

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

INTELLIGENCE D’AFFAIRES POUR LA SOUTENABILITÉ :
UNE PREUVE DE CONCEPT

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN INFORMATIQUE DE GESTION

PAR

FÁBIO EUGÊNIO DOS ANJOS BORGES

MAI 2017

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire 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.»

REMERCIEMENTS

D'abord, je tiens à remercier à Dieu pour la possibilité d'accomplir mes études au deuxième cycle et de surmonter toutes les difficultés inhérentes aux études dans une nouvelle langue et dans un nouveau pays.

Je remercie énormément à M. Albert Lejeune qui, en tant que professeur de l'Université du Québec à Montréal et directeur de recherche, m'a orienté pendant la construction de ce mémoire avec des connaissances et solutions très avancées et élaborées.

Mes sincères remerciements aux professeurs M. Mauro Oliveira de l'Institut Fédéral du Ceará – Campus Aracati (Brésil) et M. Abdellatif Obaid de l'Université du Québec à Montréal, qui ont été essentiels à mon arrivée à la maîtrise et pour le stage au Laboratoire de recherche en Télécommunication, Réseaux, Informatique Mobile et Embarquée (TRIME). Ainsi bien que les professeurs M. Joaquim Celestino Junior et M. Ricardo Holanda, de l'Université de l'État du Ceará (Brésil), pour la confiance qu'ils m'ont accordée.

Je voudrais remercier aussi à tous mes professeurs de l'UQAM pour les enseignements et les expériences partagées tout au long de ma formation, spécialement à M. Martin Cloutier qui m'a aidé à identifier mon domaine de recherche.

Enfin, je tiens à exprimer ma gratitude à ma famille (brésilienne et canadienne) et à mes amis qui étaient fiers de ma capacité à fouler ce chemin difficile.

DÉDICACE

À ma mère, Maria José, qui me suit de loin avec vos yeux gardiens.

À mon épouse Ivna Borges et nos enfants Mateus et Júlia, qui m'aident à remonter
après chaque chute.

À ma sœur Fabíola Borges, qui m'inspire à aller toujours plus loin.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	XI
LISTE DES TABLEAUX.....	XIII
RÉSUMÉ	XV
INTRODUCTION	1
CHAPITRE I	
PROBLÉMATIQUE.....	3
1.1 Le cycle de vie d'un produit.....	3
1.2 La soutenabilité.....	5
1.3 Les questions de recherche	7
1.4 Les objectifs.....	7
CHAPITRE II	
REVUE DE LA LITTÉRATURE.....	9
2.1 Les définitions de soutenabilité	9
2.2 La soutenabilité en milieu corporatif.....	10
2.3 Les approches de soutenabilité	12
2.3.1 Le standard ISO 14001.....	12
2.3.2 Sigma Guidelines	13
2.3.3 Dow Jones Sustainability Indices.....	16
2.3.4 G4 Guidelines.....	19
2.3.5 Triple Bottom Line.....	22
2.4 Les indicateurs de soutenabilité.....	24
2.5 Les rapports de soutenabilité	25
2.6 Intégration des indicateurs de soutenabilité.....	27
2.7 Technologie de l'information pour la soutenabilité.....	28
2.8 Les systèmes d'information	29

2.9	Intelligence d'affaires	31
2.10	Les types d'information	34
2.11	Les risques de l'intelligence d'affaires	35
2.12	Synthèse	36
2.13	Contribution de la recherche	38
CHAPITRE III		
	CADRE MÉTHODOLOGIQUE	39
3.1	Révision des objectifs	39
3.2	Type de recherche	40
3.2.1	Étape 1 : identification du problème et de la motivation	40
3.2.2	Étape 2 : définition des objectifs d'une solution	41
3.2.3	Étape 3 : conception et développement	41
3.2.4	Étape 4 : démonstration	42
3.2.5	Étape 5 : évaluation	42
3.2.6	Étape 6 : communication	43
CHAPITRE IV		
ÉLABORATION DE LA SOLUTION		
4.1	Lien avec la problématique	45
4.2	La structure de la méthode	46
4.3	Module 1 : sélection des indicateurs	46
4.3.1	Définir les décisions à prendre	47
4.3.2	Choisir les indicateurs	48
4.3.3	Choisir les approches de soutenabilité	48
4.3.4	Créer une stratégie d'exécution	49
4.4	Module 2 : adaptation et intégration	50
4.4.1	Identifier les sources des données	51
4.4.2	Adapter les systèmes d'information	53
4.4.3	Créer une base d'intégration de données	55
4.4.4	Transporter les données	62
4.4.5	Traiter les inconsistances	65
4.5	Module 3 : découverte de la connaissance	66

4.5.1 Comprendre les liaisons entre les faits et dimensions.....	66
4.5.2 Analyser les données.....	68
4.5.3 Supporter les prises de décisions.....	70
4.6 Les résultats attendus.....	70
CHAPITRE V	
PREUVE DE CONCEPT.....	73
5.1 Le contexte du cas.....	73
5.2 Module 1 : sélection des indicateurs.....	76
5.2.1 Définir les décisions à prendre.....	76
5.2.2 Choisir les indicateurs.....	80
5.2.3 Choisir les approches de soutenabilité.....	83
5.2.4 Créer une stratégie d'exécution.....	84
5.3 Module 2 : adaptation et intégration.....	86
5.3.1 Identifier les sources des données.....	86
5.3.2 Adapter les systèmes d'information.....	90
5.3.3 Créer une base d'intégration de données.....	91
5.3.4 Transporter les données.....	95
5.3.5 Traiter les inconsistances.....	106
5.4 Module 3 : découverte de la connaissance.....	106
5.4.1 Comprendre les liaisons entre les faits et dimensions.....	106
5.4.2 Analyser les données.....	109
5.4.3 Supporter les prises de décisions.....	116
5.5 Les résultats obtenus.....	117
5.6 Renseignements techniques.....	117
CHAPITRE VI	
DISCUSSIONS ET CONCLUSION.....	119
6.1 Discussions.....	119
6.2 Difficultés rencontrées et leçons apprises.....	123
6.3 Limites.....	124
6.4 Contribution.....	125

x

6.5 Recherches futures	125
BIBLIOGRAPHIE	127

LISTE DES FIGURES

Figure	Page
Figure 1.1 – Le cycle de vie d'un produit.	4
Figure 2.1 – Les principes du Sigma Guidelines (SIGMA Project, 2006)	14
Figure 2.2 – Recommandations pour le capital naturel (SIGMA Project, 2006).....	15
Figure 2.3 – Le cadre de gestion (SIGMA Project, 2006).	16
Figure 2.4 – Les catégories et les aspects du G4 Guidelines (GRI, 2014a).....	20
Figure 2.5 - Le Triple Bottom Line (Samajdar, 2014).....	22
Figure 2.6 – Architecture d'un Système BI (Sacu et Spruit, 2010).....	32
Figure 4.1 – Les Modules de la Méthode.....	46
Figure 4.2 – Exemple d'un modèle relationnel.	58
Figure 4.3 – Le schéma en étoile de la table de faits F_Facture.	59
Figure 4.4 – Le Schéma en flocon de neige de la table de faits Item.....	60
Figure 5.1 - Schéma en étoile des inscriptions.....	93
Figure 5.2 - Schéma en étoile de perspectives des professions.	94
Figure 5.3 - Schéma en étoile des déplacements.....	95
Figure 5.4 - ETL des programmes, facultés, campus et diplômes.	97
Figure 5.5 - ETL des inscriptions, composé par un "job".....	98
Figure 5.6 - ETL paramétré des inscriptions.....	99
Figure 5.7 - ETL de diplômes et cycles.	99
Figure 5.8 - ETL des professions.	101
Figure 5.9 - ETL de la table temporaire de professions et formations.....	102
Figure 5.10 - ETL des perspectives de professions.....	103

Figure 5.11 - ETL des régions, véhicules et déplacements ("staging").....	103
Figure 5.12 - ETLs de dimensions de véhicule, transport, déplacement et origine...	104
Figure 5.13 - ETL de déplacements et émissions de gaz.....	105
Figure 5.14 - Tableau de bord des programmes.	110
Figure 5.15 - Un rapport des inscriptions.	112
Figure 5.16 - Tableau de bord des villes.	114

LISTE DES TABLEAUX

Tableau		Page
Tableau 2.1	Les familles des indices (S&P, 2015).....	17
Tableau 2.2	Les critères de participation par familles d'indice (S&P, 2015).	18
Tableau 2.3	Comparaison des types d'information (Bara et al., 2009)	34
Tableau 5.1	Sources de données identifiées	88
Tableau 5.2	Liste des dimensions conformes.....	107

RÉSUMÉ

La dégradation de l'environnement par les activités commerciales et industrielles oblige les dirigeants des pays les plus influents du monde à fixer des objectifs de soutenabilité aux entreprises en général. La problématique de cette recherche est basée sur la difficulté d'appliquer la soutenabilité au sein des entreprises. Bien que les rapports de soutenabilité soient utilisés par les entreprises, l'intégration entre les dimensions économique, sociale et environnementale pour générer des connaissances représente un défi encore peu exploré dans la littérature. Ce mémoire comprend l'application de la soutenabilité dans les entreprises avec l'aide des approches de l'intelligence d'affaires. L'objectif est de proposer une méthode capable de sélectionner, intégrer et extraire la connaissance à partir des indicateurs de soutenabilité. La science du design a été utilisée dans le cadre méthodologique et une preuve de concept propose deux scénarios dans le milieu de l'éducation universitaire pour illustrer la faisabilité de la méthode proposée. Les résultats montrent la grande capacité de la méthode à créer une vraie intégration entre la gestion des entreprises et les questions liées à la soutenabilité. En utilisant les approches classiques de soutenabilité et d'intelligence d'affaires, cette recherche contribue à la littérature avec un cas pratique et reproductible.

MOTS-CLÉS : intelligence d'affaires, soutenabilité, entrepôt de données, indicateurs.

INTRODUCTION

La soutenabilité est présente dans les discussions entre les pays les plus influents du monde. C'est une question très pertinente et actuelle à cause de l'impact sur l'environnement : la pollution, la consommation de ressources naturelles, l'utilisation de l'énergie et de l'écosystème social. Les leaders mondiaux fixent des objectifs et des règles par rapport la soutenabilité à travers des conférences internationales comme la COP21, à Paris en décembre 2015.

Par conséquent, les entreprises essaient de s'adapter aux nouvelles règles pour réduire les effets de leurs actions commerciales et industrielles sur la planète. Les rapports de soutenabilité sont utilisés par les grandes entreprises qui visent principalement à améliorer leurs réputations auprès des clients, investisseurs et employés. Mais, la gamme d'indicateurs de soutenabilité présente dans ces rapports, cache des précieuses informations qui peuvent changer la façon de penser à soutenabilité dans les dimensions économique, sociale et environnementale.

Une façon d'appliquer la soutenabilité au sein des entreprises est l'intégration entre les indicateurs de nature environnementale et sociale avec les indicateurs corporatifs déjà connus, à travers des systèmes d'analyses de performance, par exemple les solutions d'intelligence d'affaires.

Basé sur des concepts existants depuis les années 70, le terme business intelligence (BI), connu en français comme l'intelligence d'affaires, a été créé en 1996 par le Gartner Group (Agiu, et al., 2014). BI peut être résumé comme l'utilisation de logiciels ou d'applications d'affaires pour la collecte, le traitement et l'analyse de données pour aider à la prise de décisions corporatives (Agiu, et al., 2014). En utilisant multiples

2

sources, l'intelligence est le résultat de la transformation de ces données brutes en informations utiles pour les décisions.

Cette recherche aborde le thème de la soutenabilité et propose une technique d'application dans le milieu corporatif en utilisant des principes d'intelligence d'affaires.

CHAPITRE I

PROBLÉMATIQUE

Ce chapitre aborde les difficultés rencontrées par les entreprises pour appliquer la soutenabilité et pour analyser leurs performances selon les dimensions économique, sociale et environnementale. D'abord, le cycle de vie d'un produit est expliqué pour illustrer le contexte de la problématique et ses conséquences. Ensuite, les défis de compréhension et d'application la soutenabilité dans les entreprises sont présentés.

