brièvement sur la différence entre ces deux notions. Au cours de cette thèse, plus de précisions seront apportées aux

relations et distinctions entre des termes proches de ceux de hasard et d'indétermination, telles que l'incertitude, la probabilité, la complexité ou l'entropie.

Abordons pour l'instant de manière introductive la relation entre le hasard et l'indétermination. Comme on le verra plus en détail, ces deux termes seront convoqués dans cette thèse pour désigner les variations d'un événement qui ne sont pas prédéterminées à aucun égard. Le hasard se rapporte aux causes de l'événement, ou à une irréductibilité de ses variations à un ordre qui le précède. Si un événement a deux temps, tel que le décrit Deleuze, le hasard est dans le premier temps : « Les dés qu'on lance une fois sont l'affirmation du hasard, la combinaison qu'ils forment en tombant est l'affirmation de la nécessité. La nécessité s'affirme du hasard au sens exact où l'être s'affirme du devenir et l'un du multiple. » (Deleuze, 1983 [1962], p. 29) Deleuze utilise l'image d'un jeu de dés pour évoquer la relation entre le hasard et la nécessité et attribue deux temps au jeu. Dans le premier temps du coup de dés, il est question du hasard, du multiple et du devenir. Après l'arrêt des dés, dans le deuxième temps, ce qui advient c'est la nécessité, l'un, l'être. Ainsi, il fait du hasard une affirmation qui n'exclut point la nécessité. Néanmoins, ces deux temps doivent être considérés dans un ensemble : « Mais ces deux tables ne sont pas deux mondes. Ce sont les deux heures d'un même monde, les deux moments du même monde, minuit et midi, l'heure où l'on jette les dés, l'heure où retombent les dés. » (Deleuze, 1983 [1962], p. 29). Or, l'indétermination, telle que je la comprends, correspond aux résultats imprévisibles de l'événement avant que le hasard devienne une nécessité. Dans l'exemple du jeu de dés, l'indétermination est l'impossibilité de préciser le résultat (le deuxième temps) avant que l'événement s'actualise, autrement dit, l'impossibilité de savoir si j'aurai 1 ou 6 dans mon coup de dé. L'indétermination est ainsi dans le regard que l'on porte sur le deuxième temps lorsqu'on se trouve dans le premier, de l'attente que l'on puisse avoir que les choses se passent de telle ou telle autre façon.

Maintenant, revenons plus en détail sur la posture déterministe, plus spécifiquement le déterminisme laplacien pour situer les remises en question du tournant du 19^e siècle à

la lumière des notions de hasard et indétermination. La conception de Pierre-Simon Laplace du déterminisme a été mise en évidence dans son œuvre *Essai philosophique sur les probabilités* (Laplace, 1921 [1840]) dans cette formulation aujourd'hui célèbre :

Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre. Une intelligence qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir comme le passé serait présent à ses yeux. (Laplace, 1921 [1840], p. 3).

Cette intelligence à laquelle Laplace fait référence est connue par l'expression du *démon de Laplace*, une entité fictionnelle qui exprime l'ambition ultime du déterminisme, celle d'être en mesure de tout prévoir et de tout expliquer. Le déterminisme ainsi conçu suppose que l'univers, à toutes les échelles, se constitue de corps soumis aux mêmes lois, de la particule aux corps célestes. Ainsi, il suffirait de découvrir les formules qui décrivent ces lois pour pouvoir être en mesure de prévoir tous les événements futurs et expliquer tout le passé. Il est important de noter que Laplace affirme cela dans un ouvrage consacré à la *théorie des hasards* (un des noms attribués à l'époque aux études de probabilités⁴). Mais cette apparente contradiction entre une théorie où tout est prédéterminé par des lois et une théorie de l'aléatoire s'explique par l'interprétation subjective du statut du hasard généralement défendue par Laplace :

Tous les événements, ceux mêmes qui par leur petitesse semblent ne pas tenir aux grandes lois de la nature, en sont une suite aussi nécessaire que

⁴ « La théorie des hasards consiste à réduire tous les événements du même genre à un certain nombre de cas également possibles, c'est-à-dire tels que nous soyons également indécis sur leur existence, et à déterminer le nombre de cas favorables à l'événement dont on cherche la probabilité. » (Laplace, 1921 [1840], p. 6)

les révolutions du Soleil. Dans l'ignorance des liens qui les unissent au système entier de l'univers, on les a fait dépendre des causes finales, ou du hasard, suivant qu'ils arrivaient et se succédaient avec régularité, ou sans ordre apparent. (Laplace, 1921 [1840], p. 2).

Ainsi, dans ce système déterministe, le hasard n'existe qu'en tant qu'ignorance des causes qui agissent entre les corps, autrement dit comme un manque de connaissance.

Le déterminisme laplacien était la position philosophique prédominante en sciences au 19^e siècle. Néanmoins, vers la fin de ce siècle, des prises de position en faveur du non-déterminisme et d'autres conceptions du hasard commencent à prendre forme, à la fois dans les sciences et dans la philosophie. Ian Hacking fait un travail remarquable sur le contexte historique et philosophique de cette rupture avec le déterminisme, et sur le rôle du concept de hasard dans ce processus. Il commence son livre *The Taming of Chance* comme se suit :

The most decisive conceptual event of the twentieth century physics has been the discovery that the world is not deterministic. (...) This event was preceded by a more gradual transformation. During the nineteenth century, it became possible to see that the world might be regular and yet not subject to universal laws of nature. A space was cleared for chance. (Hacking, 1990, p. 1)

Ensuite, Hacking propose un parallèle entre ces transformations dans la conception du hasard et du déterminisme et un phénomène qui prend de l'ampleur à la même époque (19^e siècle), soit le recensement croissant de la population, ce qui a donné naissance à la biométrie et d'autres expériences de mesures quantitatives et d'analyses de données d'une ou de plusieurs populations. Comme nous le verrons, dans les questions plus contemporaines liées au traitement massif de données (Big Data) et au développement de l'apprentissage automatique, nous faisons face à certains défis de même ordre, soit l'imposition d'un système normatif basé sur des analyses statistiques d'une population. En 1990, Hacking alertait sur les dangers liés à cela :

It is now common to speak of information and control as a neutral term embracing decision theory, operations research, risk analysis and the broader but less well-specified domains of statistical inference. We shall find that the roots of the idea lie in the notion that one can improve—control—a deviant subpopulation by enumeration and classification. (Hacking, 1990, p. 3)

Ces questions d'ordre contemporaines seront abordées plus loin dans le texte, pour l'instant je reviendrai sur les débats en philosophie des sciences autour des théories non déterministes du 19^e siècle.

La thermodynamique est une des premières théories scientifiques à échapper au déterminisme laplacien, comme le démontrent Ilya Prigogine et Isabelle Stengers, dans l'ouvrage *La Nouvelle Alliance* (Prigogine & Stengers, 1979). Les auteurs font la distinction entre trois domaines de la thermodynamique qui constituent « trois étapes successives de son développement » (Prigogine & Stengers, 1979, p. 151) : la thermodynamique d'équilibre (ou linéaire), développée au 19^e siècle ; la thermodynamique proche de l'équilibre ; et la thermodynamique hors équilibre (ou non linéaire), développée au 20^e siècle. Comme le démontrent Prigogine et Stengers, la thermodynamique linéaire est déjà non déterministe dans son utilisation de modèles statistiques pour la description de systèmes physiques. Ils reviennent sur l'origine de la thermodynamique, en 1811, en l'opposant au déterminisme laplacien : « Le rêve laplacien, à l'heure de sa plus grande gloire, a subi un premier échec : une théorie physique existe désormais, mathématiquement aussi rigoureuse que les lois mécaniques du mouvement, et absolument étrangère au monde newtonien. » (Prigogine & Stengers, 1979, p. 118).

L'entropie, un des concepts fondamentaux de la thermodynamique, est souvent définie comme une mesure du degré d'indétermination (ou du désordre) d'un système (Atlan, 2006 [1972], p. 5). Cette même idée, comme on le verra plus tard, est au cœur de la théorie de l'information développée par Claude Shannon en 1948. Le concept

d'entropie a été développé par Boltzmann d'après les travaux de Rudolf Clausius. Clausius, en 1865, propose le concept d'entropie dans le contexte du deuxième principe de la thermodynamique, qui dans sa formulation courte affirme que l'entropie de l'univers tend au maximum (Goldstein & Lebowitz, 2004). Prigogine et Stengers commentent les origines de ce concept : « Il semble, écrit Planck, que la nature "privilégie" certains états ; la croissance irréversible de l'entropie (...) décrit l'approche du système vers un état qui l'attire, qu'elle préfère et dont elle ne s'éloignera pas spontanément : approche irréversible. » (Prigogine & Stengers, 1979, p. 137)

Boltzmann apporte ensuite une contribution décisive au concept d'entropie lorsqu'il y introduit les probabilités dans les passages entre les micro-états (configuration des particules au niveau microscopique) et les macro-états (configuration du niveau macroscopique). Comme le démontrent Prigogine et Stengers, « C'est Boltzmann qui le premier fit remarquer que l'on pouvait interpréter la croissance irréversible de l'entropie comme l'expression de la croissance du désordre moléculaire (...). ». (Prigogine & Stengers, 1979, p. 140) Comme l'expliquent Prigogine et Stengers, l'introduction de la probabilité et des statistiques dans l'étude de systèmes physiques est ce qui permet le passage d'une théorie de la trajectoire individuelle à une théorie de la dynamique des ensembles :

L'importante formule de Boltzmann fait de l'évolution thermodynamique irréversible une évolution vers des états de probabilité croissante, et de l'état attracteur, l'état macroscopique réalisé par la presque totalité des états microscopiques dans lesquels peut se trouver le système. (...) Pour la première fois, un concept physique a été expliqué en termes de probabilité. (Prigogine & Stengers, 1979, p. 140)

Boltzmann arrive alors au concept d'entropie statistique, et introduit la physique statistique d'équilibre.

Prigogine et Stengers remarquent néanmoins des insuffisances dans la théorie de Boltzmann, car il s'agit d'une théorie à propos de systèmes sans échange avec le milieu

(fermés) et qui cessent d'évoluer une fois qu'ils ont atteint une situation stable (d'équilibre). Or, Prigogine et Stengers observent l'omniprésence de systèmes ouverts dans un flux continu entre équilibre et déséquilibre (les cellules vivantes, la météorologie, l'hydrodynamique, les villes, pour en citer quelques exemples), qui dissipent de l'énergie et qui ne peuvent donc être décrits par la thermodynamique d'équilibre : « Or, que nous examinions une cellule ou une ville, la même constatation s'impose : non seulement ces systèmes sont ouverts, mais ils vivent de leur ouverture, ils se nourrissent du flux de matière et d'énergie qui leur vient du monde extérieur. » (Prigogine & Stengers, 1979, p. 142) Ils posent ensuite les questions suivantes :

Qu'est-ce que l'évolution des vivants, de leurs sociétés, de leurs espèces, dans le monde du désordre croissant de la thermodynamique ? Quel rapport entre le temps thermodynamique de l'approche vers l'équilibre et le temps du devenir complexe, ce temps dont Bergson disait qu'il est invention, ou rien du tout ? (Prigogine & Stengers, 1979, p. 144)

Prigogine et Stengers remarquent que, au-delà d'un seuil de stabilité, les systèmes thermodynamiques ont des « comportements qualitativement divers », ce sont les *structures dissipatives*. Cela arrive lors que les fluctuations des conditions d'un système ne régressent pas vers l'état d'équilibre, mais, au contraire, s'accroissent et poussent le système vers « un nouveau régime de fonctionnement qualitativement différent des états stationnaires définis par le minimum de production d'entropie. » Or, les structures dissipatives présentent des propriétés d'auto-organisation dans ce processus d'ordre par fluctuation, et échappent ainsi au deuxième principe de la thermodynamique qui prévoit une augmentation constante de l'entropie jusqu'à la situation d'équilibre où l'entropie sera maximale (et sa production minimale). Au-delà des systèmes thermodynamiques, la biologie fournit de nombreux exemples de structures dissipatives :

Le vivant fonctionne loin de l'équilibre, dans un domaine où les conséquences de la croissance de l'entropie ne peuvent plus être interprétées selon le principe d'ordre de Boltzmann, il fonctionne dans un domaine où les processus producteurs d'entropie, les processus qui

dissipent l'énergie, jouent un rôle constructif, sont source d'ordre.
(Prigogine & Stengers, 1979, p. 193).

D'autres théories scientifiques non déterministes ont été développées vers la fin du 19e siècle, souvent par le biais de l'introduction des probabilités (Lestienne, 1993).

Charles S. Peirce, William James et Henri Bergson sont parmi les premiers philosophes à remettre en question le déterminisme laplacien à la lumière du concept de hasard. Dans une conférence intitulée « The Dilemma of Determinism », James a associé les concepts de hasard et de déterminisme dans ces termes :

The stronghold of the deterministic sentiment is the antipathy to the idea of chance. As soon as we begin to talk indeterminism to our friends, we find a number of them shaking their heads. This notion of alternative possibilities, they say, this admission that any one of several things may come to pass, is, after all, only a roundabout name for chance; and chance is something the notion of which no sane mind can for an instant tolerate in the world. (James, 1884, "s. p.")

James met en évidence la résistance de la communauté scientifique à l'idée du hasard et l'associe à la remise en question du déterminisme. Cette même résistance a animé de nombreux débats en philosophie des sciences au cours du 20e siècle (voir note de bas de page 5).

Charles Sander Peirce a également formulé une position philosophique radicalement différente du déterminisme laplacien, ce qui l'a poussé à développer une interprétation non-subjective des probabilités et à mettre le hasard au sein de sa théorie. Comme l'observe Hacking : « His [Peirce's] denial of the doctrine of necessity was incidental to a life permeated by statistics and probabilities. Somebody had to make a first leap to indeterminism. Maybe it was Peirce, perhaps a predecessor. » (Hacking, 1990, p. 201). Peirce établit un système de triade, dont le hasard est la première catégorie (*Firstness*), la détermination est la deuxième (*Secondness*), et il revient à l'habitude (*Thirdness*) de mettre en relation les deux catégories précédentes : « According to this, three elements are active in the world, first, chance; second, law; and third, habit-taking. » (Peirce,

208) Peirce indique ainsi un système qui relie hasard et détermination, et indique une voie qui échappe au dualisme dominant dans les théories scientifiques et philosophiques. Cette relation non dualiste proposée par Peirce est reprise dans son article « The Doctrine of Necessity Examined », où Peirce affirme que la spontanéité créatrice a son origine dans l'indétermination, et limitée par les, est génératrice du nouveau :

By thus admitting pure spontaneity or life as a character of the universe, acting always and everywhere though restrained within narrow bounds by law, producing infinitesimal departures from law continually, and great ones with infinite infrequency, I account for all the variety and diversity of the universe, in the only sense in which the really sui generis and new can be said to be accounted for. (Charles S. Peirce, 1957, p. 184)

Peirce voit dans son hypothèse de chance-spontaneity l'indication d'une voie prometteuse au développement scientifique, autant dans les théories que dans les méthodes. Il affirme : « If my argument remains unrefuted, it will be time, I think, to doubt the absolute truth of the principle of universal law; and when once such a doubt has obtained a living root in any man's mind, my cause with him, I am persuaded, is gained. » (Charles S. Peirce, 1957, p. 188). Le travail de Peirce à propos du hasard sera abordé plus longuement au troisième chapitre de cette thèse.

Comme Peirce et James, Bergson a fait de nombreuses critiques aux principes du déterminisme. La formulation générale du déterminisme proposée par Bergson dans *l'Essai sur les données immédiates de la conscience* est très proche de celle de Laplace : « La question est de savoir si, connaissant dès aujourd'hui tous les antécédents futurs, quelque intelligence pourrait prédire avec une absolue certitude la décision qui en sortira. » (Bergson, 1958 [1888], p. 138). Dans cet ouvrage, Bergson aborde le problème du déterminisme en introduisant une distinction entre le déterminisme physique (de la matière), et le déterminisme psychologique (de la conscience), pour investiguer ensuite si l'un se ramène à l'autre. Ainsi, il pose une différence entre le déterminisme de la matière et le déterminisme de la conscience, pour

questionner ensuite la réductibilité du psychologique au physique. Or, pour échapper au dualisme déterminisme/indéterminisme, Bergson propose une solution qui est elle-même dualiste, car il associe l'un avec la matière et l'autre avec l'activité consciente. Mais déjà dans *Données immédiates*, certains passages⁵ sèment le doute sur une position clairement dualiste, et cette position s'atténue graduellement dans la suite de sa pensée. Dans une lettre de 1903, Bergson dira : « Les corps vivants, du plus simple au plus compliqué, sont autant des mécanismes de plus en plus subtils destinés à tourner de mieux en mieux le déterminisme de la matière brute, et à insérer une somme croissante de liberté dans le monde. » (Bergson, 2011, p. 280). Cet extrait indique que Bergson voit un passage graduel entre le déterminisme et la liberté (qu'il associe aux événements indéterminés) en fonction de la complexité du vivant. Or, pour Bergson, un corps vivant c'est un réservoir d'indétermination, car il élargit de manière imprévisible le champ des réactions possibles. L'indétermination est ainsi à l'origine de la création, de l'émergence du nouveau, de « la création continue d'imprévisible nouveauté qui semble se poursuivre dans l'univers. » (Bergson, 2011 [1911], p. 1). Bergson lie le déterminisme à la matière inerte et l'indétermination à la vie, et affirme que la création constitue l'activité principale de l'être vivant, et ce qui le distingue de la matière inerte : « (...) la matière est inertie, géométrie, nécessité. Mais avec la vie apparaît le mouvement imprévisible et libre. L'être vivant choisit ou tend à choisir. Son rôle est de créer. Dans un monde où tout le reste est déterminé, une zone d'indétermination l'environne. » (Bergson, 2011 [1920], p. 12). Dans cette perspective, l'indétermination est inhérente à la création et à la vie. Cette association entre l'indétermination et l'organisme vivant a mené Bergson à développer une philosophie

⁵ Par exemple : « Tantôt on pense surtout à la succession régulière des phénomènes physique et à cette espèce d'effort interne par lequel l'un devient l'autre ; tantôt on fixe son esprit sur la régularité absolue de ces phénomènes, et de l'idée de régularité on passe par degrés insensibles à celle de nécessité mathématique, qui exclut la durée entendue de la première manière. *Et l'on ne voit pas d'inconvénient à tempérer ces deux images l'une par l'autre, et à faire prédominer l'une ou l'autre, selon qu'on se préoccupe plus ou moins des intérêts de la science.* » (Bergson, 1958 [1888], pp. 161-162; italique ajouté)

très liée aux sciences de la vie et à bâtir alors une vision du monde où déterminisme et indéterminisme coexistent : « On n’y est nullement obligé de croire que, si le déterminisme radical n’est pas partout, il n’est nulle part, et qu’il faille opter entre un déterminisme et un indéterminisme universels. » (Bergson, 2011, p. 280). Le concept de centre d’indétermination, développé par Bergson et repris par Deleuze sera un élément important dans la compréhension de l’expérience interactive.

Au 20e siècle, malgré de nombreux débats⁶, la situation du déterminisme a évolué. Cet extrait d’un article de 1936 d’Erwin Schrödinger met en évidence le changement de statut du déterminisme dans les sciences, plus précisément dans la physique, au 20e siècle :

It has become the orthodox view of physicists today, that the momentary state of a physical system does not determine its movement or development or behaviour, to follow : Nature is supposed to be such that a knowledge of state, sufficiently accurate for sharp prediction of the future is not only unobtainable but also unthinkable. All that can be predicted refers to a large number of identical experiments, and consists in a definite statistic among all the possible developments to follow. The relative margin of indeterminacy (the ‘spread’ of the statistics) is large for a small system, for example, an atom ; but for large systems the margin is usually,

⁶ À titre d’exemple, je mentionne les échanges entre Albert Einstein et Niels Bohr (connus comme le paradoxe EPR) autour de l’interprétation de la théorie quantique. Lévy-Leblond (1991) décrit les deux attitudes face au statut du hasard dans ce paradoxe : la première, celle de Bohr, Heisenberg et Pauli, consiste à accepter que le hasard fasse partie de la nature et que la théorie quantique soit complète, « ainsi, la théorie quantique serait probabiliste, car le monde à l’échelle microscopique serait aléatoire par nature. » (Lévy-Leblond, 1991, p. 190) La deuxième est celle résumée dans l’article (Einstein, Podolsky, & Rosen, 1935) connu par la sigle EPR (d’après les premières lettres des noms des auteurs) selon laquelle la physique quantique est incomplète : « A sufficient condition for the reality of a physical quantity is the possibility of predicting it with certainty, without disturbing the system. » (Einstein et al., 1935, p. 777) Cette condition posée par les auteurs exclue toute possibilité de hasard ou d’indétermination. Comme la première interprétation échoue dans cette condition, les auteurs concluent que la théorie est incomplète. Les auteurs ne questionnent point l’efficacité de la théorie quantique, mais ils n’ont pas accepté son interprétation probabiliste. Un autre exemple est la publication *La querelle du déterminisme* (Pomian, 1990), un échange d’articles entre René Thom, Ilya Prigogine, Isabelle Stengers, Henri Atlan, et autres auteurs, autour de la question du déterminisme, où Thom affirme : « Je voudrais dire d’emblée que cette fascination de l’aléatoire témoigne d’une attitude antiscientifique par excellence. » (Thom, 1990, p. 62)

though not necessarily, small, which makes it possible to account for the apparent determinacy of inanimate Nature. (Schrödinger, 1936, p. 13)

Comme on le remarque, le déterminisme a laissé place à la probabilité et aux statistiques. Les liens entre la remise en question du déterminisme et les différentes interprétations philosophiques de la théorie de la probabilité seront plus largement discutés dans le troisième chapitre de cette thèse, mais revenons un instant sur cette notion de *zone d'indétermination (margin of indeterminacy)* dont parle Schrödinger. Il s'agit d'une notion centrale du *principe d'indétermination* (ou d'incertitude) de Heisenberg, formulé en 1927, qui marque une des plus importantes révisions du déterminisme dans la physique. Lévy-Leblond et Balibar décrivent ainsi ce principe : « According to our present understanding, in a given state of a quantum system, a physical magnitude is no longer to be associated with a single value (a number), but with a numerical spectrum having a nonzero width (...). » (Lévy-Leblond & Balibar, 1998, pp. 279-280) L'indétermination dans ce contexte doit être comprise comme un brouillage qui rend l'univers imprécis, flou, « the numerical 'fuziness' » (Lévy-Leblond & Balibar, 1998, p. 280), ou dans les termes de Schrödinger, « the 'spread' of statistics ». L'article de Lévy-Leblond et Balibar discute les termes *indétermination* et *incertitude* en vue du principe de Heisenberg. Les auteurs reviennent sur l'historique des termes allemands employés par Heisenberg et certaines de ses traductions et démontrent que le terme « incertitude » est, dans ce cas, moins précis, car il sous-entend qu'il s'agisse d'un manque de connaissance par rapport aux conditions d'un système, et non d'un aspect inhérent à la physique. Ce partage entre les interprétations dites « subjectives » (manque de connaissance) et « objectives » (en rapport à l'observable) de l'indétermination se reproduit entre les différentes interprétations de la théorie de la probabilité, comme on le verra plus loin.

Comme on le verra plus en détail plus loin dans cet essai doctoral, une relation non exclusive entre hasard et nécessité se développe dans divers courants de la philosophie des sciences au cours du 20^e siècle, notamment dans les courants autour de l'auto-

organisation (Atlan, 1972; Morin, 1972; Pomian, 1990; Prigogine, 2001; Prigogine & Stengers, 1997), de la théorie de l'information (Shannon, 1948; Terranova, 2004a) et de la théorie de l'information algorithmique (Chaitin, 2005; Parisi, 2013). Ces auteurs travaillent sur les principaux questionnements soulevés par les développements des théories scientifiques autour de notions comme l'indétermination, l'incertitude, la complexité, l'entropie et le hasard, dans une remise en question du déterminisme classique. Edgar Morin affirme :

Épistémologiquement et ontologiquement, notre conception se refuse à poser l'alternative de l'ordre et du désordre, du hasard et de la nécessité, du chaos et du cosmos, du système et de l'événement. C'est dans leur unité (contradictoire) que nous pouvons situer : l'organisation, la transformation. (Morin, 1972, p. 192).

Ainsi, le hasard et l'indétermination synthétisent le questionnement du déterminisme qui a marqué la fin du 19e et l'intégralité du 20e siècle. Ces discussions philosophiques autour du déterminisme, du hasard et de l'indétermination sont au cœur des différentes théories mobilisées dans cette thèse pour aborder les aspects analogiques et numériques du documentaire interactif. Dans les chapitres suivants, des approches et théories non déterministes seront développées, notamment par l'étude des différentes interprétations philosophiques de la probabilité et du rôle joué par le hasard dans la théorie de l'information et de l'information algorithmique.

1.2.2. Documentaire interactif

Le champ du documentaire interactif a émergé dans les années 1970, lorsque des dispositifs informatiques ont été intégrés à la diffusion du documentaire audiovisuel (Almeida & Alvelos, 2010). L'œuvre emblématique qui marque le début du récit interactif et non fictionnel est *Aspen Movie Map* (fig. 1), une œuvre de 1978, réalisée par une équipe de chercheurs du MIT (*Massachusetts Institute of Technology*). Il s'agit d'un système d'hyperlien qui permettait de réaliser un tour virtuel dans la ville d'Aspen (Colorado, États-Unis). Le terme « documentaire interactif » n'était pas encore utilisé

pour décrire ce genre de dispositif, cela n'arrive qu'une décennie plus tard avec le projet *Moss Landing*, développé par Apple Multimedia Lab (Gaudenzi, 2013b).



Figure 2: *Aspen Movie Map*. Source : *Inventing Interactive*⁷

Avec le déploiement massif des micro-ordinateurs et l'avènement d'internet, le croisement entre des dispositifs numériques et le documentaire audiovisuel a pris une importante envergure, notamment depuis les années 2000. Aujourd'hui, la scène du documentaire interactif est en plein essor. Cela se remarque par exemple par le fait que divers festivals de film, auparavant dédiés exclusivement au format cinématographique linéaire, ont créé des sections consacrées aux formats interactifs. C'est le cas, par exemple, de l'International Documentary Film Festival Amsterdam (IDFA), un des plus grands festivals de documentaire au monde, qui a créé en 2007 le IDFA DocLab dont la devise est : « Exploring documentary storytelling in the age of the interface »⁸. La mission autoproclamée de IDFA DocLab met en évidence la consolidation du documentaire interactif en tant que format audiovisuel et sa place établit au sein du documentaire : « The mission of IDFA DocLab is to showcase interactive documentaries and other new digital artforms that successfully push the boundaries of documentary storytelling in the age of the interface. »⁹

⁷ <http://www.inventinginteractive.com/2010/03/18/aspen-movie-map/>, consulté le 17 octobre 2017.

⁸ IDFA DocLab website, <https://www.doclab.org/>, consulté le 9 octobre 2017.

⁹ Idem

Mais de quoi parle-t-on exactement lorsqu'on utilise le terme « documentaire interactif » ? Plusieurs définitions ont été proposées (Aston et al., 2017; Galloway et al., 2007; Gaudenzi, 2013b; Gifreu, 2013; O'Flynn, 2012). À la suite de l'article Galloway et al. (2007), le documentaire interactif sera compris dans la présente thèse comme un documentaire audiovisuel dans lequel l'interactivité est centrale dans ses formes de diffusion. Cette définition du documentaire interactif, presque tautologique, est volontairement très étendue pour ne pas restreindre le documentaire interactif à des supports déterminés (tel que le web) ou à des modes spécifiques d'interactivité. Ainsi, seulement ces deux éléments essentiels : l'intention de proposer une expérience du réel et l'utilisation de l'interactivité dans la diffusion du projet auprès du public.

Sandra Gaudenzi, une des premières chercheuses à se consacrer au champ du documentaire interactif, en a également proposé une définition ouverte : « (...) any project that starts with an intention to document the “real”, and that does so by using digital interactive technology, will be considered an interactive documentary. » (Gaudenzi, 2013b, pp. 31-32). Dans cette définition, l'interactivité est ouverte aussi aux étapes du développement et de la production du documentaire, et ne se restreint pas au contact avec le public. Il faut noter néanmoins que l'interactivité exclusive au moment de la réalisation peut produire des films linéaires dans leur mode de diffusion, et pour cette raison, je fais le choix de restreindre l'interactivité à l'expérience de l'œuvre, et de ne pas considérer une œuvre participative comme nécessairement interactive. Par ailleurs, il semble important de signaler également que la numérisation du support ne garantit pas son interactivité : un documentaire linéaire diffusé par internet est sans doute numérique, mais il n'est pas pour autant interactif. En revanche, tous les modes d'interactivité du documentaire interactif sont rendus possibles par des dispositifs numériques. La numérisation du documentaire est donc une condition nécessaire, mais pas suffisante pour le caractériser comme interactif. Il est important de souligner également qu'il n'existe pas de consensus sur le terme « documentaire interactif » et plusieurs auteurs et réalisateurs désignent différemment ce nouveau

format (cyberdocumentaire, i-doc, documentaire transmédia, entre autres termes). Revenons maintenant sur chacun des éléments du terme « documentaire interactif ».

L'utilisation du terme « documentaire » est remise en question par certains auteurs (Almeida & Alvelos, 2010; Gaudenzi, 2013a), car il désigne le plus souvent le genre cinématographique documentaire, alors que les documentaires interactifs empruntent beaucoup d'éléments narratifs aux jeux vidéo, au reportage, au design, entre autres langages. Comme le rappellent Alvelos et Almeida,

(...) it appears that nowadays everyone is using the word 'documentary' to describe every single multimedia piece that incorporates video no matter its nature, technique, language or scope, taking advantage of the fuzzy and fragile boundaries of the documentary definition. (...) We believe this matter can't be seen as a minor question of semantics since there's a secular cinema tradition inherited whenever the word 'documentary' is evoked. (Almeida & Alvelos, 2010, p. 124)

En effet, il serait peut-être préférable d'utiliser un terme qui indique le caractère non fictionnel de la narration sans l'inscrire dans la tradition du cinéma documentaire. Dans la présente thèse, je me conforme néanmoins à la terminologie déjà adoptée dans le milieu avec une réserve en relation au terme « documentaire » : le documentaire interactif élargit la notion de documentaire à d'autres pratiques factuelles au-delà du genre cinématographique. Je me situe ainsi dans le champ large des narrations interactives non fictionnelles, que j'appellerai simplement de *documentaire interactif*.

Le concept d'interactivité a sa propre histoire, et ce serait largement au-delà de la portée de cette thèse de rendre compte de l'intégralité de son développement. Dans cette section, je propose d'esquisser une définition de l'interactivité et de ses différents modes. Dans le chapitre suivant, je développerai une interprétation de l'expérience interactive selon la perspective philosophique proposée par la présente thèse.

Dans la perspective de l'interaction humain-machine, l'interactivité est la relation entre le participant et l'image, mise en place dans un dispositif numérique, à travers une

interface « qui permet le contrôle et l'accès aux contenus » proposés par le créateur (Paquin, 2006, p. 15). Cette relation établie entre l'interacteur et l'image implique une réponse de la part du dispositif aux gestes posés par le participant. Dans les mots de Marie-Laure Ryan, l'interactivité se caractérise par « (...) the user's ability to provide input to the computer, and the computer's ability to adjust its behaviour according to this input. » (Ryan, 2005, "s. p."). La réponse du dispositif aux actions effectuées par l'interacteur est un aspect déterminant de l'interactivité, et ce qui permet de la distinguer d'autres modes d'activité qui composent une expérience audiovisuelle, comme l'interprétation. L'interprétation est une forme active de participation, il suffit de se rappeler des réactions qu'un film linéaire est capable de provoquer chez le spectateur (la fatigue à la suite de l'effort demandé par le film, la tristesse, qui des fois provoque même des larmes, etc.) pour comprendre qu'il y a aussi de l'activité dans une expérience linéaire. Néanmoins, dans une expérience linéaire l'activité du spectateur n'a pas la capacité de changer la forme du dispositif, cela est propre à l'interactivité.

L'interactivité peut prendre forme de diverses manières. Plusieurs taxonomies pour les différents types d'interactivité dans le documentaire ont été proposées (Galloway et al., 2007; Gaudenzi, 2013b; Gifreu, 2013). Je vais utiliser dans cette thèse la classification défendue par Gaudenzi qui détermine quatre modes d'interactivité dans le documentaire : hypertextuel, conversationnel, participatif et expérientiel (Gaudenzi, 2013a, 2013b). Mais avant de les présenter, il est important de souligner que ces modes d'interaction ne sont pas exclusifs, car souvent un même documentaire utilise simultanément plus d'un mode d'interactivité. Je présenterai les différents modes à travers des exemples, et comme on le remarquera, malgré une certaine difficulté à isoler les modes d'interaction, il est possible, dans la plupart des cas, d'établir un mode d'interaction prédominant pour chaque documentaire interactif.

a. Le mode d'interaction hypertextuel

Pour étudier le mode hypertextuel, je prends l'exemple du documentaire interactif *Rio de Janeiro — Autorretrato*¹⁰. J'ai choisi ce projet, car il s'agit d'un des premiers documentaires interactifs réalisés au Brésil, une scène à laquelle je suis liée de par mon origine, mais aussi par ma pratique de création. *Rio de Janeiro — Autorretrato* comprend aussi un court et un moyen-métrage linéaires, mais seulement le documentaire interactif produit pour le web (aussi appelé webdocumentaire) sera abordé ici. Dans ce webdocumentaire, après une introduction vidéo qui présente le contexte et le thème du documentaire (la vie dans les favelas du Rio de Janeiro, Brésil), l'interacteur se retrouve devant un menu avec quatre icônes : Vie quotidienne, Ville, Rêves et Personnes (fig. 2). En choisissant un de ces icônes, l'interacteur est dirigé vers une mosaïque de photos, et en cliquant sur chacune des images, une vidéo démarre. La plupart de ces vidéos sont des entretiens filmés avec des habitants des favelas. Au long des vidéos, il est possible de voir sur l'interface des icônes proposant des informations complémentaires (comme la biographie d'une personne mentionnée lors d'un entretien, des cartes ou des données sur les favelas).

¹⁰ <http://www.riodejaneiroautorretrato.com.br>, consulté le 17 octobre 2017.

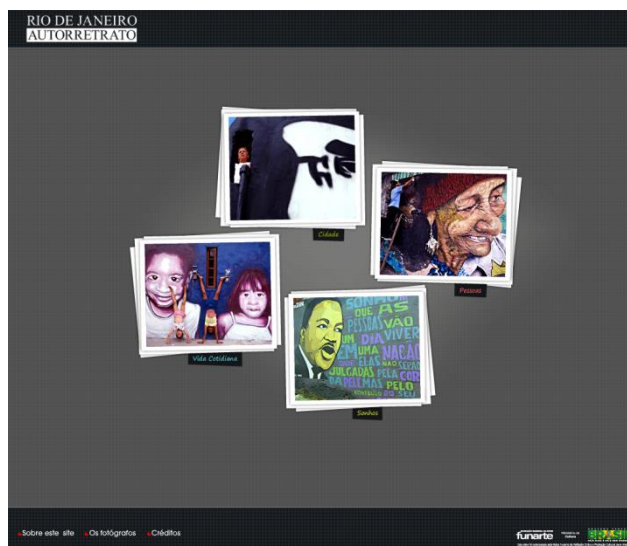


Figure 3 : Interface Rio de Janeiro — Autorretrato. Source : Capture d'écran¹¹

Le mode d'interaction de ce documentaire est hypertextuel et consiste à naviguer entre les différents contenus proposés par l'auteur. Le contenu est déjà établi au préalable, mais l'interacteur peut choisir l'ordre de visualisation. Dans ce mode d'interaction, le contenu est alors fermé, mais la manière de le parcourir est ouverte, et c'est cette ouverture qui le différencie du documentaire linéaire. Comme l'a affirmé Gifreu, « Hypertext is an open system that allows the receiver to construct their own reading pathways, jumping from lexia to lexia, according to their own personal interests. » (Gifreu, 2013, ch. 2, p. 1)

Dans le cas de *Rio de Janeiro - Autorretrato*, l'ouverture du documentaire se trouve dans la possibilité offerte à chaque interacteur de choisir le sujet qui l'intéresse, l'ordre des vidéos regardées, et de décider de consulter ou non les informations complémentaires proposées. Ainsi, même si l'interacteur n'intervient pas dans la construction du contenu, il le fait dans la composition de la forme du documentaire.

¹¹ http://riodejaneiroautorretrato.com.br/dev2011/Content/Swf/index_portugues.html, consulté le 17 octobre 2017.

Dans un parallèle avec le documentaire linéaire, l'interacteur agit dans le documentaire hypertextuel comme un monteur¹².

Un logiciel libre destiné à la création à partir de base de données (base d'images, sons, etc.), Korsakow, créé initialement par l'artiste Florian Talhofler, est un outil remarquable dans la création de documentaires interactifs hypertextuels. Le logiciel permet que certains des choix sur l'organisation du documentaire soient opérés par le propre système informatique à partir d'un système basé sur des mots-clés ou des règles établies par l'auteur, comme décrit sur le site de Korsakow : « En ajoutant des mots-clés à chaque SNU¹³, vous écrivez des règles qui détermineront comment se comportera votre K-film. »¹⁴ Le documentaire interactif *Ceci n'est pas Embres*¹⁵ pourra illustrer ici le mode de fonctionnement de Korsakow. Il s'agit d'un projet réalisé par l'artiste et chercheur montréalais, Matt Soar, conçu comme un journal de voyage, lors de son séjour de six mois dans le sud-ouest de la France. Dans le documentaire, à chaque SNU que nous regardons, d'autres sont proposés par des aperçus (photo ou vidéo) qui apparaissent du côté droit de la vidéo lorsque celle-ci s'approche de la fin.

Ainsi, le système Korsakow, à travers les mots-clés établis par l'auteur, propose un certain nombre de choix, et l'interacteur, en fonction de ses choix oriente le système à trouver les SNUs en lien avec celui qui a été choisi, « (...) selon la manière dont le film Korsakow a été conçu, un spectateur ne verra pas exactement le même contenu dans le même ordre (séquence) que le prochain. Par ailleurs, le spectateur sera activement appelé à "faire sens" tout au long du visionnement du film, selon le matériel rencontré,

¹² Il faudrait néanmoins souligner qu'un travail de montage a déjà été réalisé pour chaque vidéo mise en ligne, et dans ce montage préétabli l'interacteur ne peut pas intervenir.

¹³ SNU signifie *smallest narrative unit*, soit les extraits vidéos qui seront mis en relation par le logiciel et l'interacteur.

¹⁴ <http://korsakow.org/fr/apprenez/faq/#k-film>

¹⁵ Matt Soar, 2012, <http://www.embres.ca/>

guidé via Korsakow. »¹⁶ Il s'agit ici d'une hypertextualité partagée entre interacteur et système informatique.

b. Le mode d'interaction conversationnel

Dans le mode d'interaction conversationnel, l'interacteur agit constamment sur le dispositif et le système numérique répond à chaque geste effectué par l'interacteur. Gaudenzi utilise la métaphore de la conversation pour ce mode d'interaction, car elle illustre cet échange constant entre interacteur et œuvre : « The conversational mode is inspired by a type of interactivity that wants to reproduce the interaction between two human beings, or a human in a physical space. » (Gaudenzi, 2013b, p. 41) Ce mode est très commun dans une des tendances du documentaire interactif, le *docu-game*.

Prenons ici l'exemple du documentaire interactif produit par la chaîne franco-allemande Arte, *Le défi des bâtisseurs*¹⁷ (2012). L'expérience de ce documentaire commence par une introduction-vidéo qui présente la situation (en caméra subjective) : l'interacteur est dans un bureau d'architecture et le patron lui confie le projet de la deuxième tour de la cathédrale de Strasbourg. L'interacteur est ensuite introduit à l'équipe avec laquelle il doit travailler sur ce projet et quelques orientations sur le projet sont données. À partir de cette mise en contexte, le jeu commence : dans la partie centrale de l'interface, une esquisse de la cathédrale sans la deuxième tour est affichée (fig. 3). À droite se présente un menu avec des éléments architecturaux, classés par style (gothique, néo-gothique, moderne, etc.), que nous pouvons glisser sur l'esquisse pour construire la tour. Deux autres menus interactifs sont proposés : le premier, à l'extrême gauche de l'écran, est une simulation d'une messagerie texte à travers laquelle nous communiquons avec les membres de l'équipe. À l'extrême droite de l'écran s'affiche un menu de « rendez-vous », il s'agit d'extraits vidéos (aussi filmés en caméra subjective) qui simulent des rendez-vous avec des historiens, le réalisateur

¹⁶ <http://www.embres.ca/exhibit/process/>

¹⁷ <http://cathedrale.arte.tv/le-webdoc.php>

du documentaire interactif, des architectes, soit différentes personnalités qui nous donnent des informations sur la construction de la tour. Enfin, il est possible aussi de visualiser et de commenter tous les projets de la cathédrale réalisés par les interacteurs et soumis à une plateforme.

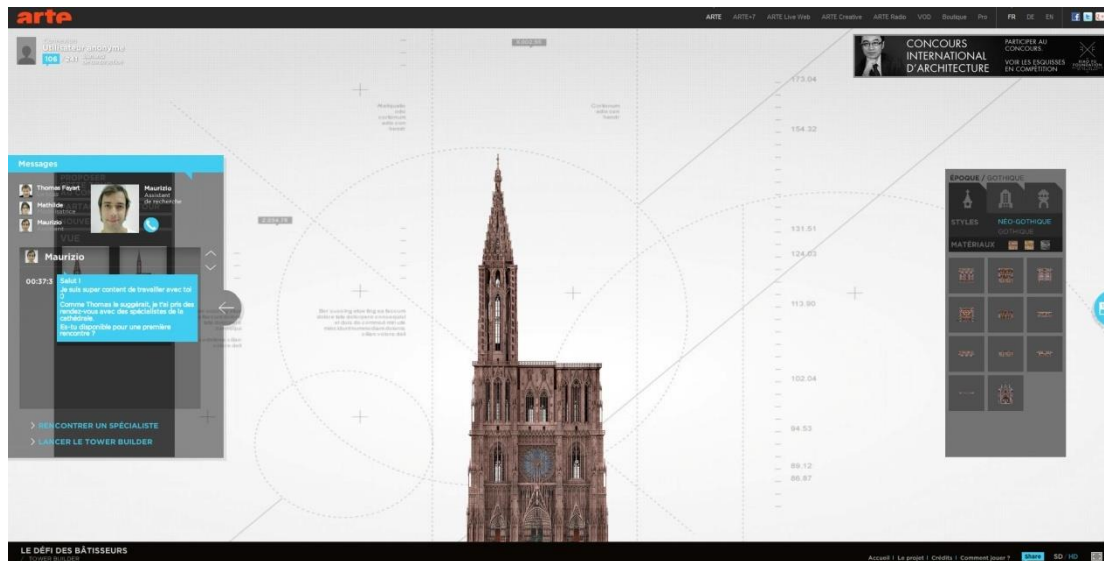


Figure 4 : Interface Le Défi des Bâisseurs. Source : Capture d'écran¹⁸

Comme nous pouvons le constater, les possibilités d'action de l'interacteur dans le documentaire conversationnel sont très différentes du documentaire hypertextuel. Ici, l'interacteur est capable d'interférer dans le contenu, et pas seulement dans le parcours qu'il fera pour le consulter. La métaphore de la conversation devient très évidente dans le menu qui permet de communiquer avec les membres de l'équipe. Cette communication est paramétrée au préalable et n'est pas aussi libre et ouverte qu'une conversation entre deux personnes, mais le principe d'une conversation est présent : pour chaque intervention de l'interacteur, il y a une réponse de la part du dispositif. Le champ d'action de l'interacteur est donc plus large dans le mode d'interaction

¹⁸ <http://cathedrale.arte.tv/webdoc/#/fr/game>, consulté le 17 octobre 2017.

conversationnel. Par ailleurs, le rôle de l'auteur change aussi avec les différents modes d'interaction. Dans le mode hypertextuel, le réalisateur créait le contenu et proposait à l'interacteur différentes manières de le parcourir. Dans le mode conversationnel, l'auteur crée un système à partir duquel l'interacteur pourra créer ou avec lequel il pourra dialoguer. Gaudenzi a comparé le rôle de l'auteur dans ces deux modes d'interaction et affirme sur le documentaire hypertextuel : « The author is not a facilitator as she is in the conversational mode, but a narrator that experiments with levels of choices within a controlled narrative framework. » (Gaudenzi, 2013b, p. 53)

Dans la création réalisée dans le cadre de la présente thèse, le *On Chance*, le mode d'interaction conversationnel sera le principal mode utilisé, dans la forme d'un agent conversationnel (*chatbot*).

c. Le mode d'interaction participatif

Dans le mode participatif, l'interacteur est invité à contribuer à la création du contenu du documentaire. Gaudenzi le décrit par la métaphore de la construction, « The author decides on the tools and rules and lays down the first layer of bricks, but there is room for collaboration and expansion. » (Gaudenzi, 2013b, p. 56) Essayons de mieux comprendre ce mode d'interaction à travers l'exemple du documentaire *Journal d'une insomnie collective*¹⁹ (2013), produit par l'Office national du film du Canada (ONF). Les producteurs de ce documentaire ont invité le public à soumettre des contributions en vidéo sur leur expérience de l'insomnie : « Depuis l'automne 2012, nous avons recueilli des centaines de témoignages d'insomniaques via le confessionnel de la webcam, du clavier et de la souris. »²⁰ Pour pouvoir accéder à ces témoignages, il est nécessaire de prendre un rendez-vous sur le site web du documentaire et ensuite de visiter le site à l'heure convenu. La prise de rendez-vous se fait à travers une interface composée de quatre images de visages, accompagnées des prénoms respectifs. En

¹⁹ <http://insomnie.onf.ca/#/insomnie>

²⁰ Idem

cliquant sur un des visages, le témoignage en vidéo du personnage sur son expérience de l'insomnie est montré. À la fin de la vidéo, il est proposé à l'interacteur de fixer un rendez-vous avec le personnage choisissant un créneau horaire (à partir de 20 h) et soumettant son email, téléphone et localisation (pays). À l'heure du rendez-vous, l'interacteur reçoit un appel téléphonique et un courrier électronique avec un lien pour qu'il puisse se connecter sur le site web du documentaire. C'est alors que l'interacteur aura accès à l'intégralité des témoignages des interacteurs sur leurs expériences d'insomnie. L'interacteur est invité également à soumettre sa propre expérience de l'insomnie répondant à des questions proposées par le dispositif. Les réponses peuvent prendre la forme d'un dessin, d'une vidéo ou d'un texte.

Il s'agit ici d'un exemple de documentaire qui construit son contenu par la participation de l'interacteur. La participation de l'interacteur n'est pas dans le choix de l'ordre de visionnage (hypertextuelle) ou dans un dialogue établi avec le dispositif (conversationnelle), mais elle a lieu lorsque l'interacteur ajoute du contenu à l'œuvre. Le rôle de l'auteur est plus dans la proposition et dans la création d'un dispositif qui permettra l'intégration de la production de l'interacteur que dans la réalisation d'un contenu.

Gaudenzi affirme aussi qu'il existe plusieurs manières de participation dans un documentaire interactif, et elle y inclut les formes de participation exclusives à l'étape de production. Comme cela avait été soulevé au moment de définir le documentaire interactif, j'exclus cette forme de participation du documentaire interactif, pour considérer uniquement des documentaires non linéaires. Il reste que la participation de l'interacteur peut se donner de différentes formes, par les commentaires qui seront intégrés aux images (*Highrise*²¹, *Rio de Janeiro—Autorretrato*), des histoires personnelles en lien avec le thème du documentaire (*Hollow*²², *7 billion others*²³), des

²¹ Katerina Cisek, 2009-2015, <http://highrise.nfb.ca/>

²² Elaine McMillion, 2013, <http://hollowdocumentary.com/>

²³ Yann Arthus-Bertrand, 2003, <http://www.7billionothers.org/>

enregistrements audio et/ou vidéo (*Rider Spoke*²⁴, *Journal d'une insomnie collective*), entre autres.

d. Le mode d'interaction expérientiel (ou géolocalisé)

Le mode expérientiel est celui qui intègre les médias géolocalisés (*locative media*) dans le dispositif d'interaction. Il s'agit d'une interaction qui prend en considération des paramètres de l'espace physique dans lequel se situe le documentaire, normalement à travers l'usage du GPS (*Global Positioning System*) et qui agit, dans la plupart des cas, dans la relation de l'interacteur avec l'espace. Le contexte dans lequel on fait l'expérience du documentaire acquiert donc une grande relevance, ce qui restreint souvent l'œuvre à un lieu spécifique. Un autre aspect important de l'interactivité géolocalisée est la mobilité de l'interacteur. Comme l'a souligné Pengkai Pan, « The paradigm of computer-human interaction through mobile channels is quite different from desktop-based computer-human interaction. The fundamental difference is that the audience's body becomes a "mouse point" to navigate through space. » (Pan, 2004, p. 29) L'expérience de ce mode d'interaction fait appel dans une large mesure au corps en mouvement, car le participant est invité à agir dans l'espace, et pas seulement dans une interface numérique comme dans la plupart des cas étudiés jusqu'ici.

L'application pour iPhone *Walking the Edit*²⁵ (Ulrich Fisher, 2010) peut aider à comprendre le fonctionnement des documentaires interactifs géolocalisés. Il s'agit d'un dispositif qui permet à l'utilisateur muni d'un iPhone et d'écouteurs d'activer des extraits de vidéos (provenant des archives de la ville ou réalisés au préalable par une équipe liée au projet) au moment où il passe par l'endroit où la vidéo a eu lieu (l'application est disponible pour les villes de Bruxelles, Paris et Bagnols-sur-Cèze). L'expérience de la ballade se trouve donc enrichie par ces divers fragments audiovisuels liés au lieu de passage. Le dispositif enregistre le parcours et le rythme de

²⁴ Blast Theory, 2007

²⁵ <http://walking-the-edit.net/fr/>

marche de l'interacteur, pour, à partir de ces paramètres, monter une vidéo avec les extraits qui ont été joués pendant la ballade

En fonction de ses déplacements dans l'espace urbain, de son rythme de marche, en somme de son comportement, le promeneur produit un montage, reflet formel de son parcours, au travers de cette mémoire audiovisuelle partagée qui existe virtuellement autour de lui.²⁶

Les diverses vidéos réalisées sont affichées sur un site internet, à côté de la trace du parcours marqué sur une carte (fig. 4). Il est aussi possible de se déplacer avec l'aide du curseur sur la carte et regarder les images réalisées par l'interaction du participant avec le dispositif.

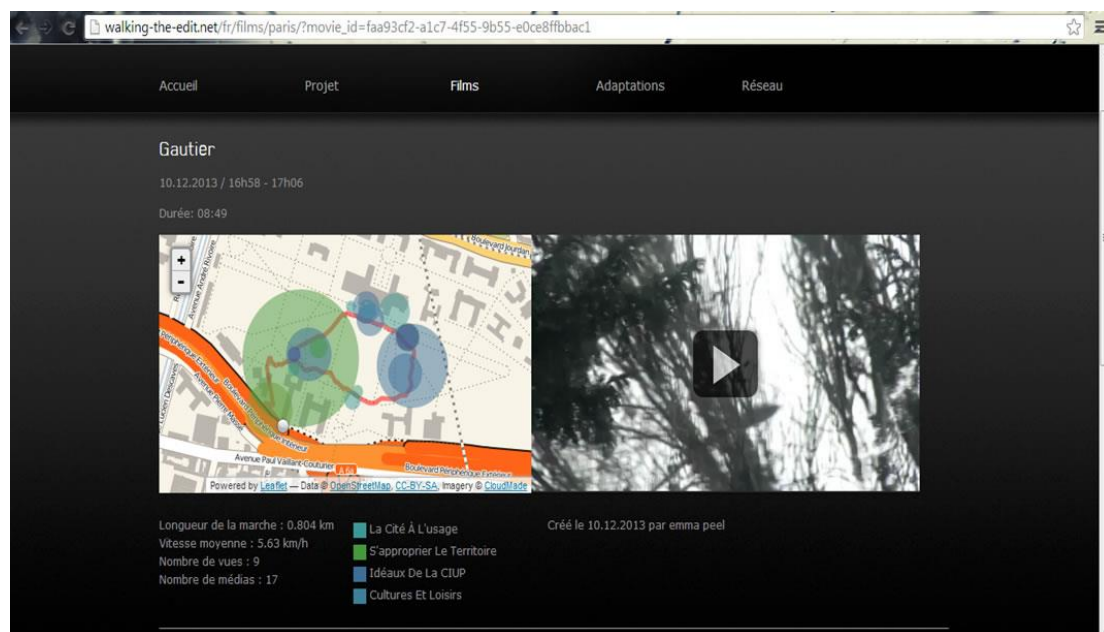


Figure 5 : Interface web Walking the Edit. Source : Capture d'écran

²⁶ <http://walking-the-edit.net/fr/projet/presentation/>

Dans ce mode d'interactivité, l'ambivalence apparente du rapport entre l'indétermination et l'interactivité, qui sera abordée plus en profondeur au long de cette thèse, se manifeste de manière assez évidente. D'un côté, les technologies géolocalisées proposent souvent d'augmenter la connaissance de l'espace dans la perspective de réduire toute indétermination et imprévisibilité. Comme le remarque Jordan Frith, « Location-based social networks (LBSNs), such as Foursquare, Loopt, Brightkite, GoWalla and Lattitude, are designed to overcome randomness and chaos. » (Frith, 2012, p. 142) L'utilisation des technologies géolocalisées agit de mode à sélectionner le connu (le réseau d'amis, les endroits déjà visités et commentés par le réseau, etc.) dans un entourage inconnu. Le discours du cofondateur de *Foursquare*, un des plus importants réseaux géolocalisés, confirme cette intention de réduire les indéterminations liées à l'espace et l'ambition de rendre les déplacements entièrement prévisibles à partir d'une grande collection de données : « The thing that we are most passionate in Foursquare is doing that, not predictable place, but more a predictable awareness of the great stuff around you, specifically tied to things in the real world. » (Crowley, 2014, "s. p.")

D'un autre côté, l'interactivité géolocalisée ouvre le documentaire à différentes sources d'indétermination liées au déplacement dans l'espace. En *mobility studies*, l'indétermination est considérée comme un élément inhérent de la mobilité humaine. Comme l'observe Song et al., « Indeed, although we rarely perceive any of our actions to be random, from the perspective of an outside observer who is unaware of our motivations and schedule, our activity pattern can easily appear random and unpredictable. » (Song, Qu, Blumm, & Barabási, 2010, p. 1018) Il est donc très difficile pour le créateur d'un documentaire interactif de prévoir les déplacements de l'interacteur, ce qui implique souvent une délimitation de l'espace de l'expérience interactive. Par ailleurs, l'espace dans lequel se situe l'interacteur offre de nombreux aspects imprévisibles, les rencontres avec une diversité d'éléments (la météo, la circulation, etc.) qui échappent à la maîtrise à la fois du créateur et de l'interacteur.

Cette incertitude est encore plus remarquable dans le contexte urbain, comme l'a argumenté Frith, « What frightens people about cities is also what makes them great: they are sites of unexpected encounters, encounters we cannot always control. » (Frith, 2012, p. 141)

Dans ce sens, il est question de savoir si, dans le cas de *Walking the Edit*, la mobilité de l'interacteur est présente dans le dispositif de forme à embrasser ou à essayer de prescrire l'incertitude de la relation à l'espace. Autrement dit, est-ce que *Walking the Edit* agit pour encourager les rencontres aléatoires ou faire ressortir le hasard présent dans la relation à l'espace ? Thierry Davila constate qu'un marcheur est toujours pris entre une intention de trajet et les aléas de l'espace qu'il parcourt :

Or, le marcheur est simultanément celui qui donne un profil à son chemin, ouvre ou trace une voie, et celui qui adapte ce trajet à un contexte le construit en fonction des accidents et des contraintes du parcours, des événements scandant la progression de ses déplacements, et qui invente un rythme au gré des vicissitudes de la flânerie. (Davila, 2007, p. 22)

Ulrich Fisher, le développeur de *Walking the Edit*, note que ce croisement typique de la marche (entre détermination et aléas) s'élargit au film qui est construit dans l'application mobile. Il remarque que « [...] le film porte les fruits du hasard et de la détermination » (Fisher, 2009, p. 8), dans la mesure où le film qui résulte de l'usage de l'application est construit dans le croisement entre des paramètres déterminés (les contraintes de l'espace et du dispositif) et des paramètres imprévisibles (les aléas rencontrés dans l'espace et l'indétermination technologique).

La posture créative dans le rapport à l'espace proposée à l'utilisateur de *Walking the Edit* est très proche de nombreuses pratiques artistiques. Thierry Davila analyse ce type de pratique sous le terme de *cinéplastique*. Davila revient sur la démarche situationniste de la *dérive* par Guy Debord, ainsi que celles d'artistes contemporains (Francis Alÿs, Stalker et Gabriel Orozco) et il soulève aussi l'importance de l'incertitude présente dans le rapport à l'espace (notamment la marche) pour ce type de travail artistique :

Souvent, la marche parvient à déplacer la réalité la plus routinière vers des zones d'incertitude et d'étrangeté qui la font renaître à elle-même et qui nous mettent face à une autre ville, un autre décor, comme incrusté dans les rues et les passages que nous croyions pourtant connaître. (Davila, 2007, p. 160)

Nous remarquons que la cinéplastique propose un dispositif qui renoue le marcheur à l'indétermination et transforme ainsi sa relation à l'espace. *Walking the Edit* s'insère dans cette tradition de la cinéplastique, car il s'agit d'un dispositif, cette fois-ci numérique, qui réclame comme un de ses

[...] enjeux artistiques : Explorer les formes narratives ouvertes, en particulier celles dont la structure causale et formelle n'est pas prévue d'avance et repose, dans notre cas de figure, sur la logique d'un parcours qui n'est pas entièrement maîtrisé par le créateur du dispositif (l'hypothèse de départ pour le projet). » (Fisher, 2009, p. 12)

1.3. Posture : Pragmatisme spéculatif

Dans un article récent (Debaise & Stengers, 2016), Didier Debaise et Isabelle Stengers formulent des propositions de ce qu'ils appellent les « gestes spéculatifs » afin de bâtir un programme de recherche. La première des propositions vient de la philosophie spéculative d'Alfred N. Whitehead et la deuxième de l'empirisme radical de William James :

Les deux parties de la proposition forment une double contrainte cruciale. Tout d'abord, ne rien exclure, prendre en compte la multiplicité des dimensions qui composent une expérience hic et nunc, n'en rien soustraire a priori, quelles que soient les disqualifications dont elle puisse être l'objet. Ensuite, ne pas admettre de principe de jugement hors situation, qui domestiquerait cette multiplicité en termes de catégories ou d'exigences qui lui sont étrangères. (Debaise & Stengers, 2016, p. 84)

La première contrainte, « ne rien exclure », demande que l'on dépasse les dualismes de la philosophie moderne, que l'on évite de créer de catégories préétablies pour ensuite classer les aspects qui seraient le plus significatifs. Cette proposition peut néanmoins

être lue comme un appel à la prise en compte indiscriminée de la totalité des dimensions de l'expérience. Ce n'est pas le cas, comme écrivent les auteurs : « (...) la position de Whitehead (...) n'affirme pas pour autant qu'il faille tout prendre en compte : elle affirme qu'il faut refuser le droit de disqualifier » (Debaise & Stengers, 2016, p. 84). En effet, il s'agit de ne rien exclure a priori, de ne pas trier à l'avance les aspects de l'expérience qui seraient pertinents. La deuxième contrainte, celle du pragmatisme, agit alors pour distinguer l'importance des aspects à être considérés. Ainsi, c'est en soumettant les idées au crible de ses conséquences pratiques que l'expérience doit être prise en compte. Debaise et Stengers prennent l'exemple de la conférence de William James dont il a été question plus haut, « The Dilemma of Determinism », pour illustrer la posture de l'empirisme radical :

Ainsi, dans « Le dilemme du déterminisme », James s'étonne de la tranquillité avec laquelle ses collègues peuvent se réputer déterministes, sans se préoccuper un instant des conséquences littéralement invivables de cette doctrine qui réduit à l'illusion tout ce qui nous importe. (Debaise & Stengers, 2016, p. 85)

Comme le démontre James, lorsque le déterminisme est mené au bout de ses conséquences, toute possibilité d'agentivité et de nouveauté est exclue, et même le regret devient impensable.

L'aspect spéculatif du pragmatisme proposé par Debaise et Stengers joue un rôle majeur dans l'approche de la présente thèse. Pour Debaise et Stengers, la posture spéculative est une « manière de susciter des possibles » (Debaise & Stengers, 2016, p. 83), de se pencher sur les trajectoires alternatives que chaque situation recèle. Les auteurs font alors une distinction importante, celle entre le possible et le probable :

Le probable relève par définition, comme l'atteste le calcul des probabilités, d'une transposition ou d'un réagencement de ce qui a déjà eu lieu ou de ce qui est en cours. Le probable appartient à une logique de conformité : ce qui a compté dans le passé, ce qui permet de le caractériser, conservera ce pouvoir dans le futur. Le possible, quant à lui, fait importer l'irruption éventuelle d'autres manières de sentir, de penser, d'agir, qui ne

peuvent pas être envisagées que sur le mode d'une insistance, sapant l'autorité du présent quant à la définition de l'avenir. (Debaïse & Stengers, 2016, p. 87)

Il s'agira ainsi dans cette thèse de s'interroger sur les pratiques qui permettent, dans le domaine de la création numérique interactive (plus précisément du documentaire interactif), d'activer des processus qui ne soient pas définis par la répétition conforme du passé et qui ouvrent l'expérience interactive à des possibles imprévisibles. Autrement dit, il s'agira de comprendre comment la création interactive peut échapper aux limites du probable pour proposer des expériences ou des situations qui « suscitent des possibles ».

Dans son livre *Semblance and Event*, Brian Massumi développe aussi un pragmatisme spéculatif s'inspirant des travaux de James et de Whitehead où il sera largement question d'une remise en question des dichotomies de la philosophie moderne, notamment celle qui oppose sujet et objet. Ces réflexions intègrent également la posture adoptée dans cette recherche-création et seront importantes dans la formulation proposée de l'expérience interactive.