1.1 Le cycle de vie d'un produit

Le processus de fabrication des produits mérite une attention majeure de l'industrie parce que le but est de toujours réduire les coûts financiers liés à la production. La Figure 1.1 propose un schéma illustratif du cycle de vie d'un produit basé sur le processus de fabrication pour expliquer la notion de soutenabilité dans trois dimensions : économique, sociale et environnementale.

Un produit peut être composé par divers sous-composants qui peuvent venir des plusieurs fournisseurs. Ces éléments sont considérés comme des intrants, qui sont des ressources naturelles ou non naturelles utilisées pour composer un produit dans le processus de fabrication. Par exemple, en considérant une pizza précuite, facilement trouvable dans les supermarchés, elle est composée par des ingrédients comme eau, sel, farine de blé, fromage, jambon, sauce tomate, etc. Mais, il faut aussi considérer les

emballages en plastique couvrant chaque pizza, les boîtes en carton qui protègent les pizzas et d'autres parties qui ne sont pas comestibles, mais qui font partie de l'ensemble du produit.



Figure 1.1 – Le cycle de vie d'un produit.

Il y a des employés impliqués dans ce processus qui, à l'aide d'équipements (mélangeur, four, emballeurs, etc.), sont responsables de la fabrication. Dans ce cas, il faut aussi considérer d'autres types de produits qui sont utilisés pour l'entretien des machines comme les lubrifiants et les filtres. L'énergie est un autre type d'intrant essentiel à tous les équipements. L'infrastructure physique couvre les meubles et les bâtiments qui hébergent les machines et les employés pendant leurs journées de travail.

À la fin du processus, les produits sont prêts pour les ventes aux supermarchés, qui atteignent les consommateurs finaux. Après la consommation de la pizza, une partie des déchets du produit est destinée aux dépotoirs, tandis qu'une autre partie est normalement envoyée au recyclage. L'avancement des ventes force le rythme de la fabrication et, ainsi, il génère des conséquences : les intrants sont consommés plus rapidement à tel point que les ressources naturelles ne peuvent pas se renouveler à la même vitesse que la consommation, les équipements deviennent obsolètes et sont remplacés, les employés tombent malades à cause du stress lié au travail, l'infrastructure physique devient insuffisante et demande plus d'espace, les déchets et les émissions de gaz augmentent les impacts sur l'environnement, etc.

1.2 La soutenabilité

Selon Khomba et Vermaak (2012), les systèmes traditionnels de rapport financier ne reflètent pas les multiples dimensions de la performance de l'entreprise en se concentrant presque exclusivement sur les mesures financières (les revenus, les coûts fixes ou liés à la production, les profits, les résultats financiers, etc.) et la rentabilité. Les objectifs stratégiques, les modèles de gestion, les processus d'affaires, les systèmes d'informations, les analyses de données et les systèmes d'aide à la décision sont, par conséquent, tous basés seulement sur la dimension économique.

Mais, l'une des définitions classiques de soutenabilité dit que le développement soutenable doit répondre aux besoins du présent sans compromettre la capacité des générations futures de répondre à leurs propres besoins (Brundtland, 1987). Alors, la soutenabilité doit s'occuper non seulement de la partie économique, mais aussi de tous les autres éléments impliqués dans le processus de fabrication.

Ainsi, la soutenabilité vise à analyser toutes les possibles conséquences de la production en utilisant trois dimensions : économique, sociale et environnementale.

Au-delà des questions économiques, la dimension sociale est basée sur la valorisation et les soins avec les collaborateurs, les clients, la société et tout qui est influencé par l'entreprise ou le produit. Enfin, la dimension environnementale vise à analyser les impacts sur la nature et les écosystèmes.

Plusieurs approches d'application de la soutenabilité aux entreprises ont été formulées : Sigma Guidelines, ISO 14001, DJSI, G4 Guidelines, Triple Bottom Line, etc. Généralement, elles sont basées sur les listes de contrôle pour évaluer le niveau de soutenabilité ou en utilisant des indicateurs pour créer des rapports de soutenabilité. Ces rapports sont normalement utilisés pour améliorer les réputations des entreprises face aux clients, employés et investisseurs ou pour s'adapter aux réglementations des gouvernements.

Cependant, sélectionner l'approche et les indicateurs de soutenabilité les plus pertinents selon le domaine d'une entreprise est une tâche difficile à gérer à cause de la variété d'options disponibles. En outre, les recommandations standards des rapports de soutenabilité ne sont pas suffisantes pour expliquer l'intégration entre les indicateurs des dimensions économiques, sociaux et environnementaux et pour interpréter les possibles relations cachées derrière ces indicateurs.

La Technologie de l'Information (TI) peut offrir quelques outils et solutions pour aider à l'application de la soutenabilité au sein des entreprises. La littérature et les approches de soutenabilité indiquent l'utilisation des systèmes d'information pour gérer les indicateurs. Mais, les approches d'analyse de données comme les outils d'intelligence d'affaires sont capables d'aider à l'intégration et à l'analyse des données. Néanmoins, l'analyse de la soutenabilité par les approches BI est encore peu connue à la fois dans les affaires et dans la littérature.

1.3 Les questions de recherche

Notre question de recherche principale, basée sur les discussions exposées au-dessus, est la suivante :

« En quoi les approches en intelligence d'affaires sont-elles utiles et pertinentes pour générer de la connaissance stratégique de soutenabilité au-delà de l'identification et de l'intégration des indicateurs économiques, sociaux et environnementaux ? »

Pour répondre à cette question, on propose la division sous la forme de questions secondaires comme les suivantes :

- Quels sont les approches et les indicateurs les plus pertinents selon le domaine d'une entreprise ?
- Comment les indicateurs sociaux et environnementaux peuvent-ils être intégrés aux indicateurs corporatifs dans le but de générer de la connaissance stratégique ?

1.4 Les objectifs

L'objectif central de cette recherche est de proposer une méthode d'application de la soutenabilité dans les entreprises, en utilisant les approches d'intelligence d'affaires, permettant de sélectionner et d'intégrer les indicateurs économiques, sociaux et environnementaux pertinents à une entreprise pour générer et utiliser des connaissances stratégiques.

Les objectifs secondaires sont :

- Identifier des types de relations pertinentes entre les indicateurs corporatifs et de soutenabilité selon le domaine et la mise au point décisionnelle d'une entreprise.

8

- Étendre le processus décisionnel pour accueillir les nouveaux indicateurs des dimensions de soutenabilité.

CHAPITRE II

REVUE DE LA LITTÉRATURE

Cette partie de notre recherche aborde les définitions et les principales publications académiques sur les concepts liés à la soutenabilité et à l'intelligence d'affaires.

2.1 Les définitions de soutenabilité

Le terme en anglais *sustainability* peut être traduit en français comme soutenabilité ou durabilité. La définition contenue dans le célèbre *Oxford English Dictionary* dit que c'est « la capacité à être maintenue à un certain taux ou niveau ». Alors, on parle d'un concept général qui n'est pas seulement lié à l'environnement et on peut l'appliquer à de différents concepts et points de vue. Plusieurs concepts de soutenabilité peuvent être trouvés dans la littérature. Mais, Gatto (1995) explique qu'il y a trois définitions majeures dans le domaine :

- Définition des biologistes : le rendement soutenu des ressources qui découlent de l'exploitation des populations et des écosystèmes.
- Définition des écologistes : l'abondance soutenue et la diversité génotypique des espèces individuelles dans des écosystèmes soumis à l'exploitation humaine ou, plus généralement, l'intervention.
- Définition des économistes : le développement économique soutenable sans compromettre les ressources existantes pour les générations futures.

Cependant, Gatto (1995) dit dans son article que les trois définitions sont incohérentes entre elles. Hansson (2010) expose aussi que le concept de soutenabilité est ambigu et souvent difficile à utiliser dans un contexte socio-technologique.

Ces concepts peuvent être classifiés principalement en deux axes de réflexion : la soutenabilité forte ou soutenabilité faible. La soutenabilité forte dit que les ressources naturelles et les choses faites par les hommes doivent être toutes préservées séparément. C'est une définition plus rigide et, normalement, défendue par les écologistes. En revanche, le concept faible est plus proche des économistes et dit qu'un développement durable doit répondre aux besoins du présent sans compromettre la capacité des générations futures à satisfaire leurs propres besoins. On peut le comprendre comme un concept plus flexible ou même laxiste, une fois que le remplacement partiel ou total des ressources naturelles par des ressources non naturelles est acceptable.

Ainsi, c'est possible de percevoir que chaque partie prenante interprète la soutenabilité selon ses propres besoins ou intérêts. Parmi les concepts de soutenabilité, le plus pertinent est celui qui a un niveau intermédiaire entre les approches radicales et qui peut gérer le conflit entre le développement économique et la préservation de l'environnement (Hansson, 2010).

Ainsi, on peut conclure que les définitions fortes et faibles de soutenabilité sont applicables et adaptées selon de différents cas. Une analyse qui combine les deux est, alors, un processus difficile à mettre en œuvre.

2.2 La soutenabilité en milieu corporatif

Selon Aras et Crowther (2009), les termes soutenabilité et développement durable sont considérés des synonymes. Plus récemment, le développement soutenable est devenu

un slogan aussi pour la Responsabilité Sociale des Entreprises – RSE (Hansson, 2010). Ce concept (en anglais, Corporate Social Responsibility – CSR) peut être compris comme un ensemble de politiques et pratiques intégrées dans les opérations commerciales et les processus de prise de décision que vise inculquer la responsabilité pour les actions actuelles et passées ainsi que les impacts futurs.

Pendant l'année de 1987, la *World Commission on the Environment and Development (WCED)* a abordé le développement soutenable comme un important concept lié à l'environnement par le *The Brundtland Report*. Selon le rapport, « le développement soutenable est un développement qui répond aux besoins du présent sans compromettre la capacité des générations futures de répondre aux leurs » (Brundtland, 1987). C'était une première approche formelle entre la soutenabilité et le milieu corporatif. Depuis lors, le thème attire l'attention des entreprises, au-delà des gouvernements et des organismes liés aux questions sociales et environnementales.

Des conférences au niveau mondial ont été faites pour discuter la soutenabilité comme, par exemple, la *Conference of Parties (COP21)*, qui a eu lieu à Paris en 2015. Les discussions entre les leaders visent à définir les objectifs de réduction des impacts sociaux et sur l'environnement de chaque pays.

Les entreprises, alors, sont impliquées dans les buts en ajoutant la soutenabilité dans leurs objectifs corporatifs. Ainsi, les stratégies corporatives ne devraient plus compromettre la capacité des autres parties prenantes d'accéder aux mêmes ressources au futur (Petrini et Pozzebon, 2009).

Alors, la soutenabilité touche le milieu corporatif en visant l'intégration entre la culture organisationnelle classique et les objectifs liés à la soutenabilité. Ceci peut être traduit comme une fusion entre les dimensions économiques, sociales et environnementales, nommée comme la triple performance ou le triple résultat (Petrini et Pozzebon, 2009).

Mais, le thème soutenabilité n'est pas habituellement intégré à la gestion d'affaires dans les entreprises, même dans celles qui considèrent la responsabilité sociale et la soutenabilité comme une partie de leurs intentions stratégiques (Petrini et Pozzebon, 2009). Généralement, la gestion de la soutenabilité est encore traitée par un secteur à part.

2.3 Les approches de soutenabilité

La littérature présente plusieurs tentatives d'appliquer la soutenabilité dans les entreprises. Mais les grands organismes globaux ont créé des approches plus complètes et connues en vigueur aujourd'hui. Munajat et Kurnia (2015) ont examiné les approches de soutenabilité les plus utilisées dans la littérature et ils ont trouvé le standard ISO 14001 et quatre guides : Sigma Guidelines, Dow Jones Sustainability Indices et G4 Guidelines.

2.3.1 Le standard ISO 14001

Selon Oxford International Dictionary, un standard est une idée ou une chose utilisée comme une mesure, une norme ou un modèle dans les évaluations comparatives. ISO 14001 c'est un standard créé en 1996 par International Organisation for Standardization (ISO), un célèbre organisme de standards. Il aborde les exigences pour créer un système de gestion de l'environnement.

Selon International Organization for Standardization (2015a), les entreprises qui utilisent ISO 14001 disent qu'il sert à démontrer la conformité avec les exigences réglementaires, pour augmenter l'engagement des employés, intégrer des questions environnementales dans la gestion des affaires, améliorer l'efficacité et réduire les

coûts pour fournir un avantage concurrentiel et encourager les fournisseurs à améliorer leurs performances environnementales.

Les résultats attendus d'un système de gestion de l'environnement, selon ISO 14001, comprennent l'amélioration de la performance environnementale, le respect aux obligations de conformité et la réalisation des objectifs environnementaux à n'importe quel type, taille ou nature d'entreprise. Ce standard peut être utilisé partiellement, mais il ne peut pas être accepté par rapport la conformité avant de remplir toutes les exigences. C'est important de définir la limite du standard : il ne précise pas les critères de performance environnementale.

2.3.2 Sigma Guidelines

Le « Sustainability Integrated Guidelines for Management », ou simplement Sigma Guidelines, a été créé en 1999 par un consortium composé par le British Standards Institution, un leader dans le domaine des standards ; le Forum for the Future, un organisme leader en bienfaisance ; AccountAbility, un organisme professionnel international de responsabilité et le Ministère britannique du Commerce et de l'Industrie (SIGMA Project, 2006).

Le but principal est d'aider les organisations à répondre efficacement aux défis posés par les dilemmes, des menaces et des opportunités sociaux, environnementaux et économiques pour construire un futur plus soutenable. La construction du guide a rassemblé les meilleures pratiques utilisées par les grandes entreprises et les organisations du secteur public, en considérant les commentaires des parties prenantes. Ainsi, le guide est un ensemble des principes qui orientent à la compréhension de la soutenabilité et un Cadre de Gestion (Management Framework) qui intègre les questions de soutenabilité dans les processus de base et de prise de décisions. Sigma est flexible et peut être utilisé selon les contextes des entreprises.



Figure 2.1 – Les principes du Sigma Guidelines (SIGMA Project, 2006)

Il y a deux principes majeurs qui orientent les recommandations du Sigma Guidelines : la gestion de cinq types de capital qui reflète l'impact général sur l'entreprise ; et le respect aux règles et normes, la transparence, la sensibilité aux parties prenantes comme un exercice de la responsabilité. Ces capitaux, selon la Figure 2.1, sont naturel, social, humain, les immobilisations du capital manufacturé et le capital financier. Les capitaux sont entrelacés et leurs limites sont souvent mélangées. Pour chaque capital, Sigma Guidelines propose une liste de recommandations pratiques comme, par exemple, les recommandations du capital naturel montrées dans la Figure 2.2.

La responsabilité (accountability), selon Sigma Guidelines, est basée sur trois éléments : la transparence, la nécessité de répondre aux parties prenantes et la

conformité volontaire aux règlements. Alors, l'union des capitaux et la responsabilité forment la base de ce guide.

- Limit and reduce over time the use of substances extracted from the Earth's crust.
- Substitute abundant minerals for those that are scarce in nature.
- Ensure that all mined materials are used efficiently within cyclic systems. Systematically reduce dependence on fossil fuels and use renewable resources instead.
- Eliminate the accumulation of human-made substances and products in nature – substitute all persistent and unnatural compounds with substances that can be easily assimilated by natural systems. Eliminate waste, re-use, recycle or re-manufacture where possible.
- Prevent the physical degradation of nature and protect and enhance biodiversity and eco-system functions.
- Draw renewable resources only from well-managed and restorative eco-systems.
- Systematically pursue the most productive and efficient use of resources and land.
- Adopt the precautionary principle in any situation that may result in the modification of nature.

Figure 2.2 – Recommandations pour le capital naturel (SIGMA Project, 2006).

Le Cadre de Gestion est un ensemble de pratiques qui visent appliquer les principes du Sigma Guidelines dans l'entreprise. Les quatre phases (Planning ; Delivery ; Monitor, Review and Report ; et Leadership and Vision) et ses sous-phases sont basées suivent le modèle « Plan-Do-Check-Act », illustré dans la Figure 2.3. Chaque phase contient une liste de propositions pour expliquer le sens de l'étape. Un exemple d'une proposition dans la phase de surveillance (monitoring) est de suivre le progrès par rapport l'estimative ; ou une autre proposition, c'est d'engager les parties prenantes internes et externes avec les rapports de soutenabilité.

Selon les créateurs, Sigma n'est pas destiné à des fins de certification, bien qu'ils croient que c'est une importante façon de garantir la crédibilité. Mais ils défendent qu'une approche de soutenabilité doive être flexible et innovante.

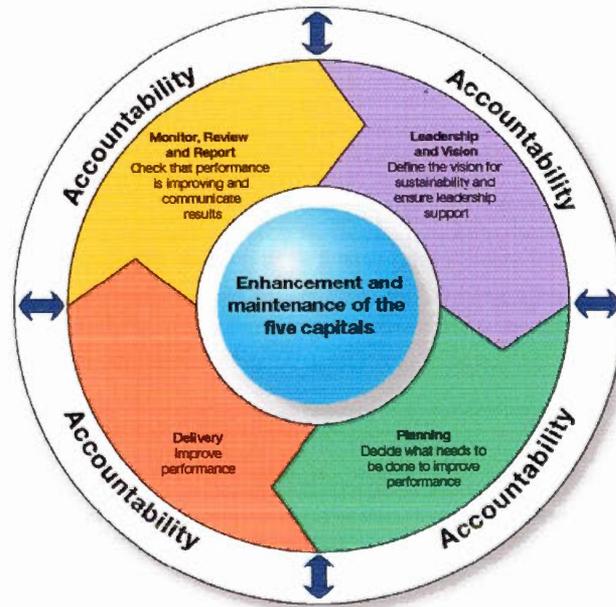


Figure 2.3 – Le cadre de gestion (SIGMA Project, 2006).

2.3.3 Dow Jones Sustainability Indices

Le Dow Jones Indices a été créé en 1999 comme une première référence mondiale en soutenabilité. En utilisant des critères objectifs de gestion des portefeuilles d'investissement, il classifie les meilleures entreprises au niveau de la soutenabilité. C'est comme une reconnaissance des efforts des entreprises qui est devenue une vraie compétition entre les géants. En effet, être parmi les leaders en soutenabilité peut rendre une réputation remarquable.

Selon S&P (2015), les indices sont le résultat de l'union de plusieurs indicateurs pour créer des données uniques à servir de comparaison entre les entreprises. Il y a des classifications ou les familles des indices qui considèrent les types de pays, la région ou le pays, selon le Tableau 2.1. Cependant, chaque division utilise de différents

univers de départ et de différents percentiles pour marquer le point "coupé" dans la sélection des entreprises les plus soutenables. Ainsi, il n'existe pas une agrégation entre les niveaux principaux.

Tableau 2.1
Les familles des indices (S&P, 2015)

Geographical Breakdown	Index
DJSI World	Dow Jones Sustainability World Dow Jones Sustainability World Enlarged Dow Jones Sustainability Emerging Markets
DJSI Regions	Dow Jones Sustainability Asia/Pacific Dow Jones Sustainability Europe Dow Jones Sustainability North America
DJSI Countries	Dow Jones Sustainability Australia Dow Jones Sustainability Canada Select 25 Dow Jones Sustainability Korea Dow Jones Sustainability Korea Capped 30% Dow Jones Sustainability Chile ¹

Le Corporate Sustainability Assessment (CSA) c'est un outil qui sert à mesurer le niveau de soutenabilité, défini par RobecoSAM, un partenaire de S&P Dow Jones Indices. À partir de cet outil, ils sont capables d'estimer le Total Sustainability Score (TSS), responsable pour classifier les entreprises candidates. Si les candidats invités répondent aux questions envoyées, ils seront évalués et selon leurs résultats, classifiés (S&P, 2015).

Pour participer au classement, les entreprises sont évaluées annuellement, selon les critères de chaque famille (Tableau 2.2). Le nombre d'invitations est variable selon le seuil constitutif existant. Le groupe des entreprises sélectionnées est nommé « Invited Universe » (S&P, 2015).

Tableau 2.2

Les critères de participation par familles d'indice (S&P, 2015).

Index	Companies Invited	Existing Constituent Threshold
DJSI World	2500 largest companies in the S&P Global BMI	Above US\$ 500 million
DJSI World Enlarged		
DJSI Emerging Markets		
DJSI Asia/Pacific		
DJSI Europe		
DJSI North America		
DJSI Australia	Entire S&P ASX 200	Above AS 100 million
DJSI Canada Select 25	Canadian companies invited to the DJSI World and DJSI North America	--
DJSI Korea Capped 30%	200 largest South Korean companies in the S&P Global BMI	Above US\$ 100 million

Kutay et Tektüfekci (2015) citent quelques exemples des indices existants dans le DJSI : « City Development Index (CDI), Environmental Sustainability Index (ESI), Environmental Vulnerability Index (EVI), Environmental Policy Index (EPI), Ecological Footprint (EF), Water Footprint (WF), Index of Sustainable Economic Welfare (ISEW), Genuine Progress Indicator (GPI), Environmentally-adjusted Domestic Product (EDP or Green EPP), Genuine Savings (GS), Wellbeing Index (WI), Human Development Index (HDI), Living Planet Index (LPI) ».

Enfin, le DJSI est une méthodologie qui utilise les indices pour générer des rapports avec les critères de soutenabilité pour les industries de plusieurs localisations géographiques ou conditions économiques (Munajat et Kurnia, 2015).

2.3.4 G4 Guidelines

Fondée en 1997, le Global Reporting Initiatives (GRI) est une organisation internationale indépendante qui aide les entreprises, les gouvernements et d'autres organisations à comprendre et à communiquer l'impact de leurs activités sur les questions essentielles de soutenabilité telles que le changement climatique, les droits de l'homme, la corruption et bien d'autres (GRI, 2013). GRI encourage les entreprises de tous les types, grandeurs ou secteurs d'activités à créer des rapports de soutenabilité. Leur mission est de permettre aux décideurs partout, grâce à nos normes de soutenabilité et réseau de multiples parties prenantes, à prendre des mesures vers une économie et un monde plus soutenable (GRI, 2013).

G4 Guidelines est un ensemble de guides qui offrent des principes des rapports, les normes de divulgations et les guides pour mettre en œuvre les rapports de soutenabilité dans les organisations (GRI, 2015a). Les deux principaux guides du G4 sont le 'Reporting Principles and Standard Disclosures' et le guide 'Implementation Manual'. Le premier est générique à tous types d'entreprises et il contient les principes, la norme de divulgation et la définition des termes clés. Le deuxième guide montre comment appliquer les recommandations du premier guide en pratique.

La structure des guides est constituée par des catégories et les aspects, présentés dans la Figure 2.4. Les catégories disponibles sont équivalentes aux dimensions de soutenabilité : économique, environnementale et sociale. Les aspects peuvent être composés par un ou plusieurs indicateurs qui seront utilisés pour mesurer la soutenabilité dans les trois catégories.

TABLE 1: CATEGORIES AND ASPECTS IN THE GUIDELINES				
Category	Economic		Environmental	
Aspects ^{III}	<ul style="list-style-type: none"> • Economic Performance • Market Presence • Indirect Economic Impacts • Procurement Practices 		<ul style="list-style-type: none"> • Materials • Energy • Water • Biodiversity • Emissions • Effluents and Waste • Products and Services • Compliance • Transport • Overall • Supplier Environmental Assessment • Environmental Grievance Mechanisms 	
Category	Social			
Sub-Categories	Labor Practices and Decent Work	Human Rights	Society	Product Responsibility
Aspects ^{III}	<ul style="list-style-type: none"> • Employment • Labor/Management Relations • Occupational Health and Safety • Training and Education • Diversity and Equal Opportunity • Equal Remuneration for Women and Men • Supplier Assessment for Labor Practices • Labor Practices Grievance Mechanisms 	<ul style="list-style-type: none"> • Investment • Non-discrimination • Freedom of Association and Collective Bargaining • Child Labor • Forced or Compulsory Labor • Security Practices • Indigenous Rights • Assessment • Supplier Human Rights Assessment • Human Rights Grievance Mechanisms 	<ul style="list-style-type: none"> • Local Communities • Anti-corruption • Public Policy • Anti-competitive Behavior • Compliance • Supplier Assessment for Impacts on Society • Grievance Mechanisms for Impacts on Society 	<ul style="list-style-type: none"> • Customer Health and Safety • Product and Service Labeling • Marketing Communications • Customer Privacy • Compliance

Figure 2.4 – Les catégories et les aspects du G4 Guidelines (GRI, 2014a).