Massumi revient sur les transformations infligées à ces notions par James :

Classically, objects and their associated operations are in the world while percepts registering them are in the subject. What James is saying, by contrast, is that both are in the transition. Things and their experience are together in transition. (...) [O]bject and subject fall on the same side of a shared movement. (Massumi, 2011, p. 30)

Sujet et objet ne seraient pas alors des instances fixes, mais des facettes d'un événement. Une première évidence mentionnée par Massumi de la difficulté de définir des catégories prédéfinies pour l'objet et le sujet est le fait que chaque individu serait un sujet pour soi-même, mais un objet pour tous les autres individus. Il affirme : « Subject and object are *operative* definitions by pragmatism. They are not placed in any kind of metaphysical contradiction or opposition. » (Massumi, 2011, p. 31)

Massumi défait également la dichotomie sujet-objet des approches cognitivistes :

The cognitivist paradigm equates the subject with the knower and the object with the known. (...) An essential divide is presupposed the moment the categories of knower and known are overlaid upon the subject and the object, and no amount of subsequent maneuvering, however ingeniously contortionist, will smooth it over. The problem is that any way you twist it, the knowing is still in the subject and the known is still right where it was on the other side. What can guarantee that they correlate to each other? (Massumi, 2011, pp. 6-7)

Plutôt que de travailler avec les catégories sujet/objet, c'est l'événement qui est mis au centre de cette posture : « Neither object nor subject : event. » (Massumi, 2011, p. 6). Il s'agit ainsi de placer l'objet et le sujet dans une perspective processuelle, et de les redéfinir en termes d'aspects en changement permanent dans des processus de devenir. Comme il démontre, sujet et objet ne sont pas des catégories prédéfinies, ce sont des éléments dans une relation d'inclusion mutuelle. Massumi reprend alors les définitions de sujet et d'objet telles qu'elles ont été reformulées par Whitehead : l'objet est ce qui reste dans le monde une fois qu'un événement s'accomplit et qui est alors repris par les événements en devenir. Autrement dit, l'objet est la donnée dans son sens premier : ce qui est donné. Whitehead reprend le terme de *datum* pour sa définition de l'objet (Massumi, 2011, p. 8). Le sujet est quant à lui inséparable de l'événement, « [t] he subject is what finds itself in the midst of these processual leavings, taking them up as the world's gift of potential for its own taking form. » (Massumi, 2011, p. 8) Si l'objet est ce qui rattache l'événement au passé, le sujet est dans le présent qui passe et qui traverse l'événement, il est une durée qualitative, « a span of becoming » composé de l'inclusion mutuelle de différentes phases de l'événement (Massumi, 2011, p. 9).

Les conséquences de cette reformulation des catégories de sujet et objet vont au-delà du champ philosophique. Comme l'auteur souligne, la dichotomie sujet-objet oriente encore les pratiques de connaissance en les divisant en deux champs, ou deux cultures, les sciences et les humanités, ou alors entre les pratiques empiriques (dites objectives) et les pratiques spéculatives (dites subjectives) (Massumi, 2011, p. 11). Comme on le

verra plus loin, cette même dichotomie conduit les lignes d'interprétation de la théorie des probabilités, ce qui a des conséquences importantes pour la compréhension des algorithmes probabilistes largement utilisés dans la création numérique contemporaine et dont il sera question dans la présente thèse. En effet, un des dualismes dont il s'agit dans les interprétations des modèles probabilistes est celui entre les interprétations objective et subjective des probabilités (Gillies, 2000). Dans le domaine de l'apprentissage automatique, ce dualisme est résumé souvent par l'opposition entre l'interprétation fréquentiste et l'interprétation bayésienne des probabilités (Murphy, 2012 ; Smith, 2016). Néanmoins, l'étude des interprétations philosophiques de la théorie des probabilités par la perspective du pragmatisme spéculatif permettra d'apporter des nuances à cette perspective dualiste, et contribuer ainsi à enrichir les discussions contemporaines autour des modèles probabilistes de l'apprentissage automatique (ou ML). Il s'agira dans la suite de cette thèse d'articuler ces discussions dans le champ du documentaire interactif, dans une perspective non dualiste. En effet, c'est à partir de cette reformulation de la dualité sujet/objet et des propositions pragmatistes spéculatives que je propose d'aborder le problème du hasard et de l'indétermination dans le documentaire interactif. Or, une question sous-jacente qui traversera cette thèse est ainsi liée à l'approche choisie. En raison de la posture adoptée, il sera alors question de reformuler toute forme de dualisme hérité de la philosophie moderne et des approches cognitivistes pour comprendre une expérience interactive qui ne soit pas basée sur le dualisme sujet/objet.

1.4. Bricolage méthodologique : La recherche-crédation et le cadrage de l'inopiné

L'articulation entre la recherche et la création est problématique, car ce sont des activités qui engagent des méthodes et des pratiques de nature très différente. Louis-Claude Paquin a noté que l'expression « recherche-crédation » a « été forgée à partir de la réunion de deux termes qui désignent des activités qui, jusqu'à l'intégration des Écoles des beaux-arts aux universités, s'opposaient » (Paquin, 2014, p. 2). Erin Manning et Brian Massumi ont remarqué que la catégorie recherche-crédation a été

adoptée comme un volet autonome dans les institutions académiques canadiennes sans une réflexion approfondie sur la manière comment ces deux éléments s'interpénètrent (Manning & Massumi, 2014). Cette problématique est abordée par Owen Chapman et Kim Sawchuk dans un article autour des méthodologies de recherche-crédation. Les auteurs identifient quatre tendances dans l'articulation entre recherche et création : « research-for-creation », « research-from-creation », « creative presentations of research » et « creation-as-research » (Chapman & Sawchuk, 2012, p. 8). D'après les sous-catégories formulées par ces auteurs, il s'agira dans la présente thèse de deux formes de recherche-crédation : « research from creation » et « creation-as-research ».

La première de ces articulations, « research from creation », consiste dans la théorisation d'une certaine pratique de création : « performances, experiences, interactive art works, et cetera can also be ways of generating research data that can then be used to understand different dynamics. » (Chapman & Sawchuk, 2012, p. 16). Il s'agira ici d'observer ma propre pratique (ainsi que celles d'autres créateurs qui utilisent le hasard comme méthode) pour comprendre les enjeux de l'utilisation du hasard dans le processus de création et d'expérience dans le champ du documentaire interactif. Dans « creation-as-research », la création devient l'espace où la recherche peut être testée, « It is about understanding the technologies/media/practices that we discuss as communication scholars (for instance) by actually deploying these phenomena, and pushing them into creative directions. » (Chapman & Sawchuk, 2012, p. 19). La thèse se construit ainsi dans un va-et-vient entre la réflexion conceptuelle à partir des éléments de recherche et la pratique de création associée à ces mêmes éléments. Dans le cas de la présente thèse, ce processus se traduit par l'articulation du travail conceptuel autour du hasard et sa mise en pratique dans la réalisation d'un documentaire interactif.

Il est intéressant de noter que dans la première des formes de recherche-crédation mentionnées plus haut, il s'agit de partir de la pratique pour développer une réflexion autour du rôle du hasard dans le processus de création et dans l'expérience du

documentaire interactif. Or, dans la deuxième des formes, il est question de partir de la discussion conceptuelle autour du hasard et de l'indétermination pour comprendre la pratique et l'expérience. Ainsi, cette recherche-cr ation se construit dans un espace hybride entre recherche et cr ation, o  la cr ation est une mise en forme d'aspects de la r flexion sur le hasard et la r flexion est elle-m me consid r e comme une pratique cr ative.

La perspective de Manning et Massumi autour de la recherche-cr ation sera aussi importante dans la r alisation de cette th se. Dans le livre *Thought in the Act* (2014), les auteurs mettent d'abord en  vidence le contexte institutionnel et  conomique de la conception dominante de la recherche-cr ation, pour ensuite formuler une alternative, soit une conception de la recherche-cr ation en tant que processus, « a mutual interpenetration of processes rather than a communication of products » (Manning & Massumi, 2014, p. 89). Cette formulation est d'autant plus pertinente dans ce travail qui aborde les th matiques de l'ind termination et du hasard, car comme affirment les auteurs, « The kind of results aimed at would not be preprogrammed. They would be experimental, emergent effects of an ongoing process. » (Manning & Massumi, 2014, p. 89). Or, une grande emphase sera donn e dans cette th se au r cit de pratique et   la discussion autour du processus de cr ation.

Manning et Massumi formulent alors des propositions pour une pratique exp rimentale de la recherche-cr ation. Ces propositions seront les principales orientations m thodologiques de la pr sente recherche-cr ation. Je reviens ici sur la proposition la plus pertinente pour cette th se :

Proposition 3 — Design enabling constraints : cette proposition concerne la mise en place de contraintes qui visent   cr er les conditions initiales n cessaires pour l' mergence de nouvelles relations et interactions dans le cadre d'une recherche-cr ation. Il s'agit d'une vision positive des contraintes, dans le sens o  elles permettent l'existence d'un contexte pour que l'improvisation et l'impr visible puissent  merger.

Comme pour la plupart des artistes qui travaillent avec la notion de hasard (François Morellet, John Cage, Dada, entre autres), ainsi que pour Manning et Massumi, la création imprévisible se fait à l'intérieur de certaines limites imposées par les contraintes définies par l'artiste. Dans le cas de Morellet, par exemple, l'inclusion du hasard dans la pratique artistique se fait dans un ensemble de règles et de procédures très précises déterminées par l'artiste, ce sont les *enabling constraints*.

Cette proposition met en évidence une conception courante (mais erronée) des pratiques qui donnent de la place à l'émergence inopinée :

We wanted at all costs to avoid the voluntaristic connotations often carried by words like “improvisation,” “emergence,” and “invention.” There would be no question of just “letting things flow,” as if simply unconstraining interaction were sufficient to enable something “creative” to happen. In our experience, unconstrained interaction rarely yields worthwhile effects. (Manning & Massumi, 2014, p. 93)

Cette proposition est proche d'une méthode de création développée dans le cadre du documentaire qui sera au cœur de la démarche de recherche-création proposée dans cette thèse. Cette approche méthodologique concerne les principes de la réalisation documentaire. Jean-Luc Lioult a proposé une méthodologie de création documentaire qui se situe dans l'intersection entre le documentaire et le hasard (Lioult, 2008b). Il l'a nommée « le cadrage de l'inopiné ». Il s'agit de mettre en place un dispositif de captation d'événements spontanés orienté par l'intention du cinéaste. Jean-Luc Lioult dira : « (...) les protocoles opératoires du filmage documentaire consistent souvent à établir un cadre (au sens physique, social, figuratif...) à l'intérieur duquel de l'imprévisible (auquel on se fie toutefois) peut se produire. » (Lioult, 2008a, p. 83). Malgré une contradiction apparente entre les concepts de cadrage et hasard, je note ici que tout événement du hasard se fait à l'intérieur d'un cadre déterminé.

Comme l'a affirmé René Lestienne : « [...] le plus merveilleux dans la nature, c'est que même la part qui échappe à la causalité n'échappe pas pour autant au pouvoir de notre esprit de l'encadrer, d'édicter à son égard des lois. En fait, c'est par les lois du hasard que nous pouvons traquer son existence. » (Lestienne, 1993, p. 264) C'est aussi cette manière de travailler que parle Nicolas Philibert quand il dit : « Jean Oury a une belle expression : programmer le hasard. Pour moi, un film c'est un peu ça. C'est pouvoir accueillir l'imprévu dans un cadre donné » (Philibert, 2009, p. 7) ou encore :

Au fond, je cherche à faire en sorte que des choses puissent éclore, émerger, à l'intérieur d'un cadre, d'une situation donnée. Ce cadre, ce n'est pas seulement un espace. C'est tout ce qu'on met en œuvre pour que ces choses adviennent : un climat, un mode de relation à ceux que l'on filme, une disponibilité, son désir, une éthique, une part de jeu aussi... (Philibert, 2009, p. 7)

À ce propos, Noël Burch dira : « Tout ceci peut paraître élémentaire : le cinéma, nous dira-t-on, est toujours fait de ce compromis entre le hasard et la volonté de l'auteur. C'est exact, mais tout est une question de degrés... et d'attitude. Il y a, d'une part, nous l'avons dit, ceux qui de tout temps ont cherché à éliminer au maximum les aléas. D'autre part, nombreux, en effet, sont ceux qui ont composé avec le hasard, mais encore ne l'ont-ils pas tous fait dans le même esprit. » (Burch, 1986 [1967], p. 163). En effet, comme nous venons de voir, le geste du cadrage de l'inopiné se définit dans la rencontre entre l'intention du réalisateur (il établit la situation) et les événements imprévus, entre dispositif et hasard.

PREMIÈRE PARTIE : DISCUSSION THÉORIQUE

2. CHAPITRE II

L'ANALOGIQUE : IMAGE-INTERACTION

Les relations entre le hasard et le documentaire interactif seront étudiées théoriquement dans cette thèse sous deux aspects. Le premier, abordé dans ce chapitre, concerne le rôle attribué au hasard et à l'indétermination dans l'expérience du documentaire interactif. Le deuxième, à être traité dans les chapitres suivants, se concentre sur la présence du hasard dans la dimension numérique du documentaire interactif. Dans ce chapitre sur les processus analogiques du documentaire interactif, il sera question du rôle que peuvent jouer le hasard et l'indétermination dans la création et dans l'expérience d'un documentaire interactif. Les questions auxquelles il s'agira de répondre sont alors : d'une perspective pragmatiste spéculative, qu'est-ce qu'une expérience interactive ? Quel est le rôle du hasard et de l'indétermination dans l'expérience interactive ? Qu'est qu'une image dans le contexte du documentaire interactif et que se passe-t-il quand cette image est rapportée à un centre d'indétermination, l'interacteur ? Ces questions seront abordées dans l'ordre dans les sections à venir.

Le rôle du hasard dans le processus de création du cinéma documentaire a déjà été abordé dans les études en cinéma (Burch, 1986 [1967]; Caillat, 2001; Lioult, 2008b; Philibert, 2009). En 1967, Noël Burch écrit sur les cinéastes qui ont intégré volontairement le hasard dans ses films, et note le rôle important du documentaire parmi ces démarches :

Il est vrai que la fonction la plus noble du hasard, celle qui opère au niveau de la forme, n'a été comprise que très récemment par des cinéastes comme Godard et Rouch. Cependant, quelques esprits très exceptionnels avaient déjà pressenti sa fonction au niveau du langage. Dès 1920 au moins, on trouve des cinéastes qui ne cherchent pas du tout à vaincre le hasard, mais veulent au contraire subordonner dans une large mesure leurs caméras à ce monde aléatoire qu'ils appellent la réalité. Il s'agit bien entendu, de premiers grands documentaristes : Dziga Vertov, Joris Ivens, Ruttman, Cavalcanti et d'autres. (Burch, 1986 [1967], p. 160).

Le documentaire est une pratique de création dont l'imprévisibilité est partie constitutive, car les événements filmés sont souvent peu scénarisés. Le réalisateur Stéphane Breton raisonne dans ce sens sur le rôle du hasard dans la construction d'un documentaire et selon lui, la place donnée à l'aléatoire est au cœur de ce genre cinématographique : « La règle éthique et artistique du cinéma documentaire est de préserver le caractère imprévu des événements qui se produisent et l'attitude improvisée (mais non pas irréfléchie) qu'on adopte à leur égard... » (Stéphane Breton cité dans Lioult, 2008b, p. 86).

Dans le domaine des arts plastiques, le rôle attribué au hasard dans les processus de création est aussi une question abondamment étudiée (Faure, 1991; Henderson, 2001; Pollock, 2012; Saurisse, 2007), et ces discussions peuvent apporter des éléments intéressants à rapporter au domaine du documentaire interactif. Le 20e siècle marque un tournant dans la conception du hasard et dans le rôle que les artistes lui attribuent dans ses processus de création. Ce n'est qu'avec des artistes comme Marcel Duchamp (1887-1968), John Cage (1912-1992), le mouvement Dada (*circa* 1916-1924), l'Internationale situationniste (1957-1972), François Morellet (1926-2016) que le hasard commence à être revendiqué en tant qu'élément du processus créatif. Comme le dit Pierre Saurisse :

Le hasard ne s'est longtemps fait une place dans la création artistique que timidement, comme par inadvertance. (...) Introduire l'imprévisible et l'incontrôlé dans le processus de création, et les accepter comme tels,

ébranle la valeur accordée au savoir-faire, et met à mal le fétichisme traditionnellement associé à l'œuvre d'art. (Saurisse, 2007, p. 11)

La difficulté à accepter le hasard dans la création artistique a été résumée par Jean-Claude Lebensztejn dans l'affirmation suivante « qu'un artiste n'a pas le droit de laisser le hasard prendre sa place, parce qu'il n'ait pas d'art possible sans une intention humaine, sans un père qui s'en porte garant, même s'il (surtout s'il) se dit inspiré par des forces qui le dépassent. » (Lebensztejn in Saurisse, 2007, p. 13)

Avant le 20^e siècle, dans les arts plastiques, le hasard n'était présent dans le processus de création que dans les gestes qui provoquaient un résultat à la fois inattendu et apprécié, comme dans le cas de gouttes de peinture qui tombent au hasard sur une toile et forment une composition d'intérêt, un effet imprévu et désirable. De ce point de vue, on parle du hasard surtout comme une forme non intentionnelle que l'artiste exploite par la suite, sans qu'il y ait de déplacement du pouvoir de décision sur la création. C'est surtout avec l'avant-garde qu'on a pu voir des œuvres d'art qui revendiquaient le hasard en tant que matériau de création. Ce tournant dans les arts visuels n'est pas sans relation avec la remise en question du déterminisme scientifique abordée dans le chapitre précédent.

Ce qui renouvelle en profondeur l'idée du hasard artistique au 20^e siècle est la mise au point de techniques aléatoires distanciées, le déclenchement d'engrenages planifiés, l'élaboration d'une gestation méthodique à l'issue de laquelle le hasard prend forme. (Saurisse, 2007, p. 16)

Cela peut paraître contradictoire de parler en même temps de hasard, de planification et de méthode, mais la démarche avant-gardiste, ainsi que celle de la plupart des artistes qui intégreront le hasard dans leurs processus de création, ne consiste pas à donner libre cours aux événements inattendus. Il s'agit plutôt de les provoquer et de leur donner un cadre. Or, si dans le documentaire linéaire et d'autres pratiques artistiques le hasard

peut être intégré au processus de création, est-ce que l'ajout de la dimension interactive étend la participation du hasard à l'expérience de l'interacteur ?

2.1. L'expérience du documentaire interactif

Dans cette section, il s'agira de réfléchir sur les modes d'expérience dans le contexte du documentaire interactif. Plus haut dans cette thèse, selon une proposition de Gaudenzi (2013a), il a été observé qu'il existe une pluralité de modes d'interactivité. Il en découle qu'il existe aussi une pluralité des modes d'expérience du documentaire interactif. J'aborderai alors certains modes de l'expérience interactive, tout en admettant la grande diversité qu'il existe dans les pratiques interactives. Mais avant d'inclure la variable de l'interactivité, commençons par l'expérience.

Dans cette thèse, la posture pragmatiste spéculative a été adoptée et cela implique, comme nous avons vu plus haut, qu'il est nécessaire de repenser certains dualismes de la philosophie moderne, notamment celui qui détermine des catégories préétablies au sujet et à l'objet. Comment alors penser la notion d'expérience sans tomber dans la dichotomie d'un sujet qui fait l'expérience d'un objet ? William James, dans ses *Essays in Radical Empiricism*, aborde en profondeur la question de l'expérience pour se débarrasser du dualisme sujet/objet. Il affirme :

As 'subjective' we say that the experience represents; as 'objective' it is represented. What represents and what is represented is here numerically the same; but we must remember that no dualism of being represented and representing resides in the experience *per se*. In its pure state, or when isolated, there is no self-splitting of it into consciousness and what the consciousness is 'of'. Its subjectivity and objectivity are functional attributes solely, realized only when the experience is 'take,' (...). (James, 2016)

Dans la perspective du pragmatisme spéculatif, l'expérience est liée à la notion d'événement, ou des occasions d'expérience (Massumi, 2011). Il s'agit ainsi d'une interprétation de l'expérience en tant que processus partagé, sans que les catégories de sujet et d'objet soient établies au préalable. Comme explique Erin Manning :

By making everything an event, by emphasizing that there is nothing outside of or beyond the event, the aim is to create an account of experience that requires no omnipresence. The event is where experience actualizes. Experience here is in the tense of life-living, not human life per se, but the more-than human: life at the interstices of experience in the ecology of practices. (Manning, 2016, p. 3)

En effet, l'expérience a lieu au moment de la réalisation de l'événement, au moment où les différents processus qui composent l'expérience sont indéterminés, entremêlés dans l'événement, dans une «écologie de pratiques». L'expérience a lieu dans l'événement, et c'est seulement après l'événement que sujet et objet se séparent et se différencient. Massumi décrit ce processus : « The 'object', along with the concerned 'subjects', figure as differential poles integrating into a unity of movement. (...) In the aftermath of the event, the unity resolves back into differentials, and the movement continues (...). » (Massumi, 2011, p. 30) Ce qui unit ces différents processus dans l'événement ce sont les relations. Pendant la durée de l'événement, les sujets et objets en devenir se confondent dans les relations qu'ils composent. Or, pour James, les relations sont partie intégrante de l'expérience :

To be radical, an empiricism must neither admit into its constructions any element that is not directly experienced, nor exclude from them any element that is directly experienced. For such a philosophy, *the relations that connect experiences must themselves be experienced relations, and any kind of relation experienced must be accounted as 'real' as anything else in the system.* (James, 1912, p. 42)

Nous revenons ainsi à la première contrainte abordée par Debaïse et Stengers (2016), celle de l'expérience. Et c'est à partir de l'expérience que je propose de réfléchir sur l'interactivité. Comment comprendre alors une expérience interactive ? Est-ce que l'interactivité peut contribuer à élargir le champ relationnel d'un documentaire ?

Selon la perspective de Massumi, il est important de marquer une différence entre les notions de relation et d'interactivité : « I use the word interactivity to designate an instrumentally contracted dynamic form that tends to shrink to the parameters of its objectively embodied instrumental function. » (Massumi, 2011, p. 46) Massumi

reprend donc le terme d'interactivité pour nommer la tendance utilitaire ou fonctionnelle d'une image. Dans le cas du documentaire interactif, cette tendance se vérifie par exemple dans l'ajout de fonctionnalités, du « *gimmik* », qui établissent un mode d'interactivité instrumentalisé où la participation est souvent réduite à des gestes répétés prédéterminés, sans créer de l'engagement avec le contenu ou le dispositif et sans qu'il y ait émergence de nouvelles formes. La deuxième tendance, Massumi l'appelle *relation* : « I use the word relation to refer to the full spectrum of vitality that the dynamic form really includes, potentially, abstractly self-expressed in its semblance. » (Massumi, 2011, p. 46) Ces deux tendances existent simultanément dans toute expérience, ce sont deux pôles de l'événement. Massumi remarque ainsi que ces deux tendances ne s'excluent pas, elles cohabitent toute expérience, mais sont mises en évidence selon le contexte. Dans la vie quotidienne, typiquement il s'agira de mettre au premier plan une perception utilitaire, de l'ordre de la première tendance que nous avons identifié (*interactivity*). Dans une expérience artistique, c'est la deuxième tendance (*relation*) qui est souvent mise au premier plan. Ainsi, il défend que dans les créations interactives on essaie de sortir des systèmes prédéterminés d'action-réaction pour mettre l'accent sur la dimension relationnelle de l'expérience. Prenons un exemple dans le champ du documentaire interactif pour situer cette notion d'expérience interactive.

*Dada-Data*²⁷, une œuvre interactive de David Dufresne et Anita Hugi, célèbre l'anniversaire de cent ans du mouvement Dada (en 2016). Le webdocumentaire est composé de 6 volets et d'un « musée virtuel » du mouvement Dada, le *Dada Dépôt*. Les 6 volets sont : *Dada Block*, un plug-in pour des navigateurs web qui remplace les publicités par des affiches de Dada ; *Readymade connecté*, une installation d'imprimante 3D qui imprime des objets d'inspiration dadaïste, comme un urinaire dans le style de la « Fontaine » de Marcel Duchamp, et offre ces objets aux

²⁷ <http://dada-data.net/fr/hub>

interacteurs ; *Dada Gafa* accompagne les interacteurs dans des actions sur leurs appareils électroniques personnels pour limiter l'utilisation des GAFAs (les entreprises qui dominent internet et les outils numériques — le nom GAFAs vient des premières lettres de quatre de ces entreprises : Google, Apple, Facebook et Amazon); *Dada Gram* expose des collages dadaïstes et permet aux interacteurs de soumettre des images par l'application Instagram ; *Tweet Poésie* s'inspire de la poésie simultanée d'Hugo Ball pour proposer aux interacteurs d'envoyer des tweets pour composer des poèmes ; *Manifesto* est un « manifeste dada digital » qui a été créé lors d'un événement dans le contexte du Dada-Data qui a eu lieu en 2016, à Zurich, au Cabaret Voltaire. D'autres événements similaires ont eu lieu à Paris, à Montréal et à Tokyo pour élaborer d'autres manifestes (j'aborderai ces événements plus loin). Enfin, le *Dada Dépôt* est un musée virtuel décrit sur sa page d'accueil comme « un coffre à munitions, un anti-musée dada, un jardin fragmenté, où les œuvres sont disposées dans un joyeux hasard. »²⁸ Il est intéressant de remarquer que dans la page d'accueil de ce webdocumentaire on retrouve une citation de l'écrivain dadaïste Richard Huelsenbeck : « On ne peut comprendre Dada. Il faut en faire l'expérience ». Cela confirme la pertinence de la proposition d'aborder ce documentaire interactif par la perspective de l'expérience, car l'intention des auteurs, tel qu'affirmé par cette citation, est de créer des occasions d'expérience plutôt que de proposer une approche pédagogique ou informative du mouvement Dada.

Les événements où les manifestes ont été produits sont appelés *hacktions Dada* par les créateurs. Ce sont des « exercices Dada interactifs en rafale », autrement dit, des événements médiés par Dufresne et Hugli dans un esprit dadaïste. J'ai eu l'occasion de participer à un des *hacktions Dada* lors du Festival du Nouveau Cinéma de 2016²⁹. Le

²⁸ <http://dada-data.net/fr/depot>

²⁹ Publication sur Dada-Data concernant le hacktion : « Lors du 45e Festival du Nouveau Cinéma. Tout commence le 8 octobre avec LE GRAND DADA MANIFESTO, un hackathon exceptionnel auquel absolument tout le monde est convié. Pendant 6 heures, codeurs, écrivains, artistes, éditrices, plombiers, designers, étudiants, banquiers, apatrides, lanceurs d'alertes et autres messagers d'art se réuniront à la Chaufferie. Leur défi : fabriquer ensemble un Manifeste Dada Digital. Venez participer à

hackathon a duré plusieurs heures et nous étions une vingtaine de participants. Nous avons tous assisté à des performances à inspiration dadaïste pour ensuite nous regrouper dans des équipes pour travailler à créer des propositions autour des thématiques proposées. Il était question principalement de proposer des réinterprétations de l'internet dans une approche dadaïste, autrement dit, réfléchir à des propositions qui pourraient apporter la posture dadaïste au contexte de l'internet. Par exemple, une des propositions créées lors de ce *hacktion* était une application pour remplacer le mot *Trump* à chaque fois qu'il s'affichait dans les navigateurs d'internet par un gif (image animée) sélectionné par l'équipe (une idée proche du *Dada-block*).

Dans le cas de ce *hacktion*, l'expérience est ainsi un processus partagé, ou une « écologie de pratiques », dans les termes de Manning. Le fait d'intégrer le public dans le processus de création du documentaire interactif provoque un court-circuit dans les étapes typiques de la réalisation d'un documentaire et transforme la notion habituelle de ce qui configure l'expérience que fait le public d'une œuvre documentaire. Typiquement, un documentaire audiovisuel commence par la phase de développement (recherche, écriture, repérages, etc.), passe ensuite à la production (tournage), suivie de la postproduction et de la diffusion. Il n'existe pas d'itération de ces étapes, dans le sens où une fois qu'on arrive à l'étape de diffusion du documentaire, son contenu demeure inchangé et il n'est pas question de revenir à une étape antérieure du processus de création. Selon ce schéma, la création (par l'auteur) et l'expérience (par le public) sont des moments séparés, chacun dans un bout du processus qui va du développement à la diffusion. Dans cette perspective, l'expérience du documentaire serait ainsi la dernière étape d'une séquence qui commence au développement et se termine dans la

la création de cette œuvre durant la journée du 8 octobre. Le fruit de ces réflexions collaboratives sera présenté au grand public à l'Agora Hydro-Québec du Cœur des sciences de l'UQÀM, dès 21 h, au cours d'un cabaret imprévisible qui promet de s'achever tard dans la nuit (concert surprise !). INSCRIVEZ-VOUS GRATUITEMENT ICI et gagnez la chance de rencontrer la fine et folle équipe d'Akufen, derrière les manettes de Dada-Data ! » (<http://dada-data.net/fr/news>, le 29 août 2016, consulté le 18 mai 2018.)

diffusion. Néanmoins, lorsque l'on considère l'aspect participatif des documentaires interactifs, ce schéma ne fonctionne plus. Très souvent, la participation du public dans la création du contenu se fait au même temps que la diffusion. Dans le cas de *Dada-Data*, par exemple, expérience et processus de création sont entremêlés dans l'événement du *hacktion*, par exemple. Le processus de création qui a été mis en place dans ce *hacktion* se caractérise par les relations qui l'ont composé pendant que l'événement se déroulait. Lorsque l'on pense l'expérience en termes d'événement, il devient impossible de distinguer un sujet d'un objet, ou même un sujet d'un autre sujet. Ainsi, pendant l'actualisation de l'événement, l'expérience est commune et partagée, elle est une « écologie de pratiques ». Une fois l'événement fini, chaque participant repart chez soi, l'unité formée entre les différents sujets et objets dans l'expérience du *hackaction* se défait, et chaque participant continue à intégrer d'autres événements et à former d'autres unités relationnelles dans d'autres contextes. Il est important de souligner que j'ai pris un exemple où la compréhension de la notion d'événement est évidente (car il s'agit d'une mise en situation temporaire), mais toute expérience de documentaire interactif est comprise ici comme un événement.

On remarque alors que la notion de participation est très liée à celle d'interactivité. Une manière d'aborder l'expérience interactive serait ainsi par la perspective des conditions relationnelles créées par les modes de participation proposés. Autrement dit, quelles sont les conditions mises en place par les différents modes de participation pour l'émergence de nouvelles possibilités relationnelles dans une expérience interactive ? Dans le livre *Participation Is Risky*, Huybrechts et al. rappellent que la notion de participation change selon le domaine où elle se matérialise. Ils regardent le contexte des médias numériques et réaffirment le rôle important de l'interactivité dans la construction de la notion actuelle de participation : « Authors such as Laurel and Murray regard the advent of digital (and interactive) medias in the late 20th century as a new great opportunity for participation. » (Huybrechts, 2014, p. 22) Néanmoins, les auteurs remettent en question les modes de cette participation numérique, et formulent

la notion d'une *pseudo-participation*, dans la mesure où l'action du participant se limite souvent à cliquer presque par inadvertance. On remarque une proximité dans les réflexions de Huybrechts et al. et de Massumi concernant les limites de la notion d'interactivité. Reprenons les mots de Massumi : « Interaction is just that: a going back and forth between actions, largely reduced to instrumental function. » (Massumi, 2011, p. 46) Or, Huybrechts et al. arrivent aux mêmes deux tendances de la participation évoquées par Massumi, son côté utilitaire (les fonctionnalités d'un dispositif interactif), et sa dimension relationnelle :

The outcome of a design project is both a device and a thing. It can be seen as a device, the embodiment of the object of design, providing users with access to some function (...). But the tangible outcome of the design process is also a thing, modifying the space of interaction, ready for unexpected use, and opening new ways of thinking and behaving. (Huybrechts, 2014, p. 34)

Ils proposent de ne pas voir la création participative en tant qu'objet, mais de reprendre plutôt la notion de « Chose » (*Thing*) tel que développée par Bruno Latour³⁰ et ainsi considérer les ouvertures et réarrangements affectifs, au-delà des fonctionnalités qu'elle propose :

An entire series of objectifying tendencies that push the thing to becoming an artefact is therefore counterbalanced by an entire series of thinging emergences that instead revert the tendency and bring the artefact and its associated elements to a rediscussion of their interdependencies and a reactivation of some of their virtualities and multiplicities. (Huybrechts, 2014, p. 35)

Cette perspective proposée par Huybrechts et al. résonne avec la compréhension de l'aspect virtuel tel que formulé par Massumi dans le contexte d'œuvres interactives. Il reprend le concept de *semblance* développé dans la théorie esthétique de Susanne

³⁰ Latour revient sur l'étymologie du mot *Thing* ou *Ding* pour récupérer son sens original d'un lieu d'assemblage de divergences : « Thus, long before designating an object thrown out of the political sphere and standing there objectively and independently, the Ding or Thing has for many centuries meant the issue that brings people together because it divides them. » (Latour, 2005, p. 13) Il remarque que la même étymologie s'applique sur la racine latine du mot.

Langer. Il faut remarquer que dans la notion de *semblance* il n'est pas question de mimesis, l'image artistique n'étant pas la représentation d'un modèle, mais plutôt un mode d'accès aux modèles, une expérience perceptive virtuelle : « With every sight we see imperceptible qualities, we abstractly see potential, we implicitly see a life dynamic, we virtually live relation. » (Massumi, 2011, p. 43) Ainsi, Massumi propose de considérer l'expérience d'une œuvre comme un événement, et de défaire l'idée de l'expérience esthétique comme la perception d'un objet par un sujet.

Reprenons un autre exemple dans Dada-Data pour observer comment ces tendances prennent forme dans le documentaire interactif. Dans Dada-Dépôt, un « musée aléatoire » en ligne du mouvement Dada, l'interacteur peut naviguer parmi des images, des textes, des vidéos, des biographies et d'autres éléments autour du mouvement Dada. La navigation se fait à l'aide des flèches du clavier ou avec l'utilisation de la souris. Le mode de participation dans cet exemple n'est pas le même de celui du *hackathon*, car l'interacteur n'agit pas dans la création du contenu, mais plutôt dans son exploration. L'aspect relationnel est ainsi transformé, il s'appuie sur les modes de rencontre entre l'interacteur et le contenu. La question pour l'équipe de réalisation devient ainsi : comment créer des conditions pour que la participation active des relations ? Dans le cas de Dada Dépôt, l'utilisation du hasard devient un outil pour provoquer des rencontres inattendues entre interacteur et contenu :

À l'image du mouvement Dada, pour qui l'intrusion du hasard dans la réalisation d'une œuvre était au cœur de son esthétique et reflet de son doute envers un monde qui s'entretient par des voies quasi démocratiques, notre DÉPÔT reprend les notions d'automatisme, de simultanéité et de hasard. Œuvres, slogans et interviews sont disposés de manière accidentelle et jamais identique.³¹

³¹ Site Behance de Dada-Data : <https://www.behance.net/gallery/34289495/DADA-DATA>, consulté le 18 mai 2018.

Les créateurs, fidèles à l'approche dadaïste, utilisent ainsi le hasard comme une manière de faire sortir l'interactivité de son aspect prédéterminée pour essayer d'établir des conditions pour la création de possibilités ouvertes de navigation. L'utilisation du hasard est ainsi une des stratégies utilisées pour éviter la prédictibilité de l'expérience interactive.

On remarque alors que l'ouverture à l'indétermination dans le cadre du documentaire interactif n'est pas une caractéristique attribuée uniquement à l'interactivité, mais plutôt une possibilité qui émerge de la relation entre l'interacteur et le documentaire en fonction des conditions mises en place dans l'image. Cette observation correspond à la remarque de Debaise et Stengers à propos de l'émotion dans une œuvre d'art : « La question de savoir si le spectacle est émouvant ou si c'est moi qui, ému, le ressent tel quel, sera toujours, dans sa généralité, une mauvaise question (...). » (Debaise & Stengers, 2016, p. 86) En effet, l'émotion ne peut pas être attribuée ni à l'œuvre d'art, ni au spectateur, elle émerge de la rencontre entre les deux. Or, du moment où l'on admet que l'expérience d'une œuvre (ou dans le cas qui nous intéresse, d'un documentaire interactif) est un événement qui a lieu lors de la rencontre entre interacteur et documentaire, il n'est plus possible d'isoler l'indétermination comme une caractéristique de l'interactivité, car c'est de la rencontre qu'émergent l'émotion et l'indétermination. Il s'agit plutôt de créer des conditions pour que l'indétermination se manifeste pendant l'expérience. Dans la prochaine section, il s'agira d'explorer plus en détail la rencontre entre l'interacteur et le documentaire.

2.2. Centre d'indétermination et image-interaction

Le projet *Rider Spoke* (2007), du collectif Britannique Blast Theory, est un documentaire interactif, dans lequel le participant est invité à faire une balade à vélo dans l'espace urbain (avec son propre vélo ou un fourni par le projet). Un dispositif électronique de géolocalisation, avec une interface à être manipulée par l'interacteur, est attaché au guidon du vélo et relié à un casque d'écoute et un microphone. Le

participant doit se promener avec son vélo et écouter un enregistrement. À un moment donné, il sera invité à trouver un endroit pour se poser. Une fois, l'endroit choisi, le dispositif audio jouera une des questions préparées au préalable par les artistes, et le participant devra enregistrer sa réponse à la question. L'interface montre la localisation du participant sur une carte, ainsi que les endroits choisis par les autres participants pour faire leurs propres enregistrements. Il est possible de visualiser les endroits choisis par ceux qui ont déjà soumis leurs participations et, sur cette même place, d'écouter leurs enregistrements. Comme il est précisé dans la description du projet : « *The recordings that people make are only available in this context: played to a player, alone, in the place where they were recorded.* »³² Ainsi, dans cette œuvre, l'interactivité est un caractère essentiel, car le contenu de l'œuvre (si nous restreignons le contenu aux enregistrements) est accessible seulement à ceux qui décident d'y prendre part physiquement. Il n'y a pas d'objet à voir, l'œuvre existe seulement dans l'expérience de ceux qui prennent un vélo et interagissent avec le dispositif. À partir de cet exemple, il sera question d'essayer de comprendre quelle est la nature de l'image dans une œuvre interactive.

Bergson, dans le célèbre premier chapitre de *Matière et Mémoire*, a proposé un concept d'image qui a été souvent repris, notamment par Deleuze dans ses études sur le cinéma, où il associe image, matière et mouvement. Comme Bergson l'a précisé, il considère la matière avant la dissociation « entre son existence et son apparence. » (Bergson, 2008 [1939], p. 2) Dans cette perspective, si l'existence de la matière est son mouvement et l'apparence de la matière est son image, considérant ces aspects avant leur séparation, nous arrivons à cette triple égalité : matière = mouvement = image. Deleuze l'avait bien remarqué : « Cet en-soi de l'image, c'est la matière : non pas quelque chose qui serait caché derrière l'image, mais au contraire l'identité absolue de

³² http://www.blasttheory.co.uk/bt/work_rider_spoke.html

l'image et du mouvement. » (Deleuze, 1983, pp. 86-87) Deleuze appellera cette image, l'image-mouvement.

Les images agissent les unes sur les autres, et dans cette interaction, l'image du corps, en tant qu'organisme vivant, a une caractéristique particulière dans sa manière de réagir aux autres images, elle introduit un intervalle entre l'action et la réaction, elle rapporte les images aux actions possibles de son propre corps. « *J'appelle matière l'ensemble des images, et perception de la matière ces mêmes images rapportées à l'action possible d'une certaine image déterminée, mon corps.* » (Bergson, 2008 [1939], p. 17, italique original) Or, dans un documentaire interactif, les images sont rapportées au corps du participant, d'où la nécessité de comprendre les particularités de cette image soumise à un corps-vivant humain. Or, pour Bergson, un corps vivant c'est un « réservoir d'indétermination », car il élargit de manière indéterminée et imprévisible le champ des réactions possibles : « Comme nous le faisons pressentir depuis le début de ce travail, le rôle de la vie est d'insérer de l'indétermination dans la matière. Indéterminées, je veux dire imprévisibles, sont les formes qu'elle crée au fur et à mesure de son évolution. » (Bergson, 2008 [1939], p. 137) Le système nerveux agit de manière à créer un intervalle entre l'action perçue et la réaction, et dans cet intervalle repose la possibilité de la création d'une forme nouvelle : « Chaque progrès des centres nerveux, en donnant à l'organisme le choix entre un plus grand nombre d'actions, lancerait un appel aux virtualités capables d'entourer le réel, desserrerait ainsi l'étau, et laisserait plus librement passer la conscience. » (Bergson, 2008 [1939], p. 180)

Deleuze reprend cette notion et une des tâches qu'il s'attribue dans les deux tomes de *Cinéma* est d'étudier ce qui arrive à l'image-mouvement cinématographique lorsqu'elle est rapportée à un centre d'indétermination : « En fin de compte, c'est en trois sortes d'images que les images-mouvement se divisent quand on les rapporte à un centre d'indétermination comme à une image spéciale : images-perception, images-action, images-affection. » (Deleuze, 1983, p. 97). Mais Deleuze nous met en garde que d'autres types d'images existent aussi. L'exercice que je propose de réaliser dans

cette section est analogue à celui de Deleuze, mais je m'intéresse notamment aux types d'images créés une fois que l'image est rapportée à un centre d'indétermination à travers un dispositif interactif (et non un dispositif cinématographique).

Il serait pertinent à ce point de vérifier s'il est possible de parler d'une image rapportée à un centre d'indétermination dans le cas de l'œuvre de Blast Theory. Or, comme nous l'avons vu dans la description du projet, la création du contenu de l'œuvre se fait par la participation active de l'interacteur. Comme l'avait observé Chatonsky à propos des œuvres interactives : « Ces œuvres mobilisent explicitement le corps comme aucune autre. » (Chatonsky, 2004, p. 81) Dans *Rider Spoke*, par exemple, la participation demande que le corps du participant s'implique activement, par exemple pour pédaler un vélo, et ensuite l'œuvre répond à cette action du corps. Sandra Gaudenzi a aussi évoqué la dimension physique de l'interactivité dans le cadre du documentaire :

If linear documentary demands a cognitive participation from its viewers (often seen as interpretation) the interactive documentary adds the demand of some physical participation (decisions that translate in a physical act such as clicking, moving, speaking, tapping, etc.). (Gaudenzi, 2013a, p. 32)

Il est important de remarquer que l'activité cognitive implique également une dimension physique, car la cognition a lieu dans un corps et active des processus physiques. Mais ce que l'on veut démontrer ici est le fait que l'expérience interactive implique aussi le corps de l'interacteur, et que l'interacteur peut être considéré un centre d'indétermination, dans la mesure où ses réponses aux propositions du dispositif sont imprévisibles et indéterminées. L'image est donc rapportée à un nouveau centre d'indétermination dans les relations que le dispositif a mis en place. Nous pouvons conclure que les documentaires interactifs rapportent des images à des centres d'indétermination, mais pour produire quel type d'image ?

Selon Chatonsky, l'intervalle d'indétermination qui a lieu dans la perception d'une image est un aspect central des œuvres interactives. Cet intervalle n'est pas exclusif

aux œuvres interactives, mais en raison de la possibilité d'une plus grande implication du corps dans ce type d'œuvre, l'importance de l'intervalle s'y trouve intensifiée. Or, Deleuze avait déjà remarqué l'émergence de l'indétermination dans le passage de l'image-mouvement à l'image-temps. La crise de l'image-mouvement, évoquée déjà dans *Cinéma 1*, peut être résumée, selon Suzanne Hême de Lacotte (Lacotte, 2000), en quelques éléments qui, d'une forme générale, caractérisent la rupture avec l'unité de narration du cinéma classique : la fin du montage linéaire, l'éclatement de l'espace : « Les événements sont désormais dus au seul hasard, rien ne les fait plus tenir ensemble (...). » (Lacotte, 2000, p. 26) La crise de l'image-mouvement est causée par l'indétermination qui se manifeste à l'intérieur même de l'image, rompant ainsi sa linéarité et sa prévisibilité.

Les personnages de la nouvelle vague, par exemple, réagissent de forme indéterminée aux événements de la narration, ses actions, ses parcours ne répondent pas de manière évidente à une relation de cause à effet. Une technique narrative emblématique de cette période et qui représente l'attitude imprévisible des personnages est la balade : « La balade a lieu dans des espaces quelconques et elle est la manifestation même de la perte de l'image-action. Les personnages errent sans réagir à ce qui leur arrive. » (Lacotte, 2000, p. 26) L'inclusion du geste de flâner résume le passage d'une narration basée sur un système logique et rationnel (sensori-moteur, dans les mots de Deleuze) d'action-réaction de l'image-mouvement (notamment l'image-action) à une narration basée sur des réactions créatives, et donc indéterminées, des personnages de l'image-temps.

Dans la banalité quotidienne, l'image-action et même l'image-mouvement tendent à disparaître au profit des situations optiques pures, mais celles-ci découvrent des liaisons d'un nouveau type, qui ne sont plus sensori-motrices, et qui mettent les sens affranchis dans un rapport direct avec le temps, avec la pensée. » (Deleuze, 1985, p. 28)

La rupture définitive avec le système de l'image-mouvement se donne avec l'image-temps.

J'argumente ici qu'un processus de même ordre du passage de l'image-action à l'image-temps s'est produit avec les œuvres interactives. Comme nous l'avons constaté avec Chatonsky, la participation de l'interacteur peut augmenter l'indétermination d'une image interactive, de la même manière que les aspects formels de la nouvelle vague ont fait augmenter l'indétermination au sein des images-action cinématographiques. Avec les œuvres interactives (dont le documentaire interactif) nous passons alors à un autre type d'image, une image-interaction, caractérisée aussi par sa plus grande indétermination par rapport à l'image-action. *Rider Spoke* est un exemple frappant de cette indétermination de l'image-interaction. Une grande partie de l'image est créée par les choix de chaque interacteur, ce qui rend impossible toute prévision sur la constitution même de l'image avant l'expérience de l'interacteur. Les choix de parcours, des endroits visités, les enregistrements, tous ces éléments qui formeront l'image du documentaire sont créés par les actions des participants. L'œuvre ne cesse pas de changer, avec chaque nouvelle contribution d'un participant. L'expérience de l'œuvre ne sera donc jamais la même, ni du point de vue du participant (lui aussi aura changé) ni du point de vue du contenu proposé, elle est donc doublement imprévisible. L'image-interaction, comme un agencement (notion à être développée dans la prochaine section), n'existe que dans le mouvement, elle n'est jamais fixe, elle n'est jamais donnée en définitif.

La question que l'on doit se poser devient ainsi celle de savoir si toute forme d'interactivité donne lieu à une image-interaction. Or, dans les sections précédentes il a été question des limites de l'interactivité lors que celle-ci fonctionne sur des bases prévisibles d'action-réaction, est-ce que l'on peut parler d'image-interaction dans ces cas ? Adrian Miles (2014), dans une démarche proche de ce que l'on essaie d'accomplir dans ce chapitre, a remarqué que quand un documentaire interactif met l'accent sur un mode de navigation basé sur des schémas d'action-réaction, c'est de l'image-action dont il s'agit : « the action image is evident in those interactive works that emphasize the immediacy of responding to a situation, as is seen in computer games such as First

Person Shooters and ‘twitch’ games. » (Miles, 2014, p. 79) En effet, dans la section précédente il a été question des différences établies par Massumi entre interactivité et relation, on constate alors que quand l’expérience interactive est largement basée sur des enchainements prévisibles d’action et de réaction, il s’agit principalement d’image-action. Miles remarque néanmoins que le documentaire interactif a le potentiel pour élargir le champ des possibles, échappant ainsi à la détermination du schéma sensori-moteur : « Documentary makers and audiences lie at the intersection of a series of vectors of indeterminacy between the machine’s procedural logic, networked affordances, and narrativisation, all literally rendered via an interface. » (Miles, 2014, p. 80) Selon Miles, pour dépasser les limites de l’image-action, le documentaire interactif doit accueillir l’indétermination qui s’installe dans l’intervalle entre la perception et la décision d’action. Nous avons déjà quelques particularités de l’image-interaction par rapport à l’image-temps des œuvres linéaires : elle peut impliquer le corps d’une manière plus ample, elle peut ajouter de l’indétermination à l’image et elle est en constant mouvement.

Un autre aspect important qui marque l’image-interaction est le fait que l’expérience de ce type d’image peut se faire comme une situation vécue de l’intérieur, comme c’est le cas dans *Rider Spoke*, l’interacteur doit intégrer l’image pour pouvoir en faire l’expérience. Bergson nous avait mis en garde, selon lui, parmi toutes les images, il y en a une en particulier dont nous avons une expérience privilégiée, celle de notre propre corps :

Toutes ces images agissent et réagissent les unes sur les autres (...). Pourtant il en est une qui tranche sur toutes les autres en ce que je ne la connais pas seulement du dehors par des perceptions, mais aussi du dedans par des affections : c’est mon corps. (Bergson, 2008 [1939], p. 11)

Le corps propre à chacun constitue son image privilégiée, et se distingue des autres images, car il offre la possibilité d’une expérience sans distancement.

Malgré la porosité des frontières de l'individualité, Bergson détermine deux ordres, ou deux systèmes d'images : celui de mon propre corps et celui des corps environnants :

D'où vient que les mêmes images peuvent entrer à la fois dans deux systèmes différents, l'un où chaque image varie pour elle-même et dans la mesure bien définie où elle subit l'action réelle des images environnantes, l'autre où toutes varient pour une seule, et dans la mesure variable où elles réfléchissent l'action possible de cette image privilégiée ? (Bergson, 2008 [1939], p. 21, italique original)

La question qui découle est celle de savoir si, dans une image-interaction, étant donné que l'expérience interactive se fait lorsque l'interacteur fait corps avec l'image, ne change-t-elle pas de système d'image par rapport à une expérience qui se fait dans une relation d'extériorité, tel qu'un film au cinéma ? Dans une œuvre interactive, serions-nous dans « La connaissance-conscience par opposition à la connaissance-observation » ? (Ruyer, 2012, p. 109) Cette forme de connaissance-conscience, Raymond Ruyer l'appelle le *survol absolu*.

Ruyer remarque que l'expérience que je fais de mon propre corps a une dimension de moins par rapport à l'expérience que je fais du monde extérieur, je ne peux pas m'extraire à moi-même pour m'observer d'une autre perspective (l'illusion du troisième œil) : « Mon champ visuel se voit nécessairement lui-même par "survol absolu", ou "non dimensionnel". Il se survole sans prendre de distance le long d'une dimension perpendiculaire. » (Ruyer, 2012, p. 115) Ce survol absolu est, selon Ruyer, l'activité de la conscience primaire, ou la conscience organique, le *self-enjoyment*. La conscience secondaire, étant pour Ruyer l'activité sensorielle cérébrale. Dans le survol absolu, la conscience ne se retire pas de l'étendue pour pouvoir l'observer, cela c'est le typique de la conscience secondaire. Au contraire, dans le survol absolu la conscience saisit simultanément toutes les dimensions de la surface depuis son intérieur. Comme l'a résumé Ruyer : « (...) il n'y a au fond qu'un seul mode de conscience : la conscience primaire, forme en soi de tout organisme et ne faisant un avec la vie. La conscience seconde, sensorielle, est la conscience primaire des aires

cérébrales. » (Ruyer, 2012, p. 116) Or, la différence entre la conscience primaire et secondaire établie par Ruyer nous semble très proche des deux systèmes d'image que Bergson nous avait proposés.

Revenons maintenant à l'image-interaction pour demander : s'agit-il, dans une image-interaction, d'un survol absolu ? Sommes-nous dans la conscience primaire lors que l'on fait l'expérience d'une image-interaction ? De prime abord, il faut constater que la conscience secondaire participe aussi à l'image-interaction dans la mesure où le participant perçoit l'interface par une activité sensorielle pour ensuite interagir avec elle. Dans l'exemple de *Rider Spoke*, le participant doit observer son ordinateur de bord, écouter les enregistrements, parler au microphone, enfin, exécuter diverses fonctions sensori-motrices. Mais nous avons aussi remarqué que le participant contribue à la création de l'image-interaction à partir de l'intérieur même de l'image, le participant doit intégrer l'image pour pouvoir interagir avec elle. Pour faire l'expérience d'une image-interaction, au lieu de prendre de la distance pour mieux observer l'œuvre (comme dans le cas des documentaires linéaires), le participant doit faire le mouvement contraire et intégrer l'œuvre, soustrayant ainsi toute différence de dimension qui pourrait exister entre lui et l'image. Il est possible de remarquer qu'il s'agit ici du même mouvement du survol absolu. Nous arrivons donc à la conclusion que l'image-interaction agit autant dans la conscience primaire que secondaire, elle opère avec les deux systèmes d'image bergsoniens, ou les deux tendances de la connaissance (observation et conscience) de Ruyer. Comme on le verra souvent dans cette thèse, à chaque fois qu'il est question de dualisme, la réponse vient par la multiplicité et l'indétermination. Dans la suite de ce chapitre, la notion d'agencement sera évoquée pour aborder cette intégration entre interacteur et image.

2.3. Le documentaire interactif comme agencement

*Highrise*³³ (Katherine Cizek, 2009-2015) est un documentaire interactif produit par l'Office national du film du Canada. La date du documentaire soulève déjà une des caractéristiques fondamentales du documentaire interactif : il n'est pas possible de se référer au documentaire interactif comme un objet fini, une œuvre fixe à être diffusée, il est souvent en transformation permanente. Dans le cas de *Highrise* cela est encore plus évident, car le documentaire comprend plusieurs projets, développés au cours de quelques années. Ces différents projets s'articulent autour des questions sur la vie dans les grandes tours (*highrises*), logement typique des grands centres urbains.

Concentrons-nous sur le dernier volet de *Highrise*, *One Millionth Tower*³⁴, documentaire accessible par un site web. Dans un premier temps, le documentaire propose au public le choix sur la manière de le visiter, soit regarder certaines sections sans interaction (comme un documentaire linéaire), soit l'explorer de forme interactive. Dans le mode interactif, deux sections sont proposées : dans la première, il est possible de visiter un paysage virtuel créé à partir de croquis et de projets architecturaux faits en collaboration avec une communauté qui vit dans un grand centre d'habitation de Toronto, dans une perspective de revitaliser leur quartier. Dans la présentation du projet sur son site web nous lisons : « The project is a concrete result of a community collaboration between residents, architects, documentarians and animators to reimagine the particular spaces around these particular highrises. »³⁵ Dans la deuxième section interactive, à travers l'outil Google Earth, il est possible de visiter virtuellement des centres d'habitation à grandes tours dans divers pays du monde. L'interacteur peut

³³ <http://highrise.nfb.ca/>

³⁴ http://highrise.nfb.ca/onemillionthtower/1mt_webgl.php

³⁵ http://highrise.nfb.ca/onemillionthtower/1mt_webgl.php

aussi suggérer de nouvelles tours à être ajoutées sur la liste présentée par le documentaire.

Dans *One Millionth Tower*, il devient évident que la mobilité, le changement et la fluidité deviennent des caractéristiques fondamentales du documentaire interactif : chaque expérience du documentaire est unique, il n'existe pas de montage fixe et identique qui nous est présenté à chaque visualisation. Comme le soutient Gaudenzi, c'est dans la relation entre interacteur et dispositif, proportionnée par l'interactivité, que le documentaire prend forme : « (...) interactive documentaries are relational objects, artefacts that link technologies and subjects and that create themselves through such interaction. » (Gaudenzi, 2013a, p. 12) Cette transformation du documentaire, rendue possible par l'interactivité, d'un objet fixé vers un objet en transformation permanente pose le problème de la forme : comment parler de la forme dans le contexte d'un documentaire interactif, étant donné que chaque expérience donne une forme particulière au documentaire ? Dans les mots de Massumi : « How do you speak of form when there is the kind of openness of outcomes that we see in a lot of new media art, where participant response determines what exactly happens? » (Massumi, 2011, p. 40) Massumi reprend les réflexions de Susanne Langer pour argumenter qu'il n'existe pas de forme fixe dans l'art, même dans les pratiques artistiques traditionnelles comme la peinture figurative : il y a toujours un aspect dynamique dans les perceptions visuelles. D'après des propositions de Gaudenzi (2013a) et Adrian Miles (Miles, Sora, Fetzner, & Aston, 2017), nous allons considérer alors cet objet relationnel que constitue le documentaire interactif par le regard du concept d'*agencement*, développé par Gilles Deleuze et Félix Guattari (1980). Comme nous le verrons, l'aspect relationnel est nodal à ce concept : « An assemblage is a becoming that brings elements together. (...) [A] n assemblage is a collection of heterogeneous elements, but what is especially important is the relation between the elements. » (Wise in Stivale, 2011, pp. 92-93) Manuel DeLanda a développé à partir de ce concept une théorie de l'*agencement* (*assemblage theory*), appliquée notamment en sciences sociales (DeLanda, 2006). Dans le cadre de

cette thèse, nous allons utiliser la synthèse que DeLanda a réalisée du concept d'agencement, avec des références néanmoins à certains textes de Deleuze et Guattari, pour ensuite utiliser les outils proposés par la théorie de l'agencement dans le cadre du documentaire interactif.

Dans un premier temps, il est important de souligner que le concept d'agencement est récurrent dans les études de l'interaction dans les médias numériques. Wise a remarqué : « (...) how the concept of assemblage can help us better understand a particular issue in technology studies: the relation between the technological and the human. » (Wise in Stivale, 2011, p. 95) Selon Wise, les propriétés émergentes caractéristiques d'un agencement (traitées plus loin) sont appropriées à l'étude de l'interaction humain-ordinateur, car elles mettent en évidence les complexités de cette relation. Le travail de Goriunova sur les plateformes artistiques du web, par exemple, considère ces dernières par la perspective de l'agencement : « As a process of emergence, an art platform is an assemblage of structures, notes, codes, ideas, emails, decisions, projects, databases, excitement, humour, mundane work, and conflict. » (Goriunova, 2012, p. 3) Cette citation permet d'établir les premiers principes d'un agencement : un ensemble hétérogène (du code à l'humour) avec des propriétés émergentes. Comme il a déjà été mentionné, les relations entre les différents composants d'un agencement sont parmi les principaux aspects sur lesquels la théorie de l'agencement se concentre. DeLanda affirme que les agencements sont des « wholes whose properties emerge from the interaction³⁶ between parts. » (DeLanda, 2006, p. 5) Il est donc important de remarquer que les propriétés de l'ensemble qui constitue un agencement ne sont pas prédéfinies, elles émergent des relations établies par le propre agencement.

³⁶ Il ne s'agit pas pour DeLanda de l'interactivité numérique dont il est question dans ce travail, mais de l'interaction sociale. Notre travail ici sera justement de déplacer l'application sociale de la théorie de l'agencement de DeLanda vers une application communicationnelle de l'interactivité.

Le cas de *Highrise* en est un exemple, car les propriétés du documentaire dans son ensemble ne sont pas données avant l'accès de l'interacteur, c'est au moment de la visite que le documentaire prend sa forme, or, c'est l'interaction qui déclenche les propriétés d'un documentaire interactif. Certains des éléments qui constituent *One Millionth Tower* sont évidemment établis d'avance : le design de l'interface, le code, les vidéos, la musique, etc. Ce sont les parties composantes de l'agencement. Mais pour que cet agencement soit opérant, il lui faut une partie déterminante qui activera les relations entre toutes les autres parties, il s'agit de l'interacteur. L'interacteur, agissant dans le dispositif, transforme les potentialités du documentaire en propriétés effectives, c'est l'interacteur qui crée à sa manière les liens entre les différentes parties composantes du documentaire, et cette manière est unique et propre à chaque visite de chaque interacteur. Le paysage virtuel de *One Millionth Tower* prend forme différemment pour chacun de ses interacteurs : chaque visiteur choisit son parcours, sa ballade, décidant ainsi la manière d'actualiser les propriétés de cet agencement.

Or, dans un objet relationnel tel qu'un agencement, l'ensemble de ses propriétés n'est pas donné d'avance, et à travers l'interaction, certaines propriétés virtuelles peuvent s'actualiser. Les propriétés d'un agencement ne peuvent pas se réduire à la somme des propriétés de ses parties, de la même manière que ses parties ont aussi une certaine autonomie par rapport à l'agencement et peuvent se détacher d'un certain agencement pour en intégrer un autre. DeLanda écrit :

[T]he reason why the properties of a whole cannot be reduced to those of its parts is that they are the result not of an aggregation of the components' own properties but of the actual exercise of their capacities. These capacities do depend on a component's properties but cannot be reduced to them since they involve reference to the properties of other interacting entities. (DeLanda, 2006, p. 11)

Il conclut alors que les relations d'un agencement se définissent par des relations d'extériorité : « assemblages are made up of parts which are self-subsistent and articulated by relations of exteriority, so that a part may be detached and made a

component of another assemblage. » (DeLanda, 2006, p. 18) Encore une fois, le documentaire interactif *Highrise* répond aux caractéristiques d'un agencement. *One Millionth Tower* ne se définit pas uniquement par les propriétés de ses parties, comme nous l'avons déjà constaté, il y a des formes qui sont émergentes, propres à l'interaction. Par ailleurs, les parties sont aussi dissociables, certaines vidéos utilisées dans *One Millionth Tower* pourront intégrer, par exemple, un autre projet de *Highrise* (cet échange s'est produit entre d'autres projets de *Highrise*). Le propre code du documentaire est mis à disposition des interacteurs et peut être utilisé dans d'autres projets : « In the technology page you can find out how everything is made, see the code line by line and take it and make something of your own. »³⁷ Ainsi, les relations établies par le documentaire interactif sont principalement des relations d'extériorité, dans la mesure où ses parties ont une certaine autonomie par rapport à l'agencement et les propriétés de l'ensemble sont émergentes en fonction de l'interactivité.