Alors, les guides G4 contiennent plusieurs combinaisons des catégories, sous-catégories, aspects et indicateurs. Mais, il y a aussi les guides de divulgation spécifiques qui sont utilisés pour aider les entreprises à choisir les indicateurs les plus pertinents à vos domaines (GRI, 2013). Les guides spécifiques disponibles pour les secteurs sont : opérations des aéroports, préparation d'aliments, médias, construction, services publics d'électricité, des mines et des métaux, organisateurs d'événements, pétrole et du gaz, les ONG et services financiers.

Pour mieux comprendre, on prend le guide ‘G4 Sector Disclosure – Food Processing’ qui montre les indicateurs indiqués spécifiquement pour mesurer la soutenabilité d’une entreprise du secteur de préparation d’aliments dans les trois dimensions. Au-delà des indicateurs génériques recommandés par le guide standard, le guide du secteur ajoute d’autres aspects et indicateurs particuliers comme les suivants (GRI, 2015b) :

- Aspect ‘Bien-être Animal’ (catégorie sociale / sous-catégorie société) : Cet aspect aborde les potentiels problèmes de bien-être animal associés à la manipulation, le traitement et la transformation des animaux (terrestres et aquatiques) spécifiquement élevés pour la production de nourriture.
 - Indicateur : Pourcentage et total des animaux soulevés et/ou transformés par espèces et types de races.
- Aspect ‘Performance Économique’ (dimension économique) :
 - Indicateur : Pourcentage de volume acheté qui est vérifié comme en conformément aux normes internationales de production responsable.
- Aspect ‘Matériaux’ (dimension environnementale) :
 - Indicateur : Volume ou poids des matériaux utilisés (le guide spécifique ajoute la recommandation d’informer au moins les matières premières liées aux ressources naturelles).

La préparation des rapports est aussi abordée par la systématique du G4. Un processus itératif pour préparer les rapports de soutenabilité, proposé par GRI (2014a), couvre les étapes suivantes :

- Comprendre les principes de la première partie du guide.
- Choisir entre les options de conformité : de base (juste les éléments essentiels) ou complète.
- Selon l’option de conformité, se préparer pour la divulgation en général.

- Se préparer pour la divulgation spécifique en considérant si le domaine de l'entreprise recommande des aspects additionnels.
- Créer les rapports de soutenabilité basés sur le Web sans oublier de notifier GRI.

Enfin, G4 Guidelines est un ensemble de guides pour orienter et gérer les rapports de soutenabilité à travers d'un processus didactique et complet, applicable à n'importe quel type d'entreprise.

2.3.5 Triple Bottom Line

En 1997, John Elkington a abordé la soutenabilité dans le livre « *Cannibals with Forks: The Triple Bottom Line of Twenty-First Century Business* » sous la forme d'un cadre qui considérait la performance corporative selon les dimensions économique, sociale et environnementale. Ce concept nommé « Triple Bottom Line » (TBL) est aussi connu comme les 3Ps, « personne, planète et profit » (Slaper et Hall, 2011). Les intersections entre les dimensions, selon Samajdar (2014), représentent les niveaux intermédiaires de soutenabilité (Figure 2.5).

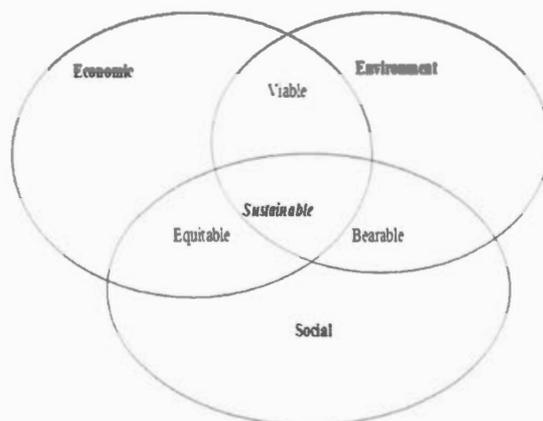


Figure 2.5 - Le Triple Bottom Line (Samajdar, 2014).

Quelques indicateurs du TBL, selon les 3Ps, sont (Slaper et Hall, 2011) :

- Personnes : taux de chômage, taux de participation à la population active féminine, revenu médian des ménages, pauvreté relative, pourcentage de la population avec un diplôme d'études postsecondaires ou d'un certificat, crimes violents par habitant, espérance de vie ajustée sur la santé.
- Planète : concentration en dioxyde de soufre, concentration des oxydes d'azote, polluants prioritaires sélectionnés, nutriments excessifs, consommation d'électricité, consommation de combustibles fossiles, gestion des déchets solides, gestion des déchets dangereux, changement dans l'utilisation des terres / couverture du sol.
- Profit : revenu personnel, coût du sous-emploi, tailles des établissements, croissance de l'emploi, répartition de l'emploi par secteur, pourcentage des entreprises dans chaque secteur, chiffre d'affaires par la contribution du secteur au produit brut de l'État.

Il n'existe pas une méthode universellement reconnue pour calculer le TBL. Mais, selon Slaper et Hall (2011), ça pourrait être considéré comme une flexibilité qui permet d'appliquer le principe dans différents contextes et profils d'entreprise, soit les business, gouvernements ou les entreprises sans but lucratif. Par contre, plusieurs approches basées sur le TBL ont été créées. Ces approches seraient un peu difficiles à les comparer à cause de leurs natures hétérogènes.

Munajat et Kurnia (2015) expliquent que les défis majeurs pour le TBL sont de comment mesurer chaque catégorie, comment trouver des données applicables aux indicateurs proposés et comment calculer la contribution d'un projet ou d'une politique de développement durable.

2.4 Les indicateurs de soutenabilité

Parmi les approches d'application de la soutenabilité dans les entreprises, les standards fournissent des listes de vérification pour comprendre la maturité de l'entreprise par rapport la soutenabilité, tandis que les guides sont conçus pour orienter les entreprises à appliquer la soutenabilité aux entreprises. La plupart des approches proposent l'utilisation des indicateurs catégorisés par dimensions pour mesurer la soutenabilité qui composeront les rapports de soutenabilité au futur.

Selon McCool et Stankey (2004), les indicateurs sont la meilleure possibilité pour révéler des intérêts publics divers, souvent contradictoires. Ils peuvent mesurer différents types de phénomènes en attribuant des valeurs numériques ou d'autres natures. Cependant, Mitchel (2006) alerte que les indicateurs peuvent induire en erreur et désinformer s'ils ne sont pas conçus avec soin et correctement utilisés.

Les nombreux indicateurs existants pour mesurer la soutenabilité n'ont pas été identifiés à l'aide d'une méthodologie explicite (Mitchel, 2006). Cela rend difficile la définition des critères de sélection des indicateurs les plus pertinents selon le domaine de l'entreprise.

McCool et Stankey (2004) disent que le choix doit considérer les contraintes de la connaissance, du coût, du temps et des programmes politiques. C'est un processus collaboratif et de poursuite qui est plus basé sur la valeur que scientifique, plus itératif que linéaire et moins privé que public.

Figge et al. (2002) présentent les concepts de 'lagging' et 'leading', concernant les indicateurs de soutenabilité. Le terme 'lagging' se réfère aux indicateurs qui sont enregistrés après un impact ou, autrement dit, ils sont reliés aux événements passés par rapport les objectifs de l'entreprise. De l'autre côté, les indicateurs du type 'leading' évaluent les activités qui se produisent avant l'impact et qui affectent la performance future des entreprises. Tanzil et Beloff (2006) expliquent que les indicateurs de

'leading' représentent comment les résultats des indicateurs du type 'lagging' devraient être atteints. Ils disent aussi que la gestion efficace de la soutenabilité nécessite de l'utilisation des concepts des indicateurs de 'lagging' et 'leading'.

Alors, les interactions entre les indicateurs de 'lagging' et 'leading' suggèrent les relations de cause-effet entre eux. Selon ces auteurs, la plupart des efforts concernant les indicateurs se concentrent dans les mesures des variables opérationnelles de la soutenabilité, comme les rapports de soutenabilité. Alors, un défi que s'impose est la découverte des indicateurs de gestion des trois dimensions de soutenabilité.

Mais, Lejeune (2017) explique que les approches classiques de soutenabilité sont conçues par des spécialistes en sciences sociales qui ont étudié le passé en utilisant des méthodes rationnelles de déduction et d'induction. Bien que ces approches soient applicables dans un contexte corporatif de base, l'auteur défend l'abduction comme une logique capable de motiver les gens à agir, au lieu de simplement imposer les indicateurs de ces approches. Il suggère une stratégie basée sur le concept de 'game storming' qui peut enrichir les processus d'identification des indicateurs et de la compréhension de la soutenabilité.

Ainsi, face à cette complexité, la sélection des indicateurs est une tâche critique et essentielle pour le succès de l'application de la soutenabilité dans les entreprises.

2.5 Les rapports de soutenabilité

Les entreprises utilisent les rapports de soutenabilité pour divulguer la performance annuelle de leurs indicateurs au grand public : les parties prenantes, les clients, le gouvernement, les activistes et toutes d'autres personnes intéressées. Des approches comme DJSI et G4 Guidelines peuvent être utilisées à ces fins.

Selon Aras et Crowther (2009), les principaux bénéfices que les entreprises cherchent avec la divulgation des rapports de soutenabilité sont :

- Amélioration de l'image de l'entreprise ou d'un produit.
- Les prestations de santé et sécurité.
- Facilité d'attirer l'investissement et le coût réduit de ces investissements.
- Amélioration des relations avec les communautés, les organismes de réglementation et les parties prenantes.
- Amélioration du moral parmi les travailleurs, conduisant à une plus grande productivité, réduction de la rotation du personnel et, par conséquent, réduction des coûts de recrutement et de formation.

Ainsi, les rapports sont considérés par les entreprises comme une façon d'améliorer leurs images devant les investisseurs, clients, employés, organismes et la communauté en général.

Searcy et Buslovich (2014) expliquent dans leur recherche avec 35 entreprises canadiennes que les principales motivations pour créer des rapports de soutenabilité sont les pressions externes provoquées par les exigences légales, les demandes des parties prenantes et les pressions concurrentielles. Les défis soulignés par les auteurs pour élaborer des rapports sont des délais serrés, des limites de ressources et des problèmes avec la collecte de données. Les intégrations entre les rapports de soutenabilité et les rapports financiers annuels des entreprises ont été mentionnées comme de potentielles utilisations des rapports en interne. Seulement 14 % des entreprises répondantes attestaient l'existence de cette intégration.

De façon similaire, Adams et Frost (2008) indiquent qu'il n'y a pas beaucoup de recherches qui abordent les processus de création des rapports de soutenabilité et la mesure dans laquelle les données recueillies sont utilisées dans la prise de décision au sein des entreprises. La relation entre les rapports et la gestion intégrée est influencée

et même limitée par les processus existants dans l'entreprise. Selon les auteurs, la responsabilité envers les parties prenantes peut être compromise quand ils perçoivent que les données de soutenabilité de l'entreprise ne sont pas positives. À ce point, l'accent par les gouvernements sur l'amélioration de la reddition de comptes entraînerait des changements qui pourraient conduire à une amélioration des performances.

Alors, c'est possible de conclure qu'il y a encore une grande difficulté pour intégrer les données de soutenabilité et financières. Ce scénario peut être causé parce que les parties prenantes intéressées à ces rapports de soutenabilité sont, en général, externes à la gestion de l'entreprise.

2.6 Intégration des indicateurs de soutenabilité

Selon Morse et al. (2001), malgré les nombreuses listes d'indicateurs proposées, un problème qui subsiste est comment les intégrer pour savoir si quelque chose est soutenable ou non. Ils disent aussi que dans certaines études, une intégration quantitative a été adoptée où les indicateurs reçoivent des valeurs numériques et sont intégrés mathématiquement pour produire une valeur de soutenabilité.

Mais, à notre avis, il y a beaucoup d'informations cachées derrière les indicateurs de soutenabilité et les relations entre elles peuvent aider les entreprises à découvrir d'autres avantages pour investir dans la soutenabilité. Par exemple, comment utiliser les indicateurs des trois dimensions pour répondre aux questions comme :

- Quels sont les impacts économiques, sociaux et environnementaux de la fabrication d'une pizza ?
- Quel est l'impact de la qualité de l'air sur la santé de mes employés ?
- Comment la productivité de mes employés est influée par l'état de santé de leurs enfants, le cas échéant ?