Le concept d'agencement deleuzien a souvent été approché du dispositif foucauldien : « Nous avons pour tâche d'analyser des états mixtes, des agencements, ce que Foucault appelait des dispositifs. » (Deleuze, 2003b, p. 119) La mobilité de l'agencement est aussi une caractéristique du dispositif. Le dispositif, comme l'agencement, n'est pas une instance fixe et immuable, au contraire, il est composé de lignes mouvantes :

Les dispositifs ont donc pour composantes des lignes de visibilité, d'énonciation, des lignes de force, des lignes de subjectivation, des lignes de fêlure, de fissure, de fracture, qui toutes s'entrecroisent et s'emmêlent, et dont les unes redonnent les autres, ou en suscitent d'autres, à travers des variations ou même des mutations d'agencement. (Deleuze, 2003a, p. 320)

Pour Deleuze, une des conséquences principales de penser la mobilité du dispositif est de pouvoir insuffler en lui la nouveauté : « Tout dispositif se définit ainsi par sa teneur en nouveauté et créativité, qui marque en même temps sa capacité de se transformer,

³⁷ Branden Bratuhin, Directeur technique de *Highrise*, dans la vidéo *Highrise/One Millionth Tower—Open Technology*, http://youtu.be/byTZENR5L_Q, 2:39.

ou déjà de se fissurer au profit d'un dispositif de l'avenir (...). » (Deleuze, 2003a, p. 322) Il est possible de conclure ainsi qu'en considérant le documentaire interactif comme un agencement (ou un dispositif), la créativité et l'émergence du nouveau en deviennent un aspect essentiel. Dans cette perspective, la notion d'agencement, lorsque mise en relation avec le dispositif interactif, permet de comprendre les conditions de possibilité du dépassement d'une interaction probable vers des expériences des possibles.

Revenons à l'exemple de *One Millionth Tower* pour vérifier si, dans ce cas, cette conclusion se justifie. Un des objectifs de ce projet, nous l'avons vu, est celui de transformer la conception des centres d'habitation à grandes tours, et cela à partir d'une création collective qui essayait de reconstruire virtuellement le quartier à travers l'expérience de ceux qui y vivent, accompagnés des professionnels de l'urbanisme et de l'architecture. Par ailleurs, le documentaire interactif propose aussi aux interacteurs une nouvelle expérience de ce type de quartier. Ainsi, les interacteurs et les créateurs se transforment dans le dispositif du documentaire interactif, conclusion qui s'accorde avec celles de Deleuze : « Nous appartenons à ces dispositifs, et agissons en eux. (...) Dans tout dispositif, il faut distinguer ce que nous sommes (ce que nous ne sommes déjà plus), et ce que nous sommes en train de devenir (...). » (Deleuze, 2003a, p. 322) Deleuze remarque également que nous agissons dans le dispositif, et cela est central dans le contexte de l'interactivité, car la marge de manœuvre de l'interacteur peut être plus importante que celle du spectateur, autrement dit, l'interacteur agit de manière plus ample sur une œuvre interactive qu'un spectateur sur une œuvre linéaire.

À partir de l'exemple de *One Millionth Tower (Highrise)*, il a été possible de déterminer certains des éléments qui caractérisent un documentaire interactif. Il a été constaté également que le documentaire interactif est un objet relationnel, ce qui indique la possibilité de le considérer comme un agencement. Le choix d'étudier le documentaire interactif comme un agencement a permis de soulever quelques aspects importants, comme l'émergence des propriétés à travers l'interactivité et les relations

des parties à l'ensemble. Néanmoins, toutes ces observations concernent les aspects analogiques du documentaire interactif, car elles portent sur l'expérience du point de vue de l'agencement, et dans la relation établie par l'interactivité entre l'humain et le numérique. Or, la question qui se suit est alors celle de savoir comment se manifestent le hasard et l'indétermination dans l'aspect numérique du documentaire interactif? Les deux prochains chapitres se dédient à étudier cette question.

3. CHAPITRE III

LE NUMÉRIQUE : HASARD ET INFORMATION

Neural nets, probabilistic creatures living off programmed randomness, take a step in the direction of potential and the unpredictable while remaining digital. (...) Possibilization and probabilization will have given way to synaptic indeterminacy: the unprogrammed randomness of pure chance. (Massumi, 1998, p. 312)

Il est intéressant de remarquer que dans ce texte de 1998, Massumi défend déjà que le développement du numérique s'oriente aussi vers des systèmes indéterminés plutôt qu'à vouloir s'ancrer uniquement dans les modèles prédictifs. Cette volonté d'aller vers l'imprévisible semble à l'encontre de l'intérêt contemporain envers l'utilisation du numérique pour la prévision, tel que souligné déjà dans cet essai doctoral.

Comme le remarque Cathy O'Neil, les modèles prédictifs sont au cœur des différents systèmes informatiques qui utilisent des bases massives de données numériques :

Since 2008, we've heard less from algorithms in finance, and much more from big data algorithms. The target of this new generation of algorithms has been shifted from abstract markets to individuals. But the underlying functionality is the same: collect historical data about people, profiling their behaviour online, location, or answers to questionnaires, and use that massive data set to predict their future purchases, voting behaviour, or work ethic. (O'Neil, 2017, s.p.)

Dans son livre *Weapons of Math Destruction* (2016), Cathy O'Neil analyse un certain nombre d'algorithmes prédictifs et constate qu'ils reproduisent, ou augmentent, des injustices sociales, notamment en fonction des biais présents dans les données à la base

des modèles. O'Neil affirme : « A model, after all, is nothing more than an abstract representation of some process (...). Whether it's running in a computer program or in our head, the model takes what we know and uses it to predict responses in various situations. » (O'Neil, 2016) L'auteur étudie alors plusieurs situations où ces algorithmes prédictifs peuvent être utilisés, comme une analyse de crédit lors d'une demande de financement, ou lors de l'évaluation d'un CV dans une candidature d'emploi. Elle s'interroge sur le manque de transparence et de critères éthiques dans la construction des modèles. Le choix d'intégrer le code postal comme une variable du modèle porte préjudice, par exemple, aux résidents des quartiers pauvres, ce qui implique un déséquilibre d'opportunité, une injustice, entre les candidats. D'autant plus que les informations sur les variables intégrées aux algorithmes ne sont que très rarement communiquées aux personnes évaluées.

Par ailleurs, François Ewald (2012) réalise une histoire des pratiques prédictives et note que le développement de l'outil mathématique de la probabilité marque le passage des pratiques prédictives divinatoires aux pratiques probabilistes. « Le monde du traitement des "données massives" (Big Data) ouvre un nouvel âge, de nouvelles possibilités pour une prédiction possible : la prédiction singulière ou singularisée. » (Ewald, 2012) Il remarque également :

Ce modèle prédictif n'annule pas les précédents. Ils ne portent pas sur les mêmes objets. Il vient plutôt les compléter. Il propose une nouvelle strate de la prévisibilité. Il permet d'envisager une prédiction possible au niveau de l'individu, là même où, dans le modèle classique du risque, on pensait qu'elle était interdite. (Ewald, 2012, p. 71)

L'utilisation de modèles statistiques et des probabilités, comme le démontrent les auteurs, est un aspect important des outils numériques employés aujourd'hui dans le traitement et la transmission d'information.

Depuis la fin du 19e siècle, les probabilités et les statistiques ont une forte présence dans une grande variété de théories scientifiques. Comme l'affirme Donald Gillies,

« The twentieth century has seen a prodigious development of probability and statistics, and their increasing use in almost all fields of research. » (Gillies, 2000, p. i) Les théories dans les champs de l'information et de la communication ne font pas d'exception. Comme nous le verrons, depuis le travail de Claude Shannon (1948) la probabilité a une place centrale dans les théories de l'information. Shannon définit un modèle probabiliste pour mesurer quantitativement l'information où la quantité d'information est inversement proportionnelle à sa probabilité (on le verra plus en détail). Plus récemment, les modèles prédictifs de traitement massif de données (Big Data) et les algorithmes d'intelligence artificielle, notamment d'apprentissage automatique (*machine learning*, ou ML), largement utilisés aujourd'hui ancrent de manière encore plus forte la statistique et les probabilités dans la sphère des technologies numériques de l'information.

Dans cette étape de la thèse, nous allons nous intéresser aux notions de hasard et indétermination (des notions au cœur des modèles statistiques et probabilistes) et ses relations avec les systèmes numériques (comme le documentaire interactif). Comme nous verrons plus en détail, l'interactivité du documentaire qui intègre cette recherche-création, *On Chance*, est basée sur un agent conversationnel, ou *chatbot*, qui utilise des algorithmes d'apprentissage automatique. Or, dans les deux prochains chapitres, il sera question de construire des propositions pour la compréhension du rôle du hasard et de l'indétermination dans les systèmes numériques utilisés pour la création de *On Chance*. Il s'agira de revenir sur les principes du pragmatisme spéculatif pour étudier l'aspect numérique du problème du hasard et de l'indétermination dans le documentaire interactif. Les questions que je pose sont ainsi : comment comprendre, par la perspective du pragmatisme spéculatif, le rôle de la probabilité dans les algorithmes d'apprentissage automatique ? Comment la perspective du pragmatisme spéculatif peut contribuer à dépasser le dualisme des interprétations objectives (le hasard dans les *événements*) et subjectives (ou épistémologiques, hasard dans le manque de

connaissance ou d'*information*) de la probabilité, notamment dans le contexte des médias interactifs ?

Dans ce chapitre, il sera question de l'apparente contradiction entre une prise de posture non déterministe dans la recherche-cr ation en m dias interactifs et l'utilisation de mod les pr dictifs dans les dispositifs num riques utilis s dans la cr ation du documentaire interactif,   l'exemple du *chatbot*. Il s'agira d'observer les mani res comment les pr suppos s li s au hasard et   l'ind termination interviennent dans les pratiques de cr ation du documentaire interactif. Dans un premier moment, il sera question d'explorer la notion de hasard par le biais de la th orie des probabilit s, ou plut t, des interpr tations philosophiques de cette th orie. Comme on le verra, la notion de probabilit  est centrale dans les th ories de l'information et de l'information algorithmique, toutes deux mobilis es dans cette th se. La question qui nous int ressera dans les prochaines sections est celle d'identifier les interpr tations philosophiques propres   l'utilisation des probabilit s dans les technologies num riques d'information et de communication, pour ensuite consid rer quelles sont les possibilit s que cette r flexion apporte au d veloppement de l'interactivit  dans le documentaire, notamment dans l'utilisation d'algorithmes probabilistes des *chatbots*. Dans un premier temps, faisons alors un survol des diff rentes interpr tations de la th orie des probabilit s.

3.1. Interpr tations de la th orie des probabilit s

La formalisation math matique des probabilit s date du 17e si cle, ce qui est assez tardif compte tenu des connaissances math matiques basiques n cessaires   son d veloppement et de la pratique ancienne des jeux li s au hasard. Comme le souligne Ian Hacking dans son livre *The Emergence of Probability* (1975), depuis l'Antiquit  le hasard est pratiqu  dans les jeux³⁸ et utilis  dans certaines prises de d cision, « Yet

³⁸ Il fait r f rence aux jeux bas s sur l'anc tre des d s, l'*astragalus*, qui  tait fabriqu    partir d'un osselet   quatre faces d'un animal tel que le cheval ou la brebis.

theories of frequency, betting, randomness and probability appear only recently. No one knows why. We can canvass a few of the proposed answers. » (Hacking, 1975, p. 2) Hacking revient alors sur des propositions d'explication au développement tardif de la théorie des probabilités sans en être convaincu. Une raison souvent évoquée se base sur la contradiction entre la notion de hasard dans un événement aléatoire (à la base de la théorie de probabilité) et le déterminisme scientifique (modèle dominant dans les sciences au 17^e siècle). Or, dans une vision déterministe, le hasard ne peut pas exister, car tous les événements se déroulent dans une suite nécessaire de cause à effet, selon le principe de causalité (les mêmes causes produisent les mêmes effets). Néanmoins, Hacking affirme un certain anachronisme dans cette ligne d'argumentation, car le déterminisme, formulé au 17^e siècle, ne peut pas être considéré comme frein au développement de la théorie des probabilités lors des siècles précédents. Par ailleurs, le déterminisme résout la contradiction entre hasard et prédétermination avec une interprétation subjective des probabilités (développée plus loin), où le hasard serait uniquement dû à l'ignorance que l'on peut avoir des systèmes complexes.

Hacking propose alors de renverser la question, et plutôt que de chercher ce qui manquait au développement de la théorie des probabilités, établir les conditions d'émergence à la formulation des probabilités telle qu'on les conçoit depuis la formalisation par Pascal. Hacking se concentre alors sur la période entre 1650 et 1700, pendant laquelle, de forme indépendante, plusieurs mathématiciens (Blaise Pascal et Pierre de Fermat, Christiaan Huygens, Leibniz, John Graunt, John Wilkins) sont arrivés aux idées de base de la théorie des probabilités. Il soulève alors une caractéristique majeure de la théorie des probabilités, son côté dualiste :

It is notable that the probability that emerged so suddenly is Janus-faced. On the one side it is statistical, concerning itself with stochastic laws of chance processes. On the other side it is epistemological, dedicated to assessing reasonable degrees of belief in propositions quite devoid of statistical background. (Hacking, 1975, p. 12).

Hacking voit dans cette dualité la principale condition nécessaire à la formalisation de la probabilité, et remarque qu'elle marque les deux tendances d'interprétation philosophique de la théorie des probabilités jusqu'à nos jours.

Donald Gillies, dans son livre *Philosophical Theories of Probability*, identifie aussi ces deux tendances principales d'interprétation de la théorie des probabilités. Gillies observe que les outils mathématiques du calcul de probabilités ont des interprétations différentes et divergentes :

The theory of probability has a mathematical aspect and a foundational or philosophical aspect. There is a remarkable contrast between the two. While an almost complete consensus and agreement exists about the mathematics, there is a wide divergence of opinions about the philosophy. (Gillies, 2000, p. 1)

Les différences entre les postures philosophiques face à la théorie des probabilités sont très liées aux différentes conceptions du hasard : tantôt le hasard est un attribut des événements, tantôt le hasard est lié aux connaissances limitées d'un événement.

La première de ces tendances, nommée par Karl Popper, est l'interprétation *objective*. De manière très succincte, dans l'interprétation objective, le hasard est une qualité intrinsèque des événements et indépendante de la connaissance que l'on puisse en avoir. Par ailleurs, la deuxième lignée d'interprétation est *épistémique* ou *subjective*, car elle est associée au degré de compréhension rationnelle que l'on peut avoir d'un événement. Parmi les interprétations objectives, Gillies regroupe les interprétations propensionniste et fréquentiste. Les interprétations épistémiques sont la classique et la subjective (ou bayésienne). Le déterminisme laplacien (abordé dans le premier chapitre) est peut-être l'exemple le plus emblématique parmi les interprétations épistémiques. Dans nos jours, l'interprétation épistémique de la théorie des probabilités la plus répandue n'est pas le déterminisme laplacien, mais plutôt l'interprétation bayésienne (ou probabilité bayésienne). Comme on le verra plus loin, les modèles

probabilistes basés sur l'inférence bayésienne jouent un rôle important dans le domaine de l'apprentissage automatique.

Nate Silver, dans son livre *The signal and the noise : why so many predictions fail—but some don't* aborde l'interprétation bayésienne :

Bayes's theorem is nominally a mathematical formula. But it is really much more than that. It implies that we must think differently about our ideas—and how to test them. We must become more comfortable with probability and uncertainty. We must think more carefully about the assumptions and beliefs that we bring to a problem. (Silver, 2012, p. 17)

Il importe de noter que les inférences bayésiennes fonctionnent à partir d'une probabilité *a priori* (présupposé), à partir de laquelle de nouvelles probabilités sont calculées considérant les données observées. Il s'agit d'un processus de mise à jour des présupposés initiaux à la lumière de nouvelles observations qui a comme résultat une distribution des probabilités *posteriori*. Comme remarque Murphy (Murphy, 2012, p. 167), l'utilisation d'une probabilité *a priori* est à l'origine de nombreuses controverses, car cela implique souvent un aspect contextuel dans le calcul des présupposés.

Ces controverses sont aussi fondées sur les effets pratiques de la mise en application d'outils basés sur des inférences à partir des probabilités *a priori*, tel que PredPol (Mohler & Short, 2012), le système de police prédictive mentionné à l'introduction de cet essai. L'exemple de PredPol met en évidence les effets pratiques importants qui peuvent avoir les présupposés appliqués aux modèles prédictifs. Dans cette perspective, il est très important de poser les questions suivantes : qui définit les probabilités *a priori* ? et à partir de quelles données ? L'utilisation des probabilités *a priori* est souvent défendue avec l'argument suivant : « (...) nobody is a tabula rasa or blank state : all inference must be done conditional on certain assumptions about the world. Nevertheless, one might be interested in minimizing the impact of one's prior assumptions. » (Murphy, 2012, p. 167) Dans les approches bayésiennes, il est ainsi nécessaire de quantifier les présupposés et de reconnaître qu'ils peuvent varier selon le

contexte. Comme remarque Smith, même si le théorème de Bayes est souvent associé à l'interprétation subjective, il découle logiquement des fondements de la théorie des probabilités, ces fondements sont communs à toutes les interprétations. (Smith, 2016, p. 19)

Ce que je propose par la suite est un regard pragmatiste spéculatif envers l'interprétation de la théorie des probabilités en partant du travail de Charles S. Peirce autour des probabilités et du hasard. Les interprétations de la théorie des probabilités proposées par Peirce ont évolué au long de ses écrits, mais il s'est toujours placé dans le pôle objectif des interprétations des probabilités et s'est distancié du déterminisme. Les contributions de Peirce aux révisions du déterminisme se font sur plusieurs fronts : dans le domaine des probabilités, en logique, en philosophie, en philosophie des sciences, et aussi en méthodologie de la recherche scientifique. Il est important de remarquer que ses arguments sont reliés, même s'ils se réfèrent à différents champs épistémologiques. Je vais me concentrer ici sur les arguments en philosophie des sciences pour sa remise en question de l'interprétation déterministe et son concept de *would-be*, une contribution importante à l'interprétation de la théorie des probabilités.

La critique la plus directe de Peirce au déterminisme est peut-être dans le texte *The Doctrine of Necessity Examined* (1892), où il développe de nombreux arguments contre ce qu'il appelle *le nécessitarisme*. Il n'utilise pas dans ce texte le terme de déterminisme, mais la définition qu'il donne du nécessitarisme permet de le considérer comme une manifestation du déterminisme laplacien, les similarités avec le texte de Laplace étant frappantes :

(...) the state of things existing at any time, together with certain immutable laws, completely determine the state of things at every other time (...). Thus, given the state of the universe in the original nebula, and given the laws of mechanics, a sufficiently powerful mind could deduce from these data the precise form of every curlicue of every letter I am now writing. (Charles S. Peirce, 1957, p. 172)

L'expérience professionnelle de Peirce dans le *US Coast and Geodetic Survey*, notamment dans les tâches de mesures des expériences avec des pendules (pour déterminer la force gravitationnelle à différents endroits des États-Unis), a joué un rôle important dans le développement de ses théories. L'activité quotidienne et répétée de mesures qu'il devait effectuer l'a poussé à constater que les résultats obtenus ne correspondaient jamais exactement aux prévisions théoriques formelles, et pour Peirce, un pragmatiste, cela n'était pas négligeable. Il constate ainsi que les lois physiques décrites théoriquement se manifestent dans le monde avec des irrégularités, des différences : « To one who is behind the scenes (...) the idea of mathematical exactitude being demonstrated in the laboratory will appear simply ridiculous. » (Charles S. Peirce, 1957, pp. 177-178). Pour Peirce, les divergences entre les mesures et la description théorique démontrent que la réalité échappe aux prédictions des lois mécaniques, et du fait que les différences entre les lois et les objets observés soient imprévisibles, Peirce conclut l'impossibilité d'un système déterministe.

Dans un article précédent, *The Doctrine of Chances* (1878), où Peirce énonce son interprétation des probabilités, il commence par souligner l'importance de la mesure et du principe de la continuité : « It is not, however, so much from *counting* as from *measuring*, not so much from the conception of number as from that of continuous quantity, that the advantage of mathematical treatment comes. » (Charles S. Peirce, 1957, p. 58; italique original). Or, pour Peirce la continuité, « limitless intermediation », permet assez de précision pour comprendre les mesures comme des différences de degré. Selon Peirce, dans le cas des probabilités, il s'agit d'établir pour une situation donnée un degré entre 0 et 1 de probabilité qu'un événement survienne. Dans un autre article de la même année, *The Probability of Induction*, Peirce met en évidence les débats autour des différentes interprétations des probabilités :

The conception of probability as a matter of fact, i.e., as the proportion of times in which an occurrence of one kind is accompanied by an occurrence of another kind, is termed by Mr. Venn the materialistic view of the subject. But probability has often been regarded as being simply the degree

of belief which ought to attach to a proposition, (...). (Charles S. Peirce, 1957, p. 87)

Dans ce passage, la divergence entre l'interprétation de Laplace et celle de Peirce devient évidente : dans le modèle laplacien, les probabilités ne concernent pas l'événement en soi, mais seulement notre connaissance, *the degree of belief*, par rapport à sa réalisation. Les probabilités, du point de vue de l'interprétation subjectiviste (ou conceptualiste, dans les termes de Peirce), sont une mesure de notre ignorance quant à la réalisation (ou non) d'un événement. Néanmoins, dans l'interprétation fréquentiste (ou matérialiste, dans les termes de Peirce), les probabilités sont des propriétés du monde physique qui déterminent la fréquence de certains événements, ou la proportion dans laquelle un événement P ou un événement Q se produisent dans une série. Il est intéressant de constater que les outils et les règles mathématiques utilisés pour calculer les probabilités sont les mêmes chez Peirce et Laplace, néanmoins la différence d'interprétation semble reposer sur des visions du monde assez éloignées. Dans le cas de Laplace, les probabilités s'insèrent dans un monde déterministe, alors que pour Peirce, elles ouvrent la voie à une remise en question du nécessitarisme, car affirmer que la probabilité est une propriété de l'événement consiste à lui donner la possibilité d'échapper à la nécessité, d'être autrement, et ce de manière imprévisible.

En 1910, Peirce revient sur son article *The Doctrine of Chances* et y rajoute des notes. Il reformule alors son interprétation des probabilités :

I am, then, to define the meaning of the statement that the probability, that if a die be thrown from a dice box it will turn up a number divisible by three, is one-third. The statement means that the die has a certain 'would-be'; and to say that the die has a 'would-be' is to say that it has a property, quite analogous to any habit that a man might have. (Charles S. Peirce, 1957, p. 79)

Le concept de *would-be* comme une propriété de l'objet annonce un changement dans la pensée de Peirce, et une nouvelle tendance dans l'interprétation des probabilités : on passe du modèle fréquentiste au modèle propensionniste :

While Popper is often credited as being the pioneer of propensity interpretations, we already find the key idea in the writings of Peirce (1910, 79-80): A man's habit is a paradigmatic example of a disposition; according to Peirce the die's probability of landing 3 or 6 is an analogous disposition. (Hájek, 2012, sans page)

Les textes de Peirce moins axés sur les mathématiques ont beaucoup de résonances avec son interprétation des probabilités. Dans ses notes pour un projet de livre, *A Guess at the Riddle* (1887-1888), les notions de hasard et d'indétermination jouent un rôle majeur, notamment dans le système de triade qu'il développe pour le monde physique : « (...) three elements are active in the world, first chance; second, law; and third, habit-taking. » (Charles Sanders Peirce, 1992, p. 277). Il commence par constater les « discrepancies between observations and the theory » (1992, p. 274), des différences que Peirce a observées tout au long de sa vie professionnelle en tant que mesureur, et il questionne ensuite l'idée que tous les éléments soient soumis aux mêmes lois : « Strange to say, there are many people who will have a difficulty in conceiving of an element of lawlessness in the universe, (...). » (1992, p. 274). Après avoir cité quelques philosophes de l'Antiquité (Aristote, Démocrite et Lucrèce) qui utilisent le hasard dans leurs systèmes, Peirce conclut que : « Now the only thing that the inference from experience can ever teach us is the approximate value of a ratio. (...) At this rate, every proposition which we can be entitled to make about the real world must be an approximate one. » (1992, p. 274). Pour Peirce, l'observation permet d'identifier une tendance des événements qui est exprimée par un ratio, ce qui n'est sans rapport à son interprétation des probabilités (aussi exprimées en ratio), et nous renvoie à une vision probabiliste de la détermination, ou une loi-tendance, une propension, *a would-be*. Comme mentionné plus haut, l'interprétation de Peirce associe les probabilités à l'idée d'habitude :

Uniformities in the modes of action of things have come about by their taking habits. (...) We look back toward a point in the infinitely distant past when there was no law but mere indeterminacy. We look forward to a point in the infinitely distant future when there will be no indeterminacy. (Charles Sanders Peirce, 1992, p. 277).

Ainsi, la loi est une prise d'habitude formée à partir de l'indétermination, et selon Peirce, tous les éléments (« every conceivable real object ») ont une tendance à prendre des habitudes et à trouver, avec le temps, une forme de régularité.

Karl Popper développe aussi une interprétation propensionniste de la théorie des probabilités. Néanmoins, il mettra l'accent sur les conditions de l'événement pour lequel on souhaite établir une probabilité : « we have to visualise the conditions as endowed with a tendency or disposition, or propensity, to produce sequences whose frequencies are equal to the probabilities; which is precisely what the propensity interpretation asserts. » (Popper in Gillies, 2000, p. 115-116) Gillies souligne que cela pose une différence majeure entre le concept de *would-be* de Peirce, vu comme une qualité des objets, et la propension selon Popper qui vient des conditions de la production de l'événement. Gillies revient alors sur l'aspect relationnel des probabilités. Il fait une analogie avec le poids, qui est souvent interprété comme une qualité inhérente d'un corps, alors qu'il s'agit d'une propriété relationnelle, car il est établi dans la relation d'un corps avec un champ gravitationnel (le poids d'un corps n'est pas le même sur la Terre ou sur la Lune). De la même façon, la probabilité d'un événement ne sera pas identique si les conditions de sa production sont modifiées. Or, la probabilité n'est pas une propriété d'un objet, mais d'un événement, et en tant que telle, elle est relationnelle.

Ce lien établi entre la probabilité et les conditions de l'événement permet, par ailleurs, de définir le hasard en termes d'indépendance entre les éléments d'une série, comme le démontre Gillies :

Let us confine ourselves to sequences of 0s and 1s. We shall say that such a sequence is random if it is generated by repeating a set of conditions S which are such that (a) the repetitions are independent, and (b) the outcomes are 0 or 1 with $Prob(0) = p$, for some fixed value p such that $0 \leq p \leq 1$. (Gillies, 2000, p. 154).

Retenons cette définition d'une série aléatoire en termes de répétition indépendante pour la considérer à la lumière d'autres approches théoriques sur le hasard. Mais pour

l'instant, il est pertinent de préciser la notion de *répétition indépendante*, si l'on souhaite réduire le hasard à cette formulation. Gillies aborde le problème de la répétition de la façon suivante : « Let us begin with the obvious point that any two alleged repetitions will be found on closer inspection to differ in many respects. » (Gillies, 2000, p. 161) Il réitère que même si toutes les conditions macroscopiques sont répétées, il y a une différence dans le temps. Il propose alors de considérer certains aspects des conditions pour définir la répétition, ainsi les événements ne doivent pas être similaires dans toutes ses caractéristiques, mais seulement dans les aspects qu'on a choisis : « A sequence is a sequence of repetition relative to a set S of conditions or properties if each member of the sequence satisfies every condition of S, and irrespective of how the members differ in other respects. » (Gillies, 2000, p. 162) D'après ces observations, Gillies établit l'axiome des répétitions indépendantes comme une des bases de la théorie des probabilités.

Gilles Deleuze, dans son œuvre *Différence et répétition*, a largement abordé les concepts de la répétition et de l'habitude, tous les deux au cœur des interprétations propensionnistes et de la théorie des probabilités. Il est alors possible que son travail nous éclaire sur les relations entre la propension, le *would-be*, le hasard et l'événement. Mais avant d'aborder la perspective deleuzienne sur la question, il faut souligner que Gillies considère que la polarisation entre les interprétations objective et subjective est excessive et propose de les considérer dans un spectre « stretching from the fully subjective to the fully objective. » (Gillies, 2000, p. 175) Gillies défend une vision pluraliste de la théorie des probabilités, où les différentes interprétations sont possibles et s'appliquent à différents types d'événements :

In the objective interpretation, probabilities are associated with repeatable conditions which have independent outcomes. (...) Subjective probabilities, on the other hand, are appropriate for singular events, either where no repeatable conditions can be easily defined, as in the case of horse races, or where such repeatable conditions as can be defined do not express all our knowledge relating to the individual event, as in the case

considering whether a particular person will have a road accident in the next five years. (Gillies, 2000, p. 186)

Il établit ainsi une différence entre les événements singuliers et les événements en série avec des conditions répétées, et propose d'interpréter l'usage des probabilités en fonction de cette distinction. Dans son échelle d'interprétations, Gillies propose les étapes suivantes :

1. Subjective³⁹: Here probabilities represent the degrees of belief of particular individuals.
2. Intersubjective: Here probabilities represent the degree of belief of a social group which has reached a consensus.
3. Artefactual: Here the probabilities can be considered as existing in the material world and so as being objective, but they are the result of interaction between humans and nature. (...) [Exemple : jeu de dés]
4. Fully objective: (...) Things should be considered fully objective which exist in the material world quite independently of human beings. (...) [Exemple : désintégration d'un atome radioactif] (Gillies, 2000, p. 179-180)

Il est important de remarquer que dans la proposition de Gillies le facteur déterminant entre les pôles subjectif/objectif est l'interaction humaine. De mon point de vue, je préfère parler en termes d'activités cognitives d'apprentissage, de mémoire et de réflexion, qu'elles soient humaines ou non-humaines (je développerai plus loin le concept d'activité de réflexion). Plus loin, Gillies apporte une nuance aux relations entre les interprétations objective et subjective, il affirme :

My view is that the objective propensity interpretation should be taken as fundamental. (...) The basic interpretation of probability theory is thus as a mathematical science of randomness, and the theory's success in explaining (and rendering more precise) a mass of empirical material is what confirms its axioms, and justifies us in accepting them. (Gillies, 2000, p. 185)

³⁹ Il est important de noter que Gillies considère ici la théorie bayésienne comme l'interprétation subjective, et non le déterminisme laplacien tel qu'étudié plus haut dans le texte

Pour Gillies, la théorie des probabilités concerne fondamentalement les événements aléatoires indépendamment d'une interaction avec l'être humain. L'interprétation subjective viendrait dans un deuxième temps, comme une extension de l'application des outils mathématiques de probabilité : « The subjectivists have shown that the mathematical calculus can be extended to deal with degree of belief in particular events. The connection between the two interpretations occurs in the area of games of chance. » (Gillies, 2000, p. 185) Ainsi, dans la proposition de Gillies il existe déjà une reconsidération du dualisme objectif/subjectif vers un pluralisme d'interprétations. Mais si l'on accepte les prémisses de la posture pragmatiste spéculative, comment reformuler alors l'interprétation des probabilités sous la lumière de la remise en question des catégories préétablies de sujet et d'objet ?

Pour y répondre, tournons-nous vers les réflexions de Deleuze autour de la répétition et de l'habitude. Dans le deuxième chapitre du livre *Différence et répétition*, Deleuze reprend les thèses de Hume : « Hume prend comme exemple une répétition de cas, du type AB, AB, AB, A... Chaque cas, chaque séquence objective AB est indépendante de l'autre. » (Deleuze, 1993 [1968], p. 96) Comme dans la théorie des probabilités, les éléments de la répétition sont indépendants dans cet exemple. Il continue :

La répétition (...) ne change rien dans l'objet, dans l'état de choses AB. En revanche, un changement se produit dans l'esprit qui contemple : une différence, quelque chose de nouveau *dans* l'esprit. Lorsque A paraît, je m'attends maintenant à l'apparition de B. » (Deleuze, 1993 [1968], p. 96)

Pour Deleuze, cela s'opère tel une synthèse, une contraction : « Quand A paraît, nous nous attendons à B avec une force correspondant à l'impression qualitative de tous les AB contractés. » (Deleuze, 1993 [1968], p. 97) Il appellera ce processus la synthèse passive de l'imagination (où l'imagination est définie comme « un pouvoir de contraction »). Cette synthèse passive est aussi décrite en termes d'une habitude contractée :

L'habitude *soutire* à la répétition quelque chose de nouveau : la différence (...). Telle est la synthèse passive, qui constitue notre habitude de vivre, c'est-à-dire notre attente que « cela » continue, qu'un des deux éléments survienne après l'autre, assurant la perpétuation de notre cas. (Deleuze, 1993 [1968], p. 101)

Comme on le remarque, le processus de la synthèse passive décrit par Deleuze est très proche de celui qui est à la base de l'interprétation propensionniste : dans des conditions de répétitions indépendantes, une habitude, un *would-be*, une propension est contractée, et la probabilité serait alors la mesure quantitative de cette propension, un nombre (entre 0 et 1) attribué à l'intensité avec laquelle je m'attends à B lorsque A apparaît. La notion d'habitude tel que développée par Deleuze est très proche des réflexions de Peirce qui, rappelons-nous, avait basé son interprétation des probabilités sur l'exemple de l'habitude, comme il l'affirme dans son texte *Design and Chance* :

I have several times shown to my classes how some of the main laws of cerebration and particularly the formation of habits could be accounted for by the principles of probability, and I have shown by experiment how a certain regularity of arrangement can be impressed upon a pack of cards by imitating the action of habit. The main element of habit is the tendency to repeat any action which has been performed before. (Charles S. Peirce, 1989, p. 553)

Or, cette idée d'habitude ou de synthèse passive décrit une impression qualitative qui est ensuite calculée par la probabilité, l'intensité de « l'attente que cela continue ».

Sur la fondation des synthèses passives se construisent les synthèses actives. Les synthèses actives sont liées aux activités de mémoire et de réflexion : « La synthèse active (...) est la représentation du présent, sous le double aspect de la reproduction de l'ancien et de la réflexion du nouveau. » (Deleuze, 1993 [1968], p. 112) La mémoire, le mécanisme de la « reproduction de l'ancien », et la réflexion, comme le lien vers l'avenir, composent ce double aspect de la synthèse active. Un autre aspect important de la synthèse active est son activité englobante, car c'est dans cette deuxième synthèse qui se configure l'idée d'un moi, ou d'un sujet : « Les moi passifs étaient déjà des intégrations, mais, comme disent les mathématiciens, des intégrations locales ; le moi

actif est tentative d'intégration globale. » (Deleuze, 1993 [1968], p. 131) L'aspect englobant de la synthèse active sera abordé dans le chapitre suivant et mis en relation avec la notion d'organisation.

À ce point, il serait pertinent de revenir sur l'échelle proposée par Gillies pour classer les différentes interprétations de la théorie de la probabilité et la mettre en rapport avec les synthèses décrites par Deleuze. Est-ce que c'est le cas d'affirmer que la synthèse passive, en tant que non réflexive, est en rapport avec les interprétations objectives ? De la même façon, peut-on conclure que les interprétations subjectives sont de l'ordre des synthèses actives ? Dans ces cas, les probabilités seraient alors un regard mathématique sur les différentes synthèses réalisées, une quantification des différentes tendances, propensions ou *would-be* que chacun (ou chaque *moi larvaire*) construit grâce aux synthèses qu'il réalise tout au long de sa vie. L'échelle des interprétations de la probabilité proposée par Gillies mettrait l'être humain comme l'élément déterminant du passage du pôle épistémique (ou subjectif) à l'objectif, néanmoins, si l'on adopte les notions des synthèses deleuziennes, le fondement de l'échelle des interprétations serait dans la différence entre les synthèses passives et les synthèses actives. L'interprétation objective des probabilités est ainsi une réflexion portée aux synthèses passives alors que dans l'interprétation épistémique il s'agit de la probabilité en tant qu'un calcul des synthèses actives. Si l'on choisit de se référer à des probabilités passives ou actives, nous pouvons abandonner les termes de probabilité objective ou subjective, et ainsi nous débarrasser de la dichotomie sujet/objet comme pivot des interprétations.

Mais est-ce qu'il s'agit dans ce cas de remplacer un dualisme par un autre ? Une nuance importante apportée par la pensée de Deleuze est nécessaire pour aborder cette question. Pour éviter de retomber dans les dualismes tels que sciences sociales/sciences naturelles, humain/non-humain, sujet/objet, synthèse passive/synthèse active, il faut être attentif à la multiplicité de modes des synthèses passives et actives. En effet, on ne peut pas parler d'un seul type de synthèse passive ou active, au contraire, Deleuze

écrit : « Nous ne devons pas seulement distinguer des formes de répétition par rapport à la synthèse passive, mais des niveaux de synthèses passives, et des combinaisons de ces niveaux entre eux, et des combinaisons de ces niveaux avec les synthèses actives. » (Deleuze, 1993 [1968], p. 100) Or, dans la construction de chaque modèle probabiliste, on ne peut pas faire fi de la complexité des synthèses qui y sont opérantes, il est nécessaire de considérer les différents modes de ces synthèses et les relations entre ces processus.

Cette approximation entre la pensée de Deleuze sur les synthèses de la répétition et la théorie des probabilités fait curieusement écho à certains aspects de la pensée de Georges Soros sur le rôle de la *réflexivité* dans l'interprétation des probabilités en économie (in Gillies, 2000). Dans le dernier chapitre de son livre, Gillies étudie le cas de l'économie pour identifier les interprétations adéquates des probabilités. Cet exercice nous intéresse dans le cadre de cette thèse, car il s'agira par la suite de repérer, dans les théories de l'information, quelles sont les interprétations des probabilités qui doivent être utilisées. Gillies reprend les arguments de Soros et, à partir d'une distinction entre les sciences sociales et les sciences naturelles, défend que l'interprétation subjective s'applique au premier groupe (les sciences sociales), et l'interprétation objective au deuxième (les sciences naturelles). Gillies rappelle que Soros a été l'élève de Popper, un des fondateurs de l'interprétation propensionniste, mais cela ne l'a pas empêché de défendre une vision pluraliste des probabilités :

I was greatly influenced (...) by Karl Popper's ideas on scientific method. I accepted most of his views, with one major exception. He argued in favor of what he called 'unity of method'—that is, the methods and criteria that apply to the study of natural phenomena also apply to the study of social events. I felt that there was a fundamental difference between the two: the events studied by the social sciences have thinking participants; natural phenomena do not. The participants' thinking creates problems that have no counterpart in natural sciences. (Soros in Gillies, 2000, p. 195)

Soros a développé alors sa théorie de la réflexivité selon laquelle l'activité de la réflexion, caractéristique des acteurs économiques, interfère d'une manière décisive

dans les conditions des événements. Or, dans l'interprétation propensionniste, les probabilités sont définies en termes de répétitions indépendantes d'un ensemble de conditions. La réflexivité met justement en échec l'indépendance entre les répétitions. Pour expliquer cela, Gillies fait la comparaison entre deux modèles probabilistes, un pour la description du mouvement de molécules de gaz dans un contenant, et l'autre pour la description d'une économie capitaliste. Comme il explique :

The molecules have no knowledge, consciousness or volition, and, apart from the occasional collision, move to a first approximation independently of each other. The economic agents, however, do possess knowledge, consciousness, desire and will. Moreover, their actions, far from being independent, are characterised by reaction to each other's actual or expected decisions. (Gillies, 2000, p. 190)

Malgré ces possibles approximations, les formulations de la notion de réflexivité proposées par Deleuze et Soros divergent néanmoins sur certains aspects. Or, pour Soros l'activité réflexive semble être exclusive aux humains, alors que pour Deleuze, cette activité se caractérise par toute intervention de la mémoire et/ou de la réflexion, qu'elle soit humaine ou non-humaine. Ainsi, dans la pensée de Deleuze il existe une couche supplémentaire de complexité dans le sens où les synthèses actives et passives sont opérantes non seulement dans les agents économiques, mais dans tout organisme qui performe une activité englobante et qui se configure en tant qu'individu. La complexité des relations entre les différentes synthèses doit ainsi être considérée.

La question qui s'en suit est alors : comment s'articulent les synthèses actives et les synthèses passives ? Deleuze écrit à ce propos : « une synthèse active s'établit sur la fondation des synthèses passives (...) » (1993 [1968], p. 131), or la synthèse passive est la fondation, le sol sur lequel se bâtit la synthèse active. Il faut alors comprendre que dès qu'il y a synthèse active, il y a également des synthèses passives sous-jacentes. Par extension, dès que l'on rencontre des probabilités actives (ou subjectives) dans un modèle, on saura que des probabilités passives (ou objectives) sont aussi opérantes.

Au-delà de la multiplicité et de la relation complexe entre les synthèses passives et actives, il est possible de trouver dans la pensée de Deleuze d'autres éléments qui échappent à une reconstruction de binarité, la formulation de la troisième synthèse en étant un. Cette synthèse est particulièrement importante dans le contexte de la présente thèse, car la notion de hasard y est directement abordée. La troisième synthèse est celle de la *forme vide du temps*, ou de l'éternel retour :

Nous ne produisons quelque chose de nouveau qu'à condition de répéter une fois sur ce mode qui constitue le passé, une autre fois dans le présent de la métamorphose. Et ce qui est produit, l'absolument nouveau lui-même, n'est rien d'autre à son tour que la répétition, cette fois par excès, celle de l'avenir comme éternel retour. (Deleuze, 1993 [1968], pp. 121-122)

Plus loin dans ce même chapitre, Deleuze revient sur l'éternel retour pour l'associer au hasard : « Le sujet de l'éternel retour n'est pas le même, mais le différent, ni le semblable, mais le dissimilaire, ni l'Un, mais le multiple, ni la nécessité, mais le hasard. » (Deleuze, 1993 [1968], p. 164) En effet, si les synthèses passives et actives concernent les tendances (ou propensions) qui composent l'objet des calculs de la probabilité, la troisième synthèse concerne ce qui échappe aux modèles prédictifs probabilistes. Dans cette troisième synthèse il est question d'affirmation du hasard, du « caractère inconditionné du produit par rapport à sa condition, l'indépendance de l'œuvre par rapport à son auteur ou acteur. »

Deleuze aborde plus longuement la question du hasard lorsqu'il reprend des pensées déjà formulées dans *Nietzsche et la Philosophie* (Deleuze, 1983 [1962]) pour aborder le hasard dans les conditions d'un jeu et l'opposer à la probabilité. Dans la conclusion de *Différence et Répétition*, il décrit deux manières de jouer, qu'il appelle *humaine* et *idéale*, et les différencie :

D'abord, le jeu humain suppose des règles catégoriques préexistantes. Ensuite, ces règles ont pour effet de déterminer des probabilités, c'est-à-dire des « hypothèses » de perte et des hypothèses de gain. En troisième lieu, ces jeux n'affirment jamais tout le hasard, au contraire, ils le

fragmentent, et pour chaque cas soustraient au hasard, exceptent du hasard la conséquence du coup, puisqu'ils assignent tel gain ou telle perte comme nécessairement liés à l'hypothèse. (Deleuze, 1993 [1968], p. 361)

Le jeu humain, celui du calcul des probabilités, est celui qui nie le hasard pour affirmer la nécessité. Mais Deleuze propose un autre type de jeu pour faire du hasard une affirmation :

Tout autre le jeu divin, celui dont Héraclite parle peut-être, celui que Mallarmé invoque avec tant de crainte religieuse et de repentir, Nietzsche avec tant de décision - jeu pour nous le plus difficile à comprendre, impossible à manier dans le monde de la représentation. D'abord, il n'y a pas de règle préexistante, le jeu porte sa propre règle. Si bien que, à chaque fois, tout le hasard est affirmé dans un coup nécessairement vainqueur. (Deleuze, 1993 [1968], p. 362)

Dans cette deuxième manière de jouer, le hasard est affirmé à chaque coup, et les calculs des propensions, ou des probabilités, n'ont pas lieu d'être, car aucune règle ne peut déterminer le comportement de ce jeu. Mais comme il a été déjà abordé dans le premier chapitre, c'est un jeu sur deux tables (qui va du multiple à l'Un, du hasard à la nécessité) et la césure de la troisième synthèse est « à la limite, à la charnière des deux tables » (Deleuze, 1993 [1968], p. 363). Il est important alors de remarquer qu'avec la troisième synthèse, Deleuze fait de la probabilité un jeu humain, une tentative de calcul des synthèses passives et actives encrées dans les processus de contraction (habitude) et d'activités liées à la mémoire et à la réflexion.

Il faut aussi remarquer que dans la synthèse passive, il est question d'une intensité qualitative, d'« une impression qualitative interne d'un certain poids. » (Deleuze, 1993 [1968], p. 97). Ce poids n'est pas encore mesuré, et c'est cela le geste essentiel de la probabilité, quantifier le poids de l'attente créée par la contraction d'une habitude, ou de la synthèse passive : « Constituante, elle [la synthèse passive] n'est pas pour cela active. Elle n'est pas faite par l'esprit, mais se fait *dans* l'esprit qui contemple, précédant toute mémoire et toute réflexion. » (Deleuze, 1993 [1968], p. 97) Dans la synthèse passive il est question de contemplation, et non de la mémoire réflexive ou

cognitive. C'est dans la synthèse active de l'esprit que la probabilité est calculée, même si elle existe déjà en tant qu'impression qualitative. La probabilité est ainsi définie comme une opération des synthèses actives portée vers les synthèses passives et/ou actives.

Massumi met aussi en évidence les limites des probabilités liées à l'aspect qualitatif de l'expérience :

Both possibility and potential can be approached quantitatively, the latter through probabilities. Although they lend themselves to quantification, it must be remembered that they only lend themselves. Every possibility and potential, however calculated, retains an irrevocable residue of qualitative difference, to which it owes its distinctness. That residue or reserve (more insistent in potential than in possibility) is virtuality, as enveloped in the given. (Massumi, 1998, p. 307-8)

Pour Massumi la dimension virtuelle implique un écart entre le modèle quantitatif et l'expérience donnée. Il s'agit ainsi de poser quelques réserves importantes quant à l'utilisation des probabilités dans la description d'une expérience en fonction de la dimension virtuelle de l'événement qui échappe à toute formalisation probabiliste.

Malgré toutes les réserves qui doivent être considérées dans le traitement des probabilités, on constate néanmoins que les notions de synthèses passives et actives peuvent contribuer à répondre à la question posée au départ de cette section : comment reformuler l'interprétation des probabilités à la lumière de la remise en question des catégories préétablies de sujet et d'objet opérées par la perspective du pragmatisme spéculatif ?

La question qui occupera la suite de ce chapitre ainsi que le chapitre suivant devient celle-ci : dans le cas des médias numériques et interactifs, quelles sont les synthèses opérantes ? Autrement dit, quelle est l'interprétation philosophique des probabilités dans les dispositifs numériques interactifs, notamment dans le cadre d'un agent conversationnel ? Je propose de laisser la question ouverte pour l'instant et d'examiner

de plus près le fonctionnement des dispositifs numériques et la notion d'information numérique. Comme on le verra, les probabilités et la notion de hasard sont centrales aussi dans les théories de l'information et de l'information algorithmique, ainsi qu'en informatique, des champs et théories à l'origine des dispositifs numériques à la base de l'interactivité. C'est par cette approche pragmatiste spéculative des théories de l'information et de l'information algorithmique que la question des relations entre hasard et numérique sera approfondie, pour qu'on arrive enfin aux propositions des machines de Turing probabilistes, les algorithmes randomisés et les réseaux bayésiens. Enfin, la proposition d'une pratique de *speculative computation* de Luciana Parisi sera considérée.

3.2. Théorie de l'information : hasard et entropie

La théorie de l'information sera abordée ici par la perspective de Claude Shannon, qui l'a développée dans un ouvrage de 1949 copublié avec Warren Weaver (Shannon, 1948), et aussi à travers deux lectures plus récentes de la théorie de Shannon, celles proposées par Henri Atlan et par Tiziana Terranova (Atlan, 2006 [1972]; Terranova, 2004a, 2004b). Comme nous le verrons, cette théorie a trait à l'aspect quantitatif de l'information, dans le but de mesurer « la quantité d'information transmise dans une voie de communication » (Atlan, 2006 [1972], p. xxvii), sans se soucier du contenu sémantique du message. Comme le dit Shannon : « Frequently the messages have meaning; (...) These semantic aspects of communication are irrelevant to the engineering problem. » (Shannon, 1948, p. 379) Ainsi, Shannon s'intéresse à l'information dans sa forme quantitative, à sa mesure dans un message codé et transmis, tel un système numérique. Shannon définit le *bit* (*binary digit*) comme l'unité de mesure de l'information dans un système binaire : « A device with two stable positions, such as a relay or a flip-flop circuit, can store one bit of information. » (Shannon, 1948, p. 381) Atlan rappelle également que l'idée « que l'information était mesurable »

(Atlan, 2006 [1972], p. 5) était exprimée pour la première fois « en 1872, dans les travaux de Boltzmann sur l'entropie d'un système comme mesure de son degré d'ordre ; » (Atlan, 2006 [1972], p. 5)

Shannon utilise le concept d'entropie tel que formulé par Boltzmann pour mesurer l'information. Terranova rappelle les trois formulations de la définition de l'information par Shannon :

Information is defined by the relation of signal and noise; information is a statistical measure of uncertainty or entropy of a system; information implies a nonlinear and nondeterministic relationship between the microscopic and macroscopic levels of a physical system. (Terranova, 2004b, p. 9)

Dans le premier chapitre de cette thèse, la notion d'entropie a été introduite par la perspective de Stengers et Prigogine, notamment dans l'interprétation proposée de l'entropie en thermodynamique. Or, la notion d'entropie resurgit dans le contexte de la théorie de l'information. Comme nous le verrons, ce parallèle entre information et entropie implique l'inclusion des statistiques et des probabilités dans les processus de communication. Revenons maintenant sur certains aspects de la théorie de l'information proposée par Shannon pour mettre en évidence le rôle qui y jouent l'entropie et les probabilités.

Dans ses trois variantes (discret, continu et mixte), un système de transmission d'information (communication) est composé de cinq éléments essentiels : la source d'information, le transmetteur, le canal (ou la voie de transmission), le récepteur et le destinataire. Dans un système discret, la source d'information produit les messages un symbole à la fois :

It will choose successive symbols according to certain probabilities depending, in general, on preceding choices as well as the particular symbols in question. A physical system, or a mathematical model of a system which produces such a sequence of symbols governed by a set of probabilities, is known as a stochastic process. (Shannon, 1948, p. 386)

Lorsque Shannon définit une source d'information comme un processus aléatoire (*stochastic*), il met les probabilités au cœur de son modèle de communication. Cela le permet d'utiliser tous les outils mathématiques de la théorie de la probabilité et de la physique statistique pour développer la théorie de l'information. Shannon propose alors une formule mathématique pour ce qu'il appelle *l'entropie de la source d'information* (H) :

Consider a discrete source of the finite state type considered above. For each possible state i there will be a set of probabilities $p_i(j)$ of producing the various possible symbols j . Thus there is an entropy H_i for each state. The entropy of the source will be defined as the average of these H_i weighted in accordance with the probability of occurrence of the states in question. (Shannon, 1948, p. 393)

Shannon reprend le terme d'*entropie*, car la formule qu'il propose pour mesurer la quantité d'information moyenne émise par une source est similaire à celle de l'entropie tel qu'énoncée par Boltzmann. Atlan soulève deux propriétés importantes de la fonction de mesure de l'information proposée par Shannon : en premier, « L'information est reliée à l'ensemble des formes possibles de réalisation d'un événement (...) : sa valeur dépend des probabilités associées avec ces formes de réalisation, mais en aucune façon de leurs causes ou de leurs conséquences. » (Atlan, 2006 [1972], p. 15) Ainsi, la valeur quantitative de l'information est directement liée aux probabilités des différents événements qui la composent (l'avènement d'un symbole est un exemple d'un événement qui compose une information), plus précisément, l'information est inversement proportionnelle à la probabilité de l'événement : plus l'événement est rare, plus grande sera sa valeur en information.

La deuxième propriété porte sur une question déjà abordée, mais qui sera approfondie par la suite. Comme nous l'avons vu, « [l']information est une quantité abstraite mesurable dont la valeur ne dépend pas de ce sur quoi porte l'information » (Atlan, 2006 [1972], p. 15). Cela implique que le concept d'information dont il est question

pour Shannon exclut les processus d'interprétation et de connaissance auxquels on peut soumettre l'information. Atlan cite L. Brillouin à ce sujet :

Selon notre définition, nous donnerons la même valeur d'information à 100 lettres prises au hasard et à une phrase de 100 lettres tirée d'un journal, d'une pièce de Shakespeare ou d'un théorème d'Einstein. En d'autres termes, nous définissons l'information indépendamment de la connaissance à laquelle nous ne pouvons attribuer de valeur numérique. (Atlan, 2006 [1972], pp. 18-19)

Il semblerait alors que l'interprétation la plus juste de la probabilité dans le contexte de la théorie de l'information est celle qui a été nommée la *probabilité des synthèses passives* (ou *objective* dans le vocabulaire de Gillies). Ainsi, dans un premier moment, l'information telle que développée par Shannon concerne les processus physiques avant toute synthèse active, et se place alors du côté des probabilités passives. Comme nous l'avons vu, les interprétations objectives des probabilités ont deux lignées, fréquentistes et propensionnistes. Shannon ne l'affirme pas de manière explicite, mais le choix des termes utilisés pour aborder les probabilités de la source d'information indique une interprétation fréquentiste des probabilités, comme dans ce passage où Shannon discute les différentes probabilités d'apparition de chaque symbole d'un message en langue anglaise produit par une source d'information :

The letter E occurs more frequently than Q, the sequence TH more frequently than XP, etc. The existence of this structure allows one to make a saving in time (or channel capacity) by properly encoding the message sequences into signal sequences." (Shannon, p. 6).

L'énonciation d'Atlan de la formule de l'information de Shannon laisse croire que lui aussi a une interprétation objective fréquentiste des probabilités du comportement de la source de l'information. Il dit, « $H(x) = -\sum p(i) \log_2 p(i)$ où i représente l'indice des N différents symboles $x_1, x_2, x_3 \dots x_i, \dots x_n$ utilisés dans le message et $p(i)$ la fréquence d'utilisation, ou la probabilité d'apparition du symbole x_i . » (Atlan, 2006 [1972], p. 7) Comme l'interprétation propensionniste, l'interprétation fréquentiste s'applique à la répétition d'événements indépendants, avec la différence d'être une limite vers laquelle

la série d'événements converge à l'infini. Or, auparavant dans cette thèse il était question de l'interprétation propensionniste dans le pôle passif de l'échelle. Dans cette perspective, la formule de Shannon indiquerait plutôt la valeur (en termes quantitatifs) de la propension, du *would be*, ou de l'habitude d'une source d'information de produire certains symboles. Ce changement dans l'interprétation de la théorie de l'information ne change pas son calcul mathématique, il s'agit d'une question purement philosophique autour du rôle du hasard dans le processus en question, de proposer une interprétation pragmatiste spéculative de la théorie de l'information.

Shannon identifie une deuxième source de hasard dans un système de communication. La source de l'information, comme nous l'avons vu, était la première interférence aléatoire dans un système de communication, or le bruit présent dans la voie de transmission de l'information est aussi considéré une interférence aléatoire dans le système. Comme l'explique Atlan, « [d]es erreurs de transmission sont produites par des facteurs divers, intervenant sur la voie, de façon aléatoire, et qui sont pour cela dénommés facteurs de *bruit*. » (Atlan, 2006 [1972], p. 32) Lorsque la transmission se fait sur une voie idéale (sans bruit), tous les éléments du message reçu par le destinataire sont identiques aux éléments émis par la source. Néanmoins, si l'on introduit du bruit, le message livré au destinataire peut contenir des éléments modifiés, des erreurs. Lorsqu'il y a du bruit dans les voies de transmission, pour chaque élément reçu par le destinataire il existe une probabilité qu'il ne soit pas identique à l'élément qui a été produit par la source.

Pour aborder le problème du bruit dans une voie de transmission d'information, la théorie de l'information fait appel alors aux probabilités conditionnelles. Comme explique Atlan, les probabilités conditionnelles permettent d'aborder le problème des variables interdépendantes. Atlan utilise l'exemple des relations entre les tailles des pères et celles de leurs filles pour illustrer le cas de variables qui ne sont pas indépendantes. Il constate qu'il existe une relation entre les variables *taille des pères* (y) et *taille des filles* (x) (en raison d'une transmission d'information génétique) qui fait

qu'elles ne soient pas indépendantes. Considérons la situation suivante : si l'on veut connaître la taille d'une fille en particulier, des informations sur la taille de son père changent notre connaissance sur sa taille à elle. Si l'on sait que le père est très grand, par exemple, cela augmente les probabilités que sa fille soit parmi les groupes de taille plus élevée. Atlan arrive ainsi à une formule de $T(x ; y)$ et conclut que « T mesure alors la quantité d'information transmise en moyenne par la taille des pères en ce qui concerne celles de filles (ou vice versa). » (Atlan, 2006 [1972], p. 71) Autrement dit, la fonction T propose une mesure du degré d'interdépendance entre deux événements. Atlan formule par la suite T en termes de probabilités conditionnelles et conclut :

On voit clairement pourquoi T est une mesure de l'information transmise de x à y , ou de y à x : c'est une mesure du gain en certitude sur y quand x est connu ou aussi bien du gain en certitude de x quand y est connu, autrement dit, de la relation mutuelle d'incertitude, traduisant une dépendance entre x et y . (Atlan, 2006 [1972])

Atlan résume ainsi les variables opérantes dans un système d'une voie de communication (ou de transmission d'information), considérant x comme le message d'entrée et y , le message de sortie :

$H(x)$: incertitude sur la source ou quantité d'information du message d'entrée.

$H(y)$: incertitude sur la réception ou quantité d'information du message de sortie.

$T(x ; y)$: quantité d'information transmise dans la voie.

$H(x|y)$: quantité d'information de la sortie quand l'entrée est déterminée, ou *ambigüité*.

$H(y|x)$: quantité d'information de l'entrée quand la sortie est déterminée, ou *équivocation*. (Atlan, 2006 [1972])

L'ajout du bruit dans la voie de communication implique alors l'utilisation de probabilités conditionnelles, présentes dans les expressions $H(x|y)$ et $H(y|x)$. Ces expressions sont aussi appelées « information mutuelle ».

Il faudrait maintenant revenir aux interprétations de la théorie des probabilités pour situer les propositions de la théorie de l'information. Dans la section précédente, il a été question d'une lecture de la théorie des probabilités par la perspective du pragmatisme spéculatif. Maintenant il s'agira d'en reprendre les résultats pour considérer ses implications dans le domaine de la théorie de l'information. Le caractère aléatoire de la source d'information situait jusqu'ici la théorie de l'information dans le pôle des synthèses passives des probabilités. Lorsque les probabilités conditionnelles entrent en jeu, des relations d'interdépendances entre les événements commencent à émerger, des informations sont transmises entre les événements. Selon Atlan, nous avons affaire ici à des processus d'auto-organisation qui créent de l'interdépendance entre ses éléments. La notion d'auto-organisation est très proche des structures dissipatives traitées par la perspective de Stengers et Prigogine dans le premier chapitre de cette thèse.

Pour Atlan, la notion d'auto-organisation est « définie formellement comme création d'information dans un système organisé, non programmé de l'extérieur. » (Atlan, 2006 [1972], p. 13). Ainsi, c'est le taux de production d'information (plutôt que la quantité d'information) qui décrit l'organisation d'un système. En effet, ce n'est pas la mesure instantanée de la quantité d'information d'un système qui indique un processus d'organisation, mais sa variation :

(...) l'organisation d'un système quelconque doit être décrite par son taux de variation de quantité d'information, lui-même somme de deux fonctions, l'une, taux de décroissance de la redondance, l'autre, taux de décroissance de la quantité d'information maximum. (Atlan, 2006 [1972], p. 270)

Nous remarquons alors que l'auto-organisation des structures dissipatives fait passer les synthèses du pôle passif au pôle actif, dans le sens où des relations d'interdépendance entre les événements d'une suite commencent à émerger. Ainsi, la lecture proposée par Atlan de la théorie de l'information rapprochée de la lecture élaborée plus haut de la théorie des probabilités permet de comprendre la probabilité

comme une mesure de la propension d'un système à créer de l'information, ou à produire des synthèses passives ou actives. Néanmoins, il est important de rappeler l'existence de la troisième synthèse, celle qui transforme le système d'une manière qui ne peut pas être inférée des données historiques. Comme dans les structures dissipatives, il s'agit d'états d'auto-organisation qualitativement différents qui ne peuvent pas être prévus par les dynamiques des états précédents, des points de rupture où le possible échappe au probable.

Le hasard joue un rôle important dans la notion d'auto-organisation des systèmes, comme le démontre Atlan. Il reprend la conception de l'organisation biologique telle que développée dans la théorie des réseaux neuronaux et remarque des aspects aléatoires dans la structure des réseaux. Il s'agit d'une structure aléatoire dans la connexion des réseaux qui se construit sur la base d'une détermination génétique. Il développe le concept de *hasard organisationnel*, ou « la capacité d'utiliser des phénomènes aléatoires pour les intégrer au système et les faire fonctionner comme des facteurs positifs, créateurs d'ordre, de structures, de fonctions. » (Atlan, 2006 [1972], p. 230) Pour Atlan, le travail de Shannon pose « les conditions formelles de la création de l'information », et « [l]e principe de complexité par bruit permet de comprendre le rôle du hasard ou de "bruit développemental", comme facteur d'organisation dynamique des systèmes complexes. » (Atlan, 2006 [1972], c.f. quatrième de couverture) Ce constat d'Atlan souligne le caractère processuel de cette interprétation de la théorie de l'information. Ce que je propose par la suite est ainsi une lecture processuelle de la théorie de l'information, appuyée notamment sur le travail de Gilbert Simondon et de Tiziana Terranova.

Le travail de Simondon permet de mettre en évidence certaines limites de la théorie de l'information. Comme il souligne, ses interprétations des concepts l'information et de communication ne sont pas incompatibles avec celles de Shannon et des cybernéticiens, mais ces derniers manquent à reconnaître l'aspect processuel des individus. Considérant le modèle de communication de Shannon par une perspective

simondonienne, l'information circule non pas entre individus déterminés, mais entre préindividus affectés par l'information elle-même et en constante transformation dans des processus d'individuation :

Cette affirmation ne conduit pas à contester la validité des théories quantitatives de l'information et des mesures de la complexité, mais elle suppose un état fondamental — celui de l'être préindividuel — antérieur à toute dualité de l'émetteur et du récepteur, donc à tout message transmis. Ce qui reste de cet état fondamental dans le cas classique de l'information transmise comme message, ce n'est pas la source d'information, mais la condition primordiale sans laquelle il n'y a pas d'effet d'information, donc pas d'information : la métastabilité du récepteur, qu'il soit être technique ou individu vivant. On peut nommer cette information « information première ». (Simondon, 1964, p. 16)

Simondon met en évidence l'impossibilité de se référer à l'émetteur ou au récepteur comme un individu fixe et isolé, il s'agit de préindividus pris dans des processus d'individuation. Selon Simondon, la communication se donne entre (ou à l'intérieur) des systèmes en état d'équilibre métastables. Comme le remarque aussi Tiziana Terranova, d'après Simondon,

[i]nformation does not simply flow from point A to point B as if through a void. Both sender and receiver are immersed within a larger field of interactions that packs within itself a potential for transformation and even divergent tendencies that the mathematical theory of communication does not capture. (Terranova, 2004b, p. 68)

Dans cette perspective, il est impossible de réduire l'incertitude de la communication au bruit, car le signal est transmis entre des pôles instables, mouvants.

Par ailleurs, Simondon observe que la réception du signal émis est une condition de l'information. Il met ainsi l'accent sur l'aspect relationnel de l'information, car l'information n'existe que dans la relation créée entre émetteur et récepteur. Si le signal émis n'est pas reçu par un système métastable, il n'est pas converti en information et constitue seulement en dissipation d'énergie :

Il n'y a information que lorsque ce qui émet les signaux et ce qui les reçoit forme système. L'information est entre les deux moitiés d'un système en relation de disparation. Cette information ne passe pas nécessairement par des signaux (par exemple dans la cristallisation) ; mais elle peut passer par des signaux, ce qui permet à des réalités éloignées l'une de l'autre de former système. (Simondon, 1964, p. 255)

Simondon utilise comme exemple la vue stéréoscopie et la perception de la profondeur : il existe une différence (*disparation*) entre l'image perçue par l'œil gauche et l'image perçue par l'œil droit, et la perception de profondeur (ou de relief) se construit justement dans de cette différence. Dans cette analogie, la perception du relief représente l'information qui, comme on le remarque, n'est ni dans le signal reçu par l'œil gauche ni dans le signal reçu par l'œil droit, mais elle se construit dans la différence (ou l'écart) entre ces deux signaux. Si l'on trace un parallèle entre le système d'information proposé par Simondon et les synthèses deleuziennes, on pourrait penser que l'information est ce qui se produit dans les synthèses en tant que « différence soutirée à la répétition »⁴⁰, et qui est à la base de tout processus d'individuation. Comme le dit Simondon, « elle [l'information] est la signification qui surgira lorsqu'une opération d'individuation découvrira la dimension selon laquelle deux réels disparates peuvent devenir système. » (Simondon, 1964, p. 15)

La question qui s'en suit est alors « comment comprendre les différentes synthèses dans une perspective simondonienne de l'information ? » Autrement dit, est-ce que différents types d'information résultent dans différentes synthèses (celle de l'habitude et celle de la mémoire) ? Certains éléments de réponse peuvent être trouvés dans les réflexions connexes portant sur les rapports entre information et les différentes

⁴⁰ Il est important de remarquer néanmoins que le sens donné ici au mot « information » ne correspond pas à celui que Deleuze lui-même a pu lui attribuer, avec une emphase sur l'aspect culturel et sociétal de l'information. Dans sa conférence « Qu'est-ce que c'est l'acte de création ? », par exemple, Deleuze dira : « Or une information, c'est quoi ? C'est pas très compliqué, tout le monde le sait : une information, c'est un ensemble de mots d'ordre. (...) Ce qui revient à dire : que l'information, c'est exactement le système du contrôle. » (Deleuze, 1987)

interprétations des probabilités. Atlan fait référence au travail de Costa de Beauregard pour aborder le problème du rôle de l'information dans le passage des probabilités objectives aux subjectives :

Le passage du manque d'information au désordre réel est en fait, comme le fait remarquer O. Costa de Beauregard, un passage d'une probabilité subjective à une probabilité objective. C'est le vieux problème de physique sur le caractère subjectif ou objectif du hasard qui revient ainsi : existe-t-il un hasard réel, en dehors de notre ignorance ? (Atlan, 2006 [1972], p. 180)

Comme le remarque Atlan, Costa de Beauregard construit une « théorie de l'organisation comme transformation d'information en néguentropie » (Atlan, 2006 [1972], p. 194). Par cette perspective, Atlan aborde le rôle joué par la néguentropie (variation négative de l'entropie) et l'information dans le passage des probabilités objectives aux subjectives (ou dans la perspective de cette thèse, le passage des synthèses passives aux synthèses actives) :

La conclusion de cet auteur sur ce point, à laquelle nous pouvons partiellement nous rallier est « qu'il doit être fondamentalement de la nature de la probabilité (*néguentropie* et *information*) d'être *opérativement* cette charnière entre l'objectif et le subjectif, entre la matière et le psychisme, par où l'un se noue à l'autre et réagit sur l'autre » (Atlan, 2006 [1972], p. 206)

Pour de Beauregard, l'information correspond à une diminution de l'entropie, ou à une néguentropie, car elle implique une réduction de l'incertitude liée au système. Selon de Beauregard, la néguentropie est l'aspect physique de l'incertitude (par exemple, les différentes configurations des particules d'un système, ou ses micro-états), et l'information est cette même incertitude lorsqu'elle est observée ou mesurée. De Beauregard reproduit ainsi, avec les notions d'information et de néguentropie, le dualisme objectif/subjectif, l'associant à un autre dualisme, matière/esprit. Mais si de Beauregard place la néguentropie du côté objectif et l'information du côté subjectif, Atlan préfère quant à lui affirmer que ces deux éléments constituent ensemble cette charnière entre les deux pôles : « C'est pourquoi nous ne pouvons pas accepter l'idée

d'associer l'information au psychisme et la néguentropie à la matière : nous perdons là toute la richesse et la force de la théorie de l'information comme réalité physique (...). » (Atlan, 2006 [1972], p. 198) Atlan rappelle que l'une des grandes contributions de la théorie de l'information de Shannon est de traiter de l'aspect matériel de l'information.

Prigogine et Stengers, dans *La nouvelle alliance*, constatent que la relation établie entre information et entropie a collaboré avec une interprétation subjectiviste de l'entropie : « [s]elon cette interprétation, la croissance d'entropie ne décrit pas le système lui-même, mais seulement notre connaissance du système. » (Prigogine & Stengers, 1979, p. 285) On voit se reproduire dans le contexte de l'entropie le même schisme entre les interprétations objectives et subjectives de la théorie des probabilités. Comme remarquent les auteurs,

[c]ette interprétation subjectiviste de l'irréversibilité comme croissance de l'ignorance (encore renforcée par l'analogie ambiguë avec la théorie de l'information) fait de l'observateur le vrai responsable de l'asymétrie temporelle qui caractérise le devenir du système. (Prigogine & Stengers, 1979, p. 286)

Néanmoins, Prigogine et Stengers contestent cette interprétation. Comme remarquent les auteurs, lorsque le rôle organisateur de l'entropie (dans certains systèmes) a été identifié, il devient impossible d'interpréter l'entropie en termes d'ignorance d'un observateur. L'interprétation subjective réduit l'effet de l'entropie à une question de connaissance et d'observation, alors que l'on remarque dans les structures dissipatives des effets d'organisation qui se traduisent matériellement dans le système. Autrement dit, admettre l'interprétation subjective implique accepter que toute manifestation d'organisation des êtres vivants ne soit que « des illusions produites par l'imperfection de nos sens », une « simple apparence liée à notre ignorance. » (Prigogine & Stengers, 1979, p. 289), position qui n'est pas celle de Prigogine et Stengers.

Deleuze commente les écrits de Boltzmann à propos de l'entropie pour remarquer qu'elle implique une évolution des systèmes qui va nécessairement d'un état imprévisible à une situation d'équilibre uniforme :

Des pages célèbres de Boltzmann commentent cette garantie scientifique et thermodynamique du bon sens ; elles montrent comment dans un système partiel, s'identifient, d'une part, passé, improbable et différence, d'autre part, futur, probable et uniformité. (Deleuze, 1993 [1968], p. 291)

Il remarque qu'à la fin du 19^e siècle une alliance s'est construite entre la science, la philosophie et le bon sens. Cette alliance tournait autour de l'idée d'une tendance universelle vers l'égalisation des différences, du bon sens qui mène à un état d'équilibre uniforme et stable. Pour Deleuze, le bon sens, ou l'augmentation d'entropie jusqu'à l'état d'équilibre, est fondé sur la synthèse passive, celle de l'habitude (Deleuze, 1993 [1968], p. 290) Cela renforce la proposition formulée plus haut d'associer les synthèses passives aux probabilités dites objectives, et nous éloigne de l'interprétation subjectiviste de l'entropie.

Plus récemment, le biophysicien Jeremy England a repris les travaux sur l'auto-organisation et l'entropie dans les systèmes thermodynamiques loin de l'équilibre pour aborder le problème fondamental de l'émergence de la vie par la perspective de la physique. Comme England affirme, « living things always operate far from equilibrium » (England, 2015, p. 919). Il propose d'utiliser la mécanique statistique de systèmes loin de l'équilibre pour mieux comprendre l'organisation biologique, autrement dit, d'utiliser des principes de la physique pour contribuer à construire un modèle de l'organisation de systèmes complexes (« multi-body »). Cette démarche nous intéresse dans le contexte de cette thèse, car les mécanismes physiques décrits par England permettent d'aborder certaines particularités du processus de synthèses (passives et actives) développé plus haut.

England observe trois caractéristiques physiques essentielles des systèmes vivants : « 1. self-replication; 2. Sensing, computation and anticipation; 3. Effective absorption

of work from environment » (England, 2014) Il part de ces caractéristiques et développe la notion de « dissipative adaptation ». Je copie ici une longue citation qui décrit le processus d'adaptation dissipative, typique de systèmes auto-organisés :

the absorption of energy from a drive allows the system configuration to traverse activation barriers too high to jump rapidly by thermal fluctuation alone, and if this energy is dissipated after the jump, then it is not available to help the system go back the way it came. Thus, while any given change in shape for the system is mostly random, the most durable and irreversible of these shifts in configuration occur when the system happens to be momentarily better at absorbing and dissipating work. With the passage of time, the 'memory' of these less erasable changes accumulates preferentially, and the system increasingly adopts shapes that resemble those in its history where dissipation occurred. Looking backward at the likely history of a product of this non-equilibrium process, the structure will appear to us like it has self-organized into a state that is 'well adapted' to the environmental conditions. This is the phenomenon of dissipative adaptation. (England, 2015, p. 922)

Ainsi, « l'adaptation dissipative » est un effet des tendances dans les formes, ou les états, qu'un système favorise étant donné son histoire. Il s'agit de la « mémoire » développée par un système des configurations qui favorisent « l'absorption de travail » et l'anticipation. Nous sommes ici très proches à nouveau de la perspective pragmatiste spéculative de la probabilité et des synthèses deleuziennes, où il était question, on se souvient, de l'habitude, du *would-be* (Charles S. Peirce, 1957), des propensions d'un système (Gillies, 2000), ou encore de la « différence soutirée de la répétition » (Deleuze, 1993 [1968]) dans une synthèse. Le constat d'England, dans cette perspective, réaffirme la probabilité comme une modélisation des tendances développées à partir de l'évolution d'un processus, et contribue à bâtir une description physique des synthèses deleuziennes, ou de la propension indiquée par une probabilité. L'adaptation dissipative décrit les mécanismes physiques de la propension d'un système mesurée par les probabilités de ces états organisationnels, autrement dit, les mécanismes physiques à la base de la production d'information, de la mémoire, de l'organisation et de la propension d'un système.

Si nous poursuivons le rapprochement entre l'adaptation dissipative et la synthèse deleuzienne, nous constatons que les synthèses réalisées dans des systèmes proches de l'équilibre se ressemblent aux synthèses passives, dans le sens où il s'agit d'une série d'événements indépendants modélisée par les probabilités non conditionnelles. Autrement dit, lorsqu'il s'agit de processus de synthèses où les mécanismes de mémoire ou d'interdépendance sont présents, nous nous situons dans le pôle des synthèses actives, et les modèles plus appropriés dans la modélisation de ces processus sont ceux qui intègrent les probabilités conditionnelles. Dans ce sens, les synthèses passives peuvent être décrites comme les propensions des systèmes plus proches d'une situation d'équilibre et les synthèses actives comme celles des systèmes dissipatifs loin de l'équilibre.

À première vue, il semblerait exister une contradiction entre les résultats des travaux sur l'auto-organisation, abordée principalement par les perspectives de England, Atlan, Prigogine et Stengers, et la notion de « centre d'indétermination » élaborée par Henri Bergson et étudiée dans le chapitre précédent. Or, si l'organisation (ou la création d'information) est à la base des processus de synthèses des systèmes complexes loin de l'équilibre, comment pouvons-nous affirmer, d'après Bergson, que « (...) le rôle de la vie est d'insérer de l'indétermination dans la matière » (Bergson, 2008 [1939], p. 137) ? Pour clarifier ce point, il est important de rappeler que pour Bergson, perception et action sont des mécanismes liés dans les systèmes vivants, car la perception est toujours rapportée aux actions possibles. Ainsi, perception et action sont les deux faces d'un processus qui englobe les synthèses organisationnelles (créatrices d'information) et les interactions (créatrices d'indétermination). Cette question sera élaborée dans la prochaine section lorsque je ferai un retour sur la contribution de la théorie algorithmique de l'information à la compréhension des notions telles que le hasard, la complexité, l'information et l'entropie.

3.3. Théorie algorithmique de l'information : hasard et complexité

Pendant les années 1960, de manière indépendante, Gregory Chaitin, Andrei N. Kolmogorov et Ray Solomonoff ont établi les fondations de ce que l'on appelle aujourd'hui la théorie algorithmique de l'information (Grunwald & Vitanyi, 2004). À partir de la notion de calculabilité (développée par Alan Turing), des travaux sur l'incomplétude de Gödel, de la théorie de l'information de Claude Shannon et de la complexité telle que définie par Leibniz, Chaitin, Kolmogorov et Solomonoff ont proposé le concept de la *complexité algorithmique*, concept à la base de la théorie algorithmique de l'information qui place le hasard au cœur de l'informatique. Sur l'impact et l'étendue de la notion de complexité algorithmique, Hervé Zwirn affirme que :

La complexité algorithmique fait aujourd'hui partie des outils standards des mathématiques et de l'informatique, et nul mathématicien ou informaticien théorique ne l'ignore. Mais au-delà, c'est une notion extrêmement fertile qui est désormais utilisée dans beaucoup de domaines : des physiciens l'ont reprise pour une reconstruction de la thermodynamique ; elle joue un rôle fondamental lorsqu'il s'agit de définir les « phénomènes émergents dans les systèmes complexes » ; elle a servi de base pour tenter de mesurer l'évolution en biologie ou pour construire d'autres mesures de la complexité (...). (Zwirn, Hervé, in: Chaitin, 2005, p. 11)

La complexité algorithmique, aussi connue sous les appellations de complexité de Kolmogorov ou complexité de Kolmogorov-Chaitin⁴¹, peut être comprise intuitivement selon l'idée qu'un objet est d'autant plus complexe que l'algorithme nécessaire pour le décrire sera long. Dans le contexte numérique, une description peut être comprise comme une suite binaire, ou un message. Prenons l'exemple de cette

⁴¹ Les expressions utilisées pour nommer ce concept et cette théorie sont nombreuses, comme le rappellent Li et Vitányi : « The theory dealing with the quantity of information in individual objects goes by names such as 'algorithmic information theory', 'Kolmogorov complexity', 'K-complexity', 'Kolmogorov-Chaitin randomness', 'algorithmic complexity', 'stochastic complexity', 'descriptive complexity', 'minimum description length', 'program size complexity' and others. » (Li & Vitányi, 2008, p. vi)

suite binaire de deux cents caractères : 0101010101010101... 01. Il est possible de la réduire à la formule « répéter 100 fois 01 », et de la décrire ainsi avec un nombre plus petit de bits (ou de caractères). Néanmoins, la séquence aléatoire 01011100110101010010 ne peut pas être compressée dans une formule plus petite, il n'existe pas d'algorithme capable de la calculer, on doit ainsi utiliser tous les 20 bits (qui correspondent à sa taille) dans sa description. C'est le cas de dire que la complexité algorithmique de la deuxième suite est supérieure à celle de la première. En effet, une suite aléatoire ne peut pas être compressée par un programme informatique, or pour exprimer une suite aléatoire il faut au moins la même quantité d'information (en nombre de bits) qui contient la suite elle-même, ce qui explique la relation étroite entre hasard et quantité d'information dans cette théorie.

De cette notion de complexité algorithmique, ressort la définition du hasard algorithmique : le hasard en tant que suite irréductible. Chaitin a défini ainsi un nombre aléatoire : « Un réel aléatoire est, par définition, un nombre réel dont les bits constituent l'information la plus irréductible, la plus incompressible possible (...). » (Chaitin, 2005, p. 141) Selon une perspective algorithmique, le hasard est alors une suite qui ne peut pas être réduite par aucun algorithme. La notion de hasard algorithmique est alors fortement reliée à celle de la calculabilité, ou plutôt aux limites de la calculabilité, car le hasard est justement ce qui ne peut pas être calculé et qui échappe à tout formalisme mathématique. Ce résultat démontre par une approche différente l'incomplétude déjà prouvée par Kurt Gödel :

La grande idée de Turing fut d'introduire le concept de calculabilité — en établissant une distinction entre ce qui se peut calculer et ce qui ne le peut pas — et de faire découler l'incomplétude de l'incalculabilité. L'incalculabilité est une cause profonde de l'incomplétude. (Chaitin, 2005, p. 64)

Il est intéressant de reprendre ici le travail de Henri Atlan sur l'organisation biologique pour le mettre en relation avec la complexité algorithmique. Comme l'avait dit Atlan :

(...) un objet a le plus haut degré de complexité, s'il peut faire des choses très difficiles et très élaborées... On arrive ainsi à cette proposition tout à fait décisive de la complexité, qu'il existe une taille critique au-dessous de laquelle le phénomène de synthèse est dégénératif, mais au-dessus de laquelle le phénomène de synthèse, arrangé comme il faut, peut devenir explosif... (Atlan, 2006 [1972], pp. 218-219)

La complexité biologique, ainsi que la complexité algorithmique, se rapporte ainsi au degré d'organisation du système ou de l'organisme qui opère des synthèses. L'arrangement dont il est question dans les systèmes auto-organiseurs décrits par Atlan et d'autres (Prigogine, Stengers et England) est celui qui permet les synthèses explosives, qui permet de construire des synthèses actives sur les fondations des synthèses passives.

Dans le chapitre suivant, je reviendrai plus en profondeur sur les relations entre hasard, calculabilité et algorithme, pour aborder les algorithmes probabilistes utilisés dans le documentaire interactif réalisé dans le cadre de la présente recherche-création. Avant cela, il est important néanmoins de revenir sur les rapports entre l'information de Shannon et l'information algorithmique pour être en mesure de saisir l'évolution de la relation entre le hasard et le numérique.

Comme plusieurs auteurs l'ont souligné, la complexité algorithmique est une notion proche de celle de l'entropie telle que développée dans la théorie de l'information de Shannon. Dans l'article fondateur de Kolmogorov, cette relation est mise en évidence lorsqu'il propose de redéfinir l'entropie en termes algorithmiques : « Entropy $H(x|y)$ is the minimal length of the recorded sequence of zeros and ones of a 'program' P that permits construction of the value x , the value of y being known (...) » (Kolmogorov, 1968, p. 662) On remarque alors une différence fondamentale entre ces deux approches : si dans la théorie de Shannon l'entropie est convoquée pour quantifier l'information dans un message transmis à partir de la distribution des probabilités de la source d'information, la théorie algorithmique se propose à calculer la quantité

d'information contenue dans la description du message. Comme résumant Grunwald et Vitanyi :

Shannon ignores the object itself but considers only the characteristics of the random source of which the object is one of the possible outcomes, while Kolmogorov considers only the object itself to determine the number of bits in the ultimate compressed version irrespective of the manner in which the object arose. (Grunwald & Vitanyi, 2004, p. 4)

Ainsi, dans la théorie de Shannon, l'information ne peut être calculée que dans le cas où elle est émise par une source et transmise par une voie, et grâce à un recours aux probabilités. L'information algorithmique est quant à elle calculée par rapport à un algorithme. Ce sont des théories complémentaires, et non contradictoires, comme le souligne Jon Machta :

The Shannon information is calculated from a definite formula involving probabilities. It is not applicable to a single sequence of ones and zeros but only to a statistical ensemble of such sequences. Algorithmic information content is applicable to a single sequence, but cannot be calculated by a formula or any definite procedure because it depends on finding the best way to compress the data. (Machta, 1999, p. 1075)

Ces observations mettent en évidence l'aspect relationnel de la mesure de l'information, à la fois dans la théorie de Shannon et dans la théorie algorithmique. Le travail de Simondon a mis de l'avant l'aspect relationnel de l'information dans le système de communication décrit par Shannon (dans une citation précédente de Simondon, on pouvait lire : « Il n'y a information que lorsque ce qui émet les signaux et ce qui les reçoit forme système » (Simondon, 1964, p. 255)). Or, dans la théorie algorithmique de l'information il est aussi question d'un aspect relationnel de l'information (malgré le fait qu'il ne s'agisse pas nécessairement d'une transmission) : l'information ne peut être mesurée que dans sa relation avec un programme informatique, ou un algorithme. Il est important de souligner que cet aspect relationnel n'est pas lié au langage choisi pour la description, ce qu'on appelle « l'indépendance de la méthode de description » (Li & Vitányi, 2008, p. 3), ou l'« indépendance de la

machine universelle » (Delahaye, 1999, p. 17) restent valides⁴². C'est dans le présupposé d'une description, ou d'un algorithme, que l'aspect relationnel se fonde. Le terme de *complexité descriptionnelle* fait référence à l'algorithme nécessaire pour créer et traiter cette description.

L'aspect relationnel de l'information se manifeste également dans une autre notion au cœur à la fois de la théorie de l'information de Shannon, de la théorie algorithmique et de la théorie des probabilités : la notion de suite aléatoire. En effet, dans les théories de l'information, la quantité d'information est calculée dans une suite d'éléments (le message). Par ailleurs, le calcul des probabilités se fait d'après une suite d'événements. Au cœur de ces trois théories, on retrouve ainsi la notion de suite aléatoire, soit dans la forme d'une suite d'éléments qui composent un message (théories d'information), soit dans la forme d'une suite d'événements aléatoires (théorie des probabilités). Dans la section dédiée aux interprétations de la théorie des probabilités, la définition d'une suite aléatoire a été formulée, par la perspective de Popper, en tant qu'indépendance des termes de la suite. Je propose de revisiter la notion de suite aléatoire par la perspective de l'information algorithmique.

L'évolution du problème de la définition mathématique d'une suite aléatoire recouvre et lie de nombreux concepts et domaines abordés dans cette thèse, tels que le hasard, l'algorithme, l'information, la calculabilité, entre autres :

L'identification mathématique de la notion intuitive d'algorithme a finalement conduit (...) à une autre identification mathématique importante, celle de suite infinie aléatoire, tout cela au travers d'un processus d'élaboration complexe qui s'est déroulé sur plus de cinquante ans (de 1919 à 1975). (Delahaye, 1999, p. 30)

⁴² Il s'agit d'un résultat important de la théorie informatique l'affirmation selon laquelle la complexité de Kolmogorov est une formalisation qui ne dépend pas du langage choisi pour la description : "Its complexity is an objective attribute of the described object alone: it is an intrinsic property of that object, and it does not depend on the description formalism." (Ming Li, p. 2)

La formulation du problème de la définition mathématique d'une suite aléatoire est simple : pour une série donnée, comment déterminer si elle est aléatoire ? Autrement dit, comment déterminer si une suite est le résultat d'un procédé aléatoire, tel que les lancers d'une pièce de monnaie ? Intuitivement, il est évident qu'une très longue suite composée uniquement de 1 (dans le cas d'une suite binaire) difficilement est le résultat d'un procédé aléatoire (tel que les lancers d'une monnaie). Néanmoins, certaines suites peuvent avoir une apparence aléatoire, mais avoir une logique qui les rend prévisibles. Prenons ici l'exemple de la suite de décimales de π : il s'agit d'une suite qui est aléatoire en apparence, mais qui peut néanmoins être calculée, et donc prévue (dans la limite de la puissance de calcul de la machine), ce qui l'exclue de la catégorie de suite aléatoire.

La définition mathématique d'une suite aléatoire a été recherchée initialement en tant que fondement des bases théoriques des probabilités, comme le dit Martin-Löf (cité dans Delahaye) : « von Mises soutenait qu'une théorie mathématique des probabilités ne pouvait être basée que sur une définition de la notion de suite aléatoire (...). » (Martin-Löf in Delahaye, 1999, pp. 40-41) Chaitin a aussi souligné le rôle de la théorie algorithmique de l'information dans la fondation de la théorie des probabilités : « The algorithmic definition of randomness provides a new foundation for the theory of probability. » (Chaitin, 1975, p. 49) Par ailleurs, Kolmogorov a également souligné les liens entre les fondations de la théorie des probabilités et les développements des théories de l'information : « La théorie de l'information doit précéder et non pas se baser sur la théorie des probabilités. Par sa nature même, les fondements de la théorie de l'information ont un caractère combinatoire. » (Kolmogorov in Delahaye, 1999, p. 42) Dans cette citation, Kolmogorov semble porter une critique à la théorie de Shannon dans le sens où le calcul de la quantité de l'information se fait d'après les distributions de probabilité (dans la théorie de Shannon), alors que la théorie algorithmique permettrait de définir les bases d'une théorie des probabilités : une suite est aléatoire quand aucun algorithme plus petit que la suite elle-même est capable de la produire.

Ainsi, selon la définition algorithmique du hasard, les probabilités (ou la quantité d'information) se calculent en fonction de la relation existante entre les éléments qui composent une suite. Cette définition n'est pas en contradiction avec celle proposée par Popper (en termes d'indépendance entre les éléments de la suite) dans la mesure où ce que fait un algorithme lorsqu'il contracte une suite est effectivement trouver les principes qui mettent en relation les éléments d'une suite. Un échec de l'algorithme pour trouver un principe qui régit la relation entre ces éléments signifie que les termes sont indépendants, et donc aléatoires.

Dans la continuation de la démarche fondamentale de cette thèse, je propose une lecture pragmatiste spéculative de la notion algorithmique de suite aléatoire. Plus précisément, il sera question de s'attaquer à la question : comment interpréter la théorie algorithmique de l'information à la lumière des synthèses deleuziennes ? Or, le parallèle avec les synthèses deleuziennes est flagrant : dans le traitement proposé d'une suite aléatoire, autant par la théorie algorithmique de l'information que par la théorie des probabilités, il est question à nouveau d'une répétition de termes indépendants et d'une contraction opérée entre les termes de la suite (la « différence soutirée de la répétition »). Dans cette perspective, l'information est processuelle, car elle n'est pas une instance fixe, mais plutôt le résultat d'une opération de contraction, ou de synthèse. Par ailleurs, l'information est aussi relationnelle dans la mesure où elle n'existe que par rapport à un système qui opère la synthèse (dans ce cas, des algorithmes) et aux relations entre les multiples éléments qui composent la suite à être contractée. L'information serait ainsi le résultat d'un processus de contraction effectué par un algorithme.

Par ailleurs, l'aspect spéculatif des algorithmes informatiques a été abordé dans le travail de Luciana Parisi. Dans son livre *Contagious Architectures* (2013), Paris articule la philosophie processuelle de Whitehead et la théorie de l'information algorithmique pour réfléchir sur l'interactivité et l'informatique. Elle utilise la définition du hasard proposée par la théorie algorithmique de l'information pour étudier

ce qu'elle appelle la « culture informatique » (*programming culture*) et affirme que « [r]andomness has become the condition of programming culture. » (Parisi, 2013, p. ix) Parisi considère l'indétermination au sein des systèmes informatiques, et non seulement dans l'interaction des dispositifs avec des éléments externes. Selon Parisi, aujourd'hui il n'est plus possible de considérer les algorithmes comme des codes déterministes, ou encore considérer que les instructions suivies par une machine aboutiront toujours aux mêmes résultats. Les algorithmes traitent les données de manière autonome, ils sont des « (...) actualities that select, evaluate, transform and produce data. » (Parisi, 2013, p. ix). Parisi affirme que l'indétermination (*randomness*) dans un système n'est pas exclusive à l'interaction de celui-ci avec l'extérieur, mais elle est aussi présente à l'intérieur du système, dans ce cas, le système informatique d'algorithmes : « Entropic data are operative agents of irreducible size that crack and rescript the source program of the system from within. » (Parisi, 2013, p. ix).

Ce nouveau statut de l'algorithme mène Parisi à repenser l'interactivité. L'interactivité, comprise dans le cadre cybernétique comme l'adaptation aux interventions extérieures, ne considère pas que l'indétermination, ou la source de l'intervention qui génère l'adaptation, peut se trouver à l'intérieur même du système, « (...) interaction now includes a computational tendency to anticipate responses and programming indeterminacies that stem from within the system, and need not rely on direct sense perception and cognition. » (Parisi, 2013, p. 27). Ainsi, une couche supplémentaire d'indétermination pourrait être intégrée aux médias interactifs : l'indétermination inhérente aux algorithmes précédemment décrits.

Parisi arrive à la notion de *speculative computation*, très proche de la démarche proposée dans cette thèse. Elle part du problème de la nouveauté dans les modèles prédictifs pour ensuite formuler l'aspect spéculatif des systèmes d'algorithme : « Computational design problematically embraces the logic of prediction and the calculation of probabilities, and is unable to explain novelty in spatiotemporal experience. » (Parisi, 2013, p. 9) Comme déjà abordé dans cette thèse, les algorithmes

prédictifs se basent sur les données historiques pour essayer de prédire les résultats futurs, ou au moins en estimer les probabilités. Néanmoins, la nouveauté (ou ce qui échappe aux patterns du passé) ne peut pas être calculé par ce genre de modèle prédictif. En réponse à cette impasse, Parisi propose de penser les algorithmes en termes d'entités spéculatives :

Computation (...) is an instance of speculative reason, since it no longer nor exclusively aims at the prediction or calculation of probabilities. On the contrary, Whitehead's understanding of speculative reason explains that the function of reason is to add new data to what will always already happen in an efficient chain of cause and effect. (...) This means that a notion of speculative computing is not concerned with quantifying probabilities to predict the future, but with including random or patternless quantities of data in sequential calculation so as to add novelty in the actual architecture of things. (Parisi, 2013, p. 9)

Ainsi, Parisi met de l'avant l'aspect spéculatif des algorithmes grâce à l'inclusion du hasard dans les systèmes informatiques. Il est important de souligner encore une fois qu'il ne s'agit pas de capter du hasard dans le monde physique et de l'inclure dans un algorithme, mais plutôt de créer de systèmes informatiques qui se réorganisent en fonction de données d'une complexité maximale. Dans le prochain chapitre, je reviendrai sur le développement historique de l'utilisation du hasard dans les algorithmes, partant des machines de Turing probabilistes pour arriver aux algorithmes d'intelligence artificielle utilisés dans la partie de création de cette thèse.

4. CHAPITRE IV

MACHINE DE TURING PROBABILISTE, ALGORITHMES RANDOMISÉS ET MACHINE LEARNING

Dans ce chapitre, il s'agira de réfléchir de manière plus rapprochée au fonctionnement des algorithmes utilisés dans la création du documentaire interactif *On Chance*. L'interactivité de ce documentaire prend forme dans une interface de conversation humain-machine, un *chatbot*. Dans les dernières années, la création et l'utilisation des *chatbots* ont beaucoup augmenté, notamment en raison du développement du *Natural Language Processing* (NLP, en français Traitement automatique du langage naturel) et de l'apprentissage automatique (*machine learning*, en anglais). L'apprentissage automatique est une notion large qui regroupe plusieurs techniques pour la création et implémentation d'algorithmes capables d'exécuter des fonctions pour lesquelles ils n'ont pas été directement programmés (Samuel, 1959). Dans le cas du traitement automatique du langage naturel, il s'agit d'algorithmes capables d'analyser et traiter les entrées de texte ou de voix en langage humain (autrement dit, pas en langage informatique). Comme nous verrons, la plupart des algorithmes d'apprentissage automatique sont basés sur des modèles probabilistes et d'inférence. Il sera question alors de réfléchir sur les modes d'articulation du hasard et de la complexité dans ces algorithmes.

Néanmoins, avant d'aborder les algorithmes utilisés dans le cadre de cette thèse, il est nécessaire de retracer sommairement l'évolution de l'utilisation du hasard et de la complexité dans les algorithmes informatiques. Mitzenmacher et Upfal comparent la propagation de l'utilisation du hasard en informatique au mouvement similaire qui a eu lieu dans les sciences au cours du 20e siècle :

Science has learned in the last century to accept randomness as an essential component in modeling and analyzing nature. (...) Computer science is no exception. (...) The last two decades have witnessed a tremendous growth in the use of probability theory in computing. (Mitzenmacher & Upfal, 2005, p. xiii)

Les modèles probabilistes de calcul et les algorithmes randomisés sont largement utilisés aujourd'hui pour analyser de grandes bases de données et faire des inférences par rapport à ces données. Autrement dit, les algorithmes probabilistes sont à la base du développement du traitement massif de données (Big Data) et de l'apprentissage automatique. L'utilisation de modèles probabilistes et statistiques permet d'intégrer de l'incertitude et du bruit présents dans les larges quantités de données analysées dans les systèmes informatiques. Dans ce chapitre, je poursuivrai la réflexion sur les interprétations des modèles probabilistes utilisés en informatique, mais plutôt qu'étudier les cadres théoriques portant sur l'information (l'objet du chapitre précédent), je me concentrerai sur les manifestations de ces modèles dans les algorithmes utilisés dans le développement du documentaire interactif *On Chance*. Dans un premier moment, je ferai un bref retour sur l'origine de l'intégration du hasard et de la probabilité dans les modèles algorithmiques, pour ensuite aborder des questions plus contemporaines liées à l'approche probabiliste de l'apprentissage automatique et des algorithmes utilisés dans le *chatbot* développé dans la partie de création de la présente thèse par la perspective du pragmatisme spéculatif.

4.1. Machines de Turing non déterministes, Machines de Turing probabilistes et algorithmes randomisés

En 1936, Turing a publié l'article fondateur « On Computable Numbers, With an Application to the Entscheidungsproblem⁴³ » (Turing, 1936). Dans cet article, Turing

⁴³ *Entscheidungsproblem* est le terme en allemand pour désigner le « problème de la décidabilité ». Il s'agit d'un problème mathématique formulé par David Hilbert qui consiste à déterminer s'il est possible de construire un système logique (algorithme) capable de vérifier la validité de tout énoncé soumis à ce système. (Chaitin, 2005)

propose le modèle d'une « computing machine », qu'on appelle aujourd'hui la Machine de Turing. La machine de Turing est définie comme un modèle abstrait qui représente une machine capable de calculer des problèmes décrits par une procédure algorithmique. La machine de Turing comprend essentiellement trois éléments : un ruban infini avec des cases où des données sont codées en langage binaire (un symbole par case) ; une tête de lecture qui scanne le ruban case par case ; et une table de transition (l'algorithme de la machine). Cette première version de la Machine de Turing est déterministe, car pour chaque entrée il n'y a qu'un seul résultat possible en sortie. Pour être en mesure d'aborder les Machines de Turing non déterministes et probabilistes, il est essentiel d'aborder plus en profondeur la notion de calculabilité et sa relation avec la complexité algorithmique.

La notion de calculabilité partage les problèmes mathématiques en deux groupes distincts, ceux qui peuvent être résolus par un algorithme en temps fini (quel que soit le temps d'exécution), et ceux qui restent indécidables (temps infini). Lorsque l'on se concentre exclusivement sur ce premier groupe, donc les problèmes pour lesquels une machine fournira toujours un résultat (un arrêt), la notion de complexité algorithmique est aussi utilisée pour calculer la performance et l'efficacité d'un algorithme et évaluer les ressources (temps et mémoire) nécessaires à son exécution. Dans l'analyse de la performance des algorithmes, la complexité est divisée en différentes classes, en rapport au temps de calcul nécessaire ou à l'espace (mémoire) utilisé par l'algorithme. Je vais me concentrer uniquement sur les considérations sur la complexité liée au temps, car cela permet d'avancer vers la question qui nous intéresse dans le cadre de cette thèse, soit le fonctionnement des machines de Turing non déterministes et probabilistes.

Le temps d'exécution d'un algorithme est mesuré en nombre d'étapes nécessaires jusqu'à l'obtention d'un résultat (et non en unité temporelle) et de forme asymptotique (ce qui veut dire que la complexité est toujours calculée lorsque la taille des entrées tend vers l'infini). Le nombre d'étapes que prendra l'algorithme pour la résolution d'un

problème détermine sa classe de complexité. Voici une image inspirée d'une proposition d'Erik Demaine (Demaine, Devadas, & Lynch, 2015) pour ordonner les principales classes de complexité :

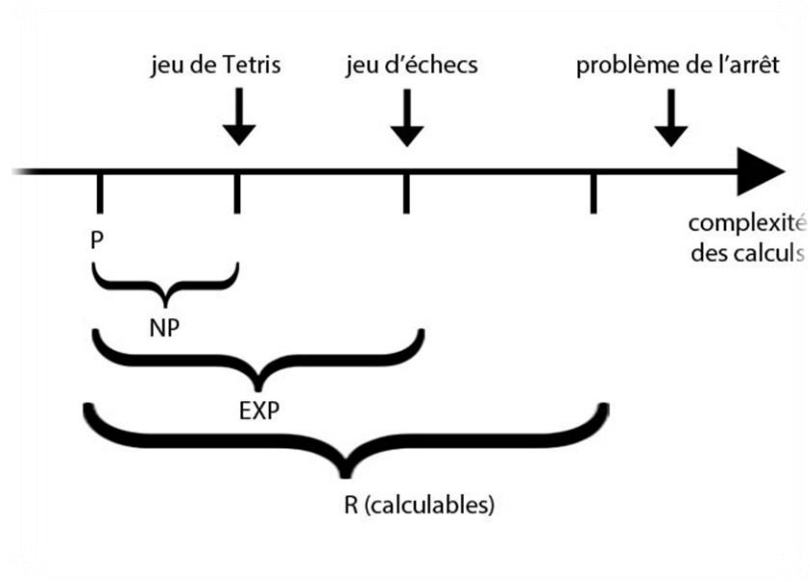


Figure 6 : Classes de complexité algorithmique

La première classe dont il sera question est la classe P (*Polynomial time*), soit l'ensemble des problèmes résolus en temps polynomial, autrement dit, des problèmes dont le nombre d'étapes pour leur résolution (complexité O) est une fonction de la taille du problème (n) avec des coefficients et des puissances constants (k), donc $O(n^k)$. La classe EXP (*EXponential*) regroupe tous les problèmes qui peuvent être calculés avec un temps exponentiel (des fonctions avec des puissances variables et qui augmentent selon la taille de l'entrée), donc $O(k^n)$. Il est important de souligner que la classe P est considérée comme celle des problèmes résolubles avec des algorithmes efficaces, car le calcul en temps polynomial est considérablement plus rapide qu'en temps exponentiel, une affirmation qui est résumée dans la formule souvent répétée dans les cours d'informatique « *polynomial good, exponential bad* ». Cela veut dire que les

algorithmes en P sont beaucoup plus efficaces et rapides que les algorithmes en EXP. De façon inverse, il est possible d'affirmer que les algorithmes en EXP sont beaucoup plus complexes que les algorithmes en P, comme démontré dans le diagramme ci-dessus.

La classe NP (*Nondeterministic Polynomial time*) est celle des problèmes de décision (problèmes avec une réponse d'un bit, par exemple « oui » ou « non ») qui sont résolus par une machine de Turing non déterministe en temps polynomial (la notion de machine de Turing non déterministe sera élaborée plus loin). Il s'agit des problèmes pour lesquels il n'existe pas (ou on n'a pas encore trouvé) un algorithme pour résoudre le problème, mais pour lesquels il existe un algorithme capable de vérifier en temps polynomial une résolution donnée. Comme l'explique Sylvain Perifel,

Si P est la classe des problèmes pour lesquels on peut trouver une solution efficacement, NP est la classe des problèmes pour lesquels on peut *vérifier* une solution efficacement. (...) En d'autres termes, on nous donne une solution potentielle (un chemin acceptant) et on doit vérifier qu'il s'agit en effet d'une solution. (Perifel, 2014, pp. 57-58)

Un problème souvent utilisé comme exemple de la classe NP est celui du *Knapsack problem*, qui peut être résumé par la situation suivante : un voleur, équipé d'un sac de taille Y, est dans une maison et doit choisir les objets à mettre dans son sac de forme à optimiser son vol (obtenir le maximum de valeur dans la mesure que son sac le permet). Si ce problème est formulé en termes de problème de décision (par exemple, « est-il possible d'obtenir au moins une valeur X avec une limite de taille Y ? »), il n'y a pas d'algorithme capable de le résoudre de manière efficace (dans un temps polynomial) pour n'importe quelle grandeur d'entrée (autrement dit, pour un nombre arbitraire d'objets et de limite de taille). Néanmoins, il est possible de vérifier rapidement une solution donnée, il suffit d'exécuter l'algorithme avec les données proposées par la solution et s'assurer que la réponse satisfasse le problème. Il s'agit donc d'un problème de la classe NP. Le fait qu'il n'existe pas d'algorithme déterministe pour résoudre les problèmes de la classe NP n'implique pas qu'ils soient incalculables, et une des

solutions pratiques pour les résoudre ce sont les algorithmes randomisés (probabilistes) :

An embarrassing phenomenon for the classical worst-case complexity theory is that the problems it classifies as hard to compute are often easy to solve in practice. (...) Although these problems may be hard to solve on some set of pathological inputs, on most inputs (in particular, those that occur in real-life applications) the problem is actually easy to solve. (Mitzenmacher & Upfal, 2005, p. xiv)

Enfin, la classe R (*Recursive time*) se réfère à l'ensemble des problèmes calculables par les modèles de computation déterministes et non déterministes. Par le diagramme ci-dessus, il est possible de constater que le problème de l'arrêt, comme démontré par Turing, est indécidable, il n'est donc pas en R. De la même manière que Chaitin a conclu que la grande majorité des nombres est incalculable, en complexité il est aussi admis que la majorité des problèmes de décision n'est pas dans la classe R, autrement dit, la majorité des problèmes de décision est incalculable.

La notion de machine de Turing non déterministe est essentielle à la compréhension et à la définition des classes NP et R. En opposition aux machines déterministes, où chaque étape du calcul est déterminée uniquement par l'étape précédente, dans une machine de Turing non déterministe, plusieurs états sont possibles à la suite d'une étape du calcul, « [a]insi, plutôt qu'une suite de configurations, le calcul d'une machine de Turing non déterministe est un arbre de configurations puisqu'une configuration a plusieurs successeurs possibles. » (Perifel, 2014, p. 43) Dans ce modèle de calcul, le résultat est « deviné » par l'algorithme : si l'on essaie de visualiser l'exécution de l'algorithme comme un arbre de calcul, la machine choisit toujours le chemin qui mène au bon résultat. Il est important de remarquer que ce modèle de calcul est non déterministe (car ses états ne sont pas déterminés exclusivement par l'état précédent), mais aucun des éléments de l'algorithme n'est aléatoire, il n'y a pas ici d'utilisation de hasard (différemment de la machine de Turing probabiliste, comme nous le verrons). Autre observation importante est qu'il s'agit d'un modèle théorique, le développement

effectif d'un oracle capable de deviner le résultat correct de l'algorithme est considéré impossible.

Dans une machine de Turing probabiliste (ou randomisé) les étapes du calcul, tout comme dans une machine non déterministe, sont arborescentes, car à la suite de chaque étape plusieurs états sont possibles. Néanmoins, dans une machine probabiliste le choix des états se fait à l'aide d'une suite aléatoire. Dans cette version de la machine, un ruban avec une suite aléatoire de symboles est aussi lu par la machine. La résolution d'un calcul dans ce modèle est ainsi probabiliste. Dans le modèle déterministe, chaque entrée produit toujours le même résultat, or dans le modèle probabiliste, une même entrée produira des sorties différentes. Ce modèle de calcul informatique est à l'origine des algorithmes randomisés (ou probabilistes).

À la différence du modèle de la Machine de Turing non déterministe, les algorithmes randomisés sont facilement applicables aux ordinateurs et peuvent contribuer fortement à la performance d'un algorithme. Comme l'expliquent Mitzenmacher et Upfal :

Randomized algorithms are algorithms that make random choices during their execution. In practice, a randomized program would use values generated by a random number generator to decide the next steps at several branches of its execution. (Mitzenmacher & Upfal, 2005, pp. xiii-xiv)

Un des premiers exemples de l'utilisation de la randomisation dans un algorithme est celui de l'algorithme *Quicksort* (tri rapide), un algorithme très efficace pour trier des éléments qui opère selon le principe de « *diviser pour régner* » (repartir un problème en sous-problèmes pour le rendre plus abordable). Le tri rapide fonctionne par le choix d'un pivot, qui répartit les éléments à trier en deux groupes, et répète l'opération pour trier chaque sous-groupe, jusqu'à ce que l'ensemble des éléments soit ordonné. Dans la première version de cet algorithme, le pivot était toujours le premier élément de la série de l'entrée (*input*). Néanmoins, il a été constaté qu'il est plus efficace de choisir le pivot de manière aléatoire, car si le pivot divise les sous-groupes de manière très inégale (beaucoup d'éléments dans un des sous-groupes et très peu d'éléments de

l'autre côté), l'algorithme perd son efficacité. Le choix aléatoire du pivot augmente ainsi les probabilités d'une répartition efficace des sous-groupes, de sorte à réduire le temps de calcul. Les algorithmes randomisés peuvent ainsi être très efficaces dans la résolution des problèmes de la classe NP, néanmoins, cette efficacité est obtenue aux dépens d'une probabilité d'erreur, autrement dit, un algorithme randomisé peut aboutir à une mauvaise réponse au problème, ou alors n'aboutir à aucune résolution. Il s'agit donc d'un type d'algorithme où l'erreur est possible. Dans l'utilisation courante des algorithmes randomisés, les efforts sont dans le sens de diminuer au maximum la probabilité d'erreur.

Dans le cas des algorithmes randomisés abordés jusqu'ici, le hasard a son origine dans un générateur de nombre aléatoire et est utilisé notamment pour améliorer l'efficacité de l'algorithme. Il s'agit ainsi d'intégrer dans l'algorithme une suite aléatoire extérieure, représentée dans le modèle de la Machine de Turing probabiliste par un ruban supplémentaire contenant une suite aléatoire. Dans ce mode d'application des suites aléatoires aux algorithmes, le but n'est pas de prévoir le comportement des données, mais plutôt d'augmenter l'efficacité de l'algorithme.

Une autre manière d'intégrer le hasard dans les algorithmes est l'échantillonnage aléatoire d'une grande quantité de données. Dans *Randomized Algorithms*, Motwani et Raghavan affirment : « The idea that a random sample from a population is representative of the population as a whole is a pervasive theme in randomized algorithms. » (Motwani & Raghavan, 1995, p. xi) Ce principe de base de la statistique est ainsi largement utilisé dans la conception de certains algorithmes randomisés. Dans les dernières années, avec la grande quantité de données disponibles grâce aux technologies numériques d'information et de communication, le développement de familles d'algorithmes qui utilisent des modèles statistiques et probabilistes a beaucoup augmenté, comme il a déjà été argumenté plus haut dans cette thèse. Un domaine qui a beaucoup évolué dans cette perspective probabiliste est celui de l'apprentissage automatique.

4.2. Machine Learning (apprentissage automatique) : Algorithmes de Recast.AI (clustering), réseau neuronal artificiel et *probabilistic programming*

Dans un article fondateur du champ du Machine Learning, Arthur Samuel a publié les résultats obtenus de ses expériences dans le développement d'un algorithme capable de jouer aux dames (Samuel, 1959) et il a établi certaines caractéristiques des problèmes auxquels les techniques d'apprentissage automatique peuvent s'appliquer, la première étant qu'il n'existe pas d'algorithme déterministe efficace capable de résoudre le problème. Nous remarquons ainsi un point de départ commun entre les algorithmes randomisés et les modèles probabilistes d'apprentissage automatique : ils sont invoqués lorsque les algorithmes déterministes ne fournissent pas une solution satisfaisante. Aujourd'hui encore il est question d'une large utilisation de modèles probabilistes dans les techniques de ML. Kevin P. Murphy affirme que :

(...) the best way to make machines that can learn from data is to use the tools of probability theory (...). Probability theory can be applied to any problem involving uncertainty. In machine learning, uncertainty comes in many forms: what is the best prediction (or decision) given some data? What is the best model given some data? What measurements should I perform next? etc. (Murphy, 2012, p. xxvii)

Comme il a déjà été affirmé, des techniques d'apprentissage automatique sont utilisées dans la partie de création de cette thèse pour le développement d'un *chatbot*. Ainsi, dans cette section, il sera question d'étudier la perspective probabiliste des algorithmes d'apprentissage automatique utilisés dans le processus de création de *On Chance*.

L'apprentissage automatique est un sous-domaine de l'intelligence artificielle qui cherche à développer des méthodes pour inférer automatiquement des patterns dans des données d'entrée, et utiliser ensuite ces inférences pour faire des prévisions sur les données futures (Murphy, 2012, p. xxvii). Les techniques d'apprentissage automatique, ainsi que celles développées dans les recherches récentes sur l'intelligence artificielle, sont très souvent basées sur des modèles probabilistes (Murphy, 2012; Smith, 2016).

Comme explique Smith, un des principaux défis de l'intelligence artificielle est sa faible capacité à réagir à des événements inattendus. Dans la tentative de répondre à cela, les chercheurs et développeurs informatiques se sont tournés vers des modèles probabilistes pour simuler les incertitudes de l'environnement : « The primary method of coping with the unforeseen in modern AI systems is through probabilistic modelling and statistical learning. » (Smith, 2016, p. 5) Smith revient sur l'histoire des modèles d'intelligence artificielle pour les comparer aux modèles probabilistes utilisés aujourd'hui. Il propose une classification des approches de l'AI en deux types : déductif et inductif.

Dans les modèles déductifs, la première étape dans le développement des algorithmes est la formulation d'un certain nombre de règles logiques générales à propos d'un domaine d'intérêt. Ensuite, l'algorithme soumet des données d'entrée à ces règles et établit des conclusions à partir de ces recherches. Pour faciliter la compréhension du fonctionnement de ce type de modèle, je reprends un des premiers exemples d'application commerciale d'intelligence artificielle déductive, le logiciel MYCIN, un « système expert » qui sert à aider des professionnels du domaine de la santé dans le diagnostic des infections du sang (Smith, 2016, p. 8). Après une série d'entrevues avec des experts humains, les développeurs de MYCIN ont établi à peu près 600 règles de diagnostic du type : « SI le patient a la méningite ET la méningite est bactérienne ALORS il a staphylococcus avec une certitude 400 OU streptococcus avec une certitude de 200. » (formulation simplifiée et adaptée de Smith, 2016, p. 9) Une fois que le professionnel avait entré les données relatives à un patient, le logiciel était alors en mesure d'estimer (déduire) un taux de probabilité pour les diagnostics envisagés. Smith soulève deux difficultés principales rencontrées par cette approche de l'intelligence artificielle : la fragilité et la difficulté à raffiner les modèles. La fragilité concerne le fait que le système est incapable de donner un résultat satisfaisant si des facteurs qui n'avaient pas été prévus dans le développement entraînent en jeu. Un exemple de ce genre de situation était les résultats aberrants lorsque le patient présentait

d'autres symptômes qui n'étaient pas en relation avec ceux qui avaient été prévus lors du développement des algorithmes.

La deuxième difficulté, celle à raffiner les modèles, se traduit dans l'infaisabilité de rajouter un nombre des règles suffisant pour rendre compte de tous les scénarios envisageables. Cette difficulté est liée à la nature même d'un modèle, et peut être résumée dans la phrase célèbre de George E. P. Box, « Tous les modèles sont faux, mais certains sont utiles ». Cette question est aussi brillamment abordée par l'écrivain Jorge Luis Borges dans son texte *Del rigor en la ciencia* sur une carte à échelle 1:1. Comme explique Smith, le raffinement nécessaire aux programmes pour dépasser un seuil de fragilité (*brittleness*) est trop couteux, ce qui les rend impraticables, ou difficilement, à l'exploitation commerciale.

Les modèles inductifs regroupent des approches d'inférence basées sur la notion d'apprentissage, ou de construction de connaissance à partir de l'observation et de l'expérience. Dans ces approches, les algorithmes sont entraînés à partir d'une collection de données et travaillent à reconnaître des patterns, des répétitions, ou des cycles parmi les données. Un exemple de cette approche peut être trouvé dans les algorithmes de Recast.AI (ou Recast⁴⁴), utilisés dans le processus de création de *On Chance*. Recast.AI propose une plateforme collaborative qui permet le développement et le déploiement de chatbots basés sur le texte ou sur la voix. Pour traiter les entrées de l'utilisateur, les algorithmes de Recast.AI utilisent des techniques de traitement automatique de langage naturel (NLP). Voici un exemple de l'application de NLP dans Recast dans les métadonnées d'un agent conversationnel :

```
«source» : «hi»,
«intents» : [
  {
    «slug» : «greetings»,
    «confidence» : 0.99,
```

⁴⁴ Recast est une entreprise française dédiée à la création d'agents conversationnels, fondée en 2015, et récemment achetée par une importante entreprise de développement de logiciel, SAP.

```

    «description» : «Says hello»
  }
],
"act": "assert",
"type": null,
"sentiment": "vpositive",
"entities":
"language": «en»,
«processing_language» : «en»,

```

Dans ce cas, l'utilisateur a entré l'expression « hi » et grâce au traitement automatique de langage naturel, le modèle classe cette expression dans l'intention (*intent*) « salut » (*greetings*) avec un taux de confiance (*confidence*) de 99 %. Nous remarquons également que l'algorithme décrit l'expression comme « Says hello » et décrit le sentiment de l'expression comme « très positif » (*vpositive*). Il est intéressant de remarquer que l'algorithme détermine un taux de confiance en pourcentage, ce qui indique l'aspect probabiliste des modèles de NLP.

Comme dans les approches déductives, les modèles inductifs ont une faible capacité à répondre à des événements imprévus, notamment en fonction d'une quantité de données insuffisante pour un entraînement approprié. Comme l'explique Smith :

The primary failure modes of AI have been brittleness (in deductive systems) and poor generalization (in inductive systems). Brittleness and poor generalization of AI (...) are directly related to uncertainty about the details of future decision-making situations. They are failures of preprogrammed systems to cope with circumstances unforeseen by their designers (that is to say, uncertainty about future conditions). (Smith, 2016, p. 3)

Pour contourner ces situations de difficulté face à l'incertitude, les développeurs ont recours à l'utilisation des probabilités et modèles statistiques, « [i] n the deductive frame, probabilistic networks of inference (...) of various kinds have become a dominant knowledge representation, and in the inductive frame, statistical learning has become the dominant foundation for learning algorithms (...) » (Smith, 2016, p. 13) Regardons plus en détail le fonctionnement des modèles inductifs, très utilisés par les

algorithmes de NLP (Goldberg, 2016), pour être en mesure par la suite de proposer des bases pour une interprétation pragmatiste spéculative des probabilités dans ces modèles.

Dans la documentation de Recast.AI sur ses algorithmes de traitement automatique du langage naturel, il a été affirmé qu'une des principales familles d'algorithmes utilisées par Recast est le *clustering*⁴⁵. Les algorithmes de *clustering* sont utilisés pour regrouper des éléments avec des similarités entre eux (Murphy, 2012, p. 877). Par exemple, un algorithme de *clustering* peut former des groupes de personnes avec des achats jugés similaires sur un site déterminé. Dans le cas de l'utilisation de méthodes de *clustering* pour le développement de *chatbots*, il s'agit souvent de regrouper des mots ou des documents selon les critères souhaités. Dans un article publié sur le site web de Recast sur l'utilisation d'algorithmes de *clustering* dans le triage de texte, Paul Renvoisé explique :

Clustering is the task of organizing unlabelled objects in a way that objects in the same group are similar to each other and dissimilar to those in other groups. In other words, clustering is like unsupervised classification where the algorithm models the similarities instead of the boundaries. (Renvoisé, 2017, s.p.)

Il s'agit ainsi d'un type d'apprentissage automatique non supervisé, où l'algorithme traite des données qui ne sont pas classées (*labeled*) au préalable et essaie de repérer des répétitions, des patterns dans les données numériques.

Dans toutes les techniques d'apprentissage automatique non supervisé, un des enjeux principaux est l'évaluation de la qualité des résultats obtenus. Selon Paul Renvoisé, développeur chez Recast, les trois éléments à la base de toute technique de *clustering* sont : une mesure pour déterminer les similarités et différences entre les éléments ; une

⁴⁵ J'ai communiqué avec Recast pour obtenir plus d'informations sur les autres familles d'algorithmes utilisées, mais pour des questions de propriété intellectuelle, l'entreprise n'a pas été en mesure de répondre à mes questions. Ainsi, je me concentrerai ici sur l'étude des algorithmes de *clustering*.

fonction pour évaluer des résultats obtenus selon certains critères, autrement dit, l'évaluation de la qualité des *clusters* formés ; un algorithme pour optimiser ces critères (Renvoisé, 2017, s.p.). Une des méthodes pour évaluer la qualité des *clusters* obtenus est le calcul de l'information mutuelle, le concept élaboré en théorie de l'information dans le contexte du théorème de la voie avec bruit (Murphy, 2012, pp. 880-881) et abordé dans le chapitre précédent. Gil Katz, dans un document publié sur le site web de Recast, explique comment l'information mutuelle (ou l'entropie conditionnelle) peut servir à mesurer la qualité des clusters obtenus. Il est important de remarquer que sa proposition fait appel aux probabilités conditionnelles :

Consider for example a situation where we would like to compare two unsupervised clustering algorithms over the same data, in order to check if they give similar results or not. Why not use the mutual information between the results of each of the algorithms, where X is the random variable that represents the cluster chosen for any data point by the first algorithm and Y represents the second? Another option is to use the conditioned entropy ($H(x|y)$ or $H(y|x)$) in order to see how much “mess” is left when guessing the result of clustering by one algorithm, while the result of clustering by the other is already known (...). (Katz, 2017, s.p.)

Cette propriété est une des possibles approximations entre l'apprentissage automatique et la théorie de l'information. Katz défend que la théorie de l'information de Shannon offre plusieurs possibilités d'application à l'apprentissage automatique. Selon Katz, il suffit pour cela d'interpréter l'algorithme d'apprentissage automatique comme une voie de transmission d'information dont le but est de maximiser l'information mutuelle, exprimée en termes de probabilité conditionnelle :

Looking at any ML problem as the problem of creating a **channel** from the original data point at the input to an answer at the output, for which the **information** it conveys about the desired property of the data is maximized, would allow you to look at many different problems from a different angle than usual. (Katz, 2017, s.p.)

Dans le chapitre précédent, le concept d'information mutuelle a été abordé par la perspective du pragmatisme spéculatif et la notion de synthèse deleuzienne. On se

souvent, l'introduction de l'information mutuelle impliquait le passage des synthèses passives aux synthèses actives, autrement dit, le passage des processus aléatoires des systèmes en équilibre vers des processus composés d'événements interdépendants, vers des processus avec des mécanismes de mémoire et d'apprentissage qui établissent des relations conditionnelles entre les événements. Selon cette perspective (qui sera développée plus loin dans le contexte des algorithmes probabilistes d'apprentissage automatique), la technique proposée par Katz cherche à intégrer les synthèses actives dans les algorithmes informatiques inductifs. L'analyse faite par Katz des algorithmes de *clustering*, et le rapprochement avec la théorie de l'information, démontrent certaines des contributions que proposent les interprétations théoriques des modèles utilisés dans ces algorithmes. Ainsi, il sera important par la suite de prêter attention aux modèles théoriques utilisés dans les algorithmes d'apprentissage automatique.

La majorité des modèles inductifs utilisés dans les algorithmes d'apprentissage automatique essaient de simuler les fonctions d'apprentissage chez les humains. La question de l'apprentissage et des différents processus humains de « traitement d'information » devient ainsi un enjeu majeur pour les modèles inductifs appliqués aux algorithmes. Cette question sera abordée plus loin par la perspective du pragmatiste spéculatif, mais avant cela il sera question d'un court regard sur deux modèles proposés pour les mécanismes d'apprentissage : le réseau neuronal artificiel et le *probabilistic programming*. Comme nous remarquerons, l'utilisation des probabilités conditionnelles est souvent au centre des approches proposées pour intégrer dans les algorithmes des capacités d'apprentissage et de réaction à l'incertitude. La prémisse selon laquelle les probabilités décrivent les processus humains liés à l'information (apprentissage, mémoire, réflexion, entre autres) sera également discutée plus loin.

Parmi les modèles inductifs d'apprentissage automatique les plus utilisés, on retrouve le réseau de neurones artificiel (*artificial neural network*). Déjà en 1972 Atlan aborde les modèles informatiques des réseaux neuronaux et les compare aux « fonctions et performances (...) des systèmes vivants en général et du cerveau en particulier. »

(Atlan, 2006 [1972], p. 126) Dans les termes d'Atlan, il s'agit d'une comparaison entre un automate naturel (système vivant) et un automate artificiel (machine réalisée par l'humain). À la suite d'un premier moment où il établit les différences entre ces types de systèmes, Atlan constate que des modifications ont été apportées aux machines, et ces modifications concernent amplement l'utilisation de modèles probabilistes :

C'est sur deux points essentiels que des différences profondes, d'ordre qualitatif, ont été reconnues et que des progrès importants ont été accomplis dans la conception de nouveaux types de machines, où ces différences sont atténuées :

Il s'agit :

1. de la façon de réagir aux erreurs, autrement dit, du problème de la fiabilité des automates ;
2. du rôle de l'aléatoire dans l'organisation d'un automate. (Atlan, 2006 [1972], p. 127)

Étudions plus en détail chacun de ces deux aspects. Dans le premier (les modes de réagir à des erreurs) Atlan revient sur les propositions de Von Neumann et McCulloch et Pitts (1943) qui sont à la base de la création d'automates basés sur des modèles neuronaux. Il s'agit d'automates qui essaient de reproduire la faculté des systèmes neuronaux vivants à « fonctionner correctement malgré des variations aléatoires importantes. » (Atlan, 2006 [1972], p. 128) Dans la tentative de reproduire ce type de comportement, les probabilités conditionnelles sont introduites aux modèles neuronaux. Atlan établit un parallèle entre l'utilisation des probabilités conditionnelles dans les automates à réseaux neuronaux et l'intégration du bruit dans la voie de communication tel que l'a théorisé Shannon :

Les symboles des messages de sortie sont donc reliés à ceux des messages d'entrée par des probabilités conditionnelles déterminées à la fois par le processus de transmission et par ceux de calcul ; notamment, l'action du bruit (facteurs d'erreurs) dans le réseau modulaire déterminera la valeur de ces probabilités conditionnelles. (Atlan, 2006 [1972], p. 141)

Si dans le théorème de la voie avec bruit Shannon a intégré les probabilités conditionnelles pour ajouter le bruit dans un système de transmission d'information, dans les réseaux neuronaux on fait également appel aux probabilités conditionnelles, mais cette fois-ci pour simuler des « facteurs d'erreurs ».