- Quelle est la relation entre l'évolution des prix des ingrédients de ma pizza et l'échauffement global ?
- Quels sont mes produits les plus soutenables ?

Les rapports de soutenabilité ne sont pas prêts à répondre ces questions directement. Il faut, d'abord, intégrer les indicateurs afin qu'on puisse comprendre leurs relations et convertir les données en informations valables.

2.7 Technologie de l'information pour la soutenabilité

La technologie de l'information (TI) joue un rôle important par rapport la soutenabilité. Seidel et al. (2015) expliquent que la TI peut être utilisée pour automatiser, informer et transformer :

- Automatiser fait référence au remplacement du travail humain par l'informatique. L'automatisation des processus d'affaires est souvent liée à l'augmentation de l'efficacité dans le contexte de la soutenabilité.
- Informer concerne à comment les parties prenantes internes et externes des divers niveaux organisationnels obtiennent des informations pertinentes. La surveillance des impacts par capteurs, l'utilisation des outils collaboratifs d'aide à la décision, les tests de soutenabilité des fournisseurs et la démocratisation de l'information sont des exemples des façons d'informer, liées à la soutenabilité.
- La TI peut transformer la nature des produits, des processus et leurs relations avec les clients ou les fournisseurs. Les systèmes de télétravail et de conférences peuvent transformer les modèles traditionnels de travail et favoriser la soutenabilité pour réduire l'espace physique nécessaire aux activités.

2.8 Les systèmes d'information

Les approches de soutenabilité qui visent à construire des rapports peuvent être mises en œuvre à travers l'adaptation des systèmes d'information existants dans les entreprises. Par exemple, le G4 Guidelines suggère une liste générique des indicateurs à utiliser et où on pourrait trouver les données pour générer les rapports de soutenabilité (GRI, 2015a). Ainsi, les systèmes d'information seraient responsables pour héberger les données liées aux indicateurs de soutenabilité.

En considérant les dimensions économique, sociale et environnementale de la soutenabilité, divers départements et systèmes corporatifs pourraient être affectés : financier, ressources humaines, contrôle de la production, ventes, achats, etc. Quelques grands fournisseurs des 'Progiciels de Gestion Intégré' (en anglais, 'Enterprise Resource Planning' – ERP) ont déjà des modules pour enregistrer les indicateurs et gérer les données et les rapports de soutenabilité. Mais, dans certains cas, les données liées à la soutenabilité pourraient exiger la création des nouveaux systèmes d'information, en particulier quand les systèmes de gestion ne considèrent que la dimension économique dans ces solutions.

Pourtant, il faut aussi considérer que l'ensemble des systèmes d'une entreprise peuvent être de différents fournisseurs, dans différentes plateformes et bases de données. Dans ce contexte, les données pour créer les rapports sont dispersées. Alors, la question qu'on se pose est de comment intégrer les données qui sont stockées dans des systèmes différents et hétérogènes et, parfois, des bases de données distinctes ?

Malhotra et al. (2013) ont mené une analyse de 30 articles, liés aux questions des systèmes d'information et soutenabilité, publiés entre les années 2008 et 2013 dans les principaux journaux de gestion de l'environnement et de systèmes d'information. Les auteurs ont identifié deux courants de recherche : les études liées à la conception, la

mise en œuvre et l'impact des systèmes d'information sur la soutenabilité (Green IS) et les recherches sur les technologies de l'efficacité énergétique et de l'utilisation des équipements (Green IT).

Au niveau de la recherche académique, ils ont constaté que la plupart des publications sur les systèmes d'information pour la soutenabilité sont basées sur la conception et l'analyse, en comparaison aux études de la science du design ou orientées vers l'impact. Bien que les systèmes d'information aient la capacité de créer des organisations plus soutenables, la littérature ne dispose pas de beaucoup d'idées sur la façon de concevoir ces systèmes et comment exécuter la transition vers le développement durable en utilisant des systèmes d'information. Enfin, ils soulignent que la recherche sur les systèmes d'information gagnerait beaucoup en importance si elle pouvait être un agent du changement plutôt qu'un rapporteur passif des pratiques utilisées.

Ahmed et Sundaram (2012) exposent que les approches de soutenabilité ne sont pas prises en charge par des systèmes d'information robustes et qu'elles ne tiennent pas compte des modèles pour les processus d'intégration. Pour cette raison, la prise de décision est en silo. Ils disent encore que les systèmes actuels qui s'occupent de la formulation, la communication, l'exécution et le suivi du concept de soutenabilité négligent l'harmonisation entre la gestion et la construction des rapports. Ils expliquent que les systèmes d'information ne sont pas conçus pour capturer des données sociales et environnementales ni pour la gestion des données de toutes les dimensions de la soutenabilité. Et, comme la pratique actuelle de la haute direction est de développer les stratégies de durabilité (la réduction des transports, la plantation d'arbres, le recyclage), sans les aligner sur la vision et les stratégies organisationnelles, les systèmes d'information ne sont pas adaptés.

Enfin, selon Hilpert et al. (2014), les entreprises qui produisent des rapports de soutenabilité dans un niveau de détail approprié sont confrontées à de grands efforts pour la collecte de données afin d'obtenir une bonne transparence. C'est-à-dire que les

systèmes utilisent des données presque brutes comme lesquelles qui viennent des capteurs. Étant donné que les systèmes d'information pour la soutenabilité ne fournissent pas de solutions pratiques pour ces interdépendances contrastées entre transparence, effort et complexité, les entreprises s'appuient sur des données secondaires (comme les moyennes calculées sur les bases de données) et n'utilisent pas ces types spécifiques de systèmes.

Alors, bien que les systèmes d'information jouent un rôle très important, quelques difficultés sont rencontrées pour gérer et intégrer les données des trois dimensions de soutenabilité. Les approches en intelligence d'affaires sont des solutions faisables pour combler ces lacunes.

2.9 Intelligence d'affaires

Une technique capable de résoudre le problème d'intégration des données hétérogènes est le Business intelligence (BI) ou, en français, l'intelligence d'affaires (IA). La littérature qui aborde l'intégration et l'analyse des indicateurs de soutenabilité en utilisant un BI est limitée. Petrini et Pozzebon (2009) considèrent un BI comme un ensemble de méthodes et d'outils qui disposent d'un grand potentiel encore inexploré et qui peut soutenir les pratiques et la gestion de la soutenabilité. Un système de BI a une importance évidente comme un canal de communication et de diffusion de l'information.

Haupt et al. (2015) affirment que les silos d'information sont les plus grands défis pour la gestion de l'information pour la prise de décision sur le développement soutenable. Ainsi, un outil de gestion intégrée de l'information, comme une approche BI, pourrait surmonter le problème.

Généralement, les gestionnaires utilisent les solutions BI pour les aider pendant la prise de décision grâce à sa nature intégrative et à la capacité de gérer et analyser des données volumineuses. Un BI a la capacité de permettre aux gestionnaires de visualiser les données dans une différente perspective, naviguer à de divers niveaux d'agrégation, lancer des requêtes à un des ensembles de données afin de découvrir de nouveaux facteurs qui influent sur des processus d'affaires et aussi pour anticiper et prévoir les changements à l'intérieur et à l'extérieur de l'organisation (Bara et al., 2009).

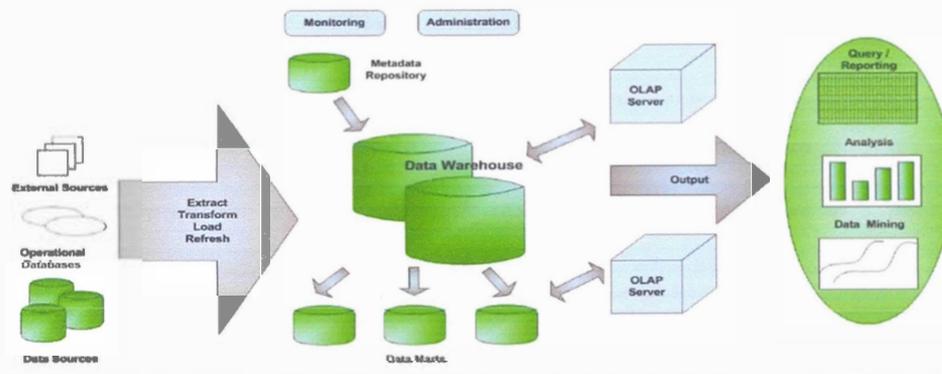


Figure 2.6 – Architecture d'un Système BI (Sacu et Spruit, 2010).

Selon l'architecture d'une solution BI proposée par Sacu et Spruit (2010), un BI intègre les données de différentes sources (bases de données des Systèmes d'Information, fichiers textes, tableurs, données de l'Internet, etc.) dans une base commune nommée Data Warehouse (DW) ou Entrepôt de Données. Le BI utilise une base de données multidimensionnelle pour intégrer les données de plusieurs bases relationnelles (Bucher et al., 2009). Cette modélisation de données est adaptée aux bases de données volumineuses et bien optimisées pour les requêtes (Yao et al., 2013, p. 3). Après avoir intégré les données au niveau physique, les outils BI permettent d'extraire des indicateurs stratégiques et d'analyser les données avec une gamme de fonctionnalités : des requêtes ad hoc, les rapports, les analyses dynamiques, les tableaux de bord, l'exploration des données, etc.

Les données qui alimentent l'entrepôt de données peuvent venir des systèmes d'information, des tableurs, des fichiers textes, médias numériques, etc. Le processus de transfert de données entre les sources et la base de l'entrepôt de données est nommé le Processus ETL : Extraction, Transformation et Charge (« Extraction, Transformation and Load »). Ce processus est un facteur critique pour le succès d'un BI. Le type de donnée le plus courant dans un processus ETL est la donnée des bases de données relationnelles. Alors, un BI joue un rôle client par rapport les données provenant des SI : les SI donnent les données transactionnelles et le BI les transforment en faits (les valeurs numériques) et dimensions (les points de vue pour agréger ou segmenter les faits).

En analysant la littérature sur le sujet BI, selon Petrini et Pozzebon (2009), les approches méthodologiques concernant la planification de l'information ne sont pas assez nombreuses et les méthodologies sont en général proposées par les fournisseurs des solutions et leurs consultants. Ces méthodes ne sont pas souvent validées par la communauté académique. En considérant l'ajout des dimensions de la soutenabilité dans une solution BI traditionnelle, il sera créé un nouveau scénario, peu exploré dans la littérature et dans la pratique.

L'utilisation des approches BI dans le contexte de la soutenabilité pose un autre défi, qui est relié aux investissements nécessaires pour ces types de solutions. Haupt et al. (2015) expliquent que les coûts pour la mise en œuvre des solutions BI impliquent les droits de licences d'utilisation des logiciels qui peuvent être très chers, les professionnels qualifiés et l'infrastructure technologique qui soutient les solutions. Dans ce cas, la justification pour investir des ressources dans un projet qui ne se traduit pas en termes financiers peut devenir un grand défi. En bref, les difficultés d'appliquer les approches BI et la soutenabilité ne sont pas vraiment techniques, mais méthodologiques et conceptuelles, comme soulignent Petrini et Pozzebon (2009).

2.10 Les types d'information

Les types d'information présentés par les systèmes d'information classiques (transactionnels) et les systèmes du type BI sont différents. Bara et al. (2009) présentent une comparaison entre les informations disponibles dans les rapports des systèmes d'informations et des solutions BI :

Tableau 2.3
Comparaison des types d'information (Bara et al., 2009)

Caractéristiques	Rapports des SI	Rapports BI
Objective	Analyse d'indicateurs qui mesure les activités courantes et internes ou les rapports quotidiens.	Optimisation des processus, analyser les principaux indicateurs de performance, prévoir les données internes et externes, mise au point interne et externe.
Niveau de décision	Opérationnel / moyen	Stratégique / haut
Utilisateur impliqué	Niveau opérationnel de la gestion	Cadres supérieurs, le niveau stratégique de la gestion
Data management	Bases de données relationnelles	Entrepôt de Données, OLAP, Data Mining
Opération typique	Rapport/analyse	Analyse
Volume des records / transaction	Limité	Énorme
Orientation des données	Record	Cube
Degré de détail	Détaillé, consolidé, pré agrégat	Agrégat
Âge des données	Actuel	Historique/actuel/prospective

De la même façon, Bara et al. (2009) indiquent les différences entre le développement des systèmes d'informations et les solutions BI :

- Orientation vers des occasions d'affaires plutôt que les besoins transactionnels.
- La mise en œuvre des décisions stratégiques, non seulement les décisions départementales ou opérationnelles.
- Analyse fondée sur les besoins de l'entreprise, ce qui est le plus important du processus.

- Processus de développement cyclique, axé sur l'évaluation et l'amélioration des versions successives, non seulement la construction et la livraison d'une grande singulière version finale.

C'est possible de percevoir que la nature et les fins des systèmes d'information et des systèmes BI sont bien différentes et spécifiques. Mais, ils sont complémentaires et, en général, interdépendants.

2.11 Les risques de l'intelligence d'affaires

Ces caractéristiques et fonctionnalités rendent au BI une puissante capacité d'analyse de données. Mais, l'effort pour développer une solution BI peut être énorme, selon l'envergure, la qualité et la complexité des données des systèmes sources, au-delà des règles d'affaires pour intégrer les données dans la base commune.

Si les systèmes sources ne sont pas encore déployés, il y a un risque qu'une mauvaise identification de réels besoins d'affaires de l'entreprise puisse provoquer des changements de ces exigences et successives révisions (Bara et al., 2009).

Bara et al. (2009) soulignent aussi un autre problème dans développement des solutions en BI : identifier et répondre aux exigences de plusieurs rôles qui sont impliqués dans le développement BI est compliqué et peut influencer le succès des solutions. Alors, si la difficulté de comprendre les exigences des rôles d'un BI existe déjà, la complexité peut augmenter énormément au moment qu'on ajoute les rôles liés à la soutenabilité.

Parmi les étapes du développement des solutions BI, l'identification des exigences est la plus importante pour le succès des projets. Il existe peu méthodologies spécifiques pour l'identification des exigences des parties prenantes et d'autres rôles impliqués et les approches proposées dans la littérature sont basées sur la collecte des exigences du développement des systèmes d'information.

Enfin, bien que les approches de BI soient une possible solution pour intégrer les indicateurs de soutenabilité, il faut aussi régler les problèmes du développement pour réduire les risques d'investir dans solutions de ce type.

2.12 Synthèse

La soutenabilité, parmi des concepts ambigus, peut être traduite comme le développement dans les dimensions économique, sociale et environnementale qui ne compromet pas l'accès aux mêmes ressources par les générations futures. D'une manière similaire, le développement soutenable est présent au sein des entreprises pour réduire les impacts de leurs activités sur les trois dimensions.

Les approches de développement de la soutenabilité sont sous la forme de standard ou comme des guides pour orienter l'application de la soutenabilité. Sigma Guidelines sera utile dans notre recherche pour orienter le processus de sélection des indicateurs selon le domaine de l'entreprise en utilisant ses principes. Il propose aux gestionnaires un changement de la culture organisationnelle pour penser à la soutenabilité. Sigma défend que les capitaux ne doivent pas être traités séparément et ils ont toujours des relations entre eux. Les approches Sigma Guidelines et le Triple Bottom Line ont des équivalences au niveau des dimensions et peuvent être intégrés.

ISO 14001 c'est un standard que vise à évaluer la conformité des pratiques environnementales d'une entreprise. Cette approche est intéressante pour comprendre les exigences réglementaires, mais selon Curkovic et Sroufe (2011), elle n'est pas connectée assez directement aux normes de performance et elle présente des risques au niveau des difficultés d'application. Comme cette recherche vise faire la mesure et l'analyse de la soutenabilité, cette approche n'offre pas les éléments nécessaires pour être considérée.

Le DJSI, Dow Jones Sustainability Index, c'est une approche de classement des entreprises selon leurs efforts concernant la soutenabilité. Cependant, le processus de sélection pour participer des évaluations est basé sur les critères économiques et focalisé dans les grandes entreprises du marché mondial. Cette limitation pourrait réduire la portée de notre recherche à cause du manque de flexibilité. La comparaison entre les membres de différentes familles n'est pas directe à cause de différents critères adoptés par famille.

Le G4 Guidelines, comme le DJSI, vise à faciliter la création des rapports de soutenabilité. Mais, le DJSI est plus axé aux grandes entreprises. Les guides du G4 sont faciles à comprendre et aident le choix des indicateurs pertinents au domaine de l'entreprise.

Le Triple Bottom Line explicite les trois dimensions et leurs indicateurs. Ainsi, est importante pour notre recherche. Mais chaque approche sera utilisée pour créer un mécanisme efficace d'application de la soutenabilité. Mais, la sélection des indicateurs à utiliser selon un domaine reste comme un point d'incertitude dans les approches.

Les rapports de soutenabilité divulguent les indicateurs pour mesurer la performance dans les trois dimensions et ils émergent comme façon d'améliorer la réputation de l'entreprise en face aux clients, investisseurs et parties prenantes. Les relations entre les indicateurs peuvent révéler des connaissances stratégiques, mais l'intégration entre les indicateurs n'est pas abordée par les approches.

L'application de la soutenabilité est recommandée par l'adaptation des systèmes d'information tandis que pour résoudre le problème d'intégration, les solutions d'intelligence d'affaires sont bien indiquées. C'est une technologie adaptée aux données volumineuses et offre des ressources avancées pour analyser les données. Mais, il apporte aussi des risques qui peuvent provoquer l'échec des projets.

Dans ce cadre, on propose une méthode capable de sélectionner et intégrer les indicateurs économiques, sociaux et environnementaux en utilisant des approches d'intelligence d'affaires pour appliquer la soutenabilité aux entreprises et, ensuite, identifier les connaissances stratégiques cachées dans les indicateurs.

2.13 Contribution de la recherche

Les contributions de ce travail sont identifiées du côté de la soutenabilité et, aussi, pour le développement de solutions d'intelligence d'affaires.

Bien qu'il existe plusieurs approches pour générer des rapports de soutenabilité au sein des entreprises, cette recherche vise utiliser l'information de ces rapports pour la prise de décisions. Une contribution théorique serait la méthode d'application basée sur les approches d'intelligence d'affaires qui peut être adaptée à plusieurs contextes. Spécifiquement, la preuve de concept appliquée au milieu de l'éducation universitaire peut être vue comme une contribution pour planter une semence envers la soutenabilité. Une contribution pratique pour la soutenabilité est l'extraction des connaissances stratégiques en considérant des indicateurs des dimensions économique, sociale et environnementale qui peut être intégrée à la gestion des entreprises.

Au niveau de l'intelligence d'affaires, une contribution théorique est le fait de mettre en évidence la capacité de l'intelligence d'affaires de promouvoir l'intégration entre les trois dimensions de soutenabilité. D'un point de vue pratique, la preuve de concept peut servir de modèle à générer les systèmes d'informations plus performants pour gérer les questions soutenables.

CHAPITRE III

CADRE MÉTHODOLOGIQUE

Cette section du document vise à décrire la méthodologie qui sera appliquée dans la recherche. La première partie fait brièvement une révision de l'objectif principal et de l'approche pour l'atteindre. Ensuite, la Science du Design (en anglais, « Design Science Research ») est décrite et adaptée au contexte de cette recherche. À la fin, une preuve de concept avec deux scénarios pratiques est présentée pour démontrer la faisabilité de la méthode proposée.

3.1 Révision des objectifs

L'objectif de notre recherche est de proposer une méthode d'application de la soutenabilité dans les entreprises, en utilisant les approches d'intelligence d'affaires, permettant de sélectionner et d'intégrer les indicateurs économiques, sociaux et environnementaux pertinents à une entreprise pour générer et utiliser des connaissances stratégiques.

La méthode proposée est composée par trois modules :

- Module 1 – choisir les approches et les indicateurs de soutenabilité les plus pertinents selon le domaine de l'entreprise.
- Module 2 – définir les étapes d'adaptation des systèmes d'information ; définir la méthodologie d'intégration ; appliquer l'intégration.

- Module 3 – identifier les relations entre les indicateurs et découvrir la connaissance des trois dimensions de la soutenabilité.

Alors, la méthode va fonctionner comme une stratégie pour comprendre les relations entre les variables économiques, sociales et environnementales. Ainsi, d'autres utilisations de ces informations pourront faire émerger de nouvelles découvertes qui influent sur le domaine de l'entreprise.

3.2 Type de recherche

Le type de la recherche choisi est la science du design (SD). Souvent utilisée dans le domaine des systèmes d'information, la Science du Design est normalement composée par six étapes, selon Peffers et al. (2007) qui seront décrites dans les prochaines sections, aussi bien que les explications de comment cette recherche remplira les exigences par rapport la SD.

3.2.1 Étape 1 : identification du problème et de la motivation

Identifier le problème est le premier pas d'une recherche et il est un facteur de motivation pour trouver une solution à la problématique par les chercheurs et lecteurs. Dans cette recherche, cette partie est présente dans le Chapitre 1, qui aborde la problématique. En bref, les problèmes identifiés sont les suivants :

- Sélectionner l'approche et les indicateurs de soutenabilité les plus pertinents selon le domaine de l'entreprise sont des tâches difficiles à gérer à cause de la variété d'options disponibles.
- Les recommandations standards des rapports de soutenabilité ne sont pas suffisantes pour expliquer l'intégration entre les indicateurs des dimensions

économiques, sociales et environnementales et pour interpréter les possibles relations cachées derrière ces indicateurs.

- L'analyse de la soutenabilité par les approches d'intelligence d'affaires est encore peu connue à la fois dans les affaires et dans la littérature.

C'est important de souligner encore que la pertinence de la soutenabilité et l'inquiétude des pays les plus grands au monde par rapport les questions d'environnement sont des facteurs qui mettent en évidence la pertinence de ce sujet au niveau mondial.

3.2.2 Étape 2 : définition des objectifs d'une solution

Les objectifs bien définis contribuent à la compréhension des possibles solutions et à analyser si elles sont faisables et efficaces. Le but principal de cette recherche est de proposer une méthode d'application de la soutenabilité au sein des entreprises en utilisant les approches d'intelligence d'affaires. À partir des objectifs, la revue de la littérature montre des concepts qui sont liés au sujet de l'intelligence d'affaires et soutenabilité, aussi bien que l'essai de recueillir les tentatives de répondre aux problèmes similaires qui se posent.

3.2.3 Étape 3 : conception et développement

La conception vise à créer un ensemble d'artefacts qui sera utilisé pour résoudre les problèmes soulignés dans la première étape et pour atteindre les objectifs. Ces artefacts peuvent être des spécifications, des systèmes d'information ou, comme dans ce cas, des méthodes et spécifications. C'est important de remarquer que l'alignement entre les objectifs et les artefacts est un facteur de succès d'une recherche basée sur la stratégie de la SD.

Ainsi, selon cette recherche, l'artefact central proposé sera la méthode dont la contribution sera l'application de la soutenabilité dans les entreprises. Des diagrammes qui décrivent les étapes des méthodes proposées feront partie de cet artefact. Le Chapitre 4 décrira en profondeur cette méthode.

3.2.4 Étape 4 : démonstration

La démonstration sert à valider l'utilisation de l'artefact créé pour assurer sa faisabilité et pertinence face aux objectifs. Une recherche basée sur la Science du Design peut utiliser de différents types de techniques pour cette étape : « expérimentation, simulation, étude de cas, la preuve de concept, ou une autre activité appréciée » (Peffer et al., 2007). Donc, cette recherche utilisera une preuve de concept pour appliquer la méthode proposée.

En utilisant une preuve de concept de deux scénarios hypothétiques avec des données réelles, approximatives et fictives, cette recherche suivra les étapes de la méthode pour démontrer comment les appliquer et expérimenter de possibles points d'amélioration. Cette preuve de concept mérite une attention spéciale parce que c'est une activité qui vise à expérimenter la méthode proposée et à démontrer la capacité des solutions d'intelligence d'affaires à trouver des connaissances liées à la soutenabilité. Le Chapitre 5 expliquera en détail la preuve de concept créée dans cette recherche.