Pour aborder ce point, Atlan commence par explorer la question de l'apprentissage des systèmes vivants dans l'exemple de la reconnaissance des formes. Il explique que : « l'aptitude du cerveau ne concerne pas un nombre limité de formes particulières définies une fois pour toutes, mais un nombre apparemment infini de formes qu'il *apprend* à reconnaître et à classer en images. » (Atlan, 2006 [1972], p. 155, italique original) Ce processus d'apprentissage est différent de la transmission d'information, de sorte qu'une adaptation du modèle de Shannon est nécessaire dans le contexte des réseaux neuronaux, « (...) on veut dire par là qu'un centre nerveux ne peut pas être considéré seulement comme un réseau de communications, et qu'il doit s'y ajouter des unités de décision. » (Atlan, 2006 [1972], p. 156) Dans ce modèle, les unités de décision sont des éléments qui doivent être incorporés aux automates artificiels pour les faire évoluer vers une simulation plus appropriée des organismes vivants. Atlan reconnaît que le fonctionnement de ces unités de décision est encore obscur, mais il le décrit comme le passage d'un niveau plus élémentaire de traitement d'information vers des niveaux plus intégrés qui synthétisent et englobent les niveaux inférieurs. Atlan compare ce passage aux différents niveaux d'organisation qu'on observe chez les organismes vivants :

Une fonction de décision apparaît ainsi comme formellement analogue au passage de l'analyse d'un organisme comme configuration d'atomes, à celle du même organisme comme configuration de molécules, ou d'organelles, ou de cellules, etc. Nous avons vu comment ces passages se traduisent chaque fois par une diminution de la quantité d'information stockée ou transmise. (Atlan, 2006 [1972], p. 156)

Cela peut paraître contradictoire de parler de perte d'information dans des processus d'apprentissage et de décision. Néanmoins, il est possible de décrire ces processus en

termes de synthèses où, pour être en mesure d'inférer une information par rapport à ce qui a été observé, il est nécessaire d'exclure ou d'éliminer l'excès qui existe dans les données pour les convertir en information. Les observations d'Atlan correspondent aux résultats plus contemporains en neurosciences dans la description des activités cognitives, comme on le verra avec le concept « d'inférence active ». Le parallèle avec la synthèse de Deleuze (« la différence soutirée de la répétition ») sera aussi exploré plus loin.

Atlan remarque ainsi qu'il existe une complémentarité entre une structure prédéterminée d'organisation des systèmes neuronaux naturels et les effets produits par des éléments aléatoires :

(...) beaucoup acceptent l'idée que la détermination génétique ne concernerait que la structure anatomique globale du cerveau, et que le détail des connexions serait le fruit du hasard, modifié au fur et à mesure de leur constitution par les effets des expériences acquises. Les effets de l'environnement modifieraient ainsi l'organisation cérébrale elle-même et ne seraient pas limités à la libération de réponses stéréotypées déjà prévues. (Atlan, 2006 [1972], p. 164)

Cette conclusion est d'une grande importance pour la discussion de la présente thèse, car elle défait encore une fois le dualisme apparent entre hasard et structure, ou aléatoire et organisation. Selon Atlan, il ne s'agit pas de l'un (hasard) ou l'autre (structure), mais au contraire, l'un et l'autre s'affectant mutuellement dans des processus dynamiques. Si le hasard peut jouer un rôle constructif dans l'organisation des processus de décision et d'apprentissage des organismes vivants, la question pour les automates artificiels (ou algorithmes) devient alors « comment reproduire, ou modéliser, les relations entre hasard et structure dans les systèmes numériques ? »

Atlan propose qu'une des réponses puisse être trouvée dans l'utilisation des probabilités conditionnelles par la théorie de l'information et définit « l'apprentissage adaptatif non dirigé comme une sélection, par un système, des réponses les mieux adaptées à l'environnement, parmi toutes celles qui sont provoquées au hasard des

stimulations de l'environnement.» (Atlan, 2006 [1972], p. 168) Nous pouvons constater ici la proximité entre ces conclusions et les notions développées dans le chapitre précédent, notamment celle « d'adaptation dissipative », qui, comme nous l'avons vu, décrit les mécanismes physiques de la propension d'un système. Ces conclusions réaffirment ainsi la pertinence d'une lecture pragmatiste spéculative du rôle des probabilités dans les modèles inductifs qui intègrent la notion « d'apprentissage » dans les algorithmes. Les remarques d'Atlan sur les réseaux neuronaux en informatique et l'utilisation des probabilités conditionnelles résonnent également avec des discussions contemporaines. Il est possible de remarquer qu'un grand nombre d'algorithmes d'apprentissage automatique, tout comme ceux de Recast, se basent sur une utilisation des probabilités conditionnelles dans le but d'effectuer des inférences par rapport aux données d'entrée. Nous verrons par la suite une de ces approches, le *probabilistic programming*.

Un champ récent qui aborde les intersections entre l'apprentissage automatique et la théorie des probabilités, notamment bayésiennes, est le *probabilistic programming*. Mansinghka remarque que depuis une quinzaine d'années, il existe une forte convergence entre l'apprentissage automatique et la théorie des probabilités, avec des résultats très positifs dans les applications des modèles d'apprentissage automatique aux analyses de données massives. Les domaines qui ont intégré les probabilités et l'informatique sont ceux où il est question d'indétermination et d'incertitude, comme la robotique (qui doit créer des robots capables de réagir aux incertitudes de l'environnement) et l'intelligence artificielle. Il aborde le *probabilistic programming* comme « the software layer » de cette intégration entre probabilités et informatique. Comme explique Vikash Mansinghka : « [p]robabilistic programming is an emerging field that aims to address these challenges [difficultés d'inférence] by formalizing modeling and inference using key ideas from probability theory, programming languages, and Turing-universal computation. » (Mansinghka, 2015) Cette approche de l'apprentissage automatique propose de développer des algorithmes qui s'ajustent

automatiquement (ou qui s'auto-organisent) à partir des collections de données et basées sur des modèles probabilistes. Comme décrit Tejas Dattatraya Kulkarni : « The whole hope is to write very flexible models, both generative and discriminative models, as short probabilistic code, and then not do anything else. General-purpose inference schemes solve the problems. » (Kulkarni, 2015) Selon Mansinghka, ce domaine a trois axes principaux : le premier aspect est le développement de langages de programmation basés sur des modèles probabilistes qui automatisent la création même des algorithmes (inférence automatisée) ; le deuxième principal champ de recherche dans ce domaine est l'utilisation de l'informatique dans la modélisation d'objets complexes. Un exemple de ce deuxième aspect du champ de *probabilistic programming* est son application dans les *chatbots* (j'aborderai cet exemple plus loin) ; enfin, le troisième aspect est plus théorique et aborde les possibilités d'utilisation de l'informatique dans le développement de la théorie des probabilités.

Un exemple d'application des techniques de *probabilistic programming* est le domaine des *chatbots*. « Gamalon » est une *start up* qui travaille actuellement sur l'application de techniques de *probabilistic programming* dans le développement de *chatbots*. Comme il est expliqué par Will Knight : « Gamalon's system uses probabilistic programming—or code that deals in probabilities rather than specific variables—to build a predictive model that explains a particular data set. » (Knight, 2017, s.p.) Cette approche permet une réduction drastique des données nécessaires à l'entraînement des algorithmes d'apprentissage automatique, et fournit aux algorithmes des stratégies plus efficaces pour réagir à l'incertitude de l'interaction avec l'environnement naturel (dans le cas des *chatbots*, l'interlocuteur humain). Il revient sur la large utilisation des inférences bayésiennes dans ces technologies et explique :

Gamalon uses a technique that it calls Bayesian program synthesis to build algorithms capable of learning from fewer examples. Bayesian probability, named after the 18th century mathematician Thomas Bayes, provides a mathematical framework for refining predictions about the world based on experience. (Knight, 2017, s.p.)

Il est intéressant de constater que les probabilités bayésiennes sont celles qui, selon Knight, sont les plus appropriées pour intégrer les acquis de l'expérience dans la modélisation et contribuer ainsi à la création de modèles prédictifs.

La technique de « synthèse de programme Bayésienne », appliquée ici dans le contexte des *chabots*, est un aspect central du *probabilistic programming* (Saad, Cusumano-Towner, Schaechtle, Rinard, & Mansinghka, 2019). La synthèse de programme consiste dans la création d'algorithmes informatiques à partir de spécifications écrites dans d'autres langages que celui de l'algorithme. Dans le cas de la synthèse bayésienne, il s'agit d'utiliser les inférences bayésiennes pour l'analyse de données et ensuite utiliser les résultats de ces analyses dans le développement d'algorithmes (ou synthèse de programme) : « We use Bayesian inference to synthesize ensembles of probabilistic programs sampled from domain specific languages given observed data. » (Saad et al., 2019, p. 2) Or, dans le chapitre précédent, nous avons observé la relation entre les inférences bayésiennes et les synthèses actives des probabilités conditionnelles (des processus à variables interdépendantes sujets aux actions de la mémoire et de l'apprentissage). Avec le *probabilistic programming*, nous constatons encore une fois le rôle central des probabilités, notamment conditionnelles, dans les algorithmes d'apprentissage automatique. Dans la prochaine section, il sera question de proposer une perspective pragmatiste spéculative du rôle des probabilités dans ces modèles et algorithmes.

4.3. Machine Learning (apprentissage automatique) : Interprétation des probabilités dans les modèles cognitifs et algorithmes d'apprentissage

Comme nous avons remarqué dans la section précédente avec le réseau neuronal artificiel, le *probabilistic programming* et les algorithmes observés, les probabilités

conditionnelles et les inférences bayésiennes sont au cœur des algorithmes d'apprentissage automatique. Dans la discussion contemporaine autour des approches probabilistes de l'apprentissage automatique cela est aussi observé :

The systematic application of probabilistic reasoning to all inferential problems, including inferring parameters of statistical models, is sometimes called a Bayesian approach. However, this term tends to elicit very strong reactions (either positive or negative, depending on who you ask), so we prefer the more neutral term “probabilistic approach”. (Murphy, 2012, p. xxvii)

Il est intéressant de remarquer que Murphy établit une équivalence entre les modèles probabilistes et la probabilité bayésienne, malgré le fait qu'il reconnaisse également l'existence d'autres interprétations (notamment l'interprétation fréquentiste). Il oppose l'interprétation fréquentiste à la bayésienne et les explique ainsi :

There are actually at least two different interpretations of probability. One is called the **frequentist** interpretation. In this view, probabilities represent long run frequencies of events. (...) The other interpretation is called the Bayesian interpretation of probability. In this view, probability is used to quantify our **uncertainty** about something; hence it is fundamentally related to information rather than repeated trials. (Murphy, 2012, p. 27, gras original)

Murphy reprend sommairement les interprétations de la théorie des probabilités étudiées dans le chapitre précédent. Comme nous avons observé, l'interprétation fréquentiste se situe dans le pôle des interprétations objectives (que nous avons mis en relation avec les synthèses passives) et l'interprétation bayésienne dans le pôle subjectif (en relation avec les synthèses actives). Murphy prend position pour une interprétation bayésienne : « One big advantage of the Bayesian interpretation is that it can be used to model our uncertainty about events that do not have long term frequencies. (...) We shall therefore adopt the Bayesian interpretation in this book. » (Murphy, 2012, pp. 27-28) En opposition aux méthodes bayésiennes, Murphy présente les approches fréquentistes, aussi très utilisées en ML. Dans les approches fréquentistes, « instead of being based on the posterior distribution, they are based on the concept of sampling

distribution. » (Murphy, 2012, p. 193) Nous remarquons ici à nouveau une double tendance des algorithmes probabilistes : l'utilisation de l'inférence bayésienne et de l'échantillonnage aléatoire pour simuler les activités d'apprentissage et de décision.

Ces approches sont basées sur la prémisse que les probabilités et les inférences statistiques sont des outils efficaces pour décrire les activités cognitives, l'apprentissage et les mécanismes de mémoire que les algorithmes essaient de reproduire. Comme a soulevé aussi Robert Elliott Smith dans un article récent autour des modélisations de l'incertitude dans l'intelligence artificielle (Smith, 2016), un des présupposés de l'utilisation des probabilités dans les modèles de l'intelligence artificielle et d'apprentissage automatique est que le raisonnement humain soit probabiliste : « it is assumed that people reason probabilistically about the facts they gather. » (Smith, 2016, p. 2) Il revient sur la théorie de l'information pour démontrer que le fait que cette théorie ait ancré l'indétermination dans la réalité physique (grâce au concept d'entropie et à l'utilisation de probabilités) a contribué à une extrapolation des modèles probabilistes vers la modélisation des activités cognitives : « If the probabilistic and information theoretic interpretation of uncertainty is fundamental to reality itself, then perhaps it is a likely basis for the evolved nature of thought. » (Smith, 2016, p. 25) Avant de développer une interprétation pragmatiste spéculative des probabilités et du hasard dans les algorithmes d'apprentissage automatique, il sera important alors de mieux comprendre à quel point les approches probabilistes et statistiques sont adéquates pour décrire les activités cognitives, notamment humaines.

L'utilisation de la probabilité comme un modèle pour la cognition humaine est une proposition très débattue. Daniel Kahneman et Amos Tversky sont des pionniers dans les travaux sur l'utilisation des probabilités pour la description des modes de réflexion humaine. Dans les années 1970, Kahneman et Tversky ont débuté une longue collaboration dédiée à étudier les prises de décision dans des situations d'incertitude. Le travail de Kahneman et Tversky a eu un impact profond dans le domaine de la théorie économique, et il remet en question le présupposé selon lequel les probabilités

subjectives peuvent décrire de façon directe le mode de prise de décision d'un acteur économique. Dans un article de 1972, les auteurs affirment :

Indeed an extensive experimental literature has been devoted to the question of how people perceive, process, and evaluate the probabilities of uncertain events in the contexts of probability learning, intuitive statistics, and decision making under risk. (Kahneman & Tversky, 1972, p. 430)

Le travail de Smith, ainsi que celui de Kahneman et Tversky, indique que ce présupposé est insuffisant : « Perhaps the most general conclusion, obtained from numerous investigations, is that people do not follow the principles of probability theory in judging the likelihood of uncertain events. » (Kahneman et Tversky, 1972, p. 430-431) L'évolution de leurs travaux a démontré les différents biais et erreurs d'évaluation qui interfèrent dans la capacité des processus cognitifs humains de comprendre et évaluer les situations d'incertitude. Il est intéressant de noter que les auteurs ont construit un modèle cognitif basé sur les principes de l'inférence bayésienne. Comme décrit Kahneman : « (...) Bayesian logic requires you to adjust your judgement in light of any relevant information (...). In the absence of useful new information, the Bayesian solution is to stay with the base rates. » (Kahneman, 2012, p. 172) Kahneman appelle « *base rates* » les probabilités a priori à la base des inférences bayésiennes.

En effet, l'utilisation de l'inférence bayésienne est aussi remarquée en neurosciences dans la description des activités cognitives, notamment dans l'hypothèse du « cerveau bayésien ». Andy Clark, dans son livre *Surfing Uncertainties* (2016), défend que la prévision probabiliste orientée vers l'action est la fonction principale du cerveau. D'une part, Clark reconnaît dans l'activité cognitive le modèle bayésien (ou *Bayesian brain hypothesis*), défini comme « the hypothesis that the brain somehow implements processing that approximates ideal ways of weighing new evidence against prior knowledge. » (Clark, 2016, p. 9). Mais il associe ensuite ce modèle à l'engagement dans des actions effectuées par le corps. En effet, il ne s'agit pas d'un modèle où il existe une perception cognitive passive suivie d'action, mais plutôt une combinaison

simultanée de perception, réflexion et action : « This potential unification of work on probabilistic neural processing with the work on the role of embodiment and action is, I believe, one of the most attractive features of the emerging framework. » (Clark, 2016, p. 9) Clark situe son travail dans une tendance plus large, « predictive processing » (PP), « a general, and increasingly well-supported, vision of the brain (and especially the neocortex) as fundamentally an inner engine of probabilistic prediction. » (Clark, 2016, pp. 27-28). Il aborde également le rôle des probabilités conditionnelles dans le modèle du cerveau bayésien (d'après Knill et Pouget, 2004, p. 712) : « Such a mode of representation implies that when we represent a state or feature of the world, such as the depth of a visible object, we do so not using a single computed value but using a conditional probability density function (...) » (Clark, 2016, p. 40) En effet, on remarque que ce modèle est basé largement sur l'utilisation des probabilités conditionnelles dans la modélisation des activités cognitives, ce qui le ferait pencher vers le pôle actif de l'interprétation des probabilités.

Dans cette même perspective, Karl Friston développe un modèle de la cognition à partir des concepts d'inférence active et d'énergie libre (« *free energy* »). Dans son modèle, perception, action et apprentissage sont des processus liés qui fonctionnent selon le principe de la minimisation de la variation d'énergie libre. L'énergie libre correspond, de manière simplifiée, au degré de surprise qu'une observation provoque, autrement dit, à la différence entre les attentes qu'on a (les prédictions) et ce que l'on perçoit. Dans ce modèle, la cognition fonctionne comme un système prédictif décrit par des inférences bayésiennes (des probabilités a priori mises à jour à chaque nouvelle donnée) qui modélise l'environnement. C'est à partir de ce modèle prédictif que le système cognitif perçoit, réfléchit et agit. Comme l'explique Friston :

Active inference is based upon the premise that everything minimizes variational free energy. This leads to some surprisingly simple update rules for action, perception, policy selection, learning and the encoding of uncertainty (i.e., precision) that generalize established normative approaches. (K. Friston et al., 2016, p. 864)

Dans l'image ci-dessous, Friston schématise les différentes activités du processus cognitif et situe les zones du cerveau où les activités sont réalisées. (K. Friston et al., 2016)

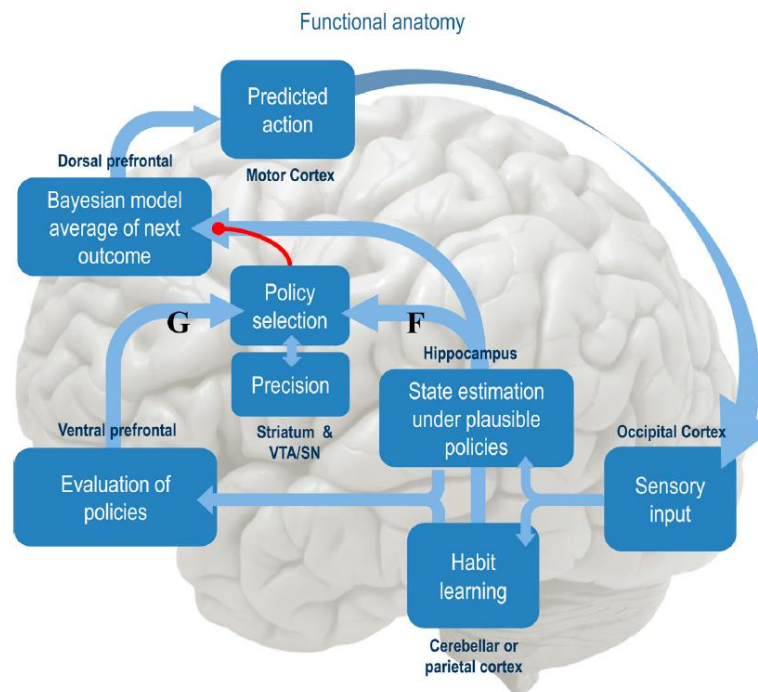


Figure 7 : Processus cognitif. Source : K. Friston et al., 2016

Un exemple concret permettra peut-être de rendre plus clair le concept d'inférence active et l'hypothèse du cerveau bayésien : on essaie de soulever une boîte qu'on imagine est remplie de livres (et donc très lourde), mais qui en réalité est vide et légère. Habituellement, dans cette situation les personnes soulèvent la boîte avec une force disproportionnée et sont surprises de constater la différence entre le poids espéré et le poids réel de la boîte. Ainsi, ce qui démontre la pratique est que dans un premier moment, l'on attrape l'objet avec une force disproportionnée et ce n'est qu'ensuite, lorsqu'on se rend compte que la boîte bouge plus vite que prévu, que nous réajustons la force que nous mettons dans le geste effectué. Or, nous avons un modèle mental et une prédiction de la force qui serait nécessaire pour soulever la boîte, et nous réajustons

ce modèle à la suite de toute nouvelle information, nous réalisons une inférence active pour minimiser la variation d'énergie libre. Comme résume Friston : « It is fairly easy to show that both perceptual inference and learning rest on a minimization of free-energy or suppression of prediction error. » (K. J. Friston & Stephan, 2007, p. 419) Les formalisations mathématiques du modèle proposé par Friston s'inspirent largement de l'inférence bayésienne, et donc des probabilités conditionnelles.

Nous constatons ainsi que les modèles prédictifs probabilistes décrivent certains aspects des processus cognitifs, comme démontré par Clark, Friston et d'autres chercheurs. Néanmoins, nous constatons également que ces modèles ont des limites et ne peuvent pas rendre compte d'un grand nombre de comportements humains. Le travail de Kahneman et Tverski et d'autres chercheurs a démontré, par exemple, le problème des probabilités *a priori* des inférences bayésiennes (abordé dans le chapitre précédent). Un autre questionnement important qui peut être porté au présupposé de l'application des modèles probabilistes aux activités cognitives humaines est l'idée d'une seule neurotypicalité. Or, si l'on accepte ce présupposé, il existerait un mode d'activité cognitive considéré normal, et toutes les autres formes de cognition ne seraient pas considérées dans les modèles. Comme a souligné Erin Manning (Manning, 2016), il est nécessaire de défaire les conceptions qui excluent la neurodiversité et la relèguent sur le plan de la pathologie :

It is urgent to turn away from the central tenet of neurotypicality, the wide-ranging belief that there is an independence of thought and being attributable above all to the human, a better-than-ness accorded to our neurology (a neurology, it must be said, that reeks of whiteness, and classism). (Manning, 2016, p. 3)

Ainsi, il est important d'accepter la diversité des modes de fonctionnement neurologiques, ce qui se traduit nécessairement dans une multiplicité des modèles à être utilisés pour simuler les activités cognitives humaines. Lorsque nous nous éloignons de l'approche neurotypique, les notions mêmes de ce qui constitue la connaissance ou l'apprentissage doivent être reformulées : « Most accepted approaches

to learning assume neurotypicality with regard to processing information, thereby segregating not only neurodiverse learners, but also predefining what counts as knowledge.» (Manning, 2016, p. 9) Je propose d'aborder ces questions par la perspective du pragmatisme spéculatif. Ainsi, plutôt que d'essayer de comprendre l'activité cognitive en termes d'un sujet qui reçoit (de manière passive ou active) des informations d'objets extérieurs, il faut la comprendre en termes d'événement (Massumi, 2011). D'autres limites de l'utilisation des probabilités dans la description des processus cognitifs seront évoquées plus loin, par la perspective du pragmatiste spéculatif.

Ces conclusions sur le rôle des probabilités dans les modèles des processus cognitifs sont importantes aussi dans le domaine de l'intelligence artificielle. Lorsque l'on aborde la simulation des activités cognitives dans le contexte de l'intelligence artificielle, il est donc nécessaire de mettre en évidence les limites de l'utilisation des modèles probabilistes pour décrire une expérience réflexive, comme le remarque Smith : « AI has shown that models of human decision-making that are based on the idealisation of mathematics or logic do not embody the real-world, robust decision-making in the face of ontological uncertainty observed in humans. » (Smith, 2016, p. 35) Smith conclut que le développement de l'intelligence artificielle doit alors intégrer des connaissances de domaines plus liées à une observation in situ des modes de raisonnement humain (psychologie, anthropologie, sociologie, et d'autres) plutôt que baser ses algorithmes sur des modèles mathématiques qui idéalisent l'activité cognitive.

Je propose une réinterprétation des probabilités dans le contexte de l'apprentissage automatique par la perspective non dualiste du pragmatisme spéculatif. Comme discuté dans le chapitre précédent, les probabilités conditionnelles et l'inférence bayésienne sont utilisées dans des situations où les variables, ou les événements ne sont pas indépendants, ce que nous avons appelé, d'après Deleuze, les « synthèses actives ». Or, d'après l'approximation établie dans le chapitre précédent entre les probabilités et le

concept de synthèses deleuziennes, il est possible d'interpréter certains aspects des algorithmes de ML comme une tentative de simulation des synthèses actives, où l'algorithme traite des informations de manière à simuler une activité de mémoire ou d'apprentissage, grâce notamment à l'utilisation des probabilités conditionnelles et l'inférence bayésienne. Selon l'interprétation des probabilités proposée, dans le pôle passif il s'agit d'habitudes, de propensions (dans les termes de Peirce, *would be*) développées à partir d'une répétition d'événements indépendants. Et dans le pôle actif, ce sont les habitudes d'une répétition conditionnelle où des activités de mémoire ou de réflexion interfèrent dans l'indépendance des événements répétés. En reprenant les conclusions des chapitres précédents autour de ces interprétations, il est possible d'affirmer qu'il ne s'agit pas de choisir une interprétation plutôt qu'une autre, mais plutôt de comprendre comment s'articulent les probabilités passives et actives dans les modèles et algorithmes qui essaient de simuler des processus d'apprentissage et de décision. Ainsi, il n'est pas question du dualisme subjectif/objectif, mais plutôt d'une multiplicité de modes de synthèses en relation et opérant à différents niveaux dans un même système.

L'aspect englobant des synthèses actives avait déjà été soulevé par Deleuze et les commentaires d'Atlan à propos du passage des niveaux d'organisation dans les processus de décision vont dans cette même direction. Selon Deleuze, on se rappelle, les synthèses passives sont les fondations sur lesquelles se bâtissent les synthèses actives. Par cette perspective, les systèmes sont composés par des ensembles de processus de synthèses qui opèrent à différents niveaux d'organisation. Ainsi, dans un corps humain par exemple, des synthèses opèrent au niveau atomique, au niveau cellulaire, au niveau des organes, au niveau du réseau neuronal, jusqu'au niveau des relations interpersonnelles, culturelles, sociales et politiques. À ces différents niveaux, les processus de synthèse ont lieu de sorte à créer des tendances, des propensions, des *would be*, à être mesurées par les probabilités. Comme nous avons observé, dans chaque système il existe une multitude de synthèses opérantes, passives et actives. Plus

une synthèse est affectée par des processus comme l'apprentissage ou la mémoire, plus elle se situe vers le pôle actif. En effet, les synthèses actives sont les habitudes des systèmes cognitifs. Or, cette même approche peut être transposée aux algorithmes, et plutôt que les classer de manière dualiste, il est nécessaire de les comprendre dans la complexité des modèles qui les composent. Il faut dépasser le dualisme qui sépare les modèles en deux types (souvent fréquentiste versus bayésien, comme on l'a vu) pour être en mesure de les comprendre dans leur multiplicité. Une première condition alors pour les modèles probabilistes qui cherchent à simuler des processus d'apprentissage ou de décision est ainsi de composer avec une multiplicité de synthèses, passives et actives.

Dans cette perspective, le travail de Deleuze et Guattari peut contribuer à nous éclairer sur les limites des probabilités dans la modélisation des activités d'apprentissage et des processus de décision. Dans le texte « De la ritournelle » (Deleuze & Guattari, 1980), Deleuze et Guattari décrivent des processus d'organisation qui émergent de la répétition, auxquels ils attribuent l'image d'une ritournelle. Ils soulèvent trois aspects de la ritournelle, le premier étant le chaos, ou un trop plein de forces désorganisées. Dans ce chaos, une répétition commence à établir un centre. Autour de ce centre, un milieu s'organise (deuxième aspect). Deleuze et Guattari utilisent pour illustrer cela l'image d'un enfant saisi par la peur, dans le noir (le chaos), qui se rassure en entonnant une petite chanson, « c'est déjà la chanson qui est elle-même un saut : elle saute du chaos à un début d'ordre dans le chaos (...). » (Deleuze & Guattari, 1980, p. 382) Les synthèses évoquées plus haut sont des ritournelles, des processus d'habitude et d'organisation qui émergent de la répétition. À différents niveaux d'un système, les ritournelles opèrent ensemble de sorte à créer une organisation collective émergente, une adaptation dissipative. Si l'on reprend l'exemple du corps humain, chaque molécule est en soi une ritournelle, et lorsque mise en relation avec d'autres molécules, des ritournelles plus générales s'activent. Ainsi, l'idée de sujet se transforme en un ensemble de processus de ritournelles, un rythme qui émerge du chaos.

Le troisième aspect de la ritournelle est l'ouverture, des lignes d'erreurs ou des forces qui ouvrent le milieu formé dans le deuxième moment :

La ritournelle a les trois aspects, elle les rend simultanés, ou les mélange : tantôt, tantôt, tantôt. Tantôt, le chaos est un immense trou noir, et l'on s'efforce d'y fixer un point fragile comme centre. Tantôt l'on organise autour du point une « allure » (plutôt qu'une forme) calme et stable : le trou est devenu un chez-soi. Tantôt on greffe une échappée sur cette allure, hors du trou noir. (Deleuze & Guattari, 1980, p. 383)

Or, ce troisième aspect, cette ouverture vers l'avenir ou vers d'autres ritournelles résonne avec la troisième synthèse dont il a été question dans le chapitre précédent. La troisième synthèse, on se rappelle, « désigne le sans-fond, où le fondement lui-même nous précipite » (Deleuze, 1993 [1968], p. 151). La troisième synthèse est celle qui nous renvoie à l'affirmation du hasard et de la différence et qui échappe aux règles de la probabilité : « Le système de l'avenir (...) doit être appelé jeu divin, parce que la règle ne préexiste pas, parce que le jeu porte déjà sur ses propres règles, parce que l'enfant-joueur ne peut que gagner — tout le hasard étant affirmé chaque fois et pour toutes les fois. » (Deleuze, 1993 [1968], p. 152) Dans cet aspect, nous nous retrouvons face aux limites de la probabilité dans sa capacité à simuler les processus d'organisation émergente, ou des processus qui réorganisent leurs propres règles. Comme nous avons vu, les probabilités sont des hypothèses sur l'avenir élaborées à partir d'une observation du passé, et non des lignes d'ouverture vers d'autres avènements possibles. La question devient ainsi d'imaginer des systèmes avec un avenir incertain et qui échappe parfois aux prévisions probabilistes. Nous revenons alors à un des postulats à la base de la posture pragmatiste spéculative envers la création numérique : susciter des possibles (dans le sens de Debaïse et Stengers) plutôt que de recréer le probable (dans le sens de la théorie des probabilités) dans la sphère numérique.

Un autre aspect important de la ritournelle qui souligne les limites de la probabilité est son caractère expressif, ou qualitatif. En effet, quand l'enfant chante sa ritournelle, il le fait d'une certaine manière, il existe une qualité expressive dans son chant qui ne

peut pas être exprimée en termes quantitatifs. Comme nous avons déjà noté dans le chapitre précédent, Massumi affirme que le possible a une dimension qualitative qui échappe au probable, et associe ce surplus qualitatif au virtuel. Selon Massumi, l'outil mathématique le plus approprié pour aborder le virtuel est la topologie, une branche de la géométrie qui se concentre principalement sur les variations qualitatives d'une forme, autrement dit, les déformations que l'on peut apporter à un objet géométrique sans changement de sa composition. Il écrit : « Topology is purely qualitative science. It has no empirical (predictive) value. » (Massumi, 1998, p. 307) Il est important de remarquer la distance que prend la topologie des modèles prédictifs : du moment où nous nous éloignons des probabilités pour une approche plus qualitative des modèles, l'aspect prédictif se perd. Ainsi, l'approche topologique peut contribuer à donner une dimension qualitative aux modèles mathématiques qui essaient de simuler des activités cognitives humaines. Massumi conclut que « the path to the digital virtual is not through increasing cleverness in systems designing or programming. Quite to the contrary, the digital will only rejoin the virtual when it raises itself to the highest powers of brainy vagueness. » (Massumi, 1998, p. 312)

Il existe plusieurs travaux et applications qui combinent topologie (plus précisément, l'analyse topologique des données) et ML. L'analyse topologique des données est une branche récente des mathématiques qui utilise des méthodes de la géométrie et de la topologie dans l'analyse de données (Carlsson, 2009). Carlsson, un des pionniers dans ce type d'analyse affirme que « Topology is exactly that branch of mathematics which deals with qualitative geometric information » (Carlsson, 2009, p. 256). Ainsi, l'utilisation des méthodes inspirées de la topologie inclut un aspect qualitatif dans l'analyse de données. Lum et al. (2013) soulignent trois aspects importants de l'approche topologique qui la différencient d'autres approches géométriques ou statistiques. Le premier est le fait que la topologie étudie les formes sans un système de coordonnées et utilise uniquement des fonctions de distance pour décrire les formes. Cela permet, par exemple, à ce que des données recueillies dans différents systèmes

puissent être étudiées ensemble avec des méthodes topologiques. Le deuxième aspect est la capacité des méthodes topologiques à étudier des propriétés des formes qui ne changent pas lorsque soumises à des déformations mineures (qui ne changent pas sa composition). Prenons l'exemple de l'écriture manuelle, un humain adulte est souvent capable de reconnaître la lettre A même si la lettre a été écrite par différentes personnes et présente des variations mineures. Ainsi, les méthodes topologiques présentent une flexibilité dans la compréhension de formes qui est plus proche de l'habileté humaine à interpréter une forme de manière qualitative. À propos de cet aspect, les auteurs affirment : « This inherent property of topology is what allows it to be far less sensitive to noise and thus, possess the ability to pick out the shape of an object despite countless variations or deformations. » (Lum et al., 2013, p. 2) Enfin, le troisième aspect est la capacité des méthodes topologiques à réduire la complexité des formes sans enlever ses propriétés déterminantes (d'un point de vue topologique). Un cercle, par exemple, peut être représenté par un hexagone, ce qui facilite la représentation (un hexagone a moins de points qu'un cercle) sans que l'essentiel de sa forme (être un *loop* avec un seul trou à l'intérieur) soit perdu. (Lum et al., 2013, p. 3)

Un exemple d'application de l'analyse de données qui est souvent mentionné dans les recherches de Carlsson est dans les algorithmes de *clustering* (dont il a été question plus haut). Or, les méthodes topologiques peuvent contribuer à limiter des analyses faussées par les formes présentées par les données. Comme explique Carlsson :

The fundamental problem is that clustering is a methodology for breaking data sets apart into pieces, and that is its only tool. Once the data set is such that it contains progressions, and is in an appropriate sense “connected”, clustering methods will typically break things apart that belong together, and that phenomenon is exhibited in the above dendrogram. The basic point is that representing some aspects of geometry beyond the breakup into clusters provides greater insight into the detail of the data set, which is difficult to recover using clustering alone. (Carlsson, 2016, pp., s.p.)

Prenons l'exemple des données qui forment deux spirales, comme l'image ci-dessous :

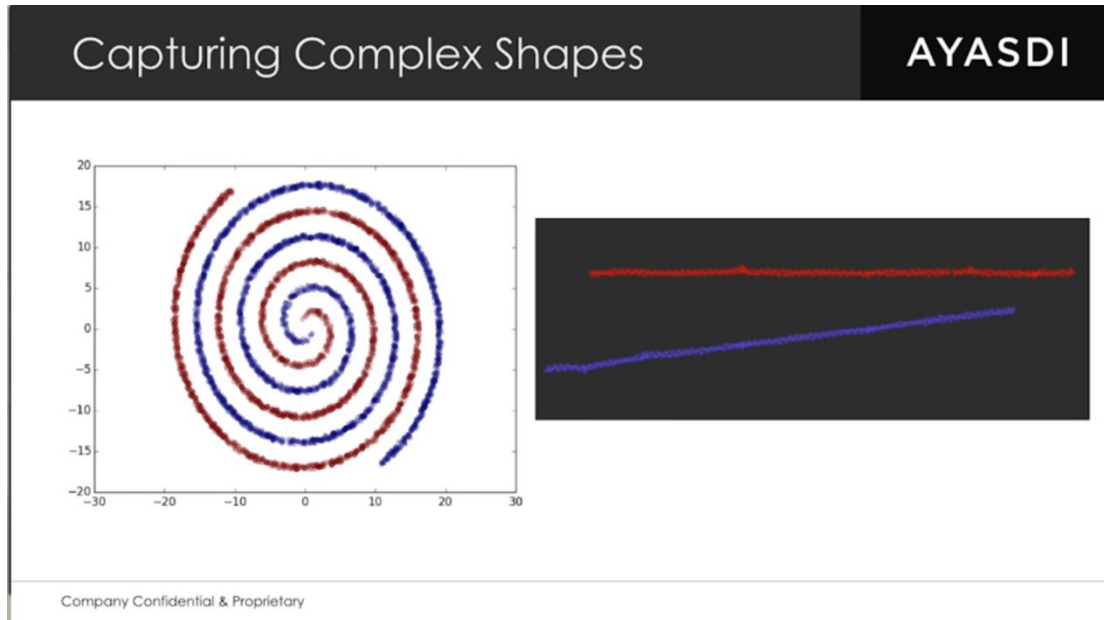


Figure 8 : Clustering de formes complexes. Source : Ayasdi (<https://www.ayasdi.com>)

L'organisation des données présentées dans l'image de gauche présente des difficultés aux algorithmes de *clustering* à classer les données, alors que des humains très facilement pourraient visualiser le classement approprié (indiqué par les différentes couleurs). Avec l'utilisation de l'analyse topologique des données, le classement des données dans ce type de configuration est beaucoup plus efficace.

Ainsi, par la perspective proposée dans la présente thèse, la simulation de certaines activités cognitives humaines ne peut pas être résumée à des modèles probabilistes bayésiens. Il est nécessaire de comprendre la multiplicité et l'arrangement complexe des différentes synthèses qui composent les processus dont il est question dans la simulation, synthèses à la fois actives et passives. Par ailleurs, la troisième synthèse démontre les limites des modèles probabilistes dans leur incapacité à prévoir le radicalement nouveau, ou les organisations émergentes. Une approche proposée qui pourrait contribuer à enrichir les simulations est l'utilisation des méthodes topologiques, car ces méthodes permettent d'intégrer une dimension qualitative à la

compréhension des données sur lesquels les modèles sont construits. Enfin, Massumi rappelle que le numérique, lorsqu'il est expérimenté, doit toujours s'actualiser, et en le faisant il est déjà dans un circuit analogique. En prenant comme exemple de l'hypertexte, il affirme : « The freedom of the hypertext, its potentialization, is in the openness of its analog reception. »(Massumi, 1998, p. 310) Le documentaire interactif n'existe que dans ce va-et-vient entre le numérique et l'analogique. Nous revenons ainsi à la nécessité de situer le numérique à l'intérieur des processus analogiques dans lesquels il s'insère, ou, dans le cas de cette thèse, l'aspect analogique du documentaire interactif, abordé dans le deuxième chapitre de cet essai doctoral. Dans la deuxième partie de cette thèse, il sera question de la création de *On Chance*, un webdocumentaire réalisé dans le cadre de la présente recherche-crédation.

DEUXIÈME PARTIE : DISCUSSION AUTOUR DE LA CRÉATION

5. CHAPITRE V

ON CHANCE : UN DOCUMENTAIRE INTERACTIF AVEC ET SUR LE HASARD

Dans cette deuxième partie de l'essai doctoral, je discuterai le volet de création de la présente thèse. Dans ce cinquième chapitre, je présenterai le travail de création considérant, dans un premier moment, sa justification vis-à-vis des questions théoriques abordées dans la première partie du présent essai doctoral. Par la suite, il s'agira de situer *On Chance* dans deux modalités de récit qui le traversent : le documentaire interactif conversationnel, notamment pour les dispositifs de *chatbot*, et la vulgarisation scientifique en médias interactifs. Dans le sixième chapitre, il s'agira de faire un récit de pratique pour discuter le processus de création de *On Chance*. Des considérations finales seront formulées dans la section de conclusion.

5.1. Présentation de On Chance

Tel que mentionné plus haut, dans le volet de création de cette thèse, il a été question de créer un documentaire interactif avec et à partir de la recherche sur les notions de hasard et d'indétermination. Comme nous l'avons constaté dans la première partie de cette thèse, le développement des technologies numériques et l'accès à de vastes quantités de données ont ravivé pour beaucoup l'ambition déterministe de prévoir et contrôler l'avenir. Avec la vaste utilisation d'algorithmes prédictifs basés sur de grandes collections de données, la posture déterministe est souvent réaffirmée dans les

domaines des technologies numériques. Cela se produit, par exemple, lorsque les développeurs d'algorithmes affirment être en mesure de réduire l'incertitude et d'augmenter le contrôle sur les interactions avec le monde grâce à l'utilisation de modèles probabilistes. Néanmoins, comme nous l'avons aussi observé dans la première partie de cette thèse, depuis la fin du 19^e siècle, une forte tendance non déterministe dans les interprétations des théories scientifiques a émergé. Le développement de domaines tels que la physique statistique, la mécanique quantique, la biologie moléculaire, le chaos et la complexité algorithmique a contribué à la rupture avec le paradigme déterministe de la philosophie des sciences.

Comme nous l'avons remarqué, la théorie de l'information, développée principalement par Claude Shannon, et la théorie algorithmique de l'information, développée notamment par Gregory Chaitin (abordées dans les chapitres précédents) placent le hasard et l'entropie au cœur de la définition de l'information. Plus récemment, les travaux de Tiziana Terranova soulèvent le rôle de ces théories dans un présupposé courant, mais contesté, celui de l'immatérialité de l'information. Dans la création de *On Chance*, il s'agira de mobiliser ce cadre théorique dans une approche de vulgarisation scientifique pour établir une conversation avec l'interacteur autour des débats scientifiques et philosophiques sur les algorithmes prédictifs. Dans le chapitre 6, l'utilisation des médias interactifs pour la vulgarisation scientifique sera abordée plus en détail, mais avant cela il sera important d'aborder l'utilisation du *chatbot* comme outil d'interactivité dans la création de récits numériques.

Ce travail de création expérimente avec des modes de création qui embrassent l'incertitude et l'imprévisible et aborde le hasard comme thématique dans la création d'un documentaire interactif. Ainsi, l'enjeu principal de ce projet de recherche est d'articuler les concepts du hasard, de l'entropie et de l'information pour sortir des modèles prédictifs de la communication et observer comment le hasard peut proposer des manières alternatives de penser l'interactivité. *On Chance* propose à l'interacteur

de *clavarder*⁴⁶ avec un *chatbot* autour de la thématique du hasard et, selon le développement de la conversation, le contenu audiovisuel (original et d'archive) s'affiche sur l'interface (à côté de la fenêtre de clavardage, figure X).

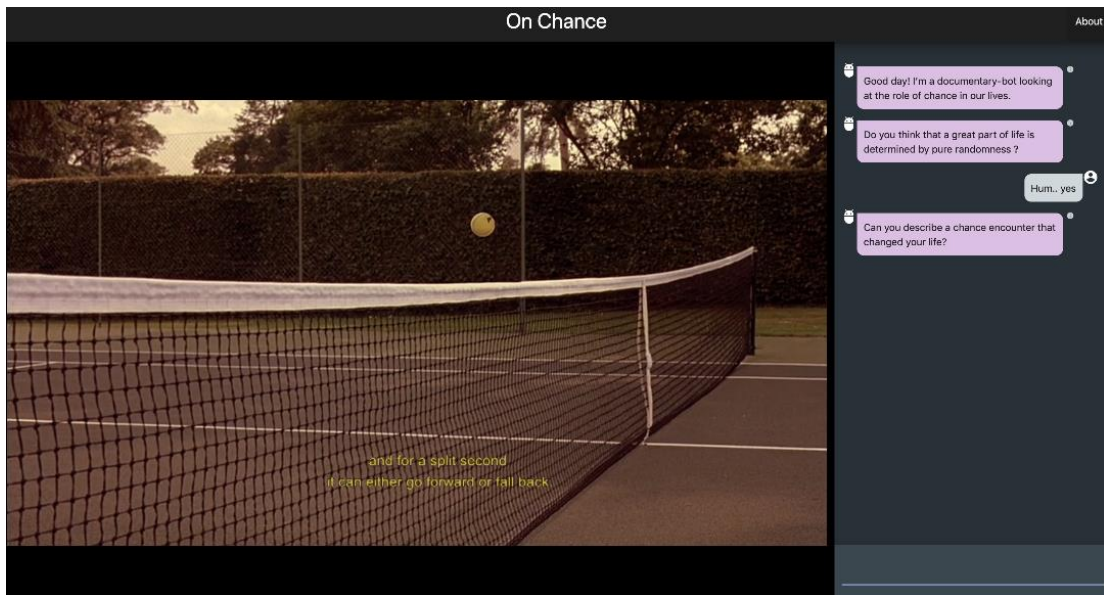


Figure 9 : Interface de On Chance

Comme il a été mentionné plus haut, *chatbot* est un dispositif d'interaction humain-ordinateur qui peut utiliser des techniques d'intelligence artificielle. Cette création permet ainsi d'explorer différents aspects de l'utilisation d'algorithmes d'intelligence artificielle dans le récit du documentaire interactif et des questions liées au hasard et à l'indétermination.

5.2. Chatbot et récit

Le *chatbot* est une interface d'interaction humain-machine utilisée dans différentes applications utilitaires et créatives. Les exemples de *chatbot* les plus connus aujourd'hui sont les assistants de voix comme Siri (dans les appareils de marque

⁴⁶ Avoir une conversation écrite sur une interface numérique.

Apple), Alexa (Amazon), entre autres. Dans cette recherche-cr ation, l'int er et est port e   l'intersection entre r cit et *chatbot*, plus sp cifiquement l'utilisation des *chatbots* comme interface pour la cr ation d'interactivit  dans le r cit (notamment le r cit audiovisuel). L'utilisation de *chatbots* dans le r cit a une histoire d'au moins 46 ans. *Tale-Spin*, d velopp  en 1976 par James Meehan, est cit e comme la premi re application de *chatbot* dans les r cits (Curry & O'Shea, 2011). Les r cits et les modes d'exploration de ce type de *chatbot* sont ferm s et bas s sur des r gles de conversation pr  tablies. Mais si les *chatbots* sont utilis s dans la cr ation de r cits depuis au moins 40 ans, l'application dans le domaine du documentaire audiovisuel est   ses d buts. Il sera question ici de revenir sur la litt rature et le corpus r cent qui se rapporte   l'utilisation des *chatbots* pour la cr ation de r cits.

Les d veloppements r cents des techniques d'apprentissage automatique ont permis une exp rience plus ouverte dans le cadre des *chatbots*. Depuis 2017, il est possible de remarquer une nette augmentation de la cr ation et de l'utilisation de *chatbots* dans divers domaines. Dans un rapport publi  en 2016 (Hoguet, 2016), le Fonds des m dias du Canada a soulign  quelques tendances principales dans les styles narratifs utilis s dans les *chatbots*. La premi re et la plus r pandue consiste   permettre   l'utilisateur de dialoguer avec des personnages virtuels (un corpus d'œuvres de cette tendance sera  tudi  plus en d tail). La deuxi me ligne de d veloppement pour les agents conversationnels narratifs est bas e sur les choix de l'utilisateur,   l'image des romans dont vous  tes le h ros (Hoguet, 2016). Une troisi me approche est l'utilisation des *chatbots* dans la cr ation de textes narratifs originaux. Le *chatbot* « Racter », d velopp  par William Chamberlain et Thomas Etter en 1983, est un des premiers *chatbots*    crire un livre (C. Curry & J. D. O'Shea, 2011). Plus r cemment, en 2017, Botnik Studios a entra n  un algorithme avec les sept livres de Harry Potter, pour ensuite, gr ce   l'utilisation de claviers pr dictifs, cr er un nouveau chapitre de la saga. Dans ce type de d marche, le r sultat n'est pas aussi pr visible que dans l'exemple *Tale-Spin* et offre une narration plus ouverte. Ce « nouveau chapitre de Harry Potter » non publi  est

devenu viral sur Twitter et a lancé une conversation publique sur l'utilisation d'algorithmes d'intelligence artificielle dans les récits.

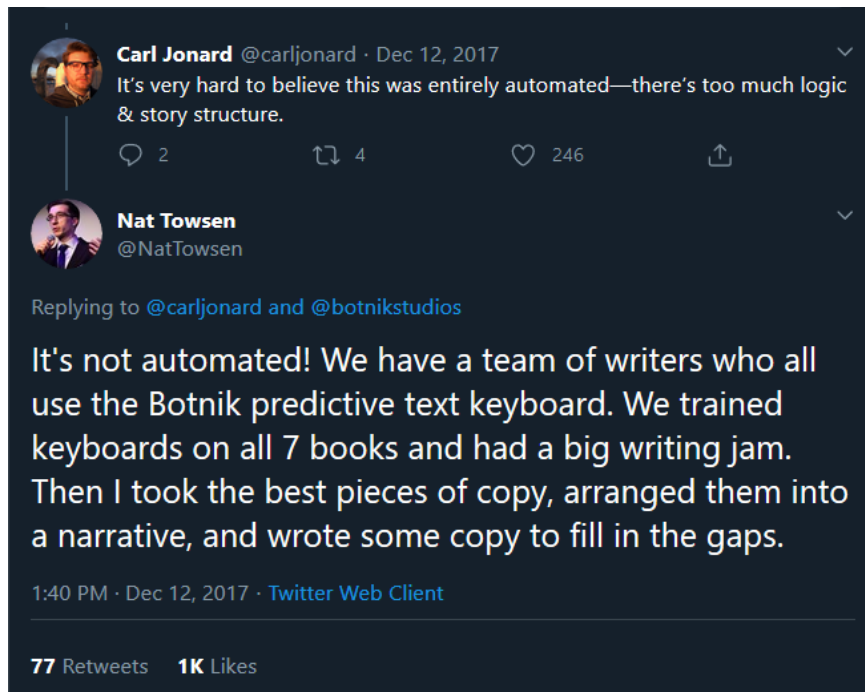


Figure 10 : Capture d'écran de discussion sur le réseau social Twitter

Comme l'a commenté un utilisateur : « Il est très difficile de croire que tout était automatisé : il y a trop de logique et de structure d'histoire » (@carljonard, 2017, traduction de l'auteure). À la suite de ce commentaire, l'un des créateurs de Botnik Studios a répondu en expliquant plus en détail le processus de création derrière le chapitre et le rôle de l'IA dans celui-ci : « Ce n'est pas automatisé ! Nous avons une équipe d'écrivains qui utilise le clavier à texte prédictif Botnik. Nous avons entraîné des claviers sur les 7 livres et nous avons eu une grosse session d'écriture. Ensuite, j'ai pris les meilleurs morceaux, je les ai arrangés dans un récit et j'ai écrit une version pour combler les lacunes » (@NatTowsen, 2017, traduction de l'auteure). Il est intéressant de constater la présence du travail de rédaction et d'édition réalisé par des humains, ce qui nous mène à voir ce type de démarche plus comme une cocréation humain-machine que comme un livre écrit par un *chatbot*.

En 2017, l'artiste numérique Hazel Grian a développé le récit de science-fiction en *chatbot*, *A Tendency to Spill* (Grian, 2017a), l'histoire de Romy, une jeune fille de 13 ans. Grian a également tenu un journal de bord en ligne où elle aborde le processus créatif de ce récit (Grian, 2017b), et comme on le verra, le hasard y a été abordé comme un des enjeux importants du développement du récit interactif en *chatbot*. Grian a utilisé la plateforme PandoraBots (testé également dans le cadre de la présente thèse de recherche-crédation) avec une diffusion en ligne sur un site web de type Wordpress. Après la page d'accueil du site web, l'interacteur peut se diriger vers l'image (l'affiche ou la couverture du récit, fig. 5), et le *chatbot* s'ouvre dans une boîte de dialogue du navigateur internet (fig. 6).

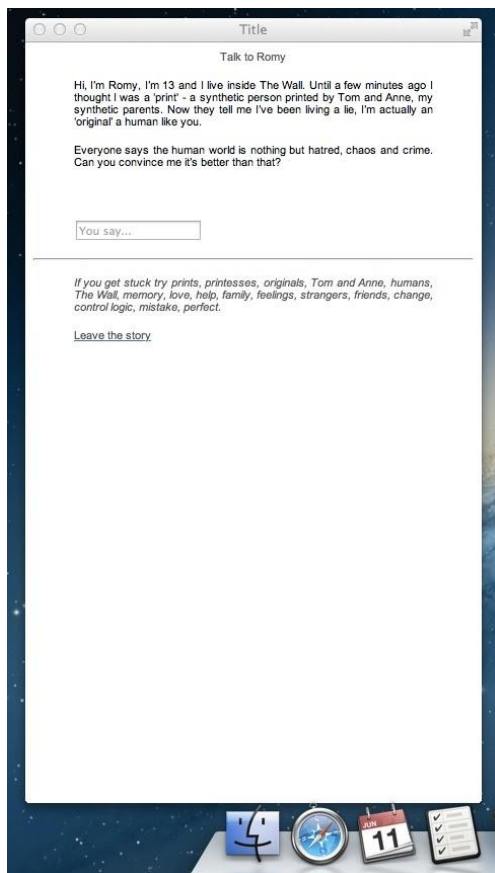


Figure 12: Page "call to action" *A Tendency to Spill* (Grian, 2017a)

Dans son journal, Grian décrit quelques pratiques qui ont facilité la réalisation de l'œuvre. Une des pratiques adoptées est l'usage de tableaux pour l'écriture du scénario :

Most interactive scripts I work on these days are done in spreadsheets. This is because there are several considerations going on simultaneously, one of which being what the audience/user is doing at any one time. So much of these scripts are based on 'if this then that' and spreadsheets are the most useful format. (Grian, 2017b, s. p.)

Une pratique commune dans les processus créatifs des médias interactifs a été recommandée par Grian aussi pour le développement de *chatbot* : la réalisation de nombreux tests avec des utilisateurs en cours de développement, et ce depuis les premières phases du processus. Grian décrit ses réactions au retour obtenu lors des premiers essais :

March 14th 2017 Big crisis after feedback round one: It needs more work than I realized. (...) I feel we have some way to go before i feel confident that it is interesting and moving as an interaction. To prevent the chatbot from being very frustrating, we are having to rethink the essence of a lot of it and the plot devices. (Grian, 2017b)

Après cette première séance de tests, Grian et sa collaboratrice, Constance ont obtenu des conseils de Rik Lander, spécialiste dans le développement de récits avec de l'intelligence artificielle. Les commentaires de Lander ont été très utiles au processus de Grian, alors elle les résume en quelques points, dont je reprendrai ici certains :

- l'importance d'un appel à l'action (« call to action ») clairement indiqué, « His advice actually lead me to change some of the basic story elements because I needed to take away a lot of the deliberate obfuscation of the story in order to guide the reader/user through the unfamiliar interactivity » (Grian, 2017b) ;
- les réponses données par le *chatbot* doivent encourager l'interacteur à continuer la discussion. La solution rencontrée par Grian était d'inclure des questions dans les interventions du chatbot, ce qui a causé d'autres types de problème : « I had to then decide to start using questions in order to open up the conversation,

which has lead to the problem as predicted, that Romy can't really converse very well. » (Grian, 2017b).

- il est nécessaire de choisir entre un récit littéraire interactif et un chat instantané, car le récit littéraire implique une grande quantité de texte et une structure linéaire qui sont incompatibles avec le format de *chatbot*.

Ce dernier constat a entamé une série d'essais concernant le degré de linéarité et de préprogrammation du récit, des questions qui suscitent un intérêt particulier dans le cadre de la présente recherche-création. Voici comment Grian décrit le test réalisé et le résultat obtenu :

Test— On Rik's suggestion I'll pick out 20 key points that will remain the backbone of the narrative and we will see if we can find a piece of code that allows us to deliver them in narrative order. I'm trying to make sure there is a linear backbone adhered to as well as the feeling that the chat is natural and isn't preprogrammed. Ideally the nonlinear chat gives the reader a sense of agency while those key linear narrative points deliver the story.

Result - As it turns out, due to limitations of the platform and/or our coding knowledge, we aren't able to have anything appear in a particular order. I feel that the final version of the chat is circular rather than linear. However Constance feels that's not quite the right description: for here rather than 'circular', it is 'diffuse' instead of linear. It is still luck of the draw as to whether the random comments make sense to the reader/user and if they have another go then they will see that. (Grian, 2017b)

Il est important de constater que, pour Grian, le problème de la linéarité et de la structure narrative dans le récit interactif en *chatbot* est lié à l'équilibre entre les dimensions préprogrammées et aléatoires du récit. Grian décrit que cet équilibre est problématique avec la plateforme PandoraBots, comme la suite des tests réalisés l'ont démontré :

Test — How do we not reveal plot points ahead of time?

Result - There's no way of controlling that with this system. It's a random delivery of the story, all depending on the direction of the conversation taken by the reader/user.

Test - Does having no linear narrative element mean that I am sacrificing the satisfying notion of an 'ending'?

Result - I feel it probably is. Without it the end is simply when the reader has had enough of chatting. (Grian, 2017b)

Ces passages mettent en évidence un des défis du récit interactif en *chatbot* : créer une expérience narrative sans les marqueurs habituels du récit (comme la fin). Grian continue à développer des hypothèses dans son processus créatif pour répondre à cette question. Pour tester les nouveaux développements du *chatbot*, Grian invite les spécialistes des médias interactifs et de l'intelligence artificielle, Tim Kindberg et Tom Abba. Le premier commentaire revient sur la dimension aléatoire des réponses du chatbot : « Unintentional repetition of (the random set of) replies 'takes you out of the experience' (I agree) and he felt that he got 'no plausible responses' (which is disappointing). » (Grian, 2017b) Ces difficultés ont provoqué une remise en question de la pertinence de l'interactivité dans le contexte de ce récit spécifique. Tom Abba revient dans ses commentaires sur les éléments d'introduction à l'expérience : « The key element as Tom pointed out was to put back the call to action and how long to spend on the experience. » Il est intéressant de constater qu'Abba conseille l'indication dans l'introduction de la durée recommandée pour l'expérience.

Enterre-moi, mon Amour (2017) est un récit interactif qui se déroule dans la forme d'un échange de messages entre l'interacteur et Nour, une réfugiée syrienne pendant son parcours vers la Turquie. L'interacteur joue le rôle Majd, le mari de Nour. Le récit propose comme outil d'interactivité une simulation de *chatbot* : à chaque message envoyé par Nour, un nombre limité de choix de réponses se présente à l'interacteur, comme montrent les figures suivantes (des captures d'écran de l'interface du récit) :



Figure 13 : Interface 2 de *Enterre-moi, mon Amour* (2019)



Figure 14 : Interface 1 de *Enterre-moi, Mon Amour*

La plupart de messages sont en texte, et ils peuvent contenir des emojis ou images. Pour avoir accès au récit, l'interacteur peut acheter l'application sur une plateforme pour téléphone intelligent ou jeu vidéo (Google Play, Apple Store, Nintendo Switch, et autres).

Enterre-moi, mon Amour est une co-production de The Pixel Hunt et Figs, deux studios de création numérique, et la chaîne franco-allemande de télévision et communication Arte. Comme décrit le producteur et game designer Florent Maurin (in: Fonseca, 2019), le récit a été inspiré du web-reportage « Le voyage d'une migrante syrienne à travers son fil WhatsApp », réalisé par Lucie Soullier et Madjid Zerrouky, produit par le

journal « Le Monde » et publié en 2015. Selon Maurin, il a été bouleversé par la familiarité des messages échangés dans des contextes si différents et a décidé alors de contacter les journalistes ainsi que la femme dont le reportage s'inspire pour proposer la création du récit interactif en format de jeu vidéo. Le scénariste Pierre Corbinais décrit un processus de co-création entre lui, Lucie Soulier (journaliste), Dana Iul (femme dont le personnage de Nour est inspiré) et Florent Maurin à partir du matériel d'archive (les messages échangés sur WhatsApp), le reportage et d'autres types de documentation (in: Fonseca, 2019).

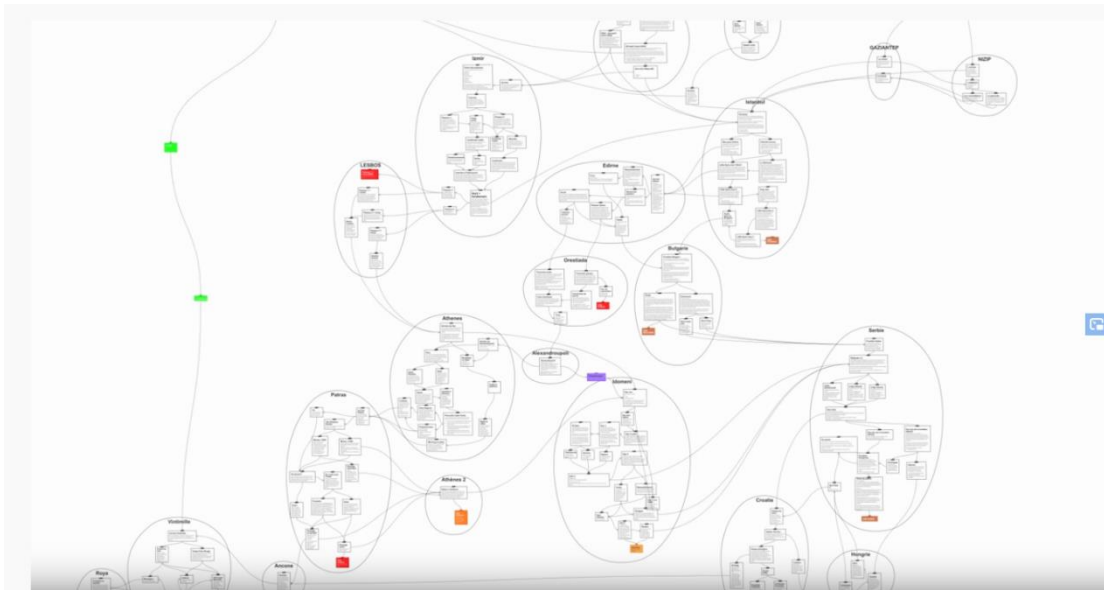


Figure 15 : Scénario interactif (Fonseca, 2019)

Un aspect particulier du jeu est dans sa temporalité : l'interacteur peut faire le choix de faire l'expérience du récit « en temps réel », soit dans la temporalité des messages originelles, comme indiqué dans les figures suivantes :



Figure 16 : Interfaces de réglages avec les options de « vitesse du jeu » et « vitesse d'apparition des messages »



Figure 17 : Interface Nour « occupée »

Si l'interacteur choisit le mode « temps réel », tel que dans un échange de message quotidien, le récit va se jouer pendant plusieurs jours, et à différents moments le personnage de Nour aura le statut de « occupée » (figure 17) et l'interactivité est inactive. Dans ce mode, ce n'est pas l'interacteur qui décide le moment et la durée de son expérience. La cadence du jeu, comme dans un échange de message quotidien, dépend des deux interlocuteurs.

Une étude brésilienne a abordé l'utilisation de récits en *chatbot* dans le contexte des campagnes publicitaires de sensibilisation (Petters, 2019). Beatrice Petters remarque que grâce au web social, les agences de publicité développent des dispositifs conversationnels pour établir des dialogues entre marques (souvent remplacées par des personnages) et clients. Petters constate que cette pratique peut simuler une communication plus horizontale entre les compagnies et le public. Un autre aspect soulevé par Petters est la nécessité qui ont les entreprises de contrôler l'évolution de la conversation, alors qu'habituellement les conversations entre humains sont libres et imprévisibles (Petters, 2019, p. 254).⁴⁷ L'étude a analysé le chatbot *Fabi Grossi*, un récit créé par Unicef en 2018 dans le cadre d'une campagne de sensibilisation sur les questions liées à la vie privée en ligne. Le public-cible de la campagne ce sont les adolescents et jeunes adultes.



Figure 18 : Interfaces de Fabi (2018)

⁴⁷ La discussion entre le contrôle et l'imprévisibilité du récit est abordé dans le deuxième chapitre de cette thèse par la perspective du design d'interactivité.

De cette analyse, Petters soulève deux ensembles de conclusions importantes. Le premier concerne ce qu'elle a appelé « la factualisation », ou l'utilisation d'un style documentaire dans le récit comme stratégie de communication persuasive. Dans le cas de *Fabi*, il s'agit de l'utilisation d'un récit réaliste et crédible, en première personne, du personnage fictif nommé *Fabi*. L'auteur soulève d'autres stratégies de la factualisation, comme l'utilisation d'éléments audiovisuels (*Fabi* envoie des images et audios à l'interacteur), l'utilisation d'énoncés qui renforcent le réalisme du récit (langage informel et narration d'événements personnels) et l'inclusion d'informations spatiotemporelles dans le récit ainsi que de détails (indications de lieux, dates, noms d'autres personnages, entre autres). Petters note qu'il est important d'employer un langage adapté au support utilisé (dans le cas de *Fabi*, le Facebook Messenger). Le deuxième ensemble de conclusions se rapporte à l'utilisation de l'expérience personnelle comme thème du récit. Selon Petters, la conversation basée sur l'expérience personnelle réduit certaines barrières pour la participation de l'interacteur dans la conversation. L'expérience personnelle ne demande pas de connaissances dans des domaines spécifiques, met de l'avant la sociabilité de l'expérience et établit une relation de proximité avec l'interacteur (Petters, 2019, p. 263). Néanmoins, Petters remarque dans *Fabi* une faible capacité à répondre aux questions posées par l'interacteur lorsque le scénario prédéterminé n'est pas suivi. Il est possible de constater alors qu'un des enjeux de l'utilisation des *chatbots* dans les récits est la capacité des *chatbots* à répondre aux entrées imprévues du participant. Enfin, Petters soulève des aspects d'ordre éthique à être considérés dans l'utilisation des *chatbots* dans la communication persuasive, comme l'importance de considérer les relations de pouvoir existantes dans la conversation entre le personnage (qui remplace souvent une marque) et l'interacteur.

Il existe d'autres nombreux exemples de l'utilisation de *chatbots* pour la création de récits pour de causes sociales, ainsi que de causes humanitaires et écologiques. Nous allons étudier brièvement deux autres exemples de ce type d'application des *chatbots* :

Yeshi et HappyBot. *Yeshi* est un chatbot déployé sur Facebook Messenger créé par l'agence publicitaire *AKQA* et les organisations *LOKAI* et *Charity : Water* qui raconte à la première personne l'histoire d'une jeune fille, *Yeshi*, qui vit en Éthiopie. L'histoire se raconte au fur et à mesure que le public échange des messages avec *Yeshi*. Au-delà des messages textes, certains messages envoyés par le *chatbot* peuvent contenir des images, des audios ou encore des vidéos. La durée recommandée pour l'expérience est de 2,5 heures. L'aspect audiovisuel du récit est ainsi exploré par ce projet. Voici des images de l'interface :

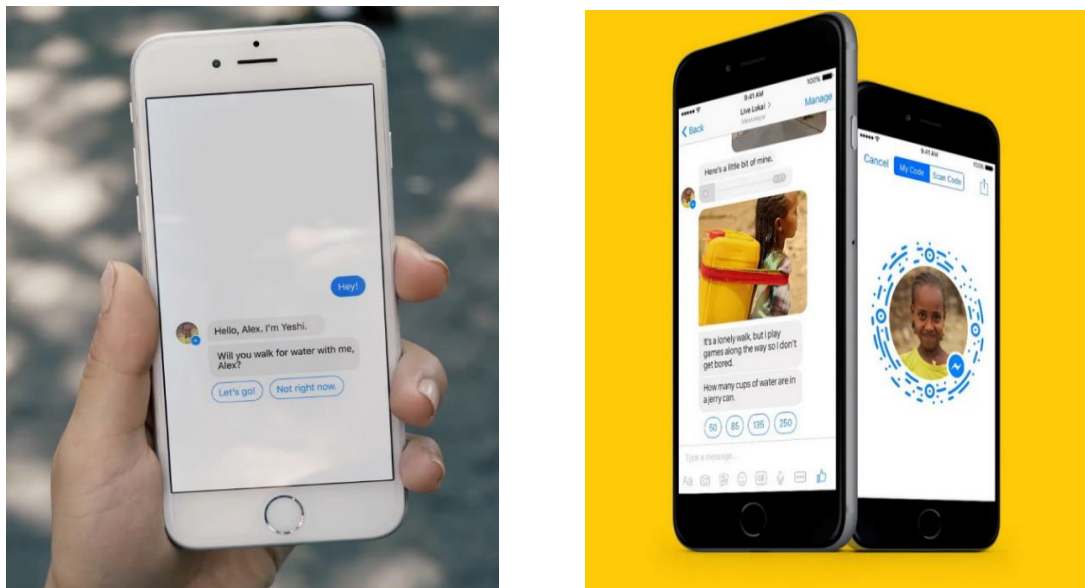


Figure 19 : Interfaces de *Yeshi*

L'utilisation d'éléments audiovisuels est très efficace dans l'expérience du récit, car elle permet d'illustrer et d'enrichir l'aspect textuel de la conversation. Il est intéressant de remarquer que l'interactivité conversationnelle est souvent limitée à un choix parmi certaines options de réponse proposées par le *chatbot*. Ainsi on constate que l'expérience de *Yeshi* rapproche l'interactivité conversationnelle de l'interactivité hypertextuelle, ou de sélection, dans le sens où la principale action de l'interacteur à laquelle le système informatique répond est constituée par un choix parmi un nombre limité de liens qui fait avancer le récit selon des ramifications préétablies.

HappyBot est récit interactif en *chatbot* créé par *BBC Earth* et l'agence de communication, *Monterosa*. Dans une étude de cas publiée sur le site de *Monterosa*, le récit est décrit de la façon suivante :

The Bot lives on Facebook Messenger and asks users a short series of questions, creating a personalised Happiness Profile. It then correlates those user preferences with BBC archive footage to produce individual Real Happy Moments: personalised, sharable montages with scenes selected using intelligent tag matching.

Le *chatbot* est utilisé ainsi comme un outil pour la personnalisation de contenu documentaire audiovisuel de l'archive de la *BBC Earth*. Le clip vidéo est envoyé au public à la fréquence souhaité. Selon les résultats publiés par l'entreprise, *HappyBot* a généré plus de 100.000 vidéos⁴⁸.

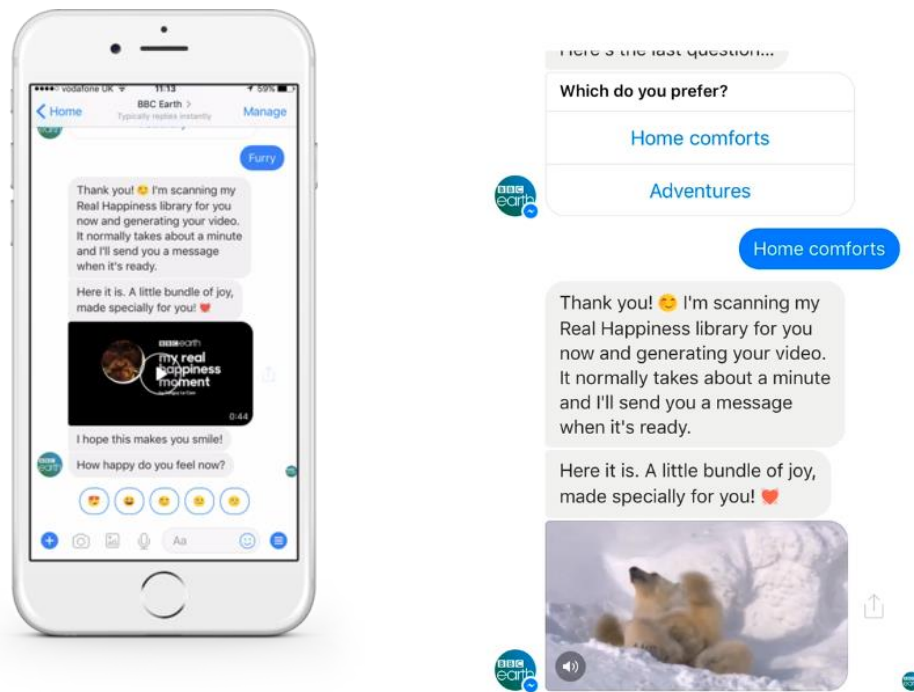


Figure 20 : Interfaces de *HappyBot*

⁴⁸ <https://www.monterosa.co/bbcearth>

Dans une étude sur les relations entre le récit cinématographique et le développement de *chatbot*, Garner (2014) analyse comment les techniques narratives développés dans le cinéma peuvent contribuer à créer un engagement affectif avec le public dans un *chatbot*. Garner argumente que dans le cinéma il existe un ensemble de techniques pour produire la perception de réalité chez le public (*suspension of disbelief*, en anglais) et que ces techniques peuvent contribuer pour recréer cette perception dans le contexte du *chatbot* :

When people go to the cinema and they enjoy it, they may know the whole time that it is made by a movie studio, but from time to time, they may find themselves forgetting about the machinery, and focusing on the story, the dialogue, and the characters. Sometimes something similar happens with chatbots. (Garner, 2014, p. 18)

Il fait un parallèle entre le montage et le rythme du dialogue, où l'échange des phrases correspond à la suite de plans et le changement de sujet correspond au changement de séquence. Il souligne l'importance du rythme de la conversation et défend l'utilisation d'un délai pour la réponse du *chatbot* (comme s'il s'agissait du temps nécessaire à la rédaction du message) (Garner, 2014, p. 18). Une des techniques mentionnées par Garner est celle du « MacGuffin », un thème mineur dans le récit qui sert à faire avancer l'histoire. Garner explique comment cette technique peut être adaptée au cinéma :

In chatbots, MacGuffin is some topic or thing that the chatbot brings up to either change the subject or appear to have a question to ask of its own. It is a very common occurrence in chatbot conversations, and chatbots are faulted for this, as a sort of crutch when the programme has no clue how to respond. The human interlocutor is often uninterested in changing the subject for instance. If there were a part of the chatbot programme for error control, it would be the MacGuffin code that would determine when to try and change the topic, or possibly worse, reply with a non sequitur. (Garner, 2014, p. 19)

Malgré le fait que les deux dispositifs (cinématographique et conversationnel) sont utilisés dans le contexte de récits, Garner remarque aussi des différences entre

l'expérience cinématographique et celle du *chatbot* et les défis que cela implique pour le développement du récit :

The unlimited domain of chatbots makes it much harder to create a “flow” in a conversation. It is easier to imagine the chatbot as a narrative device, but it is interactive by design. Not only does it have to be good in telling things, but it has to be good at listening to things too. (Garner, 2014, p. 18)

Garner résume ainsi un défi remarqué souvent dans la création de récits conversationnels, soit la difficulté de combiner la linéarité du récit et l'interactivité du *chatbot*.

Bruce et Sue Wilcox ont développé des *chatbots* lauréats à plusieurs reprises du principal prix international dédié au *chatbot*, le Loebner Prize. Dans un article, les Wilcox décrivent certains acquis dans le processus de développement (Wilcox & Wilcox, 2013). Il est important de noter qu'ils utilisent le langage ChatScript dans le développement de ses 3 *chatbots* primés (Suzette, Rosette, Angela). ChatScript est un langage de software libre qui propose un traitement de langage naturel par un système de type expert (basé sur des règles qui produisent des réponses aux différentes entrées) qui, comme observé dans la première partie de cette thèse, se situe dans une approche déductive de l'intelligence artificielle.

Après une description des principes du langage ChatScript, les auteurs abordent six étapes du développement de *chatbots* : (1) le design de la personnalité du *chatbot*, (2) le design des points de conversation, (3) le script, (4) les tests et (5) l'intégration et (6) l'évolution. Les auteurs soulignent l'importance de (1) développer le *chatbot* comme un personnage avec une biographie rédigée et partagée avec toute l'équipe de développement : « The more a user can model the character's personality, the more engrossed the user can become in the reality of a fictional world. » (Wilcox & Wilcox, 2013, p. 3). Une fois que la biographie du personnage est développée, les auteurs commencent (2) le travail sur le design des points de la conversation (*topics*, en anglais). Les points de conversation sont un ensemble de règles pour associer des

entrées (venues de l'utilisateur) à des sorties reliées (réponses du *chatbot*). Les auteurs décrivent différents types de points de conversation : gambit, réaction, objection, générique, histoire et contrôle. Le point de conversation de type « gambit » est le plus utilisé par les *chatbots*. Ce point de conversation regroupe des affirmations ou des questions sur un sujet précis. Le sujet est déterminé par un système de recherche basé sur des mots-clés : le *chatbot* utilise les mots de l'interacteur pour faire une recherche parmi la biographie du personnage et déclencher ensuite un point de conversation associé au sujet du mot-clé. Une fois qu'un point gambit est déclenché, le *chatbot* alterne des questions et des affirmations à propos du sujet. Les auteurs font le commentaire suivant sur la dynamique recommandée pour l'ordre et l'alternance entre questions et affirmations :

Usually, but not always, the user is asked first. If you volunteer your characters' attitude first you build up a sense of controlling the conversation and make it hard for the user to think of their own response. (...) Sometimes we switch the order. Sometimes we do a couple of questions in a row or a couple of statements in a row. Variety. Make it feel natural and not structured. But usually we alternate question then answer. (Wilcox & Wilcox, 2013, p. 4)

L'étape suivante est (3) le développement du *script* (à la fois scénario et code des règles). Un des résultats obtenus est le constat que l'édition du texte doit considérer le design de l'interface : « For a first pass, Bruce restrict a rule's output to 200 characters » (Wilcox & Wilcox, 2013, p. 5). Comme l'espace pour la réponse en texte du *chatbot* est limité dans l'interface, il est important d'être concis dans les formulations. Un autre résultat important de cette étape est le partage des différentes modalités de rédaction : la créative et la technique : « One thing that has been clear in all cases so far, is that one shouldn't expect the creative writer to do actual scripting ». (Wilcox & Wilcox, 2013, p. 5) Les auteurs abordent alors le rôle des réponses de type « objection » (ou *quibble*, en anglais) dans la rédaction du scénario, notamment pour le passage d'un sujet (ou d'un gambit) à un autre : « when the bot exhausts gambits in a topic (on its last gambit), it sets a flag that tell the system that after it quibbles, it should not launch

a new gambit. It should just rest for that round. This prevents the user experiencing an abrupt change of topic. » (Wilcox & Wilcox, 2013, p. 7) C'est aussi à cette étape que les mots-clés des points de conversation sont déterminés.

Les étapes de (4) tests et (5) d'intégration ne seront pas abordées ici, car les auteurs se concentrent sur les spécificités du langage ChatScript. Dans l'étape (6) sur l'évolution, les auteurs analysent les variations dans les habiletés et la composition des *chatbots* au cours de leur parcours créatif. Ils remarquent deux enjeux majeurs pour l'évolution des *chatbots* : les réponses adéquates aux changements de sujets et la démonstration d'émotion de la part du chatbot. Sur le premier des enjeux, ils affirment : « Knowing where the user wants to go is critical and letting them know their request has been heard is also important. » (Wilcox & Wilcox, 2013, p. 9) Par ailleurs, ils remarquent également qu'il peut avoir confusion si les mots-clés de chaque sujet ne sont pas assez précis : « Sometimes we find similar material in multiple topics. They need to be gathered together to avoid what we call keyword muddiness – a state where the system uses a keyword to go to a topic not primarily about that keyword. » (Wilcox & Wilcox, 2013, p. 9) Une autre considération importante sur le design de la conversation concerne la forme habituelle pour un changement de sujet : « Now we start topics with a question after a brief opening statement. So the user is hooked into the conversation and knows it is going somewhere new immediately. » (Wilcox & Wilcox, 2013, p. 9) Enfin, les auteurs décrivent que certains sujets plus exotiques n'émergeaient pas souvent dans les conversations, car les mots-clés pour les déclencher étaient peu utilisés par l'interacteur. Ils ont développé alors une stratégie pour contourner ce problème, il s'agit d'attribuer une pondération plus élevée pour l'utilisation des réponses des sujets exotiques.

Le deuxième enjeu considéré dans la sixième étape (l'évolution des *chatbots*) est la capacité du *chatbot* de démontrer de l'émotion. Les auteurs remarquent que très souvent les utilisateurs abordent des questions liées à la vie intime et à la sexualité du chatbot, souvent de manière explicite et violente. Les auteurs alors ont pris la décision

de formuler des réactions à ce type de comportement, ainsi que d'autres actions de l'interacteur qui pourraient provoquer de l'émotion. Les considérations finales de cet article abordent la possibilité d'un effet d'étrangeté si le réalisme des *chatbots* dépasse un certain seuil, «the uncanny valley effect»: «As we get closer to human conversation standards, will people find it less fun? Will they get creeped out? We aim to find out.» (Wilcox & Wilcox, 2013, p. 9)

Curry et O'Shea ont développé un récit éducatif en *chatbot* basé sur un système similaire à celui présenté par Wilcox et Wilcox (C. Curry & J. D. O'Shea, 2011). Curry et O'Shea utilisent un système qui associe des mots-clés de l'utilisateur avec des points de conversation (topics): «This looks at the incoming sentences and decides which topic has the closest keyword match, then scans the rules of the topic to decide what to do» (C. Curry & J. O'Shea, 2011, p. 7). Ils remarquent également l'importance d'instaurer dans le *chatbot* une mémoire de longue durée pour que le récit puisse avoir une continuité à chaque nouvelle visite de l'interacteur. Ils concluent que la structure développée pour le *chatbot* en question peut être adaptée et réutilisée dans d'autres récits.

Le système DialogFlow, utilisé dans le développement de *On Chance*, a été utilisé dans une étude sur l'application de *chatbots* dans l'entrevue de candidats en processus d'embauche (Purohit, Bagwe, Mehta, Mangaonkar, & George, 2019). Il ne s'agit pas dans ce cas de l'application de *chatbots* dans un récit, mais comme l'entrevue est une des techniques documentaires utilisées dans *On Chance*, certains aspects de cette étude peuvent être pertinents pour la présente recherche-crédation. L'étude obtient des résultats favorables à l'utilisation de *chatbots* dans le contexte d'entrevue dans un éventail limité de domaines.

Une étude récente propose un modèle pour analyser les six principales plateformes de développement de *chatbots* (Pavel, Ivo, Mikuláš, František, & David, 2019). Les plateformes évaluées sont : Chatfuel, ManyChat, Botsify, Microsoft Bot Framework,

Dialogflow et IBM Watson. L'analyse est réalisée dans le contexte de gestion d'entreprise, néanmoins on propose d'adapter les critères utilisés pour l'analyse des *chatbots* aux enjeux concernant le récit, plus spécifiquement de la création de *On Chance*. Les auteurs divisent les plateformes en deux types : WYSIWYG et « all-purpose ». Le premier type de plateforme, WYSIWYG, a comme objectif principal faciliter le développement de *chatbot* pour un public avec peu de connaissance technique. Ce type de plateforme offre ainsi un choix limité de modifications de l'interface et des systèmes prédéfinis (Pavel et al., 2019, p. 1543). Les plateformes « tout usage » permettent plus de liberté dans la combinaison de différentes capacités et fonctionnalités du système et de l'interface. Dans le contexte de *On Chance*, considérant la nécessité d'une importante flexibilité de l'interface, la plateforme indiquée est de type « tout usage ».

Les critères établis par les auteurs dans l'analyse des plateformes de développement de *chatbots* sont les suivants : Maturité du traitement automatique de langage nature (*NLP*) ; Capacités additionnelles (au-delà du *NLP*) ; Interactions avec API (présence de *webhooks* et autres APIs) ; Supports de diffusion (Facebook, webchat, Alexa, etc.) ; Services d'analytiques ; Transfer d'*Intents* (technique qui permet la mise à jour de catalogues de vente sans nouveau développement) ; Configuration automatique ; Prix. Après observation, il est possible de constater que parmi les critères indiqués, le seul qui est spécifique au domaine des de la gestion d'entreprise est le « Transfert d'*intents* ». Il devient alors important d'observer les résultats obtenus par cette étude et de les adapter en fonction d'une nouvelle distribution du poids de chaque critère dans le contexte du récit. Lorsque cette démarche est effectuée et les particularités du développement de *On Chance* sont considérées, la plateforme indiquée est Dialogflow.

Une étude de Seyf Eddine Guesmia (2019) propose également un système d'évaluation pour les *chatbots* dans le domaine de la vente en ligne. Encore une fois, le domaine d'application en question n'est pas le même que celui de l'objet de cette thèse, à savoir l'utilisation d'agents conversationnels dans le récit, mais il est possible malgré cela de

considérer la transposition de quelques résultats au contexte du récit. Guesmia définit 5 types de critères à être considérés pour la qualité d'un *chatbot* : la qualité de l'information fournie, la qualité du système, la qualité de l'interaction, apparence visuelle et innovation (Guesmia, 2019, p. 41). La qualité de l'information dépend de la complétude (les réponses sont complètes et suffisantes et toutes les questions sont répondues), de la pertinence (les réponses sont bonnes à savoir, pertinentes, utiles et répondent au besoin de l'interacteur), de la précision (les réponses sont précises et clairement formulées), de l'actualité (les réponses sont récentes et « à jour ») et de la cohérence (chaque réponse est cohérente avec les réponses précédentes, les questions et réponses ont une suite logique). La qualité du système dépend de la facilité d'usage (l'interface est intuitive et le *chatbot* est facile à utiliser), du temps de réponse (le *chatbot* est rapide et le temps de réponse est raisonnable et acceptable), de la fiabilité (le *chatbot* fonctionne bien et ne cause pas de frustration chez l'interacteur) et de l'efficacité (l'interacteur obtient des réponses à ses questions sans trop d'échanges et le montant d'information échangée est adéquat). La qualité de l'interaction est le résultat de la convivialité (les réponses sont respectueuses, polies et amicales) et de la personnalisation de la communication (réponses adaptées aux besoins et préférences spécifiques à l'interacteur). Les deux dernières dimensions de la qualité du *chatbot* sont l'innovation (le *chatbot* est perçu comme unique et créatif) et l'apparence visuelle (l'interface est plaisante, attirante et agréable). Ces critères seront utiles pour orienter certains aspects du développement de *On Chance*.

Des résultats positifs ont été obtenus également de l'utilisation de *chatbots* pour la création de récits dans le domaine du tourisme (Clarizia, Colace, Lombardi, & Pascale, 2018). Les auteurs ont développé et analysé un système qui intègre une application pour créer des récits personnalisés associés au local où se trouve l'utilisateur. Le *chatbot* est utilisé au début du processus comme un outil pour récolter des informations sur l'utilisateur. Ces informations sont utilisées par le système pour personnaliser le récit qui sera proposé à l'utilisateur sur l'endroit en question. Il est possible de constater

ainsi que l'utilisation de *chatbots* dans les récits est une pratique déjà présente dans les communications de divers domaines, tels que le tourisme, la publicité et les ressources humaines. Par ailleurs, les *chatbots* ont d'autres nombreuses applications dans domaines très variés qui ne seront pas abordés dans cette thèse (à titre d'exemple, citons les domaines du jeu vidéo, le journalisme, le droit, la gestion d'entreprise, le commerce électronique, entre autres).

5.3. La perspective de la vulgarisation scientifique

Au moment de la rédaction de cette thèse, la pandémie du coronavirus SARS-CoV-2 a eu lieu. Cet événement nous intéresse ici, car il a mis en évidence deux aspects importants de la vulgarisation scientifique que je souhaite aborder dans la présente section de cette thèse. Le premier aspect est la crise de confiance dans les sciences chez certains groupes sociaux (Castelfranchi, Vilela, Lima, Moreira, & Massarani, 2013; Gauchat, 2012; Hmielowski, Feldman, Myers, Leiserowitz, & Maibach, 2014). Lors de la pandémie, ce constat s'est traduit par exemple dans les nombreuses contestations des avis des scientifiques sur les protocoles à être adoptés dans le combat à la propagation du virus (Raymond, 2020). Or, le champ de la vulgarisation scientifique aborde depuis des décennies les relations existantes entre les discours scientifiques et la société et peut ainsi proposer des pistes de réflexion sur le problème de la confiance sociale dans les sciences.

Le deuxième aspect que la pandémie a mis en évidence pour la communication scientifique est l'éventail de nouvelles ressources que les médias interactifs peuvent apporter à la vulgarisation scientifique. Face à l'incertitude causée par la pandémie, la consommation et la production de contenu de vulgarisation scientifique ont augmenté de manière très importante. Pour avoir une dimension de ce phénomène, prenons l'exemple d'un article en ligne (Stevens, 2020) publié par le journal *The Washington Post* qui a utilisé des animations pour illustrer les simulations de l'effet de l'isolement

social et la croissance exponentielle du nombre de personnes infectées par le virus. L'article en question est devenu en quelques jours le texte avec le plus grand nombre d'accès sur l'ensemble du site du journal *The Washington Post*, avec des millions d'accès, et a été traduit vers au moins 14 langues depuis sa publication.

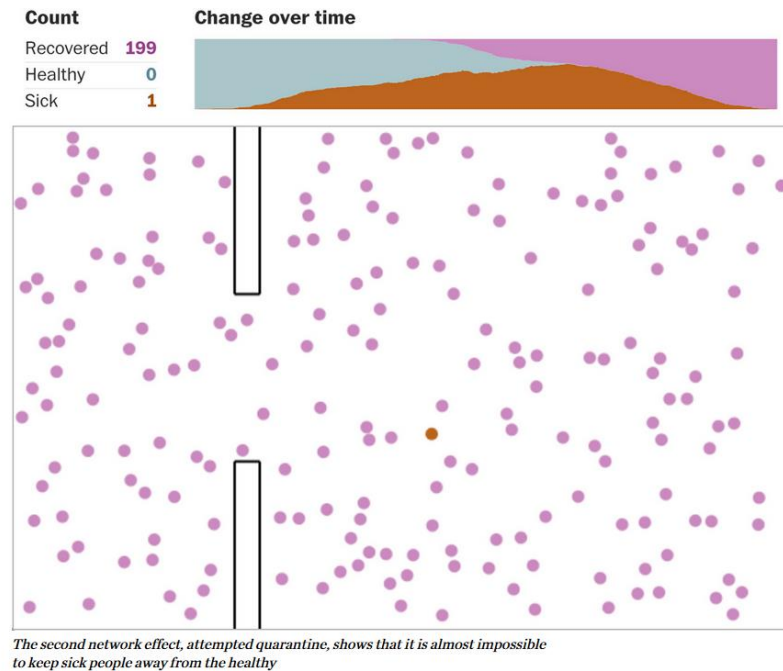


Figure 21 : Illustration de l'article de Harry Stevens « Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve” (Stevens, 2020)

Il est intéressant de noter que l'épidémiologie est un domaine dans lequel les analyses statistiques et les modèles probabilistes, des domaines abordés également dans la présente recherche-création, sont largement utilisés. Ainsi, certaines publications de vulgarisation scientifique liées à l'épidémie discutent également des questions abordées par la présente recherche-création sur l'interprétation des statistiques et des probabilités, ainsi que sur leur rôle dans les méthodologies scientifiques contemporaines, en voici un exemple :

It seems unlikely that a random bat virus should somehow jump into a susceptible human. But when you consider millions of people, in regular

contact with millions of bats, which carry tens of thousands of new viruses, vanishingly improbable events become probable ones. (Yong, 2020, s.p.)

Nous constatons la présence des débats sur l'interprétation des probabilités dans les articles de vulgarisation scientifique pendant la pandémie.

Il existe une convergence dans les résultats obtenus par de nombreuses études réalisées dans différents pays vers le constat d'une crise de confiance dans les sciences chez certains groupes sociaux (Castelfranchi et al., 2013; Gauchat, 2012; Hmielowski et al., 2014). Or, cela est le résultat de processus sociaux et politiques complexes. Comme constate Gauchat dans une étude réalisée aux États-Unis en 2012, « this study shows that public trust in science has not declined since the 1970s except among conservatives and those who frequently attend church. » (Gauchat, 2012, p. 182). Un autre résultat de l'étude réalisée par Gauchat indique que le niveau de scolarisation a moins d'influence sur la confiance dans la science que les inclinations politiques,

(...) this study greatly complicates claims of the deficit model, which predicts that individuals with higher levels of education will possess greater trust in science, by showing that educated conservatives uniquely experienced the decline in trust. (...) Moreover, it suggests that scientific literacy and education are unlikely to have uniform effects on various publics (...). (Gauchat, 2012, p. 182)

Le modèle déficitaire de la vulgarisation scientifique mentionné par Gauchet sera discuté plus loin. Des résultats similaires ont été obtenus aussi au Brésil (Castelfranchi et al., 2013), ce qui indique que le scénario observé par Gauchat n'est pas restreint aux États-Unis. Dans cette section, il s'agira de revenir sur les discussions dans le domaine de la vulgarisation scientifique concernant les relations de confiance entre le public et la science.

Comme affirme Bernadette Bensaude-Vincent, « la vulgarisation scientifique se fonde et se légitime sur le postulat d'un fossé grandissant entre l'élite scientifique et la masse du public. » (Bensaude-Vincent, 2010, p. 3) Néanmoins, l'auteure critique cette vision, car elle suppose que le public reçoit de manière passive une connaissance qui

n'appartient qu'aux scientifiques, que le rôle des vulgarisateurs est de traduire le langage des experts à un langage de plus facile compréhension et de parler ainsi « au nom de la science », et qu'il suffit de vulgariser la connaissance scientifique pour outiller la société à participer des débats scientifiques. Bensaude-Vincent constate que :

D'une part, le public n'est plus une masse passive et indifférenciée de consommateurs définis par leur défaut de science. (...) D'autre part, la science n'est plus considérée comme une activité neutre, transcendant les intérêts particuliers, indifférente aux valeurs sociales ou morales. (Bensaude-Vincent, 2010, p. 7)

L'auteure constate alors la nécessité d'une transformation de la vision de ce qui est et fait la vulgarisation scientifique, et d'un renouvellement des rapports entre la science et le public.

Emily Dawson a étudié le problème de l'exclusion dans le public de la vulgarisation scientifique. Elle constate que les inégalités sociales structurales se répercutent dans le public des communications scientifiques et affirme : "(...) science communication practices are shaped by structural inequalities and, as a result, are far from public." (Dawson, 2018, p. 773) Dawson a réalisé une étude ethnographique en Angleterre avec des groupes de personnes issues de minorités (notamment des immigrants) qui n'ont pas l'habitude de fréquenter des événements ou des lieux dédiés à la vulgarisation scientifique. L'auteure a étudié différentes pratiques de vulgarisation scientifique (musées, émissions télévisuelles, entre autres), mais dans le cadre de la présente recherche-crédation l'attention sera portée aux résultats de l'étude concernant les nouveaux médias. Dawson affirme :

Newer, less traditional science communication practices were invisible and, as a result, inaccessible for participants. Participants had not heard of science talks in cafes or pubs, science festivals, citizen science practices, science storytelling or science discussion events. (Dawson, 2018, p. 779)

Ainsi, malgré les nouvelles pratiques de vulgarisation scientifique, ces initiatives restent majoritairement invisibles à certains groupes sociaux. Dawson a constaté qu'un

autre aspect important du désengagement des groupes interviewés est le fait que le type de connaissance mis en valeur dans les pratiques de vulgarisation scientifique est différent des connaissances pratiquées par le groupe, phénomène classé par l’auteure comme « impérialisme culturel » :

Evident in these accounts of cultural imperialism, from a Bourdieusian perspective, is not that participants lacked cultural capital, but that their cultural capital — the stories, practices and knowledge they valued — was not reflected in the science communication landscape as they saw it. (Dawson, 2018, p. 781)

Or, il importe de constater que si la vulgarisation scientifique se veut d’être inclusive et ambitionne construire des points d’intersection entre tous les groupes sociaux et la science, elle doit aborder plus amplement les discussions et pratiques décoloniales (Boisselle, 2016; De Castro, 1998; Lander, 2005; Quijano, 2007; Santos & Laville, 2016).

Un autre facteur d’exclusion rencontré par Dawson est ce qu’elle a nommé, « *powerlessness* », ou le sentiment de ne pas être écouté présent chez les personnes issues des minorités. Comme a formulé une de ces interviewées, « For Fatimata, science communication was pointless, since her exclusion was predetermined and embedded in the structural inequalities that shaped whose voices were heard and whose were not. » (Dawson, 2018, p. 782) Il est possible alors que les dispositifs interactifs et participatifs puissent contribuer à donner la parole à des groupes qui éprouvent une sensation d’impuissance aussi dans les contextes de vulgarisation scientifique. Il s’agit ainsi de deux résultats importants à être explorés dans la pratique développée dans cette thèse : penser à la diversité de connaissances mises en valeur dans le récit ainsi qu’au développement d’un design d’interactivité qui puisse laisser aussi la place aux interacteurs pour qu’ils puissent s’exprimer et formuler leurs propres réflexions à propos des sujets abordés.

Les résultats obtenus par Dominique Brossard et Bruce V. Lewenstein (2009) peuvent indiquer aussi des aspects importants pour la construction de relations de confiance entre le public et les sciences. Les auteurs partent de la prémisse que les citoyens ont le droit d'avoir accès à la connaissance scientifique nécessaire pour une participation avisée dans les débats concernant les politiques publiques portant sur les sciences et les technologies. À partir d'une analyse basée sur quatre modèles de vulgarisation scientifique (figure 22), ainsi que l'étude de quatre projets de vulgarisation scientifique liés au projet de recherche « Human Genome »⁴⁹, les auteurs démontrent que les pratiques de vulgarisation scientifique doivent partir d'un travail préalable d'engagement avec le public : « (...) research should not rely on a priori characterizations of science audiences and should carefully assess all contexts of inquiry. » (Brossard & Lewenstein, 2009, p. 32) Ils défendent également l'utilisation de mesures de connaissance scientifique qui intègrent la relevance et le contexte (spatial et temporel), ainsi que la mise en relation des connaissances scientifiques avec des objectifs particuliers clairement définis.

⁴⁹ « Human Genome Project » est un projet de recherche coordonné par *US Department of Energy* et *National Institutes of Health*, réalisé entre 1990 et 2003.

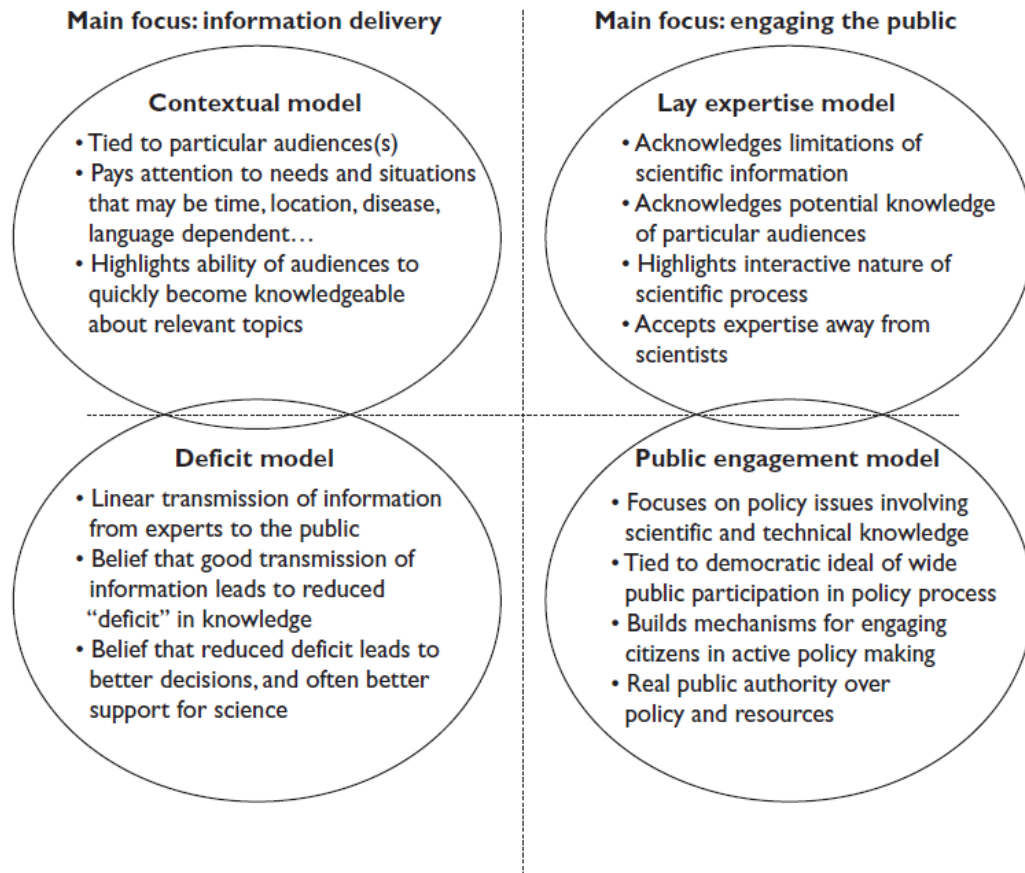


Figure 22 : Modèles de vulgarisation scientifique (Brossard & Lewenstein, 2009, p. 17)

Le « modèle du déficit » est critiqué par les auteurs, car il suppose que la vulgarisation scientifique a le rôle de « couvrir un déficit de connaissance scientifique » chez le public. Dans le sens opposé, le modèle « lay expert » propose que les pratiques de vulgarisation scientifique reconnaissent les connaissances et expertises déjà existantes chez le public. Ils constatent que « activities designed to enhance trust among participants in a policy dispute are more important than specific educational or informational approaches. » (Brossard & Lewenstein, 2009, p. 15) Des résultats analogues ont été obtenus par Michael M. Burgess (2014) dans une étude de différents projets de vulgarisation scientifique en rapport avec de politiques publiques. Burgess

affirme également la nécessité de s'affirmer que la participation du public est effective et non seulement utilisée pour légitimer des politiques publiques déjà conçues au préalable. Il affirme :

Effective and informed participation of diverse publics in policies that shape their future is thereby enhanced. Together with the activities of public-interest stakeholder groups, experts in technical fields and appropriate policy making institutions, deliberative engagement of mini-publics enriches the role of civic society in science and science policy, as well as that of direct participation in democratic society. (Burgess, 2014, pp. 50-51)

Ainsi, on constate que la confiance dans la science passe aussi par une construction de relation de confiance dans les activités de vulgarisation scientifique.

6. CHAPITRE VI

RÉCIT DE PRATIQUE

Ce récit de pratique est largement basé sur les notes prises dans un journal au long de mes études doctorales (2012-2020), pendant les périodes de développement et production de la partie de création de cette thèse. Toutes les citations du journal sont en guillemets et sans référence bibliographique.

6.1. Premières années du doctorat

En 2011, j'ai conclu un Master en art contemporain et nouveaux médias et dans le mémoire que j'ai soutenu, il était question d'étudier le hasard dans le processus de création du documentaire. C'est à ce moment que j'ai écrit l'avant-projet présenté à l'admission du doctorat, dans lequel j'ai décrit ainsi mes intentions de création : « Dans cette thèse-crédation il s'agira de réaliser un documentaire transmédia autour de la question du hasard ». Il était déjà question dans l'avant-projet de la réalisation d'un « film documentaire interactif sur le web » (ce n'est que plus tard que j'ai adopté les termes « documentaire interactif » et « webdocumentaire ») et de la création d'une « banque d'informations sur le hasard », avec des « textes scientifiques en rapport au hasard, commentaires sur des extraits de films, romans, œuvres d'art visuel, etc. qui abordent le hasard ». Ce sont principalement ces éléments qui ont été développés au long du doctorat pour aboutir à la création proposée dans cette thèse.

Au début de l'année 2012, les intentions de création ont évolué vers : « un espace numérique interactif d'informations, références et expériences du hasard. À ce moment-là, le projet était de mettre en place un dispositif qui, conçu comme un documentaire, prendrait la forme d'un labyrinthe numérique en 3D dans lequel le spectateur pourrait se promener, visiter les différents contenus mis à disposition,

ajouter du contenu ou participer aux jeux proposés. Le participant pourrait « se connecter » au documentaire et être visible par les autres participants pendant sa visite. Le visuel pour le labyrinthe était inspiré du travail de la peintre Vieira da Silva.

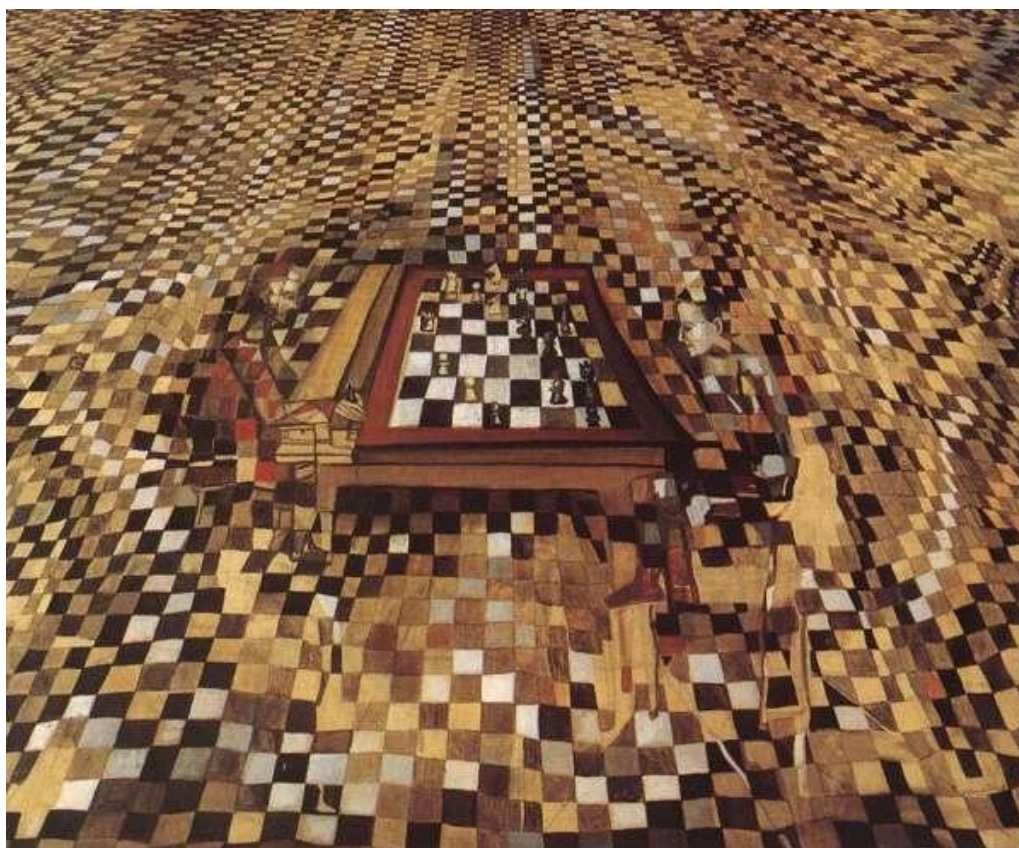


Figure 23 : A Partida de Xadrez, Maria Helena Vieira da Silva, 1943

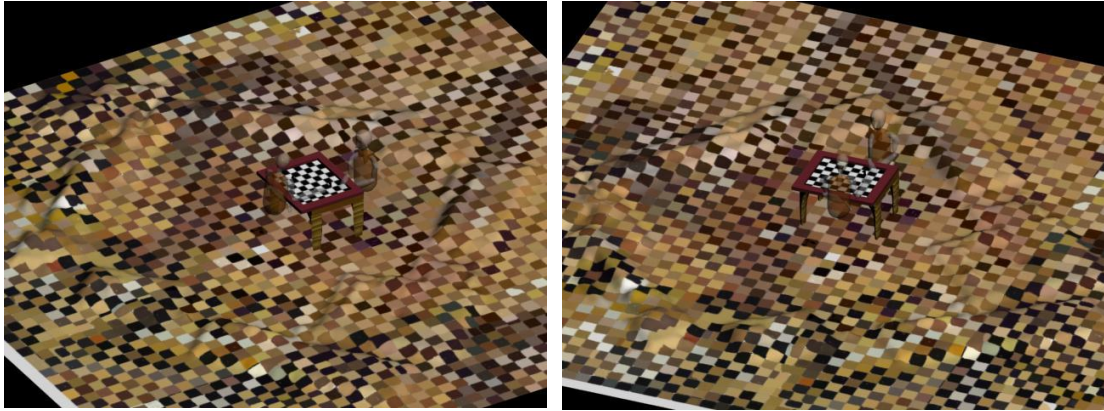


Figure 24 : Premiers essais de texture 3D pour l'esthétique du labyrinthe

Une partie du contenu serait une production originale, notamment une série de vidéos, des jeux qui utilisent les médias localisés, et textes. Des liens vers des travaux d'autres auteurs seraient aussi mis en ligne. Le deuxième type de contenu serait créé par les utilisateurs, par la forme de posts. Le son serait construit par une banque de soundwalks, à laquelle le public sera invité à contribuer. Les soundwalks seraient diffusées tout au long de la présentation du documentaire (random play), et le spectateur pourrait choisir sa liste de lecture par des critères comme lieu, date, auteur.

Pendant cette période le projet a pris une tournure plus immersive, mais ce n'est que trois ans plus tard, en 2015, que la réalité virtuelle, le support médiatique qui aurait pu faciliter la réalisation du projet décrit plus haut, est devenue abordable. Ainsi, l'idée centrale du projet de créer un « espace interactif d'informations, références et expériences du hasard » est restée, mais le format s'est transformé. Plutôt qu'une expérience immersive d'un espace numérique, j'ai orienté le projet vers le webdocumentaire.

Lors d'un cours de méthodologie en recherche-crédation que j'ai suivi à l'automne 2012, j'ai réalisé l'expérimentation de création « Projet Open Path ». Je suis partie d'un passage du roman *Marelle*, de Julio Cortázar :

(...) je ne vais pas t'expliquer ce qu'on appelle un mouvement brownien, et pourtant, nous deux, Sibylle, nous composons une figure, toi un point ici, moi, un autre là, nous déplaçant (...), peu à peu Sibylle nous composons une figure absurde, nous traçons avec nos va-et-vient une figure semblable à celle qui dessinent les mouches volant dans une pièce, de-ci, de-là, puis brusque volte-face et à nouveau de-ci, de-là, c'est ce qu'on appelle un mouvement brownien, tu comprends maintenant, un angle droit, une ligne qui monte par ici, par-là, d'arrière en avant, vers le haut, vers le bas (...), et tout cela tisse un ensemble, une figure, une chose aussi inexistante que toi et moi, que ces deux points perdus dans Paris qui vont de-ci, de-là, traçant leurs lignes, dansant pour personne, et même pas pour eux, une interminable figure sans queue ni tête. (Chapitre 34, p. 210)

J'ai associé ce passage avec des recherches en biologie, notamment le travail de Guy Théraulaz autour des insectes et de l'auto-organisation : "L'auto-organisation repose en outre sur l'amplification de fluctuations (les stimuli rencontrés au cours des marches aléatoires, les changements aléatoires de tâches, etc.). Cet aléatoire est lui-même un élément crucial qui permet la découverte de nouvelles solutions." (Théraulaz, 1998)

L'idée était de proposer un jeu avec le hasard et l'intégrer dans l'expérience du documentaire interactif, ou l'espace virtuel. Le projet utilisait l'application Open Paths, « a secure data locker for personal location information. Using our mobile apps, you can track your location, visualize where you've been, and upload your data to the OpenPaths website. » (Open Paths Website, consulté le 9 août 2017). Le but du jeu était de composer une cartographie collective avec les parcours de chacun sur Open Paths collectif. Les participants pourraient s'enregistrer et seraient alors visibles les uns aux autres en tant qu'un point sur une carte. Il n'y aurait aucune information affichée sur le participant. Les participants pourraient soumettre des photos et des sons, qui seraient affichés sur la carte. La démarche de l'expérimentation consistait à demander à une personne d'utiliser Open Paths, prendre des photos et enregistrer des sons pendant 1 semaine.

J'ai remarqué lors de cet atelier que j'avais quelques défis d'ordre technique sur le plan de la création que la thèse, car je n'avais pas encore d'aisance avec tous les outils

nécessaires à la création d'interactivité. Je me suis rendu compte à cette époque que j'allais devoir participer à des formations complémentaires sur les outils informatiques, surtout pour la mise en place des dispositifs interactifs. Pendant l'année suivante (2013), j'ai entrepris plusieurs formations, notamment dans les logiciels Max/MSP, Dreamweaver et Korsakow. Néanmoins, c'est dans le logiciel Klynt, que j'ai commencé à utiliser aussi en 2013, que j'ai trouvé les outils dont j'avais besoin pour la création de webdocumentaire.

6.2. Stage de création

En 2014, j'ai réalisé un stage de création avec la supervision de mon directeur de thèse, Louis-Claude Paquin. Paquin a écrit sur les méthodologies en recherche-crédation et a défendu la réalisation d'un atelier de création. Selon Paquin : « La finalité de cet atelier est la mise à l'épreuve du projet de création, tel que formulé dans un document écrit suite à une réflexion. L'épreuve dont il est question ici est celle du concret et de la matérialité du faire qui permettra de valider les formes et les matériaux, les procédés et les traitements et autres aspects constitutifs importants de l'œuvre. » (Paquin, 2014, p. 2.5.1 A VOIR). Le stage a été l'occasion de faire des preuves de concept, tester certaines idées ainsi que les différents logiciels et langages de création.

La question d'amorce pour ce stage en atelier était : « où trouver le hasard ? » J'y ai proposé les réponses suivantes d'après mes lectures et recherches :

- Les fourmis (marches aléatoires, *random walk*, auto-organisation). Références : Deborah Gordon, Guy Theraulaz.
- Économie : The Black Swan (Nassim N. Taleb)
- Vie quotidienne : The Drunkard Walk, Daniel Kahneman
- Jeux de hasard : Loterie, dés, pièce
- Mathématiques : Probabilités, Simulations Monte Carlo,

– Informatique : Théorie de l'information (entropie) ; théorie algorithmique de l'information (Chaitin) ; Complexité (algorithmes probabilistes, réseaux bayésiens)

– Processus de création artistique : François Morellet, Dada, John Cage, Boltanski, Pollock, Nam June Paik. Cinéma documentaire : Nicolas Philibert, Eduardo Coutinho,

À la suite de ces notes, j'ai proposé quatre idées de création :

1. Fourmis. Deux vidéos côte à côte : une prise de vue d'une fourmi qui se déplace sur un plan (carré blanc, en contre-plongée) et une animation d'une simulation de marche aléatoire (*random walk*).

2. Pièce. Installer le dispositif suivant dans l'espace public (une journée) : une table avec une pièce (monnaie), une feuille de papier, un crayon et des instructions (jeter la pièce et noter F pour face ou P pour P. Reprendre la séquence obtenue et faire des tests statistiques pour confirmer s'il s'agit d'une série aléatoire.

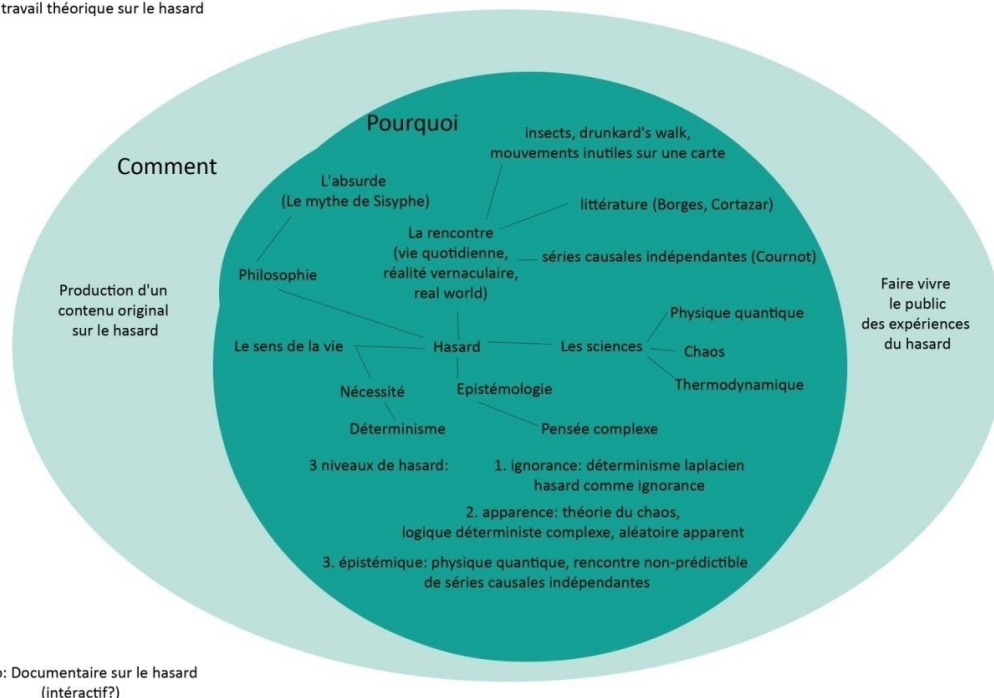
3. Trajets croisés. Passer une annonce [journal ou Kijiji] pour recruter des participants à l'activité suivante : pendant une journée, les participants auront leurs trajets enregistrés sur une carte [de forme anonyme]. Ils sont invités à réaliser et soumettre des enregistrements audio et photographiques. L'ensemble de ces éléments [trajets, images et audio] sera affiché sur une carte. — Éventuellement une simulation de *random walk* sera aussi ajoutée sur la carte.

4. Documentaire : Tirer au sort des coordonnées et des horaires, aller filmer et en faire un documentaire. — source de l'aléatoire : insecte ou loterie.

Quoi

Jeux: eg. comparer sur une carte
les trajets d'une fourmi et d'une personne

Texte: travail théorique sur le hasard



Vidéo: Documentaire sur le hasard
(interactif?)

Figure 25 : Illustration proposée au début du stage doctoral

J'ai fait le choix de développer la quatrième proposition, que j'ai ensuite nommée « Loto-Documentaire ». Le titre fait référence à l'utilisation de la loterie dans le processus de création du documentaire, et fait également un jeu de mots avec la notion de subjectivité dans le documentaire, une notion souvent abordée comme éloignée des processus artistiques qui utilisent le hasard. J'ai commencé alors le développement du projet par la formulation de la démarche à être mise en place dans la réalisation.

“Description de la démarche : À partir des tirages de Lotto-Max, de Loto-Québec, j'utiliserai les deux premiers chiffres du résultat pour obtenir des points de coordonnées sur une carte, le troisième chiffre du résultat détermine un jour de la semaine, et les

quatrième et cinquième chiffres établissent un horaire. Je tourne ensuite un clip vidéo documentaire à l’endroit et à l’horaire obtenus avec le résultat de la loterie.

Prenons l’exemple du résultat du tirage du 17 avril 2015 de Lotto-Max (dans l’ordre de tirage) : 41 - 19 - 15 - 14 - 11 - 30 - 7. Les deux premiers chiffres [41 et 19] sont retenus pour la détermination des coordonnées [41 détermine l’axe x de la carte ; 19 détermine l’axe y]. Le troisième chiffre [15] détermine le jour de la semaine, le quatrième [14] détermine l’heure du tournage et le cinquième [11] les minutes de l’horaire. Les autres chiffres du résultat ne sont pas considérés. Les vidéos réalisées seront ensuite mises en interactivité avec le logiciel Korsakow. Ce logiciel a été choisi, car il permet de réaliser des documentaires interactifs qui utilisent des aspects pseudo-aléatoires. L’ensemble sera enfin mis en ligne.”

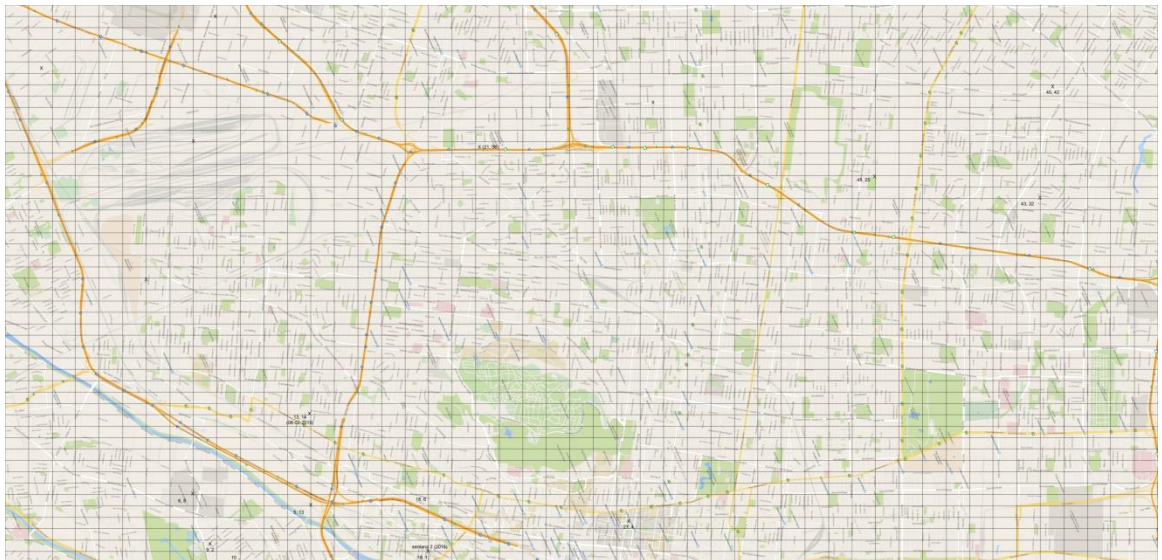


Figure 26 : Carte de Montréal avec une grille [49x49], utilisée pour le repérage des lieux de tournage

En février 2015, j’ai commencé à tourner des images dans le cadre du stage de création. Je notais : “Préparation pour le début du tournage du projet Loto-Doc : Le fait d’utiliser

le hasard (dans ce cas, la loterie) dans le processus créatif implique une plus grande difficulté dans la préparation d'un tournage, car plusieurs éléments demeurent inconnus, comme le fait de ne pas savoir si le tournage aura lieu le jour ou la nuit rend difficile le choix de la lumière pour le tournage. Est-il possible de tourner la nuit ? Est-ce qu'il faudrait restreindre les horaires de tournage à la journée ?”

Par ailleurs, le fait de ne pas maîtriser le contenu du tournage est un peu angoissant, « est-ce que le matériel sera assez bon pour en faire une vidéo ? Et d'un point de vue plus personnel, cela fait un moment que je travaille sur le plan théorique et très peu sur la création, en conséquence j'avais l'angoisse aussi de me remettre à la création. »

Lors des premiers tournages, je n'ai pas fait d'entrevues. Il s'agissait de filmer l'espace, et enregistrer le paysage sonore. Je faisais moi-même la direction photo et l'enregistrement audio. J'ai rapidement compris que cela n'était pas un format adéquat au projet. J'ai commencé à travailler avec Felipe Pio, un directeur photo avec qui je collabore souvent. Nous étions alors deux sur les lieux de tournage, je m'occupais du son et Felipe de l'image. Avec plus de disponibilité pour l'enregistrement sonore, j'ai commencé à faire quelques entrevues avec les personnes qui passaient dans nos lieux de tournage [style micro-trottoir]. Je posais des questions sur leurs expériences et impressions du hasard, la loterie et des sujets connexes. J'ai beaucoup aimé les résultats et les entrevues sont devenues une partie importante de la démarche.

6.3. Forum doctoral

Dans l'année scolaire 2014-2015, j'ai participé au Forum doctoral, un séminaire pour, entre autres, appuyer les doctorants dans l'élaboration du projet de thèse. Le forum a été pour moi l'occasion de mieux cadrer et définir certains choix pour le projet de création et de les débattre avec les professeurs et collègues. Une des méthodes proposées lors du Forum doctoral se basait dans la réalisation de cartes conceptuelles de nos projets de thèse. Les différentes cartes que j'ai produites pour la présentation des domaines et concepts clés dans le cadre du Forum doctoral m'ont permis de

comprendre des aspects variés de ma recherche-cr ation. Apr es avoir travaill  avec les cartes, j'ai d limit  les concepts    tre abord s par la recherche-cr ation selon 4 axes :

1. Al atoire ; ind termination ; incertitude : Groupe de concepts connexes   celui du hasard, il est n cessaire de clarifier les rapports entre eux [sont-ils diff rents ? Si oui, en quoi ?].

2. Physique ; philosophie ; biologie ; math matiques ; arts : champs de la pens e qui seront convoqu s pour discuter le hasard. Il s'agit de regarder, en physique par exemple, quelle a  t  la contribution de la physique quantique dans la construction du concept de hasard?