3.2.5 Étape 5 : évaluation

L'objectif principal de la méthode est d'appliquer les trois dimensions de soutenabilité dans les prises de décisions des entreprises en utilisant l'intelligence d'affaires. L'évaluation sera faite à travers des observations sur l'application de la méthode en deux scénarios fictifs dans la preuve de concept. L'idée est d'attester que la méthode

serait capable d'ajouter les dimensions sociale et environnementale à la dimension financière pour permettre l'utilisation de la soutenabilité. L'intégration des indicateurs des trois dimensions est un deuxième défi à accomplir. À la fin, la possibilité de la découverte de la connaissance serait une tâche plus complexe et audacieuse.

Les approches BI sont déjà bien connues et leurs capacités ont été démontrées par diverses publications académiques. De l'autre côté, les études concernant la soutenabilité ont révélé les approches formelles et matures pour la création des rapports. Alors, la difficulté de trouver des experts dans ces deux domaines en même temps a été un obstacle. Pour surmonter ce problème, la preuve de concept a compris la mise en œuvre effective de la solution avec des données cohérentes et très proches de la réalité. La rigueur et les détails sur la conception et la mise en œuvre ont été priorités pour assurer la reproductibilité de cette recherche par d'autres chercheurs. Des recherches futures pourront avancer sur ce travail pour trouver une manière plus efficace d'évaluation. Le Chapitre 6 abordera les discussions et la conclusion.

3.2.6 Étape 6 : communication

La communication est une étape que, selon Peffers et al (2007), sert à divulguer les informations par rapport l'application de la SD à la recherche. Dépendamment de quel type de publication sera utilisé, un format de structure est recommandé. Dans ce cas, le format d'un mémoire a été utilisé.

CHAPITRE IV

ÉLABORATION DE LA SOLUTION

Cette section du document vise à proposer une solution pour la problématique définie dans cette recherche. La première partie explique brièvement la liaison entre la problématique et la méthode qui sera proposée comme une solution pour appliquer la soutenabilité dans les entreprises. Ensuite, les trois modules et leurs étapes seront exposés, aboutissant à une section qui discute des résultats attendus après l'application de la méthode.

4.1 Lien avec la problématique

La difficulté des entreprises de considérer la soutenabilité selon les dimensions économique, sociale et environnementale dans les prises de décisions stratégiques cerne la problématique de cette recherche. Bien que les approches célèbres de soutenabilité fournissent plusieurs types d'indicateurs, la sélection de ceux qui sont le plus pertinents au profil de l'entreprise est encore un défi. La nature intégrative de ce scénario suggère l'utilisation des stratégies comme l'intelligence d'affaires. Mais, son application pour favoriser la soutenabilité est encore peu explorée dans la littérature.

Face à ce contexte, la proposition de la solution par ce document se présente sous la forme d'une méthode d'application de la soutenabilité au sein des entreprises, profitant de l'utilisation de l'intelligence d'affaires pour résoudre cette problématique.

4.2 La structure de la méthode

Cette recherche aborde des thèmes complexes dans deux différents domaines : la soutenabilité et l'intelligence d'affaires. La Revue de la Littérature montre que chaque thème est bien développé séparément, mais des applications pratiques de l'intelligence d'affaires pour gérer la soutenabilité sont encore peu utilisées.

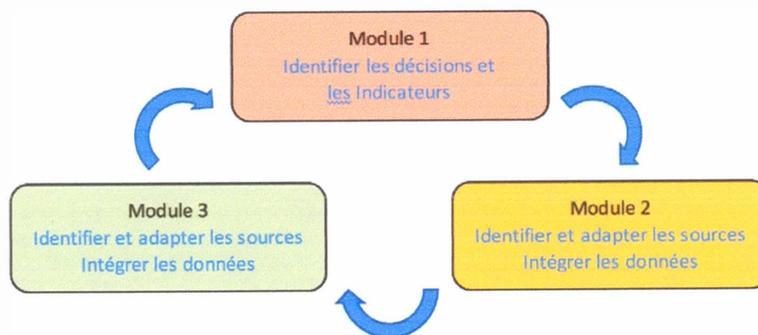


Figure 4.1 – Les Modules de la Méthode.

Ainsi, pour rapprocher ces deux concepts sans augmenter la complexité de compréhension, cette recherche distribue les étapes de la méthode dans trois modules majeurs : la sélection des indicateurs les plus pertinents selon le domaine de l'entreprise ; l'adaptation et intégration des indicateurs ; et la connaissance intégrée des trois dimensions de soutenabilité. La Figure 4.1 présente les modules qui seront abordés dans les prochaines sections.

4.3 Module 1 : sélection des indicateurs

Le premier module couvre les activités préliminaires pour sélectionner les indicateurs afin de générer de la valeur dans le sens de la création de connaissance. Après avoir

défini les décisions à prendre, les indicateurs pertinents pour l'entreprise seront sélectionnés parmi diverses options existantes dans les approches classiques de soutenabilité. À la fin, selon la portée du projet, une stratégie d'exécution sera aussi proposée pour gérer la complexité du projet. Chaque étape de ce module sera détaillée dans les prochaines sections.

4.3.1 Définir les décisions à prendre

En analysant les approches classiques de soutenabilité, comme G4 Guidelines, c'est possible d'identifier des ensembles d'indicateurs et des instructions sur la façon de les appliquer. Toutefois, ces approches sont utilisées principalement pour générer les rapports de soutenabilité. Mais, la méthode proposée dans ce document vise à extraire des connaissances à partir des trois dimensions pour supporter les processus décisionnels.

Alors, avant de choisir les indicateurs parmi ceux disponibles dans les approches de soutenabilité, il faut comprendre quelles décisions seront prises par les gestionnaires et soutenues par ces indicateurs. Cela signifie qu'au lieu de fournir un ensemble d'indicateurs qui pourront être utiles aux processus décisionnels, les décisions seront modélisées et, ensuite, les indicateurs seront élaborés selon leurs besoins.

Un avantage de cette approche est que tous les efforts seront liés à un processus décisionnel de l'entreprise. Si ce processus change, même avant que les indicateurs soient disponibles, les impacts peuvent être estimés et le projet sera réadapté plus tôt. Un autre point est que moins d'effort sera appliqué pour mettre en œuvre une solution, grâce à l'univers d'indicateurs plus limité et objectif.

Le format de la modélisation des processus décisionnels peut varier selon le domaine de l'entreprise et la technique préférée, mais l'essentiel est de considérer le processus

décisionnel comme le point de départ du projet d'intégration. C'est important de remarquer que si plusieurs processus sont identifiés, c'est possible de créer un ordre de priorité, classée par l'importance des décisions selon la stratégie de l'entreprise.

4.3.2 Choisir les indicateurs

Chaque processus décisionnel représente un point de départ pour la découverte des indicateurs qui aideront les gestionnaires dans la prise de décisions. Après avoir défini ces décisions, il faut identifier les métriques qui permettent d'effectuer la mesure des éléments à analyser. Par exemple, si une décision considère le volume de ventes, la valeur d'une facture est un indicateur qui permettra de classer les factures et, ainsi, analyser leurs comportements selon de différentes perspectives.

Selon les approches de soutenabilité, plusieurs indicateurs sont distribués dans les dimensions économique, sociale et environnementale. Les approches peuvent interpréter différemment deux indicateurs qui sont apparemment similaires. Dans ce cas, il faut comprendre les rôles de chaque type d'indicateur pour éviter les erreurs d'interprétation.

Le choix des indicateurs est essentiel pour le succès d'un projet qui intègre des dimensions de soutenabilité. Si un indicateur choisi ne présente pas des liaisons avec les autres, il sera difficile de trouver une façon d'évaluer les dimensions ensemble. Les liaisons sont identifiées à travers des perspectives communes aux concepts.

4.3.3 Choisir les approches de soutenabilité

La littérature présente des approches de soutenabilité célèbres, créées par de grands organismes, qui rassemblent des indicateurs des trois dimensions. Ces approches sont souvent utilisées pour créer des rapports de soutenabilité par les grandes entreprises.

Les rapports, en général, sont basés sur une seule approche pour assurer que les informations sont cohérentes et complètes.

Cependant, l'objectif de cette recherche n'est pas de générer des rapports de soutenabilité, mais d'aider aux entreprises à ajouter d'autres dimensions à l'analyse économique traditionnelle. Ainsi, le but est de choisir les indicateurs de soutenabilité les plus pertinents selon les décisions à prendre. De cette façon, la méthode permet d'utiliser des petites parties d'une ou, même, de plusieurs approches de soutenabilité. Cela veut dire que le mélange des approches ne compromet pas le résultat final, mais il faut faire attention aux significations des concepts pour ne pas confondre leurs sémantiques. Donc, les approches fonctionnent comme des sources des indicateurs indépendants pour la méthode.

En résumé, la méthode est flexible par rapport le choix de l'approche de soutenabilité qui sera utilisée. Il est possible de mélanger de différentes approches et même les changer après la conception. Alors, la rigueur de choix n'est pas aussi critique comme dans le cas des rapports intégrés de soutenabilité.

4.3.4 Créer une stratégie d'exécution

L'application des dimensions de soutenabilité peut mener à une augmentation importante du nombre des variables qui seront prises en compte dans les processus décisionnels. Dans ces cas, la complexité de la mise en œuvre peut devenir un risque pour le succès du projet. Ainsi, c'est important de réfléchir à une stratégie d'exécution qui permet d'orienter le projet d'intégration des dimensions de soutenabilité vers les résultats et pour contrôler les risques impliqués.

Au lieu de créer un grand projet avec de nombreux indicateurs, cette méthode recommande une stratégie avec un petit ensemble d'indicateurs qui seront responsables

pour bâtir le lien entre la dimension économique classique et les dimensions sociale et environnementale. En utilisant des cycles itératifs, les indicateurs seront ajoutés à chaque nouvelle itération. Cette action propose un graduel changement culturel de l'entreprise pour considérer d'autres implications de ses activités dans la société ou dans la nature.

Une gestion agile de projets est bien adaptée à cette situation pour sa capacité de rapidement présenter des résultats qui seront, ensuite, évolués constamment jusqu'au niveau de précision attendue par les prises de décisions. Mais, ce type de projet agile nécessite aussi que les mises en œuvre soient autant flexibles et fortement changeables. Par exemple, les routines de transport de données devraient être automatisées pour garantir une bonne évolution d'un projet. Alors, même si les cycles initiaux sont simples, le projet doit être préparé toujours pour les prochaines évolutions.

Cette stratégie d'exécution n'est pas rigide et elle recevra de plusieurs adaptations. À chaque cycle, ce plan pourra changer pour envisager de nouveaux besoins ou pour modifier ceux déjà existants. Alors, c'est important de réviser les buts et les cycles en fonctions de ces modifications.

4.4 Module 2 : adaptation et intégration

Le Module 1 a donné des recommandations pour choisir des approches de soutenabilité et des indicateurs sélectionnés selon les décisions d'une entreprise. Le Module 2 s'occupe de définir comment les données seront collectées et gérées : identifier les sources des données liées aux indicateurs, faire les adaptations dans les systèmes d'information pour stocker les nouvelles données, appliquer l'intégration entre eux et définir les démarches pour les transporter et gérer les possibles inconsistances.

4.4.1 Identifier les sources des données

Les données qu'une entreprise utilise pour prendre des décisions ne sont pas toujours créées par elle-même. Ces données peuvent avoir une source d'origine externe comme un organisme, une autre entreprise, le gouvernement, une université, etc. Les sources sont responsables pour la gestion des données ou pour seulement les maintenir en faisant des répliquions à partir d'une source originale. Par exemple, Statistique Canada est un organisme du Gouvernement du Canada qui est responsable pour concentrer des diverses statistiques par rapport plusieurs types de sujet et de différentes sources.

Alors, les données peuvent venir des sources internes ou externes à l'entreprise. Par exemple, un indicateur qui mesure la consommation d'énergie pendant la fabrication d'un produit, par exemple, est lié à la production et, alors, est considéré comme interne à l'entreprise. D'autre part, l'information des impacts globaux de la production d'énergie sur l'environnement est générée par des organismes compétents, qui sont à extérieur de l'entreprise. Dans le cas de la soutenabilité, plusieurs organismes de recherche offrent des renseignements des dimensions économiques, sociales et sur l'environnement.