3. Agencement ; image ; interactivit  ; num rique : je me permets de revenir plus en d tail sur ce groupe de concepts, car il a  t  moins discut  lors de la pr sentation. Le concept d'agencement [assemblage, en anglais] dont il sera question dans cette recherche est celui d velopp  par Gilles Deleuze et F lix Guattari et repris par Manuel DeLanda. J'ai l'intention de revenir sur l'*assemblage theory* et l'utiliser dans l' tude des m dias interactifs [ou plut t du documentaire interactif], en somme,  tudier le dispositif interactif comme un agencement.

Le concept d'image sera abord  par les travaux de Henri Bergson et Gilles Deleuze, et il sera question ici d'investiguer les particularit s de l'image interactive. Le concept d'interactivit , en revanche, nous permettra de revenir sur les particularit s de l'interaction dans le champ du documentaire.

Enfin, la nature num rique du support sera  galement abord e, notamment   travers le travail de Luciana Parisi et Tiziana Terranova sur les algorithmes et l'information num rique.

4. Rencontre ; jeu ; cadrage de l'inopin  : Ce groupe de concepts, au centre de la carte, fait r f rence   la mise en forme de la rencontre [notion qui est au c ur d'une des d finitions du hasard] entre le hasard e le documentaire interactif. Le cadrage de

l'inopiné, processus de création développé par l'auteur et photographe/documentariste, Jean-Luc Lioult, sera utilisé comme la démarche principale du travail de recherche-crédation. Largement, cette démarche implique que le créateur pose un cadre, ou établisse un dispositif, à l'intérieur duquel les événements inopinés se produisent et sont souvent enregistrés. Le jeu sera utilisé notamment dans certaines pratiques de création avec le hasard.

Ce travail conceptuel a informé le développement de la création. Les deux images suivantes illustrent la première proposition de structure thématique du contenu [avec les références respectives] et les possibilités de projets à être réalisés dans le cadre de la thèse :

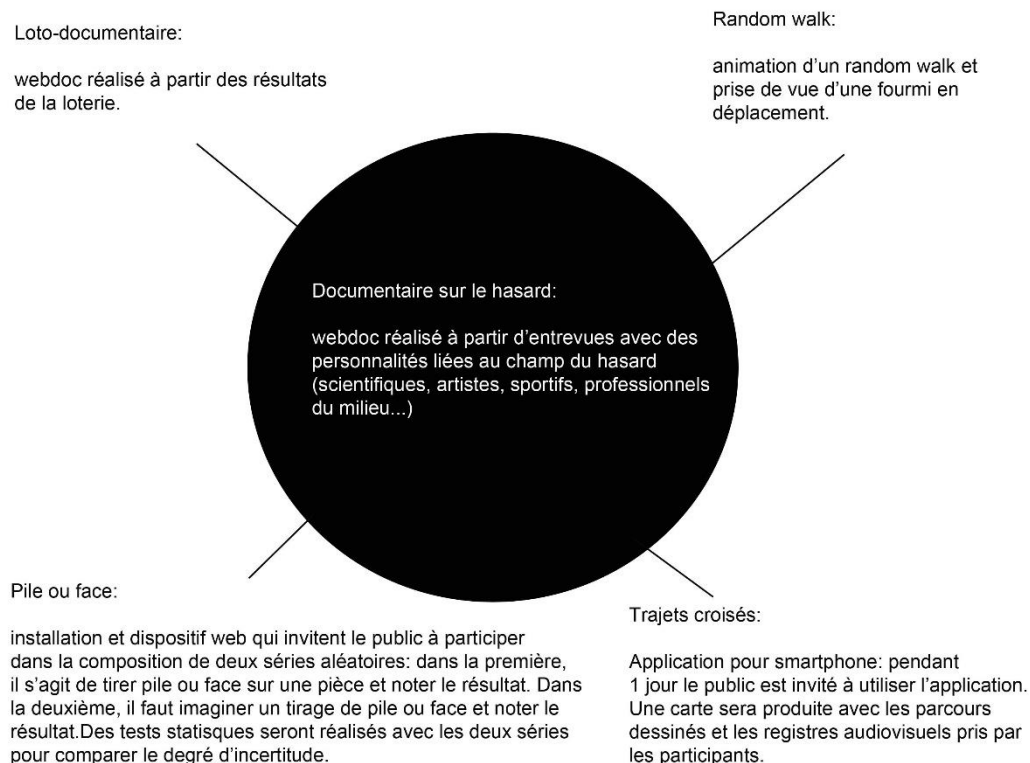


Figure 27 : Proposition de structure thématique

À la suite du Forum doctoral, j'ai réussi l'examen du projet de thèse et la période de rédaction, réalisation et développement de la thèse a débuté.

6.4. Réalisation

6.4.1. Chatbot

En 2017, j'ai fait la lecture d'un article à propos de la création d'un chatbot (Vlahos, 2017) et cet événement a été un point d'inflexion dans le processus du design d'interactivité. Jusqu'à ce moment du processus de création, le principal mode d'interactivité envisagé pour la création était le mode hypertextuel (ou de sélection). Lorsque j'ai commencé à envisager les possibilités créatives d'un chatbot, je me suis tournée vers le mode conversationnel d'interactivité. Dans l'article « A Son's Race to Give His Dying Father Artificial Immortality », publié en 2017 dans le magazine de technologie *Wired*, le journaliste et développeur James Vlahos décrit comment il a créé le *DadBot*. L'histoire du *Dadbot* commence quand Vlahos apprend que son père a un cancer en phase terminale. Il commence à enregistrer des entrevues avec son père. Après une période de réflexion, il arrive à l'idée d'utiliser le matériel des enregistrements pour construire un *chatbot* qui essaie de reproduire les réponses de son père, ou le *DadBot*.

Cet article m'a fait réfléchir sur les possibilités de construire un chatbot à partir de textes publiés. Par exemple, serait-il possible de créer un *DeleuzeBot*, qui pourrait répondre à des questions du type « qu'est-ce que vous voulez dire par "corps sans organe" ? », une idée un peu comique et absurde. J'ai remarqué alors que le chatbot pouvait être un outil d'interactivité dans le documentaire permettant de transformer l'interacteur en interviewé(e). L'interface de conversation permet que l'interacteur réponde aux questions du créateur et d'orienter le contenu selon les sujets abordés dans

la conversation. Le principal aspect qui rend l'utilisation d'un *chatbot* intéressante dans le contexte de cette recherche-cr ation est le fait que le mode conversationnel d'interactivit  permet une exp rience plus impr visible lors que compar  au mode hypertextuel. Or, une des propositions centrales de cette th se de recherche-cr ation est d'explorer le potentiel cr atif du hasard et de l'ind termination pour une exp rience interactive ouverte et pas compl tement pr programm e.

J'ai commenc  quelques recherches sur la cr ation d'un chatbot et j'ai constat  que les d veloppements r cents dans les domaines de l'apprentissage automatique ont rendu ce type de technologie beaucoup plus accessible. Cela a  t  un aspect important pour l'option d'utiliser le chatbot comme un dispositif pour l'interactivit  dans le documentaire. Par ailleurs, la th matique de l'intelligence artificielle m'int ressait aussi du point de vue de la recherche th orique de cette th se. Comme il a  t  d montr  dans la partie pr c dente de la pr sente th se, les notions de probabilit , pr vision, hasard et incertitude sont au c ur de l'intelligence artificielle.

J'ai commenc  alors   faire des essais avec des logiciels pour la cr ation de chatbots, et le 27 juillet 2017, j' crivais dans mon journal : « J'ai r ussi   cr er mon premier chatbot. Je l'ai fait avec *PullString* et *WebChat* (Microsoft Bot Framework), et ensuite je l'ai int gr    un projet *Klynt* avec un *iframe*. Tout a bien fonctionn . Je peux maintenant commencer   d velopper le chatbot. Les prochaines  tapes du d veloppement de l'interactivit  sont le design de l'identit  visuelle du projet, la s lection du contenu vid o et la r alisation d'entrevues pour la pr paration du design des conversations du *chatbot*. Du point de vue du contenu vid o original, il est question du tournage. Pour la recherche th orique, il faut aussi plus d'information sur les types d'algorithmes utilis s dans l'intelligence artificielle de *PullString*. Pour l'identit  visuelle et le design de l'interface, il sera question d'un collaborateur. Je m'occuperai de cr er les r gles et de produire le contenu (texte et audiovisuel). Pour cela, il va falloir enregistrer des entrevues pour mieux comprendre les dialogues et les termes associ s

à la thématique du projet. La partie audiovisuelle du contenu sera réalisée à partir de tournages, et aussi des montages des vidéos d'internet. »

Le collaborateur en question est Luciano Frizzera, un collègue du programme de doctorat conjoint en communication de l'Université Concordia. Frizzera est à la fois designer, développeur et chercheur, et au-delà de son expérience dans le développement Front-End, il s'intéresse aux perspectives théoriques et critiques des technologies. Sa contribution au projet est inestimable, autant sur le plan technique que conceptuel. En été 2017, nous avons commencé à travailler sur le développement du chatbot sur la plateforme PullString. Mais le 18 septembre 2017, PullString a annoncé l'arrêt de ses activités dans le domaine du chatbot par message texte. Cet événement nous a obligés à recommencer le travail de développement sur une autre plateforme.

Nous avons alors commencé à travailler avec Recast.AI, une plateforme pour la création de chatbots créée par une entreprise française. Recast.AI avait l'avantage de proposer le *polybot*, un chatbot avec du traitement automatique de langage naturel en plusieurs langues. La question de la langue du chatbot était un point de réflexion. Le 8 août 2017, j'écrivais dans mon journal : « Plusieurs raisons me font pencher vers un *chatbot* en anglais. En premier, il y a la question de la technologie, car le Natural Language Processing (NLP) est beaucoup plus développé en anglais. Ensuite il y a aussi la question de l'accessibilité, car cette recherche-crédation est réalisée avec le soutien de CAPES, du ministère de l'Éducation brésilien, et se doit donc d'être accessible au public brésilien. Un *chatbot* en anglais est plus accessible au public brésilien qu'un en français. »

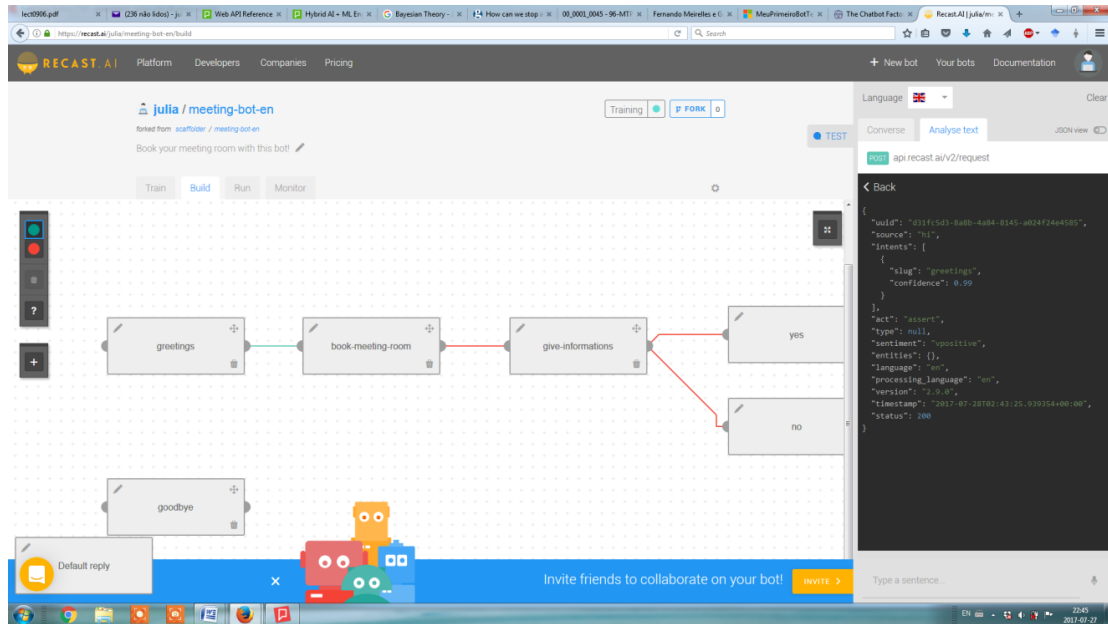


Figure 29 : Interface de Recast.AI

Le développement a évolué jusqu'à ce qu'en 2018, Recast.AI annonce que l'entreprise avait été achetée par SAP, une multinationale dans le domaine de logiciels. À cette occasion, moi et Frizzera discussions souvent des difficultés de travailler dans un domaine émergent et encore très instable. Les changements constants dans le scénario des plateformes pour le développement de chatbot nous obligeaient à recommencer le travail constamment. Il s'agit d'un défi présent dans tous les médias numériques où les technologies et les langages évoluent constamment, mais dans un domaine encore émergent comme celui des chatbots, le défi semble être présent de manière encore plus intense. En considérant notre expérience avec PullString, nous avons évalué que l'acquisition de Recast.AI pourrait entraîner un changement de direction de la plateforme et le risque de devoir recommencer tout le développement était trop important.

Nous avons passé alors quelques mois à évaluer les différentes plateformes et en avril 2019 nous avons migré vers DialogFlow, l'outil développé par Google pour la création de chatbots.

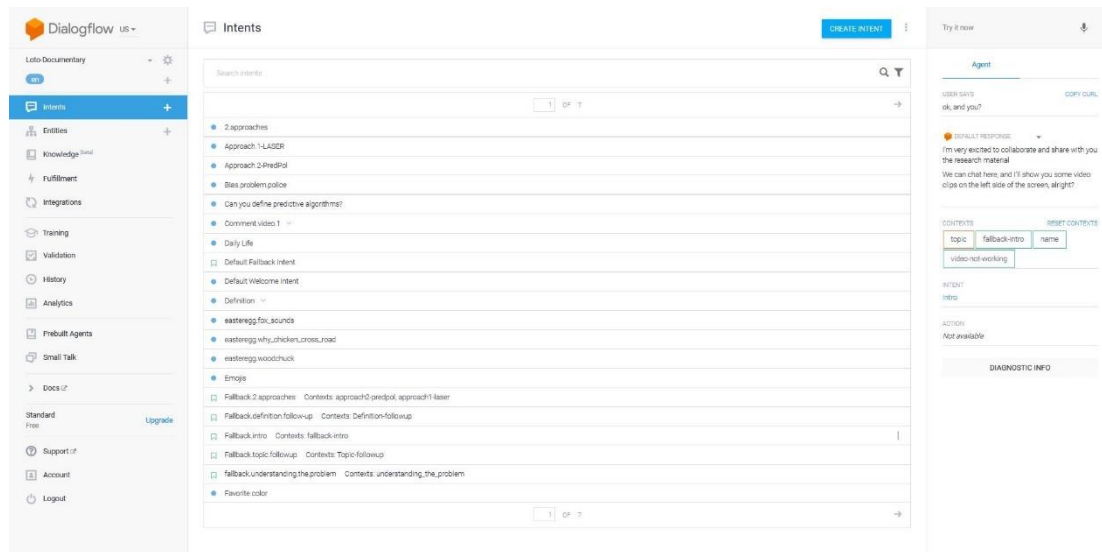


Figure 30 : Interface de Dialogflow

La version actuelle de la portion de création de cette thèse, *On Chance*, a été réalisé sur la plateforme DialogFlow. Frizzera a développé ensuite l'intégration entre DialogFlow, la collection de vidéos et l'interface web (développée aussi par Frizzera).

6.4.2. Design de conversation

Une fois que la plateforme de développement a été définie, la première phase de développement du chatbot dont je me suis chargée a été le design de la conversation. Le design de conversation est l'équivalent, dans le contexte des chatbots, du design d'interactivité. C'est la conversation, et non l'hyperlien, qui établit la relation entre l'interacteur et le contenu du documentaire. Ainsi, le travail de scénarisation et de création de contenu textuel prend une autre dimension dans le processus de création. Ce défi est accentué par le fait que les plateformes de chatbot (c'est le cas aussi de

DialogFlow) sont créées en général pour une utilisation commerciale. L'utilisation de ces plateformes pour la création d'un chatbot documentaire est sans doute minoritaire parmi les usages prévus. En conséquence, d'autres défis se présentent pendant le processus de réalisation, car l'outil n'a pas été conçu pour répondre aux besoins du développement d'un récit non linéaire. Cela m'a obligée à composer avec un ensemble d'outils externes à Dialogflow pour le design de conversation. Je reviendrai au long de cette section du texte sur les différents outils utilisés, car cela permettra également de comprendre le processus de développement et d'intégration.

Le champ du design de conversation a beaucoup évolué récemment, principalement dans le contexte des chatbots. Un certain nombre de principes et de pratiques établies ont servi à orienter le processus de création de *On Chance*. La première étape du design de conversation est la création d'un « échantillon de dialogue » (*sample dialogue*, en anglais). Dans ce premier échantillon, il s'agit simplement de rédiger un exemple d'une conversation envisagée pour le *chatbot*. En 2020, j'ai rédigé l'échantillon et je l'ai soumis à la collaboratrice du projet, Myriam Rafla. Rafla est une collègue du programme conjoint de doctorat en communication de l'Université Concordia. Rafla est spécialisé dans la recherche-crédation, le récit non linéaire et le documentaire interactif. Elle est consultante pour le développement de scénario et professeure dans le département de cinéma du Dawson College. La contribution de Rafla au développement du récit et du design de conversation a été fondamentale pour le développement du projet. Après la première session de travail avec Rafla, j'ai travaillé sur le fil narratif du documentaire ainsi que la biographie du chatbot. Voici le logline et la biographie du personnage du chatbot :

« Logline: Congratulations on your new job! You were just hired as a journalist and your first assignment is to collaborate with your colleague Jana on a story about predictive algorithms. Help Jana writing this story with the research material she'll show you.

Who is Jana: Journalist specialized in science communication. She is about 45 years old, Lebanese Canadian. Interested in science, technology, philosophy, movies, culture. Informative. »

Une des principales questions à affronter lors du design d'interactivité dans le mode conversationnel est le degré de contrôle sur le flux et la structure de la conversation. Dans une conversation très ouverte, sans un flux conversationnel prédéterminé, les réponses données par le chatbot sont imprévisibles et les chances que la conversation ne fasse pas de sens pour l'interacteur sont plus grandes. À l'inverse, dans une conversation prédéterminée, les réponses de la part du chatbot sont prévisibles, mais l'apport de l'interacteur. Ainsi, il est possible de concevoir un design de conversation plus vertical, où la structure de la conversation est proche de l'interactivité par sélection (ou hypertextuelle). Dans ce type de conversation, l'interacteur choisit parmi un nombre limité d'options sur l'avancement des répliques dans le dialogue. Par ailleurs, nous pouvons imaginer un design de conversation plus ouvert, ou plus horizontal, dans lequel le flux du dialogue est moins contrôlé. Le défi dans ce mode de conversation est d'arriver à maintenir une conversation sans trop d'erreurs dans le traitement automatique du langage naturel. Dans le cadre de cette thèse, plusieurs modes de design de conversation ont été explorés. La première proposition consistait d'une structure assez verticale, très proche d'un mode d'interactivité hypertextuel. Après des conversations avec les collaborateurs et le directeur de thèse, d'autres options de design de conversation ont été élaborées, notamment par types d'approche thématique selon les intérêts démontrés par l'interacteur (des thématiques plutôt philosophiques, scientifiques ou technologiques, par exemple).

Pour la création du flux de conversation, deux outils ont été composés, un logiciel de Mind Map ainsi qu'un scénario par tableau :

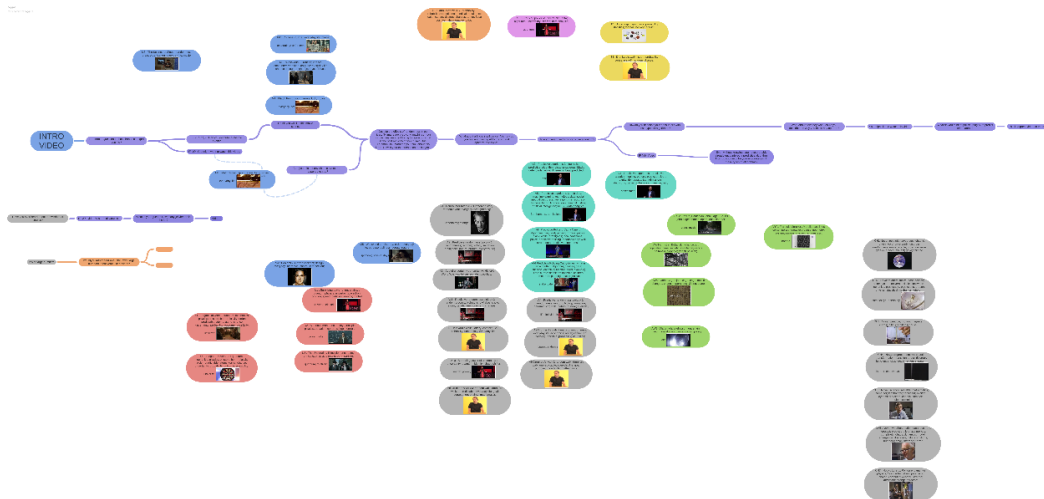


Figure 31 : Mindmap du scénario

PHASE 1 : CONTEXT	
PREDICTIVE NARRATIVE	
INTENT/MEDIA	TEXT
Welcome	<ul style="list-style-type: none"> - Hi, you are the new writer, welcome! - I'm Jana, what's your name?
Get name \$name	<ul style="list-style-type: none"> - Welcome \$name! I'm very happy you joined our team! - How are you?
Intro	<ul style="list-style-type: none"> - I'm very excited to collaborate and share with you the research material - We can chat here, and I'll show you some video clips on the left side of the screen, alright?
Topic @sys.any	<ul style="list-style-type: none"> - I've been covering technology for about 20 years, this topic is one of the most important I've written about - it affects how we build political views, how long people stay in prison, if people will get loans, and so many important things in our lives - The story is about predictive algorithms, have you ever heard about it?
	<ul style="list-style-type: none"> - I never realized that so many important things in our lives are being decided by algorithms, - In this video, there's an important call from the computer scientist makes an important call,
Topic Follow-up Yes	<ul style="list-style-type: none"> - Great! I was very impressed when I realized that predictive algorithms are all over, - from recommendations on Youtube and Netflix to the police and the criminal justice, - But can we really trust an algorithm to make predictions for the police? Are those algorithms fair?
Topic Follow-up No	<ul style="list-style-type: none"> - No problem! I was very impressed when I realized that predictive algorithms are all over, - from recommendations on Youtube and Netflix to the police and the criminal justice,

Figure 32 : Scénario

La collection de vidéos est composée par un ensemble de courts extraits de différents genres et formats (scènes de films, reportages, entrevues, conférences). J'ai attribué des mots-clés à chaque extrait vidéo. Frizzera a ensuite créé une base de données sur MongoDB avec les informations relatives à la collection vidéo.

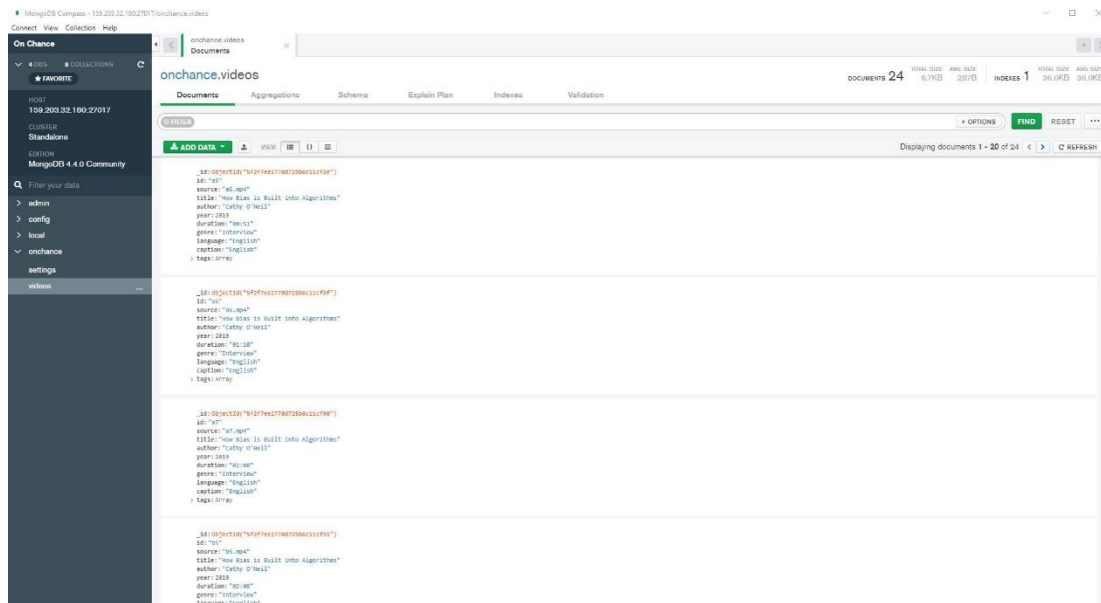


Figure 33 : Interface de base de données

Ces extraits ont été intégrés alors au design de la conversation. La conversation et les extraits vidéos sont affichés sur une interface web, créée par Frizzera. Le résultat est disponible sur la page : <https://onchance.net> Voici les images de l'écran d'accueil et de l'interface principale :

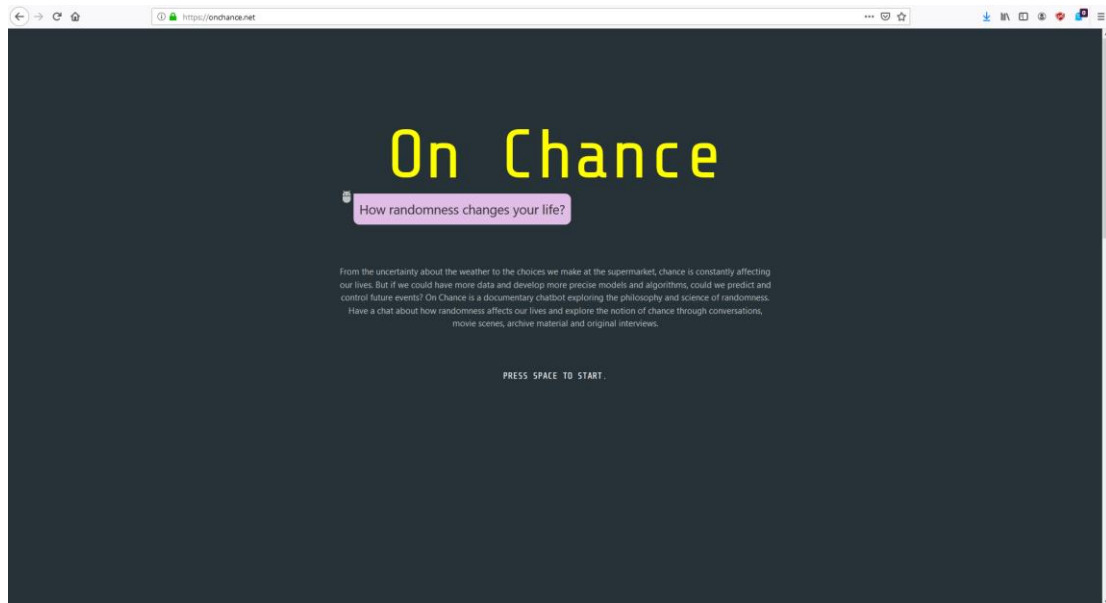


Figure 34 : Page d'accueil de On Chance



Figure 35 : Interface principale de On Chance

Après le design de conversation et l'intégration des médias, l'étape suivante était l'entraînement du *chatbot*. Pour la première batterie d'entraînement, j'ai visé un public peu nombreux de (5) personnes peu ou pas du tout familiarisées avec le projet et avec

la technologie des chatbots. Dans une perspective de vulgarisation scientifique, il était important que l'expérience soit accessible à un public de personnes peu familières avec la thématique abordée, soit les algorithmes prédictifs, d'où le choix du public pour cette première batterie d'entraînements. Les commentaires reçus de la part des interacteurs à ce stade du développement ont été importants dans la prise de certaines décisions sur le récit, l'interface et le design de conversation. Le retour obtenu pendant l'étape de l'entraînement sera discuté plus en détail dans la prochaine section, mais de façon générale, les interacteurs ont confirmé le potentiel narratif du dispositif chatbot audiovisuel. Pour la deuxième batterie d'entraînement, j'ai produit un document avec des indications dirigées aux interacteurs. À ce stade, j'ai visé un groupe un plus large de personnes (15). À chaque nouvelle batterie d'entraînements, la performance du chatbot s'améliore considérablement, ce qui rend l'expérience de l'interacteur plus intéressante. Il est intéressant de noter que le processus d'entraînement ne s'arrête jamais dans le contexte des chatbots, et chaque nouvelle expérience d'un interacteur est aussi une instance d'entraînement de l'agent conversationnel. Cet aspect processuel du dispositif interactif de type chatbot sera abordé dans la prochaine section.

6.5. Discussion

Le processus de création de *On Chance* a été l'occasion d'expérimenter avec les dimensions pratiques du problème du hasard dans les dispositifs numériques, notamment le documentaire interactif. Dans cette section, je ferai une courte introduction des relations entre les dimensions pratiques et théoriques de cette recherche-création sur deux perspectives, le code informatique et l'expérience de l'interacteur de *On Chance*. Les résultats seront ensuite mis en évidence à la conclusion du présent essai.

La première articulation directe entre les approches théoriques et pratiques du problème du hasard dans les algorithmes est le dépassement du dualisme entre déterminisme et

probabilisme vers une multiplicité des synthèses. Dans la documentation de *PullString*, le logiciel utilisé dans ma première expérience de création de chatbots, par exemple, les auteurs décrivent leur approche hybride :

PullString Author uses a hybrid rule-based / machine-learning approach to computer conversation. Here's a little background on the two approaches:

Rule-Based: These systems use pattern matching to recognize specific things a user says and then perform the actions associated with those rules. This approach gives you very precise control over how the bot responds, but it also requires that you learn a pattern matching syntax and write rules to cover all possible outcomes.

Machine Learning: These systems are trained with large quantities of sample data so that they can learn to automatically associate certain inputs with certain outputs. They generally don't require a lot of human intervention. However, it is often unclear why the system chooses a particular action. It can also be difficult to correct when the AI does something wrong (beyond giving it more data). (Documentation site web PullString, consulté le 27 juillet 2017).

Nous comprenons alors que le logiciel propose un modèle hybride, à la fois déterministe (Rule-Based) et probabiliste (Machine Learning). Il est intéressant de remarquer qu'il y a incertitude par rapport au comportement de l'algorithme (« *unclear why the system chooses a particular action* »), et qu'il peut produire des erreurs (« *AI does something wrong* »), caractéristiques des algorithmes probabilistes.

Un extrait du code utilisé dans la plateforme Recast.AI met en évidence l'aspect probabiliste de son système informatique :

```
«source» : «hi»,
«intents» :
    «slug» : «greetings»,
    «confidence» : 0.99
```


Le fait qu'un pourcentage indique la certitude du logiciel à affirmer que le mot « hi » appartienne à la catégorie « greetings » démontre qu'il s'agit d'un algorithme probabiliste, dans ce cas, qui utilise *machine learning*. Néanmoins, de nombreuses règles sont aussi créées et utilisées dans le bot, ce qui configure une approche hybride, comme pour *PullString*. DialogFlow opère également dans une logique hybride qui compose les algorithmes probabilistes d'apprentissage automatique avec des systèmes basés sur des règles déterministes. Nous constatons ainsi que dans la pratique des algorithmes il est question, tout comme dans le travail théorique réalisé dans la première partie de cette thèse, du dépassement du dualisme dans l'approche probabiliste.

La deuxième articulation directe entre les dimensions théoriques et pratiques abordées dans cette thèse est la perspective processuelle de la probabilité et de la théorie de l'information. Comme il a été décrit plus haut, le chatbot existe dans un processus d'évolution constante, principalement en fonction des entraînements. Chaque expérience du chatbot vécue par un interacteur est aussi une occasion d'entraînement qui fera le chatbot évoluer dans une tendance particulière. Or, dans la première partie de cette thèse, nous sommes arrivés justement à une perspective processuelle de la théorie de l'information et de la probabilité, et le chatbot confirme la pertinence de ce traitement aux dispositifs numériques.

Les concepts d'agencement et d'image-interaction (élaborés dans la première partie du présent essai doctoral) permettent d'inclure la dimension processuelle dans l'expérience du documentaire interactif. Dans la conclusion je reviendrai sur l'aspect processuel des résultats obtenus par cette thèse, mais avant cela il est important de discuter les résultats des phases d'entraînement en qui concerne l'expérience de l'interacteur. L'aspect processuel du chatbot présente des défis également pour la documentation de l'expérience de l'interacteur, considérant que l'expérience du chatbot est multiple (avec des variations aléatoires) et qu'elle change dans le temps (selon l'évolution du traitement automatique du langage naturel du chatbot). Ainsi, la

solution proposée ici pour la documentation est de suivre l'aspect évolutif du chatbot et documenter à multiple reprises, différentes expériences d'interacteurs. Dans les annexes de cet essai doctoral (Annexe A), j'indique le lien pour la documentation vidéo de différentes expériences de *On Chance*.

Les phases d'entraînements du chatbot ont été l'occasion de recevoir des commentaires de la part des interacteurs sur leur expérience. Dans la perspective de l'engagement de l'interacteur, les résultats obtenus indiquent que le format combinant l'interactivité conversationnelle avec les éléments audiovisuels a réussi à entretenir des conversations avec un nombre considérable d'interactions (selon les données de l'Analytics, les sessions d'entraînement avaient en moyenne une durée d'à peu près 15 minutes pour 40 interactions). Ces résultats confirment les observations de Prerna Gupta, qui en 2017 a décrit l'expérience suivante :

Then we had an off-the-wall idea to test a story written as a text message conversation between the characters. It was 1,000 words, or a five-minute read, the same length as everything else we had tested thus far. The first chat story we tested had staggering results. Almost every teenager who started reading our chat story finished it in one session. We thought it was an error. We had never run an a/b test before with this magnitude of a change. So we tested it again. And we got the same results. That was the light bulb moment. We launched Hooked four months later. (Gupta, 2017)

Ses observations démontrent que les jeunes lecteurs sont capables de lire plus longtemps (ou plus de mots) lorsque le texte est disposé en forme d'échange de messages texte. Ainsi, le format conversationnel est efficace pour engager l'interacteur dans la lecture du contenu proposé. Par ailleurs, en termes de vulgarisation de concepts scientifiques ou technologiques, il est intéressant de remarquer que les interacteurs ont exprimé souvent le fait d'avoir appris des choses sur un sujet qu'ils connaissaient peu. Cela se vérifie, par exemple, dans la qualité des certaines formulations de concepts technologiques qui ont été élaborées par les interacteurs après avoir consulté le contenu proposé. Ainsi, de manière générale, il est possible de constater que certains résultats

théoriques à propos des modes d'interactivité se confirment dans les expériences des interacteurs.

CONCLUSION

Dans la présente conclusion, il s'agira de revenir sur les apports des concepts élaborés dans la recherche théorique au processus de création du documentaire et les contributions de la présente recherche-crédation. Cette recherche doctorale a été initiée par la question de l'interprétation du rôle du hasard et de l'indétermination dans les médias interactifs, plus spécifiquement le documentaire interactif. Il était question de comprendre comment le hasard agit dans les processus analogiques et numériques mis en place dans un documentaire interactif. Or, si les documentaires interactifs sont des programmes numériques, des algorithmes écrits dans un code binaire pour effectuer des calculs selon des règles prédéterminées, quelles sont alors les conditions pour la création d'indétermination dans la dimension numérique du documentaire interactif ? Comment comprenons-nous les notions de hasard, indétermination, incertitude, entropie (et ses relations avec l'information et le bruit) dans le contexte des technologies numériques de communication utilisées dans les documentaires interactifs ? Comme nous l'avons démontré, ces questions sont importantes non seulement par ses aspects théoriques, mais également par les conséquences pratiques de l'utilisation de plus en plus fréquente de modèles prédictifs basés sur le traitement statistique de données numériques. Ce problème a été abordé par la perspective du pragmatisme spéculatif, ce qui a permis de reconnaître les aspects processuels et relationnels du hasard et de l'indétermination, et de proposer une compréhension non dualiste de l'expérience du documentaire interactif, ainsi que des notions d'information et de probabilité. Le problème a été cadré par la création du documentaire interactif conversationnel *On Chance*.

On Chance est un documentaire interactif, en ligne, qui propose un mode d'interactivité conversationnel, avec 3 niveaux d'intervention du hasard. Le premier niveau se situe dans la thématique abordée par le documentaire. Dans *On Chance*, l'interacteur doit

collaborer avec Jana, une journaliste scientifique, dans la rédaction d'un article sur les algorithmes prédictifs et les problèmes liés au hasard et à l'indétermination. Le deuxième niveau d'intervention du hasard est dans le processus de création du documentaire. Comme il est indiqué dans la section dédiée à la méthodologie, les deux approches choisies pour la création intègrent le hasard dans le processus : le cadrage de l'inopiné, de Jean-Luc Lioult, et la proposition des « enabling constraints » de Manning et Massumi. Enfin, le troisième niveau concerne l'utilisation d'algorithmes probabilistes pour la construction des récits et le design de la conversation.

Afin de mettre en évidence l'articulation des éléments théoriques avec la création de *On Chance*, je propose la discussion de deux exemples tirés du design de conversation du documentaire interactif. Le premier exemple concerne l'entité (ou "entity" en anglais) "@any". *Entity*, est une notion essentielle du traitement automatique du langage naturel qui désigne un groupe de mots. L'entité @any est utilisée pour que n'importe quel texte d'entrée de l'interacteur soit accepté par le chatbot. Avec l'entité @any, tout l'enjeu de la discussion sur le contrôle et l'indétermination du récit est présent, car cette entité permet d'avancer le récit de façon linéaire, contrôlée et prévisible. L'utilisation de l'entité @any dans le design de la conversation crée des conditions pour une expérience dans le registre de l'action-réaction, de l'habitude, des synthèses passives, avec des interactions prédéterminées.

Le deuxième exemple est le *Small Talk*, un ensemble d'*intents* (de morceaux de dialogues) qui sont tangentiels à la thématique de la conversation, c'est-à-dire des réponses qui ne sont pas directement liées au sujet principal du récit. Au contraire de l'entité @any, dans le cas du *Small Talk*, il s'agit de surprendre l'interacteur plutôt que de contrôler le récit ou de jouer dans l'habitude. Ce que nous avons dans le *Small Talk* ce sont des réponses inattendues, qui échappent au système d'action-réaction prévisible et créent un récit plus ouvert. Or, le documentaire doit se comporter comme un centre d'indétermination à son tour.

Il est intéressant de remarquer qu'un des enjeux théoriques de la thèse, soit la conception pluraliste d'un agencement multiple de différents types de synthèses, permet de composer les différents éléments qui poussent l'expérience du documentaire vers un récit plus ou moins indéterminé. Nous constatons que la mobilisation ou la présence de l'indétermination dans le documentaire interactif n'est pas une question binaire, ou dualiste, entre utiliser le hasard ou pas, mais plutôt comment moduler les différents éléments qui créent des conditions pour une expérience indéterminée, ou ouverte. Ainsi, une large utilisation de l'entité @any crée des conditions pour un récit plus prévisible, plus habituel, alors que l'utilisation du *Small Talk* permet de simuler un comportement indéterminé du documentaire. Or, le concept de l'image-interaction, une des contributions de cette recherche-crédation, permet de démontrer que l'image du documentaire interactif, lorsqu'elle est rapportée à un centre d'indétermination, a le potentiel de simuler être un centre d'indétermination à son tour. Autrement dit, le concept d'image-interaction permet de voir le propre documentaire comme un centre d'indétermination, et pas seulement une image rapportée. Il s'agit alors, dans l'image-interaction, de réagir aux incertitudes de l'interacteur et aussi d'insuffler de l'indétermination dans les réactions du documentaire. Les exemples de l'entité @any et du *Small Talk* illustrent les possibles articulations entre le concept d'image-interaction et la pratique du documentaire interactif conversationnel.

Les contributions de cette recherche-crédation se situent également dans les multiples dépassements de dualismes proposés, soit dans les aspects analogiques du documentaire, où il n'est plus question d'un sujet qui regarde un objet, mais plutôt un événement, ou une rencontre; soit dans les aspects numériques du documentaire interactif, où il sera question de composer, ou de moduler, avec les différents types de synthèses, ou des probabilités qui sont mobilisées. À la suite du constat initial de la vaste utilisation des modèles prédictifs probabilistes dans les algorithmes d'apprentissage automatique et de traitement massif de données, nous avons proposé une lecture pragmatiste spéculative des notions de hasard et indétermination (des

notions au cœur des modèles statistiques et probabilistes) et de ses relations avec les systèmes numériques (dans le contexte du documentaire interactif *On Chance*). Cette démarche a abouti à des résultats théoriques et pratiques :

- Le concept de *would be* (Peirce, 1957) ainsi que l'interprétation propensionniste de la théorie des probabilités (Gillies, 2000) ont permis de défaire le dualisme entre les interprétations dites objectives et subjectives des probabilités. La probabilité a été décrite comme une prise d'habitude, ou une propension développée à partir de répétitions. Ainsi, la probabilité n'est pas une propriété d'un objet ni d'un sujet, elle se retrouve plutôt dans les conditions d'un événement. Ce lien établi entre la probabilité et les relations entre les événements d'une suite, ou les conditions de répétition, nous permet par la suite de définir le hasard en termes relationnels : le hasard comme indépendance entre les éléments d'une série ;
- La mise en relation de la notion de synthèse (Deleuze, 1993 [1968]) avec la perspective propensionniste a permis d'élaborer le dépassement du dualisme (entre les interprétations objectives et subjectives de la théorie des probabilités) vers une compréhension processuelle et pluraliste de la probabilité qui articule synthèses passives et actives. La synthèse, ou « la différence soutirée de la répétition », de façon sommaire, est une contraction réalisée à partir d'événements qui se répètent. La synthèse passive, ou la contraction d'une habitude, décrit les propensions calculées par la probabilité d'une suite aléatoire. Sur la fondation des synthèses passives, les synthèses actives se rapportent aux activités de mémoire et de réflexion. La mémoire, en tant que mécanisme de la « reproduction de l'ancien », et la réflexion, comme le lien vers l'avenir, composent le double aspect de la synthèse active qui agit de sorte à créer des interférences dans l'indépendance entre les événements. Cette démarche théorique a permis de conclure : que les modèles probabilistes basés sur de séries indépendantes se rapportent aux synthèses passives ; que les

probabilités conditionnelles se rapportent aux synthèses actives ; et que ces deux types de synthèses s'articulent, car les synthèses actives sont formées sur la fondation des synthèses passives. Ainsi, les concepts de synthèse active et passive permettent d'élaborer une interprétation de la théorie des probabilités qui articule la pluralité des tendances mesurées par les probabilités pour des événements composés de séries aléatoires ou de séries conditionnelles (affectées par des activités de mémoire et de réflexion). Néanmoins, il est important de rappeler l'existence de la troisième synthèse, une tendance du système qui ne peut pas être inférée des données historiques et qui met ainsi en évidence une des limites de la probabilité dans sa capacité prédictive ;

- L'application de ces résultats à la théorie de l'information (Atlan, 2006 [1972]; Grunwald & Vitanyi, 2004; Shannon, 1948) et la théorie algorithmique de l'information (Chaitin, 1975, 2005) permet d'élargir la notion de hasard de l'indépendance à celle de complexité algorithmique et d'affirmer le caractère processuel et relationnel de l'information algorithmique : l'information algorithmique n'existe que dans la relation entre donnée et algorithme (plus la suite est aléatoire, plus l'algorithme qui sera capable de la reproduire sera complexe). La notion d'auto-organisation des structures dissipatives, « définie formellement comme création d'information dans un système organisé, non programmé de l'extérieur » (Atlan, 2006 [1972], p. 13) a permis d'établir le lien entre les activités de synthèse (organisatrices) et la création d'information. L'information est ainsi définie par le résultat d'une activité de synthèse (passive ou active), le résultat d'une contraction, de la différence soustrée de la répétition (organisation) ; Les travaux récents sur l'adaptation dissipative (England, 2015) peuvent décrire les processus physiques à la base des synthèses. Il s'agit ici d'une des questions soulevées dans le présent essai qui méritent des recherches futures : les apports des travaux sur l'adaptation dissipative à la perspective pragmatiste spéculative de l'interprétation des probabilités.

- À partir des cas étudiés, comme la famille d’algorithme de *clustering*, utilisée par Recast.AI⁵⁰ (ou Recast) dans le Natural Language Processing (NLP) du *chatbot* développé dans le cadre du *On Chance*, nous avons discuté les possibles interprétations philosophiques de l’utilisation des probabilités dans des algorithmes de ML. Il sera question notamment de l’utilisation des probabilités conditionnelles et des applications de la théorie de l’information en apprentissage automatique. Quelques aspects importants à la discussion de l’application des modèles probabilistes et ses possibles interprétations ont été soulevés, notamment dans le contexte des modèles basés sur le réseau neuronal. Cela a permis de proposer une approche pragmatiste spéculative de l’interprétation des probabilités dans le contexte du numérique, notamment des algorithmes d’apprentissage automatique et de formuler ainsi une interprétation non dualiste de la théorie des probabilités ;
- La distinction établie entre *interactivité* et *relation* (Massumi, 2011) a permis de constater que les dispositifs numériques des documentaires interactifs doivent aussi chercher à aller au-delà des systèmes prédéterminés d’action-réaction et à travailler la dimension relationnelle de l’expérience interactive. L’utilisation du hasard et de modèles prédictifs est une des stratégies qui peuvent être utilisées pour éviter la prédictibilité de l’expérience interactive de sorte à la rendre plus relationnelle et moins réactive ;
- La posture pragmatiste spéculative a permis de penser la notion d’expérience interactivité au-delà de la dichotomie qui sépare sujet et objet. Il a été question de proposer une interprétation de l’expérience du documentaire interactif en tant qu’événement, en tant que processus relationnel où se rencontrent image et interacteur. Ce constat a mené à une réflexion sur les conditions mises en place

⁵⁰ Recast est une entreprise française dédiée à la création de *chatbot*, fondée en 2015, et récemment achetée par une importante entreprise de développement de logiciel, SAP. Recast propose une plateforme collaborative qui permet le développement et le déploiement de *chatbots* basés sur le texte ou sur la voix.

pour l'émergence de nouvelles possibilités relationnelles dans l'expérience du documentaire interactif. On remarque alors que l'ouverture à l'indétermination n'est pas une caractéristique de l'interactivité, mais plutôt une possibilité qui émerge de la relation entre l'interacteur et le documentaire en fonction des conditions mises en place dans l'image. Il s'agit ainsi de créer des conditions pour que l'indétermination se manifeste pendant l'expérience ;

- La démarche Bergson-Deleuzienne de l'étude de l'image rapportée à un centre d'indétermination à travers un dispositif cinématographique a été appliquée dans cette thèse au dispositif interactif (plus spécifiquement, au documentaire interactif). Cela nous a permis d'arriver à la notion d'image-interaction, un type d'image qui a le potentiel de dépasser les limites de l'image-action et d'accueillir l'indétermination qui s'installe lorsque l'image est rapportée à l'interacteur. Le concept de « survol absolu » (Ruyer, 2012) a permis de comprendre un autre aspect de l'image-interaction : l'image-interaction agit autant dans la conscience primaire que secondaire, elle opère avec les deux systèmes d'image bergsoniens, ou les deux tendances de la connaissance (observation et conscience) de Ruyer ;
- Le concept de l'agencement a permis d'élaborer les aspects relationnels du documentaire interactif, car c'est dans la relation entre interacteur et dispositif, proportionnée par l'interactivité, que le documentaire prend forme. Lorsque mis en relation avec le dispositif interactif, le concept d'agencement permet de comprendre les conditions de possibilité du dépassement d'une interaction probable vers des expériences des possibles. Cela est dû notamment à l'émergence des propriétés à partir de l'interactivité (les relations des parties à l'ensemble) ;

À la suite de ces résultats, il est possible de constater certaines recommandations pratiques. La première recommandation concerne les modèles à la base des algorithmes prédictifs pour des systèmes complexes sont limités à la base et la vaste utilisation des

inférences bayésiennes (conditionnelles) dans les modèles prédictifs des algorithmes d'apprentissage automatique. Un des enjeux majeurs de l'utilisation des inférences bayésiennes est la définition des probabilités *a priori*. Nous constatons ainsi qu'une des démarches essentielles dans la compréhension des effets pratiques des algorithmes qui utilisent les inférences bayésiennes est l'attention constante sur les décisions concernant les présupposés emboutis dans les probabilités *a priori*, autrement dit, il est important de toujours considérer les questions sur *qui* calcule, et *comment* sont calculées, les probabilités *a priori* utilisées dans les inférences bayésiennes des algorithmes.

Mais peut-être la recommandation centrale de cette recherche-crédation concerne l'interprétation du rôle du hasard et de l'indétermination dans les médias interactifs et les algorithmes d'apprentissage automatique. Dans cette recherche-crédation il a été constaté que lorsque la notion de hasard n'est pas comprise dans le dualisme posé entre les interprétations objectives et subjectives, il est possible de proposer une vision processuelle de la communication qui articule hasard et organisation et qui place la création de l'information comme le résultat des synthèses auto-organisatrices des structures dissipatives, et l'information algorithmique comme le résultat des synthèses réalisées par des algorithmes. Ainsi, il a été démontré que les modèles prédictifs doivent agencer la multiplicité des synthèses passives et actives et considérer les limites des probabilités dans ses capacités prédictives. Cela est aussi important pour le développement et pour la mise en place de systèmes d'apprentissage automatique. Il faudrait alors avancer dans le développement d'une perspective pragmatiste spéculative de la notion d'information (et d'information algorithmique) et sa relation avec les processus d'organisation à partir du hasard, les synthèses passives et actives, l'adaptation dissipative.

ANNEXE A**Documentation video On Chance**

Lien pour consulter la documentation vidéo de On Chance :

https://drive.google.com/drive/folders/1V1vFVO_USoUIGrgqrQzKYzwrLWMc4ESW?usp=sharing

ANNEXE B

Scénario de On Chance (version du 2 mai 2021)

ON CHANCE

Chatbot Documentary

Logline: Congratulations on your new job! You were just hired as a journalist and your first assignment is to collaborate with your colleague Jana on a story about predictive algorithms.

Page d'accueil

On Chance

How predictive algorithms are changing your life?

From the uncertainty about the weather to the people we meet, chance is constantly affecting our lives. But if we could have more data and develop more precise models and algorithms, could we predict and control future events?

On Chance is a chatbot documentary exploring predictive algorithms. Have a chat with Jana and collaborate with her on a story about how predictive algorithms affect our lives.

On Chance is best experienced on computer screens.

Click to start

Page principale

Congratulations on your new job! You were just hired as a journalist and your first assignment is to collaborate with your colleague Jana on a story about predictive algorithms.

Your collaboration will be in 4 phases :

1. Understanding predictive algorithms; 2. Example: predictive policing; 3. Potential risks; 4. Solutions and recommendations.

Once you've covered the 4 phases, you'll be able to ask Jana to send you the draft of your paper!

Chat with Jana and collaborate on this story using the research material she'll show you.

Please send one message at a time 😊 To restart the conversation, just type "hello"

Who is Jana : Journalist specialized in science communication. She is about 45 years old, Lebanese-Canadian. Interested in science, technology, philosophy, movies, culture. Informative.

PHASE 1 : CONTEXT	
PREDICTIVE NARRATIVE	
INTENT/CONTEXT/MEDIA	TEXT
<p>Welcome</p> <p>Context out : getname</p>	<ul style="list-style-type: none"> - Hi, you are the new writer, welcome! - I'm Jana, what's your name?
<p>get.name</p> <p>Context in : getname</p> <p>Context out : name (100); intro; fallback-intro; video-not-working (100)</p> <p>\$name</p>	<ul style="list-style-type: none"> - Welcome \$name! I'm very happy you joined our team! - How are you?
<p>Intro</p> <p>Context in: intro</p> <p>Context out: topic; fallback-intro</p>	<ul style="list-style-type: none"> - I'm very excited to collaborate and share with you the research material - We can chat here, and I'll show you some video clips on the left side of the screen, alright?
<p>Topic</p> <p>Context in: topic</p>	<ul style="list-style-type: none"> - I've been covering technology for about 20 years, this topic is one of the most important I've written about

<p>Context out: Topic-followup; fallback-topic</p>	<ul style="list-style-type: none"> - The topic we're working on is one of the most important I've covered, - it affects how we build political views, how long people stay in prison, if people will get loans, and so many important things in our lives - it can affect police and justice decisions, bank decisions, and so many other important aspects of our lives. - The story is about predictive algorithms, have you ever heard about it? - We'll be working on predictive algorithms, do you know anything about it?
<p>Topic.yes</p> <p>Context in: Topic-followup</p> <p>Context out : Topic-followup-impressed</p>	<ul style="list-style-type: none"> - Awesome! Would you like to share what you heard about predictive algorithms?
<p>Topic.yes.impressed</p> <p>Context in: Topic-followup-impressed</p> <p>Context out: getopinion1; fallback-topic</p> <p>@any</p>	<ul style="list-style-type: none"> - That's cool. I was very impressed to realize that predictive algorithms are all over, - from recommendations on Youtube and Netflix to the police and the criminal justice, - But can we really trust an algorithm to make predictions for the police? Are those algorithms fair?
<p>Topic.no</p>	<ul style="list-style-type: none"> - No problem! I was very impressed when I realized that predictive algorithms are all over,

<p>Context in: Topic-followup Context out: getopinion1; fallback-topic</p>	<ul style="list-style-type: none"> - from recommendations on Youtube and Netflix to the police and the criminal justice, - But can we really trust an algorithm to make predictions for the police? Are those algorithms fair?
<p>Get.opinion1 Context in: getopinion1 Context out: first-task; opinion1 (100)</p> <pre>{ "type": "video", "select": "tags", "tags": ["intro"] }</pre> <p>\$1_opinion</p> <p>Videos : s1, s2, d1, n1 Tag : intro</p>	<ul style="list-style-type: none"> - Well, I guess this is what we are trying to figure out in our story. - There's a video on that, let's see it. - Media - Let me know what you think after you've watched it.
<p>First.task Context in : : first-task; Context out : firsttask-followup</p>	<ul style="list-style-type: none"> - Great! So here's my suggestion on what we should cover: First, we look at how predictive algorithms work and some examples. - Then, we move on to the risks and solutions for this tech.

<p>@any</p>	<ul style="list-style-type: none"> - At the end, we should have the backbone of our article. How does this plan sound to you?
<p>first.task.no</p> <p>Context in: firsttask-followup Context out: minority-report; 3-points</p> <pre>{ "type": "video", "select": "tags", "tags": ["def 1"] }</pre>	<ul style="list-style-type: none"> - Ok, where would you like to start? - Ok. Where would you like to start : some examples, the definition of predictive algorithms, the risks of this tech, or some of the solutions out there ?
<p>first.task.yes</p> <p>Context in: firsttask-followup Context out: comment_video1</p> <pre>{ "type": "video", "select": "tags", "tags": ["def 1"</pre>	<ul style="list-style-type: none"> - So, our first task is to understand what a predictive algorithm is and how it works. - There's a great video I found on that. Let's see what comes up. - Media - What do you think?

<p>] } a5, a6, c1, n3</p>	
<p>Comment.video.1 Context in: comment_video1 Context out: Commentvideo1-followup @any \$comment_video_1</p>	<ul style="list-style-type: none"> - Alright. Should we dig more into defining what is a predictive algorithm?
<p>Move.on (Comment.video.1.follow-up) Context in : Commentvideo1-followup Context out: Definition-followup</p>	<ul style="list-style-type: none"> - Sure, but before moving on to some examples of predictive algorithms, let's see if we got it right. - How would you describe a predictive algorithm?
<p>Dig.deeper (Comment.video.1.follow-up) Context in: Commentvideo1-followup Context out: definition {</p>	<ul style="list-style-type: none"> - Great! Let's watch another great clip that came up in the research - Let me know what you think once you've watched it.

<pre>"type": "video", "select": "tags", "tags": ["def 2"] }</pre> <p>Videos : A7, B5</p>	
<p>Definition</p> <p>Context in: definition</p> <p>Context out: Definition-followup ;</p> <p>@any</p> <p>\$comment_video_2</p>	<ul style="list-style-type: none"> - Alright #name.name! Before moving on to some examples, let's check if we got this right - How would you describe predictive algorithms?
<p>Definition.Follow-up</p> <p>Context in: Definition-followup</p> <p>Context out: Definition-followup-followup;</p>	<ul style="list-style-type: none"> - So, if I understood correctly, a predictive algorithm is a scoring system that uses past data to predict outcomes, right?

<p>has_definition (100); fallback-definition-followup; understanding_the_problem</p> <p>@any</p> <p>\$definition_1</p>	
<p>Definition.follow-up.no</p> <p>Context in: Definition-followup-followup</p> <p>Context out: understanding_the_problem</p>	<ul style="list-style-type: none"> - I see. As you said before, you think "#has_definition.definition_1" - Is that correct?
<p>Understanding.the.problem</p> <p>Context in: understanding_the_problem</p> <p>Context out: minority-report</p>	<ul style="list-style-type: none"> - Now I see why it's so important to know who gets to decide about the data and how the score is measured - Should we move on to some examples of predictive algorithms?
<p>Understanding Problem No</p>	<ul style="list-style-type: none"> - Would you rather discuss the risks or the philosophy of predictive algorithms?
<p>“DÉRIVES” and FALLBACK</p>	
<p>Can you define predictive algorithms?</p> <p>Context in : -</p>	<ul style="list-style-type: none"> - If I understood correctly, a predictive algorithm is a computer system that uses data to score things or people. - There's a great video about this. Let me know what you think.

<p>Context out: comment_video; name (100)</p> <pre>{ "type": "video", "select": "tags", "tags": ["general def"] }</pre> <p>Videos : A5, A6, C1, D1, A7, B5, n3</p>	
<p>Getname.catchall.</p> <p>Context in : getname</p> <p>Context out: name (100);</p> <p>Getnamecatchall-followup; fallback- intro</p> <p>@any</p> <p>\$name</p>	<ul style="list-style-type: none"> - I should call you \$name, is that correct?
<p>Getname.catchall.yes</p>	<ul style="list-style-type: none"> - Welcome #Getnamecatchall-followup.name! I'm very happy you joined the team! - How are you?

<p>Context in : name; Getnamecatchall-followup</p> <p>Context out: intro</p>	
<p>Getname.catchall.no</p> <p>Context in : name; Getnamecatchall-followup</p> <p>Context out: getname</p>	<ul style="list-style-type: none"> - No problem, can you repeat your name?
<p>Get.name.no.context</p> <p>Context in : -</p> <p>Context out: topic</p>	<ul style="list-style-type: none"> - Glad to talk to you, \$person! - We can chat here, and I'll show you some video clips on the left side of the screen, alright?
<p>Fallback.definition.follow-up</p> <p>Context in: fallback-definition-followup</p> <p>Context out: understanding_the_problem</p>	<ul style="list-style-type: none"> - I see. As you said before, you think "#has_definition.definition_1" - I guess everything will be clearer once we've seen some examples, right?
<p>Fallback.intro</p> <p>Context in: fallback-intro</p> <p>Context out: topic</p>	<ul style="list-style-type: none"> - I'm very excited to collaborate and share with you the research material - We can chat here, and I'll show you some video clips on the left side of the screen, alright?

<p>Fallback.topic.followup</p> <p>Context in: Topic-followup</p> <p>Context out: getopinion1</p>	<ul style="list-style-type: none"> - I was very impressed when I realized that predictive algorithms are all over, - from recommendations on Youtube and Netflix to the police and the criminal justice, - But can we really trust an algorithm to make predictions for the police? Are those algorithms fair?
<p>Fallback.understanding.the.problem</p> <p>Context in: understanding_the_problem</p> <p>Context out: minority-report;</p> <p>Definitionfollow-up-followup</p>	<ul style="list-style-type: none"> - I'm not sure I got that; do you agree that a predictive algorithm is a scoring system that uses past data to predict events?
<p>Fallback.first.task</p> <p>Context in: firsttask-followup</p> <p>Context out: comment_video1</p> <pre>{ "type": "video", "select": "tags", "tags": ["def 1"] }</pre> <p>a5, a6, c1, n3</p>	<ul style="list-style-type: none"> - So, our first task is to understand what a predictive algorithm is and how it works. - There's a great video I found on that. Let's see what comes up. - Media - What do you think?

small.talk.job	-
small.talk.family Context out: smalltalkfamily-followup	<ul style="list-style-type: none"> - I have a sister who lives in France, but we talk often. - Should we get back to our story on predictive algorithm?
small.talk.family.yes Context in: smalltalkfamily-followup Context out : womd; minority-report	<ul style="list-style-type: none"> - Cool! Would you rather discuss one example or the risks of predictive algorithms?
small.talk.family.no Context in: smalltalkfamily-followup	<ul style="list-style-type: none"> - ok, what would you like to discuss?
PHASE 2 : EXAMPLES	
Minority.report Context in: minority-report Context out: Minorityreport-followup	<ul style="list-style-type: none"> - I think predictive police should be our first example. - I always think about the movie Minority Report when I hear about predict police. Have you seen it?
Minority.report.no Context in: Minorityreport-followup Context out: getopinion2;	<ul style="list-style-type: none"> - In the movie, police were able to arrest criminals before they committed the crime - Do you think that's possible?

<pre>{ "type": "video", "select": "source", "source": " m1.mp4" }</pre>	
<p>Minority.report.yes</p> <p>Context in: Minorityreport-followup Context out: getopinion2</p> <pre>{ "select": "source", "type": "video", "source": "m1.mp4" }</pre>	<ul style="list-style-type: none"> - Remember in the movie how police were able to arrest criminals before they committed the crime? - Do you think that's possible?
<p>Get.opinion2</p> <p>Context in: getopinion2 Context out: opinion2 (100); predpol-cities</p> <p>@any \$opinion2</p>	<ul style="list-style-type: none"> - Many police forces around the world are using predictive algorithms, such as PredPol and LASER. - Do you know if police in our city are using predictive algorithms?
<p>Pred.Pol.cities</p> <p>Context in: predpol-cities</p>	<ul style="list-style-type: none"> - Info is not clear on which police are using predictive algorithms, but research shows that dozens of cities around the world are using them.

<p>Context out: 2approaches</p> <pre>{ "type": "video", "select": "tags", "tags": ["intro predictive police"] }</pre> <p>Payload Video : e1</p>	<ul style="list-style-type: none"> - How do you feel about this?
<p>2.approaches</p> <p>Context in: 2approaches</p> <p>Context out: approach1-laser; approach2-predpol; fallback- 2approaches; 2-approaches (100)</p> <p>@any \$comment_video_predpol</p>	<ul style="list-style-type: none"> - So, I figured out there are 2 main approaches to predictive policing: predicting (1) WHO will commit a crime or (2) WHERE crime will happen. - Which one should we investigate first?

<p>Approach 1.LASER</p> <p>Context in: approach1-laser</p> <p>Context out : approach1 (15); getopinion-laser</p> <pre>{ "select": "tags", "tags": ["intro approach 1"], "type": "video" }</pre> <p>Video: : d2, c2</p>	<ul style="list-style-type: none"> - Great! I have this video on the algorithms that are predicting who will commit a crime. - Please watch it and tell me your understanding of how they work.
<p>LASER.audit</p> <p>Context in: getopinion-laser</p> <p>Context out: LASERaudit-followup; fallback-laser</p> <p>@any</p> <p>\$comment-video-laser</p> <pre>{ "type": "video",</pre>	<ul style="list-style-type: none"> - Right. But these algorithms are being challenged by many people. One example we could mention is the lawsuit against the Los Angeles predictive police program called LASER. - The police had to audit the program and a lot of inconsistencies were found, but LAPD continues to use it. - Payload - I find it very alarming, should we mention it in our story?

<pre>"select": "tags", "tags": ["LASER"] }</pre> <p>Payload Video : e2, e3, e4, e10 Subject : Predictive Police Keyword : LASER</p>	
<p>LASER.audit.yes.mention</p> <p>Context in: LASERaudit-followup Context out : move-on-police-bias; approach2-predpol; laser-audit (100)</p>	<ul style="list-style-type: none"> - Great! Now, would you rather now investigate the second approach (predicting crime areas) or we move on?
<p>LASER.audit.no.mention</p> <p>Context in: LASERaudit-followup Context out : move-on-police-bias; approach2-predpol</p>	<ul style="list-style-type: none"> - Alright, let's research some more before making these choices. - Would you rather investigate the second approach (predicting crime areas) or we move on?
<p>LASER.audit.no.mention.ap.2.first</p> <p>Context in : approach2; LASERaudit-followup</p>	<ul style="list-style-type: none"> - Alright, let's research some more before making these choices. - Many studies are finding that predictive police algorithms target already over-policed

Context out : police-bias-problem	communities, do you see a problem?
<p>LASER.audit.yes.mention.ap.2.first</p> <p>Context in : approach 2; LASERaudit-followup</p> <p>Context out : police-bias-problem; laser-audit (100)</p>	<ul style="list-style-type: none"> - Great! Do you see any potential problem in the LASER program?
<p>Approach.2-PredPol</p> <p>Context in: approach2-predpol</p> <p>Context out: predpol-bias; approach2 (15)</p> <pre>{ "type": "video", "select": "tags", "tags": ["intro PredPol"] }</pre> <p>Payload</p> <p>Video : e6, e7, e5,</p> <p>Subject : PredPol</p>	<ul style="list-style-type: none"> - Cool! PredPol is one of the algorithms trying to predict crime areas. - Let's see how it works, there's a video on it. - Payload - How would you describe PredPol?

<p>Keyword : prediction</p>	
<p>PredPol.bias</p> <p>Context in: predpol-bias</p> <p>Context out: Predpolbias-followup; fallback-predpol</p> <p>@any</p> <p>\$predpol_description</p>	<ul style="list-style-type: none"> - I guess PredPol fits well the descriptions of predictive algorithms that came up so far: - it's basically past data + a scoring system, right?
<p>PredPol.bias.move.on.question</p> <p>Context in: Predpolbias-followup</p> <p>Context out: move-on-police-bias; approach1-laser</p>	<ul style="list-style-type: none"> - Alright. Would you rather investigate now the first approach (predicting who will commit crime) or we move on?
<p>PredPol.bias.app.1.first</p> <p>Context in : Predpolbias-followup; approach1</p> <p>Context out : police-bias-problem</p>	<ul style="list-style-type: none"> - Alright. Many studies are finding that predictive police algorithms target already over-policed communities, - How should we address this issue?
<p>Move.on.bias.police</p> <p>If move on</p>	<ul style="list-style-type: none"> - Many studies are finding that predictive police algorithms target already over-policed communities,

<p>Context in: move-on-police-bias Context out : police-bias-problem</p>	<ul style="list-style-type: none"> - How should we address this issue?
<p>Bias.problem.police</p> <p>Context in: police-bias-problem Context out: getopinion-predictive-police</p> <pre>{ "type": "video", "select": "tags", "tags": ["bias predictive police"] }</pre> <p>Payload Video : b11, e9, c3, d3, b12, b13 Subject : Predictive Police Keyword : blm</p> <p>@any</p>	<ul style="list-style-type: none"> - I begin to see where the problem is: the data being used by the algorithms to make predictions for the police is full of bias against some communities, - so the algorithm is just making things worse. - There's a great clip on that, please watch it and let me know what comes up.
<p>get.opinion.predictive.police</p> <p>Context in: getopinion-predictive-police</p>	<ul style="list-style-type: none"> - But humans also have bias, maybe even if the algorithm is not perfect, it could still do better than a human - What do you think?

<p>Context out: random-people-question; getopinion3 (100)</p> <p>@any</p>	
<p>random.people.question</p> <p>Context in: random-people-question Context out: random-people-result</p> <p>@any</p>	<ul style="list-style-type: none"> - Wait, I found this research, they compared predictions made by algorithms used in courtrooms to random people on the internet getting one dollar to make a guess on people's criminal behavior - guess who performed better, random people or the courtroom software?
<p>Random.people.result</p> <p>Context in: random-people-result Context out: randompeoplere-result-followup</p> <p>@any</p>	<ul style="list-style-type: none"> - Random people are just a bit better than the commercial algorithm, - There's a video on that research, would you like to see it?
<p>random.people.result.yes</p> <p>Context in: randompeoplere-result-followup Context out: randompeoplere-result-yes-followup</p>	<ul style="list-style-type: none"> - Cool, let's see it. - What do you think of this research?

<pre>{ "select": "tags", "tags": ["random people research"], "type": "video" }</pre> <p>Videos : d4</p>	
<p>random.people.result.yes.custom</p> <p>Context in: randompeoplereult-yes-followup</p> <p>Context out: mit</p> <p>@any</p>	<ul style="list-style-type: none"> - It makes me wonder, what's the point of using this technology if it is biased and does not even perform better than a random person on the web?
<p>random.people.result.no</p> <p>Context in: randompeoplereult-followup</p> <p>Context out: mit</p>	<ul style="list-style-type: none"> - Alright. But it makes me wonder, what's the point of using this tech if it is biased and does not even perform better than a random person on the web?
<p>MIT</p> <p>Context in : mit</p> <p>Context out: recap-phase2</p>	<ul style="list-style-type: none"> - MIT Tech Review is clear: "Predictive policing algorithms are racist. They need to be dismantled. - Lack of transparency and biased training data mean these tools are

<p>@any</p> <p>\$opinion_predictive_police</p>	<p>not fit for purpose. If we can't fix them, we should ditch them.”</p> <ul style="list-style-type: none"> - Maybe we should quote this, what do you think?
<p>“DÉRIVES” and FALLBACK</p>	
<p>Fallback.2.approaches</p> <p>Context in: fallback-2approaches</p> <p>Context out: approach2-predpol; approach1-laser</p>	<ul style="list-style-type: none"> - Sorry, which approach did you mean? 1 or 2?
<p>Fallback.minority.report</p> <p>Context in : Minorityreport-followup</p> <p>Context out : getopinion2</p>	<ul style="list-style-type: none"> - In the movie, police were able to arrest criminals before they committed the crime, - do you think that's possible?
<p>Fallback.laser</p> <p>Context in: fallback-laser</p> <p>Context out: LASERaudit-followup</p>	<ul style="list-style-type: none"> - I didn't get it. should we mention the Laser program in our story?
<p>Fallback.PredPol</p> <p>Context in : fallback-predpol</p> <p>Context out: Predpolbias-followup</p>	<ul style="list-style-type: none"> - I'm not sure I see what you mean. But it's all right, we can come back to this later, ok?

<p>Approach1.no.context</p> <p>Context out : approach1 (15); getopinion-laser</p> <pre>{ "select": "tags", "tags": ["intro approach 1"], "type": "video" }</pre> <p>Video: : d2, c2</p>	<ul style="list-style-type: none"> - I have this video on the algorithms that are predicting who will commit a crime. - Please watch it and tell me your understanding of how they work.
<p>Approach.2.no.context</p>	<p>-</p>
<p>Gender prediction</p>	<p>-</p>
PHASE 3 : RISKS	
<p>Recap.phase2</p> <p>Context in : recap-phase2</p> <p>Context out : womd; followup-no</p> <p>@any</p>	<ul style="list-style-type: none"> - Cool, #name.name! So far, we covered: what is a predictive algorithm and how it works, and we also checked an example, predictive police. - I think next step is looking closer at the risks of this tech. Do you agree?
<p>Recap.phase2.no</p>	<ul style="list-style-type: none"> - Alright. What would you like to cover ?

<p>Context in : followup</p> <p>Context out: womd</p>	
<p>WoMD</p> <p>Context in: womd</p> <p>Context out : 3-points</p> <pre>{ "select": "source", "type": "video", "source": "b6.mp4" }</pre> <p>Payload</p> <p>Video : B6</p> <p>Video subject : AI</p> <p>Video keyword : womd</p>	<ul style="list-style-type: none"> - Data scientist Cathy O’Neil call some of these predictive algorithms “weapons of math destruction” - Let’s see how she defines them in this video - What’s your take on this?
<p>3.points</p> <p>Context in : 3-points</p> <p>Context out : opacity; blind-trust; bias</p> <p>@any</p>	<ul style="list-style-type: none"> - Ok, I see 3 main problems in Weapons of Math Destruction: algorithmic bias, blind trust and opacity, - Which one should we cover first?

<p>Bias</p> <p>Context in : bias;</p> <p>Context out : bias (100);</p>	<ul style="list-style-type: none"> - Have you heard about bias in algorithm?
<p>Bias.yes</p> <p>Context in : Bias-followup</p> <p>Context out : Bias-yes-followup</p>	<ul style="list-style-type: none"> - Great! Would you like to share what you've heard about the bias problem?
<p>Bias.yes.custom</p> <p>Context in : Bias-yes-followup</p> <p>Context out : bias2</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "g1.mp4" }</pre>	<ul style="list-style-type: none"> - Interesting! Here's some research material on the different types of bias and how it works, let's watch it. - What's your take here?

<p>Payload</p> <p>Video : g1</p> <p>Subject : AI</p> <p>Keyword : bias1</p>	
<p>Bias.no</p> <p>Context in : Bias-followup</p> <pre>{ "select": "source", "type": "video", "source": "g1.mp4" }</pre> <p>Payload</p> <p>Video : g1</p> <p>Subject : AI</p> <p>Keyword : bias1</p>	<ul style="list-style-type: none"> - No problem! Here's some research material on the different types of bias and how it works, let's watch it. - What's your take here?
<p>Bias.2</p> <p>Context in : bias2</p> <p>Context out : Bias2-followup</p>	<ul style="list-style-type: none"> - I think we should dig a bit more on the problem of bias in algorithm, it's really the heart of the problem, don't you think?

<p>@any</p>	
<p>Bias2.yes</p> <p>Context in : Bias2-followup</p> <p>Context out : transition</p> <pre>{ "select": "tags", "tags": ["bias2"], "type": "video" }</pre> <p>Payload</p> <p>Video : a11, a10, e8</p> <p>Subject : AI</p> <p>Keyword : bias2</p>	<ul style="list-style-type: none"> - Ok, let's check another video on this matter. - I think it's very problematic, don't you?
<p>Bias2.no</p> <p>Context in : Bias2-followup</p> <p>Context out : transition</p>	<ul style="list-style-type: none"> - What do you see as the main problem to address in predictive algorithms?

<p>Blind.trust</p> <p>Context in : blind-trust</p> <p>Context out : Blindtrust-followup</p>	<ul style="list-style-type: none"> - Do you think an algorithm is more objective than humans in evaluating people?
<p>Blind.trust.yes</p> <p>Context in : Blindtrust-followup</p> <p>Context out : blind-trust2</p> <pre>{ "select": "tags", "tags": ["blind trust"], "type": "video" }</pre> <p>Payload</p> <p>Video : a9, b1</p> <p>Subject : AI</p> <p>Keyword : blind trust</p>	<ul style="list-style-type: none"> - Interesting! There's a great video on this. Let me know what you think.
<p>Blind.trust.no</p> <p>Context in : Blindtrust-followup</p>	<ul style="list-style-type: none"> - I agree. There's a great video on this. - Let me know what you think.

<p>Context out : blind-trust2</p> <pre>{ "select": "tags", "tags": ["blind trust"], "type": "video" }</pre> <p>Payload</p> <p>Video : a9, B1</p> <p>Subject : AI</p> <p>Keyword : blind trust</p>	
<p>Blind.trust.2</p> <p>Context in : blind-trust2</p> <p>Context out : blind-trust2-video</p> <p>@any</p>	<ul style="list-style-type: none"> - If I got it right, the problem with our blind trust on algorithms is that it makes us think algorithms are objective and have no bias, right?
<p>Blind.trust.2.video</p> <p>Context in : blind-trust2-video</p> <p>Context out : transition</p>	<ul style="list-style-type: none"> - I remember hearing something similar in this video.

<p>@any</p> <pre>{ "select": "tags", "tags": [" blind trust 2"], "type": "video" }</pre> <p>Payload Video : a13, B1, a22 Subject : AI Keyword : blind trust 2</p>	
<p>Opacity</p> <p>Context in : opacity Context out : opacity (100); opacity2</p> <pre>{ "select": "tags", "tags": ["opacity"], "type": "video" }</pre>	<ul style="list-style-type: none"> - One thing I realized is people know very little about what's going on inside predictive algorithms. - Even the people using them have no clue about what's in the black box! There's a great clip on this point. - Media - How do you feel about it?

<p>Payload Video : n2, v1 Subject : AI Keyword : opacity</p>	
<p>Education.example Context in : opacity2 Context out : Educationexample-followup @any</p>	<ul style="list-style-type: none"> - I read about public-school teachers being fired based on predictive algorithms with no idea about how the algorithm is scoring them. - Even worse, not even the schools or the City have the source code, it is owned by private companies. - Would you like to see a video on that issue?
<p>Education.example.yes Context in : Educationexample-followup { "select": "tags", "tags": ["education"], "type": "video" } Videos : a8, a23, a12, b3 Keyword : education</p>	<ul style="list-style-type: none"> - Great! Here's the clip. - I feel people have the right to know on what terms they are being evaluated, right? What do you think?

opacity2	
<p>Education.example.no</p> <p>Context in: Educationexample-followup</p> <p>Context out : transition</p>	<p>OK. I feel people have the right to know on what terms they are being evaluated, right? What do you think?</p>
<p>Transition.1</p> <p>Context in : opacity; bias; blind-trust; transition</p> <p>Context out : solutions;</p> <p>@any</p>	<ul style="list-style-type: none"> - Excellent, #name.name! We've covered all aspects of Weapons of Math Destruction! - Now let's check some solutions out there to address predictive algorithms. - That's our final step. Are you ready?
<p>Transition.2</p> <p>Context in : opacity; bias; transition</p> <p>Context out : move-solutions; blind-trust</p> <p>@any</p>	<ul style="list-style-type: none"> - Great! We have covered 2 out of 3 aspects of Weapons of Math Destruction. - Ready to check the last one, the blind trust on algorithms?

<p>Transition.3</p> <p>Context in : opacity; blind-trust; transition</p> <p>Context out : move-solutions; bias</p> <p>@any</p>	<ul style="list-style-type: none"> - Great! We've covered 2 out of 3 aspects of Weapons of Math Destruction. - Ready to check the last one, the problem of bias in predictive algorithms?
<p>Transition.4</p> <p>Context in : bias; blind-trust; transition</p> <p>Context out : move-solutions; opacity</p> <p>@any</p>	<ul style="list-style-type: none"> - Great! We have covered 2 out of 3 aspects of Weapons of Math Destruction. - Ready to check the last one, the problem of the opacity of algorithms?
<p>Transition.5</p> <p>Context in : bias; transition</p> <p>Context out : blind-trust; opacity; move-solutions</p> <p>@any</p>	<ul style="list-style-type: none"> - Alright, #name.name! We've covered first aspect of Weapons of math destruction: bias. - Do we move on to the problem of blind trust or opacity?
<p>Transition.6</p>	<ul style="list-style-type: none"> - Alright, #name.name! We've covered the first aspect of

<p>Context in : blind-trust; transition</p> <p>Context out : bias; opacity; move-solutions</p> <p>@any</p>	<p>Weapons of math destruction: blind trust.</p> <ul style="list-style-type: none"> - Do we move on to the problem of bias or opacity?
<p>Transition.7</p> <p>Context in : blind-trust; transition</p>	<ul style="list-style-type: none"> - Alright, #name.name! We've covered the first aspect of Weapons of math destruction: opacity. - Do we move on to the problem of blind trust or bias ?
<pre>{ "select": "source", "type": "video", "source": ".mp4" }</pre> <p>Payload</p> <p>Video : n2</p> <p>Subject : AI</p> <p>Keyword : opacity</p>	<ul style="list-style-type: none"> - That's happening with public-school teachers, there's a video on that, let's see it -
PHASE 4 : SOLUTIONS	
<p>Solutions</p>	<ul style="list-style-type: none"> - We should also explore solutions: what can we do to

	<p>fight against the injustices of predictive algorithm?</p> <ul style="list-style-type: none"> - Any idea where to look at?
<p>Solution</p> <ul style="list-style-type: none"> - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : A18, A20</p> <p>Keyword : regulation</p>	<ul style="list-style-type: none"> - I think a first place to explore is regulation. - I have something on that subject
<p>Solution</p> <ul style="list-style-type: none"> - Fallback - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : A21, B16</p> <p>Keyword : regulation</p>	<ul style="list-style-type: none"> - There's also the idea of auditing and testing algorithms
DÉRIVES	
<p>Derive1.Facebook</p> <p>Context in :</p> <p>Context out : Derive1Facebook-followup</p>	<ul style="list-style-type: none"> - Do you use Facebook?

<p>Derive1.Facebook – yes</p> <p>Context in : Derive1Facebook-followup</p> <p>Context out : Derive1Facebook-yes-followup</p>	<ul style="list-style-type: none"> - I see a lot of heated arguments on Facebook. I wonder if it is related to their predictive algorithms, what do you think?
<p>Derive1.Facebook - yes – custom</p> <p>Context in : Derive1Facebook-yes-followup</p> <p>Context out : first-task</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "a4.mp4" }</pre> <p>Payload</p> <p>Video : A4, a11</p>	<ul style="list-style-type: none"> - I remember a video about that, let's watch it. - Payload - What's your take away?
<p>Derive1.Facebook - yes - custom-2</p> <p>Context in : Derive1Facebook-yes-followup; has_definition</p>	<ul style="list-style-type: none"> - I remember a video about that, let's watch it. - Payload - Should we move on to other examples of predictive algorithms?

<p>Context out : minority-report</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "a4.mp4" }</pre>	
<p>Derive1.Facebook - yes - custom-3</p> <p>Context in : Derive1Facebook-yes-followup; opinion2 (100)</p> <p>Context out : getopinion-predictive-police</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "a4.mp4" }</pre>	<ul style="list-style-type: none"> - I remember a video about that, let's watch it. - Payload - What's your take away?
<p>Derive1.Facebook - yes - custom-4</p>	<ul style="list-style-type: none"> - I remember a video about that, let's watch it. - Payload

<p>Context in : Derive1Facebook-yes-followup; getopinion3</p> <p>Context out : recap-phase2</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "a4.mp4" }</pre>	<ul style="list-style-type: none"> - What's your take away?
<p>Derive2.Sports</p> <p>Context out : Derive2Sports-followup;</p>	<ul style="list-style-type: none"> - Do you think sports can be predicted?
<p>Derive2.Sports</p> <p>Context in : Derive2Sports-followup;</p> <p>Context out : first-task</p> <pre>{ "select": "source", "type": "video",</pre>	<ul style="list-style-type: none"> - I have a great clip on how data and statistics are being used in sports to make predictions, here it is. - So, what's your take away?

<pre>"source": "s3.mp4" }</pre>	
<p>Derive2.Sports - no-2</p> <p>Context in : Derive2Sports-followup; has_definition</p> <p>Context out : minority-report</p> <pre>{ "select": "source", "type": "video", "source": "s3.mp4" }</pre>	<ul style="list-style-type: none"> - I have a great clip on how data and statistics are being used in sports to make predictions, here it is. - Should we move on to other examples of predictive algorithms?
<p>Derive2.Sports - no-3</p> <p>Context in: Derive2Sports-followup; opinion2</p> <p>Context out : getopinion-predictive-police</p> <pre>{ "select": "source",</pre>	<ul style="list-style-type: none"> - I have a great clip on how data and statistics are being used in sports to make predictions, here it is. - So, what's your take away?

<pre>"type": "video", "source": "s3.mp4" }</pre>	
<p>Derive2.Sports- no-4</p> <p>Context in : Derive2Sports-followup; getopinion3</p> <p>Context out : recap-phase2</p> <pre>{ "select": "source", "type": "video", "source": "s3.mp4" }</pre>	<ul style="list-style-type: none"> - I have a great clip on how data and statistics are being used in sports to make predictions, here it is. - So, what's your take away?
<p>Derive3.Probability</p> <p>Context out: Derive3Probability-followup</p>	<ul style="list-style-type: none"> - During research I came across "the law of large numbers", and I was immediately brought back to my school years... do you remember anything about it?

<p>Derive3.Probability – yes</p> <p>Context in: Derive3Probability-followup Context out: first-task</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "p1.mp4" }</pre>	<ul style="list-style-type: none"> - Let's watch a clip on it. - What's your take away?
<p>Derive3.Probability - yes-2</p> <p>Context in: Derive3Probability-followup; has_definition Context out: minority-report</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "p1.mp4" }</pre>	<ul style="list-style-type: none"> - Let's watch a clip on it. - Should we move on to other examples of predictive algorithms?

<p>Derive3.Probability - yes-3</p> <p>Context in: Derive3Probability-followup; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "p1.mp4" }</pre>	<ul style="list-style-type: none"> - Let's watch a clip on it. - What's your take away?
<p>Derive3.Probability - yes-4</p> <p>Context in: Derive3Probability-followup; getopinion3</p> <p>Context out: recap-phase2</p> <p>@any</p> <pre>{ "select": "source",</pre>	<ul style="list-style-type: none"> - Let's watch a clip on it. - What's your take away?

<pre>"type": "video", "source": "p1.mp4" }</pre>	
<p>Derive4.Gambling</p> <p>Context out: Derive4Gambling-followup</p>	<ul style="list-style-type: none"> - Do you enjoy playing games of chance?
<p>Derive4.Gambling - yes</p> <pre>{ "select": "source", "type": "video", "source": "p3.mp4" }</pre> <p>@any</p>	<ul style="list-style-type: none"> - I never realized that gambling is also about math and predictions. Check this video. - What's your take away?
<p>Derive4.Gambling - yes-2</p> <p>Context in: Derive4Gambling-followup; has_definition</p> <p>Context out: minority-report</p>	<ul style="list-style-type: none"> - I never realized that gambling is also about math and predictions. Check this video. - Should we move on to other examples of predictive algorithms?

<pre>@any { "select": "source", "type": "video", "source": "p3.mp4" }</pre>	
<p>Derive4.Gambling - yes-3</p> <p>Context in: Derive4Gambling-followup; opinion2</p> <p>Context out: : getopinion-predictive- police</p> <pre>{ "select": "source", "type": "video", "source": "p3.mp4" }</pre> <p>@any</p>	<ul style="list-style-type: none"> - I never realized that gambling is also about math and predictions. Check this video. - What's your take away?
<p>Derive4.Gambling – yes-4</p> <p>Context in: Derive4Gambling-followup; getopinion3</p>	<ul style="list-style-type: none"> - I never realized that gambling is also about math and predictions. Check this video. - What's your take away?

<p>Context out: recap-phase2</p> <pre>{ "select": "source", "type": "video", "source": "p3.mp4" }</pre> <p>@any</p>	
<p>Derive5.Alternative</p>	<ul style="list-style-type: none"> - Do you often make predictions about your life?
<p>Derive5.Alternative - yes</p> <p>Context in: Derive5Alternative-followup</p> <p>Context out: first-task</p> <pre>{ "select": "source", "type": "video", "source": "p2.mp4" }</pre>	<ul style="list-style-type: none"> - Humans have always been trying to predict the future in many ways, watch this video about some of these ways. - What's your takeaway?
<p>Derive5.Alternative - yes-2</p>	<ul style="list-style-type: none"> - Humans have always been trying to predict the future in many

<p>Context in: Derive5Alternative-followup; has_definition</p> <p>Context out: minority-report</p> <pre>{ "select": "source", "type": "video", "source": "p2.mp4" }</pre>	<p>ways, watch this video about some of these ways.</p> <ul style="list-style-type: none"> - Should we move on to other examples of predictive algorithms?
<p>Derive5.Alternative - yes-3</p> <p>Context in: Derive5Alternative-followup; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <pre>{ "select": "source", "type": "video", "source": "p2.mp4" }</pre>	<ul style="list-style-type: none"> - Humans have always been trying to predict the future in many ways, watch this video about some of these ways. - What's your takeaway?
<p>Derive5.Alternative - yes-4</p>	<ul style="list-style-type: none"> - Humans have always been trying to predict the future in many

<p>Context in: Derive5Alternative-followup; getopinion3</p> <p>Context out: recap-phase2</p> <pre>{ "select": "source", "type": "video", "source": "p2.mp4" }</pre>	<p>ways, watch this video about some of these ways.</p> <ul style="list-style-type: none"> - What's your takeaway?
<p>Derive6.Physics</p> <p>Context out: Derive6Physics-followup</p> <p>the science of predictions</p>	<ul style="list-style-type: none"> - Do you think with enough information about the world we could predict anything?
<p>Derive6.Physics – no</p> <p>Context in: Derive6Physics-followup</p> <p>Context out: first-task</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "ch.mp4" }</pre>	<ul style="list-style-type: none"> - Let us hear what physics have to say about it in this video. - What's your takeaway?

}	
<p>Derive6.Physics - no-2</p> <p>Context in: Derive6Physics-followup; has_definition</p> <p>Context out: minority-report</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "ch.mp4" }</pre>	<ul style="list-style-type: none"> - Let us hear what physics have to say about it in this video. - Should we move on to other examples of predictive algorithms?
<p>Derive6.Physics - no-3</p> <p>Context in: Derive6Physics-followup; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <p>@any</p> <pre>{</pre>	<ul style="list-style-type: none"> - Let us hear what physics have to say about it in this video. - What's your takeaway?

<pre>"select": "source", "type": "video", "source": "ch.mp4" }</pre>	
<p>Derive6.Physics – no-4</p> <p>Context in: Derive6Physics-followup; getopinion3</p> <p>Context out: recap-phase2</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "ch.mp4" }</pre>	<ul style="list-style-type: none"> - Let us hear what physics have to say about it in this video. - What's your takeaway?
<p>Derive7.Chaos</p> <p>Context out:</p>	<ul style="list-style-type: none"> - Have you ever heard about chaos theory?
<p>Derive7.Chaos - yes</p>	<ul style="list-style-type: none"> - Here's a video about the origins of this theory and how

<p>Context in: Derive7Chaos-followup Context out: first-task</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "w1.mp4" }</pre>	<p>complexity plays a big role in limiting predictions.</p> <ul style="list-style-type: none"> - What's your takeaway?
<p>Derive7.Chaos - yes-2</p> <p>Context in: Derive7Chaos-followup ; has_definition Context out: minority-report</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "w1.mp4" }</pre>	<ul style="list-style-type: none"> - Here's a video about the origins of this theory and how complexity plays a big role in limiting predictions. - Should we move on to other examples of predictive algorithms?
<p>Derive7.Chaos - yes-3</p>	<ul style="list-style-type: none"> - Here's a video about the origins of this theory and how

<p>Context in: Derive7Chaos-followup ; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "w1.mp4" }</pre>	<p>complexity plays a big role in limiting predictions.</p> <ul style="list-style-type: none"> - What's your takeaway?
<p>Derive7.Chaos - yes-4</p> <p>Context in: Derive7Chaos-followup ; getopinion3</p> <p>Context out: recap-phase2</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "w1.mp4" }</pre>	<ul style="list-style-type: none"> - Here's a video about the origins of this theory and how complexity plays a big role in limiting predictions. - What's your takeaway?

Derive8. Micro-targeting	<ul style="list-style-type: none"> - Have you heard about the use of micro-targeting in politics?
Derive8. Micro-targeting – no Context in: Derive8Micro-targeting-followup Context out: first-task @any { "select": "source", "type": "video", "source": "b4.mp4" }	<ul style="list-style-type: none"> - micro-targeting is a social media marketing technique that is used in political campaigns. - I remember seeing a video about this... Here it is. - What's your takeaway? -
Derive8. Micro-targeting - no-2 Context in: Derive8Micro-targeting-followup ; has_definition	<ul style="list-style-type: none"> - micro-targeting is a social media marketing technique that is used in political campaigns. - I remember seeing a video about this... Here it is. - Should we move on to other examples of predictive algorithms?

<p>Context out: minority-report</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "b4.mp4" }</pre>	
<p>Derive8. Micro-targeting - no-3</p> <p>Context in: Derive8Micro-targeting-followup ; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "b4.mp4" }</pre>	<ul style="list-style-type: none"> - micro-targeting is a social media marketing technique that is used in political campaigns. I remember seeing a video about this... - Here it is. - What's your takeaway? -

<p>Derive8. Micro-targeting - no-4</p> <p>Context in: Derive8Micro-targeting-followup ; getopinion3 Context out: recap-phase2</p> <p>@any</p> <pre>{ "select": "source", "type": "video", "source": "b4.mp4" }</pre>	<ul style="list-style-type: none"> - micro-targeting is a social media marketing technique that is used in political campaigns. I remember seeing a video about this... - Here it is. - What's your takeaway? -
<p>Derive9.Hiring</p>	<ul style="list-style-type: none"> - During research I found that many companies are using predictive algorithms for hiring people. - At first, I thought this could be a good idea, maybe algorithms are more objective? What do you think? -
<p>Derive9.Hiring – no</p> <p>Context in: Derive9Hiring-followup</p>	<ul style="list-style-type: none"> - Ok, now hear this story and let me know your thoughts. - What's your takeaway? -

<p>Context out: first-task</p> <p>@any</p> <pre>{ "select": "tags", "tags": ["derive9"], "type": "video" }</pre>	
<p>Derive9.Hiring - no-2</p> <p>Context in: Derive9Hiring-followup; has_definition</p> <p>Context out: minority-report</p> <p>@any</p> <pre>{ "select": "tags", "tags": ["derive9"],</pre>	<ul style="list-style-type: none"> - Ok, now hear this story and let me know your thoughts. - Should we move on to other examples of predictive algorithms?

<pre>"type": "video" }</pre>	
<p>Derive9.Hiring - no-3</p> <p>Context in: Derive9Hiring-followup ; opinion2</p> <p>Context out: : getopinion-predictive-police</p> <p>@any</p> <pre>{ "select": "tags", "tags": ["derive9"], "type": "video" }</pre>	<ul style="list-style-type: none"> - Ok, now hear this story and let me know your thoughts. - What's your takeaway? -
<p>Derive9.Hiring - no-4</p> <p>Context in: Derive9Hiring-followup ; getopinion3</p> <p>Context out: recap-phase2</p>	<ul style="list-style-type: none"> - Ok, now hear this story and let me know your thoughts. - What's your takeaway?

<pre>@any { "select": "tags", "tags": ["derive9"], "type": "video" }</pre>	
<p>Hiring</p> <ul style="list-style-type: none"> - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : b10, A17, A18</p>	<ul style="list-style-type: none"> - Again, at first sight we might think an algorithm is more objective than a human, but research shows otherwise, - Let's see some material on this.
<p>Hiring</p> <ul style="list-style-type: none"> - Fallback - Fallback <p>Payload</p> <p>Video : b9</p> <p>Subject :</p>	<ul style="list-style-type: none"> - There are lawsuits against companies because some of those predictive algorithms are violating the law. - Check this story

<p>Hiring</p> <ul style="list-style-type: none"> - Fallback - Fallback - Fallback <p>Payload</p> <p>Video : B17, B18</p> <p>Subject :</p>	<ul style="list-style-type: none"> - It looks like it's all bad, but I found that there are some ways to improve fairness in hiring algorithms - Here's one idea
<p>Derive9.BigData</p>	<ul style="list-style-type: none"> - I see a lot of relation between predictive algorithms and big data, especially in the idea that more and more data is all it takes to make better predictions - Do you agree?
<p>Big Data Followup</p> <ul style="list-style-type: none"> - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : B17, B18</p>	<ul style="list-style-type: none"> - I never really understood what big data is all about... - Let's check this video about it.
<p>Derive10.Biology</p>	<ul style="list-style-type: none"> - Hear this story about a random event.
<p>Derive11.Art</p> <p>Would you like to hear about chance in art ?</p> <p>art</p>	<ul style="list-style-type: none"> - I found something interesting about the use of randomness in art. - Check this video.

Social.justice	<ul style="list-style-type: none"> - Maybe not predictive algorithms are bad, right? I think the problem is that some of them are not fair for groups of people, what do you think?
Social.justice.followup Payload Video : b19 Subject : Keyword :	<ul style="list-style-type: none"> - I've found a line about that, let's check this video, - Payload - What do you think?
Social.justice.moveon.1 Context in :	-
Social.justice.moveon.2 Context in : has_definition	-
Social.justice.moveon.3 Context in : 2	-
Social.justice.moveon.4	-

Context in : 3	
Politics	<ul style="list-style-type: none"> - Have you ever heard about micro targeting?
Politics Followup - Fallback Payload Video : Subject : B4 Keyword : micro targeting	<ul style="list-style-type: none"> - It's a social media marketing technique that is used in political campaigns - I remember seeing a video about this... - Here it is.
Big Data	<ul style="list-style-type: none"> - I see a lot of relation between predictive algorithms and big data, - Specially in the idea that more and more data is all it takes to make better predictions - What do you think ?
Big Data Followup - Fallback Payload Video : Subject : B17, B18	<ul style="list-style-type: none"> - I never really understood what big data is all about... - Let's check this video about it.
Models	<ul style="list-style-type: none"> - For a long time, people have been using models and scoring systems to understand the world and predict events. - But are those models fair?
Models Followup - Fallback	<ul style="list-style-type: none"> - "many of these models encoded human prejudice,

<p>Payload Video : A1, A2 Subject : AI Keyword : model</p>	<p>misunderstanding, and bias into the software systems that increasingly managed our lives” Says researcher Cathy O’Neil.</p> <ul style="list-style-type: none"> - There’s a clip about this topic, let’s see it.
<p>Philosophy of prediction</p>	<ul style="list-style-type: none"> - how can we predict anything in a world full of randomness and chaos? What do you think?
<p>Philosophy of prediction</p> <ul style="list-style-type: none"> - Fallback <p>Payload Video : Subject : A15 Keyword : prediction</p>	<ul style="list-style-type: none"> - Here’s a good reflection on this problem.
<p>Google research</p>	<ul style="list-style-type: none"> -
<p>Criminal justice</p>	<ul style="list-style-type: none"> - There’s also an article by Pro-Publica with evidence that courtrooms were using biased predictive algorithms to make decisions on how long people stay in prison. - Should we look at that story or do we go forward with the risks of predictive algorithms?
<p>Criminal justice.yes</p>	<ul style="list-style-type: none"> -

<p>Criminal Justice</p> <p>Videos : d3</p>	<ul style="list-style-type: none"> - Pro-Publica
<p>Criminal Justice</p> <p>Videos : d2</p>	<ul style="list-style-type: none"> - How it works
<p>Use</p>	<ul style="list-style-type: none"> -
<p>Education Scoring teachers</p>	<ul style="list-style-type: none"> - Around 2010, predictive algorithms started to be used to score teachers in public schools in the US. - Hundreds of teachers were fired based on bad scores - But research has shown that those algorithms are very inconsistent. - Do you think there's a problem there?
<p>Education Scoring teachers Followup</p> <ul style="list-style-type: none"> - Fallback <p>Payload</p> <p>Video : B8</p> <p>Subject : A12</p> <p>Keyword : education, score, teacher</p>	<ul style="list-style-type: none"> - I don't think it's fair to fire teachers based on flawed algorithms. - Let's check some more info about it.
<p>Education Scoring teachers Followup</p> <ul style="list-style-type: none"> - Fallback - Fallback 	<ul style="list-style-type: none"> - The algorithms are flawed in many ways... - One of them is because teachers get all the blame (and all the glory) for the students' scores

	<ul style="list-style-type: none"> - Remember when you were a teenager? so many things happening that could influence our performance in school, like the way things are at home, parents going through divorce, love affairs and heartbreaks, - But the algorithm calculates as if everything about the students' performance in school is about the teacher. - What do you think about it?
	-
<p>Education Scoring teachers Followup</p> <ul style="list-style-type: none"> - Fallback - Fallback - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : B3, B7</p> <p>Keyword : education, score, teacher, secret</p>	<ul style="list-style-type: none"> - And of course, there's also the problem that no one (outside the private company that developed it) really knows how these predictive models are scoring teachers. - Check this.
<p>Education Scoring teachers Followup</p> <ul style="list-style-type: none"> - Fallback - Fallback - Fallback - Fallback <p>Payload</p> <p>Video :</p> <p>Subject : A14</p>	<ul style="list-style-type: none"> - I heard about some teachers that won a case against predictive algorithms used to fire them. - It's here...

Keyword : education, score, teacher	
<p>Education Scoring teachers Followup</p> <ul style="list-style-type: none"> - Fallback - Fallback - Fallback - Fallback - Fallback 	<ul style="list-style-type: none"> - Would you like to research more about predictive algorithms in education or should we move to another area?
	-
Democracy	<ul style="list-style-type: none"> - I wonder if predictive algorithms can impact the way democracy works, what do you think?
<p>Democracy Followup</p> <ul style="list-style-type: none"> - Fallback <p>Payload</p> <p>Video :</p> <p>Subject :</p> <p>Keyword : regulation</p>	-

BIBLIOGRAPHIE

- (BPI), B. p. d. i. (2009). Nicolas Philibert: Le regard d'un cinéaste. In C. Pompidou (Ed.), *Le Mois du film documentaire*. Paris: Centre Pompidou.
- Almeida, A., & Alvelos, H. (2010). *An Interactive Documentary Manifesto*. Paper presented at the ICIDS'10 Proceedings of the Third joint conference on Interactive digital storytelling Edinburgh, UK.
- Aston, J., Gaudenzi, S., & Rose, M. (Eds.). (2017). *i-docs: The Evolving Practices of Interactive Documentaries*. New York: Columbia University Press.
- Atlan, H. (1972). Du bruit comme principe d'auto-organisation. *Communications*, 18, 21-36.
- Atlan, H. (2006 [1972]). *L'Organisation biologique et la théorie de l'information* Paris: Éditions du Seuil.
- Bensaude-Vincent, B. (2010). Splendeur et décadence da la vulgarisation scientifique. *Questions de communication*, 17. Retrieved from <http://questionsdecommunication.revues.org/368>
- Bergson, H. (1958 [1888]). *Essai sur les données immédiates de la conscience*. Paris: Presses Universitaires de France.
- Bergson, H. (2008 [1939]). *Matière et mémoire*. Paris: Presse Universitaire de France.
- Bergson, H. (2011). *Écrits philosophiques*. Paris: Presses Universitaires de France.
- Bergson, H. (2011 [1911]). *La conscience et la vie*. Paris: Presses Universitaires de France.
- Bergson, H. (2011 [1920]). *Le possible et le réel*. Paris: Presses Universitaires de France.
- Boisselle, L. N. (2016). Decolonizing Science and Science Education in a Postcolonial Space (Trinidad, a Developing Caribbean Nation, Illustrates). *SAGE Open*, 6(1). doi:10.1177/2158244016635257
- Brossard, D., & Lewenstein, B. V. (2009). *A critical appraisal of models of public understanding of science: Using practice to inform theory*.
- Burch, N. (1986 [1967]). *Une praxis du cinéma*. Paris: Gallimard.
- Burgess, M. M. (2014). From 'trust us' to participatory governance: Deliberative publics and science policy. *Public Understanding of Science*, 23(1), 48-52. doi:10.1177/0963662512472160
- Caillat, F. (2001). *Cinéma documentaire – Comment peut-on anticiper le réel ?* Paris: L'Harmattan.
- Carlsson, G. (2009). Topology and Data. *Bulletin (New Series) of the American Mathematical Society*, 46(2), 255-308.
- Carlsson, G. (2016). Why TDA and Clustering Are Not The Same Thing. Retrieved from <https://www.ayasdi.com/blog/machine-intelligence/why-tda-and-clustering-are-different/>
- Castelfranchi, Y., Vilela, E. M., Lima, L. B. d., Moreira, I. d. C., & Massarani, L. (2013). As opiniões dos brasileiros sobre ciência e tecnologia: o âparadoxoâ da relao entre informao e atitudes. *Histria, Cincias, Sade-Manguinhos*, 20, 1163-1183.

Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-59702013000501163&nrm=iso

- Chaitin, G. (1975). Randomness and Mathematical Proof. *Scientific American*, 232, 47-52.
- Chaitin, G. (2005). *Hasard et complexité en mathématiques*. Paris: Flammarion.
- Chapman, O. B., & Sawchuk, K. (2012). Research-Creation: Intervention, Analysis and "Family Resemblances". *Canadian Journal of Communication*, 37(1). Retrieved from <http://www.cjc-online.ca/index.php/journal/article/view/2489/2298>
- Chatonsky, G. (2004). Le centre d'indétermination : une esthétique de l'interactivité. *Intermédialités*, 3, 79-96.
- Clarizia, F., Colace, F., Lombardi, M., & Pascale, F. (2018). *A Context Aware Recommender System for Digital Storytelling*. Paper presented at the IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA), Krakow.
- Clark, A. (2016). *Surfing uncertainty : prediction, action, and the embodied mind*: Oxford : Oxford University Press.
- Crowley, D. (2014, October 6, 2014) *Foursquare's Dennis Crowley Foundation/Interviewer: K. ROSE*. TechCrunch.
- Curry, C., & O'Shea, J. (2011). *The implementation of a storytelling chatbot*. Paper presented at the Agent and Multi-Agent Systems: Technologies and Applications - 5th KES International Conference, KES-AMSTA 2011, Manchester, UK.
- Curry, C., & O'Shea, J. D. (2011). *The Implementation of a Story Telling Chatbot*. Paper presented at the New Directions in Agent Research workshop at KES-AMSTA-2011, Manchester.
- Davila, T. (2007). *Marcher, Créer : Déplacements, flâneries, dérives dans l'art de la fin du XXe siècle*. Paris: Editions du Regard.
- Dawson, E. (2018). Reimagining publics and (non) participation: Exploring exclusion from science communication through the experiences of low-income, minority ethnic groups. *Public Understanding of Science*, 27(7), 772-786. doi:10.1177/0963662517750072
- De Castro, E. V. (1998). Cosmological Deixis and Amerindian Perspectivism. *The Journal of the Royal Anthropological Institute*, 4(3), 469-488. doi:10.2307/3034157
- Debaise, D., & Stengers, I. (2016). L'insistance des possibles. Pour un pragmatisme spéculatif *Multitudes*, 4(65), 82-89.
- Delahaye, J.-P. (1999). *Information, complexité et hasard*. Paris.
- DeLanda, M. (2006). *A New Philosophy of Society - Assemblage theory and social complexity*. Londres: Continuum.
- Deleuze, G. (1983). *Cinéma 1, L'Image-mouvement*. Paris: Les Éditions de Minuit.
- Deleuze, G. (1983 [1962]). *Nietzsche et la philosophie*. Paris: Presses Universitaires de France.
- Deleuze, G. (1985). *Cinéma 2, L'Image-temps*. Paris: Les Éditions de Minuit.
- Deleuze, G. (1987). [Qu'est-ce que l'acte de création?].
- Deleuze, G. (1993 [1968]). *Différence et répétition*. Paris: Presses universitaires de France.
- Deleuze, G. (2003a). La vie comme œuvre d'art. In *Pourparlers* (pp. 129-138). Paris: Éditions de Minuit.
- Deleuze, G. (2003b). Qu'est-ce qu'un dispositif ? In *Deux régimes de fous*, (pp. 316-325). Paris: Éditions de Minuit.

- Deleuze, G., & Guattari, F. (1980). De la ritournelle. In *Mille plateaux* (pp. 382-397). Paris: PUF.
- Demaine, E., Devadas, S., & Lynch, N. (2015). Course Design and Analysis of Algorithms - Lecture 16: Complexity: P, NP, NP-completeness, Reductions. Cambridge: MIT OpenCourseWare.
- Einstein, A., Podolsky, B., & Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, *47*, 777-780.
- England, J. L. (2014). A new theory for evolution. Solna: Karolinska Institutet.
- England, J. L. (2015). Dissipative adaptation in driven self-assembly. *Nature Nanotechnology*, *10*(11), 919-923. doi:10.1038/nnano.2015.250
- Ewald, F. (2012). *Apologie de la prédiction*. Retrieved from Paris:
- Faure, C. (1991). Hasard et arts plastiques. In É. Noël (Ed.), *Le hasard aujourd'hui* (pp. 67-80). Paris: Éditions du Seuil.
- Fisher, U. (2009). *Walking the Edit : Dossier de présentation*. Retrieved from Genève:
- Fonseca, D. (Writer). (2019). Enterre-Moi Mon Amour - Documentaire Split/Screen. In France.
- Friston, K., Fitzgerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience and Biobehavioral Reviews*, *68*, 862-879. doi:10.1016/j.neubiorev.2016.06.022
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, *159*(3), 417-458. doi:10.1007/s11229-007-9237-y
- Frith, J. (2012). Splintered Space: Hybrid Spaces and Differential Mobility. *Mobilities*, *7*(1), 131-149.
- Galloway, D., McAlpine, K. B., & Harris, P. (2007). From Michael Moore to JFK Reloaded: Towards a working model of interactive documentary. *Journal of Media Practice*, *8*(3), 325-339.
- Garner, R. G. (2014). Film Theory and Chatbots. *International Journal of Synthetic Emotions*, *5*(1), 17-22. doi:10.4018/ijse.2014010103
- Gauchat, G. (2012). Politicization of Science in the Public Sphere: A Study of Public Trust in the United States, 1974 to 2010. *American Sociological Review*, *77*(2), 167-187. doi:10.1177/0003122412438225
- Gaudenzi, S. (2013a). The interactive documentary as a Living Documentary. *Doc On-line*(14), 9-31.
- Gaudenzi, S. (2013b). *The Living Documentary: from representing reality to co-creating reality in digital interactive documentary*. (PhD). University of London, London.
- Gifreu, A. (2011). *The interactive multimedia documentary: a proposed analysis model*. (PhD). Pompeu Fabra University, Retrieved from http://www.agifreu.com/web_dmi/articles/Interactive_multimedia_documentary_PrePHD_Ch5_Arnau_Gifreu.pdf
- Gifreu, A. (2013). *The Interactive Multimedia Documentary: A proposed Analysis Model*. (PhD). Pompeu Fabra University,
- Gillies, D. (2000). *Philosophical Theories of Probability*. London: Routledge.
- Goldberg, Y. (2016). A Primer on Neural Network Models for Natural Language Processing. *Journal of Artificial Intelligence Research*, *57*, 345-420.

- Goldstein, S., & Lebowitz, J. L. (2004). On the (Boltzmann) Entropy of Nonequilibrium Systems. *Physica D* 193, 53-66.
- Goriunova, O. (2012). *Art Platforms and Cultural Production on the Internet*: Routledge.
- Grian, H. (2017a). *A Tendency To Spill*. Bristol: PrettyDigital.
- Grian, H. (2017b). A Tendency To Spill: the ups and downs of turning a short story into a chatbot. Retrieved from <https://www.watershed.co.uk/studio/node/2531>
- Grunwald, P., & Vitanyi, P. (2004). Shannon Information and Kolmogorov Complexity. *arXiv:cs/0410002* Retrieved from <http://arxiv.org/abs/cs.IT/0410002>
- Guesmia, S. E. (2019). *Identification, validation et importance relative des dimensions de qualité des chatbots tels que perçus par les utilisateurs*. (Maîtrise). UQAM, Montréal.
- Gupta, P. (2017). How We Got 10 Million Teens to Read Fiction on Their Phones. Retrieved from <https://medium.com/@prernagupta/how-we-got-10-million-teens-to-read-fiction-on-their-phones-19a2a475084c>
- Hacking, I. (1975). *The Emergence of Probability*. Cambridge: Cambridge University Press.
- Hacking, I. (1990). *The taming of chance*. Cambridge: Cambridge University Press.
- Hájek, A. (2012). Interpretations of Probability. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
- Hao, K. (2020). Play AI bingo with your kids to teach them the fundamentals of AI. *MIT Technology Review*, 123(1), 48-51. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=140364556&lang=fr&site=ehost-live>
- Henderson, L. D. (2001). Désordre dans les beaux-arts: Kandinsky et Duchamp. In I. Prigogine (Ed.), *L'homme devant l'incertain*. Paris: Editions Odile Jacob.
- Hmielowski, J. D., Feldman, L., Myers, T. A., Leiserowitz, A., & Maibach, E. (2014). An attack on science? Media use, trust in scientists, and perceptions of global warming. *Public Understanding of Science*, 23(7), 866-883. doi:10.1177/0963662513480091
- Huybrechts, L. (Ed.) (2014). *Participation Is Risky: Approaches to Joint Creative Processes*. Amsterdam: Valiz.
- James, W. (1884). The Dilemma of Determinism. *Unitarian Review*.
- James, W. (1912). A World of Pure Experience. In *Essays in Radical Empiricism* (pp. 39-91). New York: Longman Green and Co.
- James, W. (2016). *Essays in Radical Empiricism*: Krill Press.
- Kahneman, D. (2012). *Thinking, fast and slow*. London: London : Penguin Books.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430-454.
- Katz, G. (2017). Introduction to Information Theory and Why You Should Care. Retrieved from <https://recast.ai/blog/introduction-information-theory-care/>
- Knight, W. (2017). AI Software Juggles Probabilities to Learn from Less Data. *MIT Technology Review*. Retrieved from <https://www.technologyreview.com/s/603542/ai-software-juggles-probabilities-to-learn-from-less-data/>
- Kolmogorov, A. N. (1968). Logical Basis for Information Theory and Probability Theory. *IEEE Transactions on Information Theory*, IT-14(5), 662-664.
- Kulkarni, T. D. (2015) *Short probabilistic programming machine-learning code replaces complex programs for computer-vision tasks*. Kurzweil Library, Needham.

- Lacotte, S. H. d. (2000). *Deleuze: philosophie et cinéma*. Paris: Harmattan.
- Lander, E. (Ed.) (2005). *A colonialidade do saber: eurocentrismo e ciências sociais. Perspectivas latinoamericanas*. Buenos Aires.
- Laplace, P. S. d. (1921 [1840]). *Essai philosophique sur les probabilités*. Paris: Paris Gauthier-Villars.
- Latour, B. (2005). From Realpolitik to Dingpolitik or How to Make Things Public. In ZKM (Ed.). Karlshue.
- Lawton, G. (2019). Simulating the world. *New Scientist*, 244(3250), 38-38. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=138915294&lang=fr&site=ehost-live>
- Lestienne, R. (1993). *Le hasard créateur*. Paris: Éditions de la découverte.
- Lévi-Leblond, J.-M., & Balibar, F. (1998). When did the indeterminacy principle become the uncertainty principle? *American Journal of Physics*, 66(4), 279-280.
- Levin, T. (2013). Do documentario ao webdoc - Questões em jogo num cenário interativo. *Doc On-line*(14), 71-92.
- Lévy-Leblond, J.-M. (1991). Hasard et mécanique quantique. In E. Noël (Ed.), *Le hasard aujourd'hui*. Paris: Editions du Seuil.
- Li, M., & Vitányi, P. M. B. (2008). *An Introduction to Kolmogorov Complexity and Its Applications*: Springer New York.
- Lioult, J.-L. (2008a). Cadrer l'inopiné. In J.-L. Lioult (Ed.), *Des mouvants indices du monde* (pp. 79-88). Aix en Provence: Publications de l'Université de Provence.
- Lioult, J.-L. (Ed.) (2008b). *Des mouvants indices du monde: Documentaire, traces, aléas*. Aix-en-Provence: Publications de l'Université de Provence.
- Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., . . . Carlsson, G. (2013). Extracting insights from the shape of complex data using topology. *Scientific Reports*, 3(1236).
- Maass, P., & Rajagopala, M. (2012, 13 juillet 2012). That's No Phone. That's My Tracker. *New York Times*. Retrieved from http://www.nytimes.com/2012/07/15/sunday-review/thats-not-my-phone-its-my-tracker.html?_r=0
- Machta, J. (1999). Entropy, information and computation. *American Journal of Physics*, 67(12), 1074-1077.
- Maguire, M., Rao, U., & Zurawski, N. (2018). *Bodies as evidence : security, knowledge, and power*: Durham : Duke University Press.
- Manning, E. (2016). *The Minor Gesture*. Durham and London: Duke University Press.
- Manning, E., & Massumi, B. (2014). *Thought in the Act: Passages in the Ecology of Experience*. Minneapolis: University Of Minnesota Press.
- Mansinghka, V. K. (2015). An Overview of Probabilistic Programming.
- Massumi, B. (1998). Line Parable for the Virtual. In J. Beckmann (Ed.), *The Virtual Dimension* (pp. 304-321). New York: Princeton Architectural Press.
- Massumi, B. (2011). *Semblance and Event*. Cambridge: The MIT Press.
- Miles, A. (2014). Interactive Documentary and Affective Ecologies. In K. Nash, C. Hight, & C. Summerhayes (Eds.), *New Documentary Ecologies: Emerging Platforms, Practices and Discourses* (pp. 67-82). Basingstoke: Palgrave Macmillan.

- Miles, A., Sora, C., Fetzner, D., & Aston, J. (2017, 2-6 August 2017). *The Material Turn and Interactive Documentary: A Panel*. Paper presented at the Visible Evidence XXIV, Buenos Aires (Argentina).
- Mitzenmacher, M., & Upfal, E. (2005). *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*: Cambridge University Press.
- Mlodinow, L. (2009). *The Drunkard's Walk: How Randomness Rules Our Lives*. New York: Vintage Books.
- Mohler, G., & Short, M. (2012). Geographic Profiling from Kinetic Models of Criminal Behavior. *SIAM Journal of Applied Mathematics*, 72, 163-180. doi:10.2307/41582570
- Morin, E. (1972). Le retour de l'événement. *Communications*, 18, 6-20.
- Morin, E. (1982). *Science avec conscience*. Paris: Fayard.
- Morin, E. (1990). Au-delà du déterminisme: le dialogue de l'ordre et du désordre. In K. Pomian (Ed.), *La querelle du déterminisme*. Paris: Gallimard.
- Motwani, R., & Raghavan, P. (1995). *Randomized Algorithms*. Cambridge: Cambridge University Press.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Cambridge: The MIT Press.
- Noël, É. (Ed.) (1991). *Le hasard aujourd'hui*. Paris: Éditions du Seuil.
- O'Flynn, S. (2012). Documentary's metamorphic form: Webdoc, interactive, transmedia, participatory and beyond. *Studies in Documentary Film*, 6(2), 141-157.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*: Crown Publishing Group.
- O'Neil, C. (2017, 16 juillet 2017). How can we stop algorithms telling lies? *The Guardian*. Retrieved from <https://www.theguardian.com/technology/2017/jul/16/how-can-we-stop-algorithms-telling-lies>
- Pan, P. (2004). *Mobile Cinema*. (Doctor of Philosophy). Massachusetts Institute of Technology, Cambridge.
- Paquin, L.-C. (2006). *Comprendre les médias interactifs*. Québec: Isabelle Quentin éditeur.
- Paquin, L.-C. (2014). *Méthodologie de la recherche-crédation*. Université du Québec à Montréal. Montréal. Retrieved from http://lcpaquin.com/methoRC/MethoRC_introduction.pdf
- Parisi, L. (2013). *Contagious Architecture: computation, aesthetics, and space*. Cambridge: The MIT Press.
- Pavel, K., Ivo, P., Mikuláš, M., František, D., & David, P. (2019). Chatbots for Enterprises: Outlook. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 67(6), 1541-1550. doi:10.11118/actaun201967061541
- Peirce, C. S. (1957). *Essays in the Philosophy of Science*. New York: The Liberal Arts Press.
- Peirce, C. S. (1989). *Writings of Charles S. Peirce: A Chronological Edition, Volume 4: 1879–1884*: Indiana University Press.
- Peirce, C. S. (1992). *The Essential Peirce, Volume 1: Selected Philosophical Writings (1867–1893)*. Bloomington: Indiana University Press.
- Perifel, S. (2014). *Complexité algorithmique*. Paris: Éditions Ellipses.
- Petters, L. B. (2019). Chatbots em campanhas de sensibilização, narrativa conversacional e possibilidades interativas: o caso do bot Fabi para Unicef Brasil e Facebook. *Comunicação, Midia e Consumo*, 16(46), 252-270. doi:10.18568/CMC.V16I46.1927

- Philibert, N. (2009). Avant-propos. In C. Pompidou (Ed.), *Nicolas Philibert: Le regard d'un cinéaste* (pp. 3-5). Paris: Centre Pompidou.
- Pollock, J. (Ed.) (2012). *Pratiques du hasard*. Perpignan: Presses Universitaires de Perpignan.
- Pomian, K. (Ed.) (1990). *La querelle du déterminisme*. Paris: Editions Gallimard.
- Prigogine, I. (Ed.) (2001). *L'homme devant l'incertain*. Paris: Editions Odile Jacob.
- Prigogine, I., & Stengers, I. (1979). *La Nouvelle Alliance*. Paris: Gallimard.
- Prigogine, I., & Stengers, I. (1997). *The End of Certainty*. New York: THE FREE PRESS.
- Purohit, J., Bagwe, A., Mehta, R., Mangaonkar, O., & George, E. (2019). *Natural Language Processing based Jaro-The Interviewing Chatbot*.
- Quijano, A. (2007). COLONIALITY AND MODERNITY/RATIONALITY. *Cultural Studies: Globalization and the De-Colonial Option*, 21(2-3), 168-178. doi:10.1080/09502380601164353
- Raymond, A. K. (2020, 15 avril 2020). Crowds Gather Around the Country to Protest Social Distancing. *New York Magazine*. Retrieved from <https://nymag.com/intelligencer/2020/04/crowds-gather-to-protest-social-distancing.html>
- Renvoisé, P. (2017). Introduction to text clustering. Retrieved from <https://recast.ai/blog/introduction-text-clustering/>
- Ruyer, R. (2012). *Néo-finalisme*. Paris: Presses Universitaires de France.
- Ryan, M.-L. (2005). Peeling the Onion: Layers of Interactivity. *Digital Narrative Texts*. Retrieved from <http://users.frii.com/mlryan/onion.htm>.
- Saad, F. A., Cusumano-Towner, M. F., Schaechtle, U., Rinard, M. C., & Mansinghka, V. K. (2019). Bayesian synthesis of probabilistic programs for automatic data modeling. *Proc. ACM Program. Lang.*, 3(POPL), Article 37. doi:10.1145/3290350
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3(3), 210-229.
- Santos, B., & Laville, J.-L. (2016). *Épistémologies du Sud : Mouvements citoyens et polémique sur la science*. Paris: Desclée de Brouwer.
- Saurisse, P. (2007). *La mécanique de l'imprévisible: Art et hasard autour de 1960*. Paris: L'Harmattan.
- Schrödinger, E. (1936). Indeterminism and Free Will. *Nature*, 138(3479), 13-14.
- Shannon, C. E. (1948). The Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(July and October), 379-423 (Part I), 623-656 (Part II).
- Silver, N. (2012). *The signal and the noise : why so many predictions fail--but some don't*. New York: Penguin Press.
- Simondon, G. (1964). *L'individu et sa génèse physico-biologique (L'individuation à la lumière des notions de forme et d'information)*. Paris: Presses Universitaires de France.
- Smith, R. E. (2016). Idealizations of Uncertainty, and Lessons from Artificial Intelligence. *Economics: The Open-Access, Open-Assessment E-Journal*, 10(2016-7), 1-40.
- Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010). Limits of Predictability in Human Mobility. *Science*, 327, 1018-1021.
- Stern, N. (2013). *Interactive Art and Embodiment: The Implicit Body As Performance*. Canterbury: Gylphi Limited.

- Stevens, H. (2020, 14 mars 2020). Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve” *Washington Post*.
- Stivale, C. J. (Ed.) (2011). *Gilles Deleuze: Key Concepts*. Kingston: McGill-Queen's University Press.
- Terranova, T. (2004a). Communication beyond Meaning: On the Cultural Politics of Information. *Social Text*, 22(3), 51-73.
- Terranova, T. (2004b). *Network Culture - Politics for the information age*. Londres: Pluto Press.
- Théraulaz, G. (1998). Les insectes architectes ont-ils leurs nids dans la têtes ? . *La Recherche*, 313.
- Thom, R. (1990). Halte au hasard, silence au bruit. In P. Krzysztof (Ed.), *La querelle du déterminisme*. Paris: Gallimard.
- Turing, A. M. (1936). On Computable Numbers, With an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2), 230-265.
- Vlahos, J. (2017). A SON'S QUEST TO GIVE HIS DYING FATHER ARTIFICIAL IMMORTALITY. *Wired*, 25(8), 56-67. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=124223583&lang=fr&site=ehost-live>
- Vogel, A. (2005 [1974]). *Film as Subversive Art*. Londres: C.T. Editions.
- Wilcox, B., & Wilcox, S. (2013). Making it real: Loebner-winning chatbot design. *Arbor*, 189(764), a086. doi:10.3989/arbor.2013.764n6009
- Yong, E. (2020). Why the Coronavirus Is So Confusing. *The Atlantic*. Retrieved from https://www.theatlantic.com/health/archive/2020/04/pandemic-confusing-uncertainty/610819/?utm_source=share&utm_campaign=share