Les sources internes sont logiquement plus faciles à accéder et à gérer parce que les processus de création et de manutention sont bien connus et contrôlés par l'entreprise elle-même. Si bien planifiées, les données ne souffrent pas des problèmes de continuité parce que le fournisseur interne, en général, sait déjà qui utilise les données et les changements sont plus prudents. Mais, c'est important de remarquer que les sources internes sont plus spécifiques et, peut-être, elles n'ont pas une portée plus large. Au niveau des coûts, la manutention et les mises à jour des données de sources internes sont plus coûteuses pour l'entreprise à cause de l'infrastructure et des processus d'affaires impliqués.

Les données de sources externes sont notamment moins contrôlées par les entreprises que les reçoivent, surtout s'il n'y a pas un accord d'approvisionnement pour assurer la

disponibilité et la qualité des données. Quelques problèmes d'interprétation des concepts ou des données peuvent créer un élément risqué. Par contre, les données externes sont maintenues par d'autres entreprises. Cela veut dire que les coûts liés à cette étape seront des sources externes. Mais, l'information n'est pas toujours libre et elles sont vendues selon leurs qualités et pertinences.

La gestion des changements des données par les sources est un thème très important. Si un ensemble de données est mis à jour, il faut utiliser des mécanismes pour identifier les changements et, le cas échéant, les enregistrer pour créer des historiques. La fréquence des mises à jour doit être sur la même base temporelle que ce que les décisions exigent. Par exemple, si un gestionnaire prend des décisions chaque mois, mais les données sont mises à jour chaque trimestre, peut-être que la décision sera moins précise à cause de la différence de période.

Trouver de bonnes sources de données est une activité cruciale pour le succès des décisions. Dans certains cas, les sources peuvent limiter l'exactitude des informations qui supportent les prises de décisions. Par exemple, si une source externe fournit des renseignements sur les émissions de gaz de la ville de Montréal et un utilisateur a besoin de cette information par arrondissement, possiblement, il devrait faire des approximations avec les techniques statistiques.

Soit avec les sources internes ou externes, le plus important est de comprendre que la qualité des données est essentielle pour prendre de bonnes décisions. La fiabilité et la rigueur par rapport la création et manutention sont des points à considérer aussi sur la qualité des sources de données. Si une source est peu fiable, les risques des décisions basées sur leurs données sont plus critiques. Des façons d'atténuer ces risques sont : trouver des sources comme des organismes officiels (universités, gouvernement, etc.), comprendre des méthodes de calcul et avoir l'accès aux étapes pour être capable de reproduire les résultats pour les vérifier de temps en temps.

Il faut régulièrement tester l'accès à la source pour assurer la disponibilité des données. Ces tests peuvent empêcher que les données inconsistantes soient considérées dans les décisions. Les stratégies de tests varient selon le type de données à travailler et la complexité d'interprétation d'eux.

4.4.2 Adapter les systèmes d'information

Après avoir choisi les indicateurs pertinents selon les décisions à prendre et avoir identifié leurs sources, souvent, quelques données peuvent ne pas exister ni à l'extérieur ni à l'intérieur de l'entreprise. Parfois, même les PGI – 'Progiciels de Gestion Intégrée' (en anglais, ERP – Enterprise Resource Planning) – ne couvrent pas tous les besoins des entreprises et, alors, d'autres systèmes d'information sont développés ou achetés. Autrement dit, ils sont de nouvelles données et il faut découvrir une façon de les gérer et de fournir des mécanismes d'accès à ceux qui ont besoin d'y accéder. Par exemple, comment serait possible de collecter et enregistrer les valeurs d'un indicateur qui mesure la quantité de déchets envoyés au recyclage ?

La nouvelle donnée peut être enregistrée dans plusieurs formats : dans un fichier texte structuré (XML, CSV, JSON, etc.), dans une feuille de données, dans une base de données (relationnelle ou non relationnelles), etc. Le choix de stockage dépend de la nature des données, du nombre de personnes qui vont les manipuler, la fréquence de modifications, la complexité des associations entre les données, les canaux et limites d'accès aux données. Quel que soit le format, il y a des logiciels qui s'occupent de gérer les données. Un fichier texte ou une feuille de données peut être manipulé par des logiciels de plusieurs fournisseurs et types. Encore, l'enregistrement dans les bases de données est souvent fait par les systèmes d'information.

Les systèmes d'information aident les entreprises à gérer les données liées à ses processus d'affaires. Parmi les options possibles, ils sont une bonne stratégie pour

manipuler l'information soit dans les bases de données soit dans toutes les autres options de format. Ils sont des logiciels qui ont l'objectif d'administrer les données liées à l'exécution des certaines tâches du monde réel en utilisant l'informatique. Ils sont des ensembles de programmes qui servent à traiter de l'information, selon quelques opérations basiques : créer, mettre à jour, lire, supprimer et, aussi, une caractéristique importante de créer des associations entre les données.

Parmi les types de systèmes d'informations, ceux du type transactionnel sont liés aux processus opérationnels d'une entreprise. Ils sont basés sur le concept d'une transaction qui signifie l'enregistrement d'une opération de l'entreprise, comme la vente d'un produit, l'admission d'un employé, le paiement d'une facture, etc. Ainsi, ils sont conçus pour résoudre des problèmes spécifiques, par exemple, contrôler la production des produits, pour aider à la gestion des ressources humaines, surveiller la production des déchets, etc.

Probablement quelques ajustements doivent être faits dans la structure des systèmes d'information déjà existants dans les entreprises pour qu'ils puissent gérer aussi les nouvelles données. Les adaptations seraient au niveau des bases de données et aussi dans les interfaces qui gèrent ces données. Ces ajustements ne sont pas toujours faciles à faire et dans certains cas les licences d'utilisation des logiciels ne permettent pas de les effectuer. Dans le dernier cas, la création d'un autre système d'information figure comme une possible solution.

Pour clarifier avec un exemple, supposez qu'une entreprise du secteur d'alimentation veuille répondre à la question sur la quantité de ses déchets envoyée au recyclage. L'entreprise d'alimentation va seulement séparer les matériaux à recycler, mais la collecte et le transport seront faits par une entreprise de recyclage qui assume tous les coûts liés à cette procédure. Alors, si l'entreprise d'alimentation souhaite connaître la quantité des déchets envoyés au recyclage, elle pourrait adapter ses logiciels qui contrôlent la production pour enregistrer les restes de chaque type de matériel (carton,

papier, plastique, etc.) utilisé dans le processus de fabrication pour calculer sa performance d'utilisation des matériaux.

Dans le cas précédent, l'adaptation a été faite au niveau des systèmes d'information déjà existants dans l'entreprise. Mais, si les circonstances du logiciel indiqué empêchent des adaptations, une nouvelle source, en général interne, peut être créée pour gérer les nouvelles données dans différents formats.

La décision sur quels systèmes d'information seront adaptés ou créés varie selon le contexte de l'entreprise : les logiciels existants, les types de licences, les coûts d'adaptation, l'urgence, les délais disponibles, etc. C'est intéressant de remarquer que les approches de soutenabilité comme G4 Guidelines offrent des recommandations sur quels genres de systèmes d'information devraient accueillir les données selon l'indicateur choisi.

Après avoir identifié les sources déjà existantes et les adapter ou créer de nouvelles sources internes, comme les systèmes d'information, toutes les données liées aux indicateurs de soutenabilité auront leurs localisations connues. Bien que quelques données n'existent pas encore, à ce moment, les places où elles seront enregistrées sont déjà disponibles.

4.4.3 Créer une base d'intégration de données

Les données qui seront utilisées dans les prises de décisions sont, en général, distribuées par de diverses sources internes et externes. Cette dispersion peut causer des difficultés par rapport la gestion d'accès, de doublons, des mises à jour et de disponibilité des données. Un autre problème avec la dispersion des données est comment les associer parce qu'elles sont disséminées dans de différentes sources.

Les formes d'accès aux données et leurs disponibilités sont très importantes pour la conception des solutions d'intégration. Bien que la technologie offre des possibilités d'accéder directement à chaque source à la demande, les coûts de traitement et les ressources qui seront dépensés à chaque utilisation deviennent des facteurs critiques pour la disponibilité des données. Un autre point est la qualité de l'accès en temps réel qui peut varier entre chaque source.

Supposez les cas où deux logiciels non intégrés gèrent des données liées aux clients : un logiciel de contrôle financier et un autre qui s'occupe de la production. Comme ils sont maintenus de façon indépendante, peut-être que quelques informations sont dupliquées dans leurs bases de données, même si elles sont des sources internes. Donc, un client qui est stocké dans la base A peut avoir un état différent d'un même client enregistré dans la base B. Autrement dit, c'est toujours possible que les sources non intégrées soient incohérentes entre elles. D'autres types de données souvent répliquées sont les listes des pays et régions, des produits, des employés, des départements de l'entreprise, etc. Une question importante dans ce cas est : comment identifier les inconsistances et quelle source contient-elle la bonne information ?

C'est important de remarquer que la portée de cette recherche ne comporte pas de possibles processus de synchronisation des données entre les deux sources de l'exemple ci-dessus. Alors, pour résoudre ou atténuer ces problèmes, cette recherche recommande une stratégie d'intégration des données afin de réduire la dépendance aux sources et de garantir la disponibilité de l'information au sein de l'entreprise.

Au niveau de l'architecture physique, cette base de données est interne et centralisée. Elle va recevoir toutes les données des sources impliquées pour effectuer cette intégration. Comme une autre copie des données est faite, cela nous amène à penser qu'une nouvelle redondance sera créée. Un autre problème avec cette stratégie est la quantité de données qui peut augmenter énormément, selon les volumes présents dans les sources.

Ainsi, pour résoudre ces problèmes et pour ne pas tout simplement créer une base de données relationnelle et normalisée, cette approche sera basée sur les concepts d'un Entrepôt de Données (en anglais, Data Warehouse – DW). Comme abordé dans la Revue de la Littérature, un Entrepôt de Données est bien adapté pour effectuer l'intégration et pour gérer les données volumineuses, grâce au changement de la structure des bases de données en utilisant une modélisation multidimensionnelle au lieu d'un modèle relationnel, largement utilisé dans les systèmes d'information transactionnels. Cette base centrale sera responsable aussi pour fournir des moyens d'éliminer les doublons et de standardiser les formes d'accès aux données.

Deux concepts clés de la modélisation multidimensionnelle sont les tables de faits et les tables de dimensions. Un fait est une donnée qui mesure un concept de façon quantitative comme, par exemple, la valeur totale est un fait qui peut mesurer une facture. Une dimension est une perspective à partir de laquelle est possible d'analyser les faits. Un exemple de dimension est le client qui permet de filtrer les valeurs des factures qui lui appartiennent. Dans ce cas, un type d'agrégation (la somme, la moyenne, la valeur maximale, etc.) est appliqué sur les valeurs de l'ensemble de factures du client.

Un schéma est un concept de ce type de modélisation qui rassemble une table de faits et des tables de dimensions, sous lesquelles les faits peuvent être analysés. Il y a deux types principaux de schémas qui concernent à la modélisation multidimensionnelle : le Schéma en Étoile (en anglais, « Star Schema ») et le Schéma en Flocon de Neige (« Snowflake Schema »). Le premier est composé par une table de faits et des associations « one-to-many » avec chaque dimension. Alors, dans ce schéma, il n'y a pas des relations entre les dimensions. D'autre part, un Schéma en Flocon de Neige est une extension du Schéma en Étoile qui permet d'avoir les liaisons entre les dimensions.

En général, un schéma abrite qu'une seule table de faits et ses dimensions, mais, il existe encore un autre type nommé Schéma de Constellations de Faits (« Fact

Constellation Schema » ou « Galaxy Schema »). La différence est que ce schéma peut avoir plusieurs tables de faits qui sont liées par des dimensions conformes (communes à au moins deux schémas). Ces trois schémas peuvent exister en même temps dans une modélisation multidimensionnelle. Les impacts se trouvent dans la performance et la flexibilité du modèle.

Pour mieux comprendre la transition des données entre les modèles relationnel et multidimensionnel, un exemple a été construit. La Figure 4.2 représente un modèle relationnel d'une base de données transactionnelle qui permet d'enregistrer les factures d'un client, composées par des éléments de divers produits. Chaque facture est émise dans un état ou province d'un pays. Les clients sont classifiés en catégories d'importance selon leurs historiques d'achats. Ce type de base de données relationnelle est souvent alimenté par les systèmes d'information transactionnels, qui figurent dans cet exemple comme une source interne.

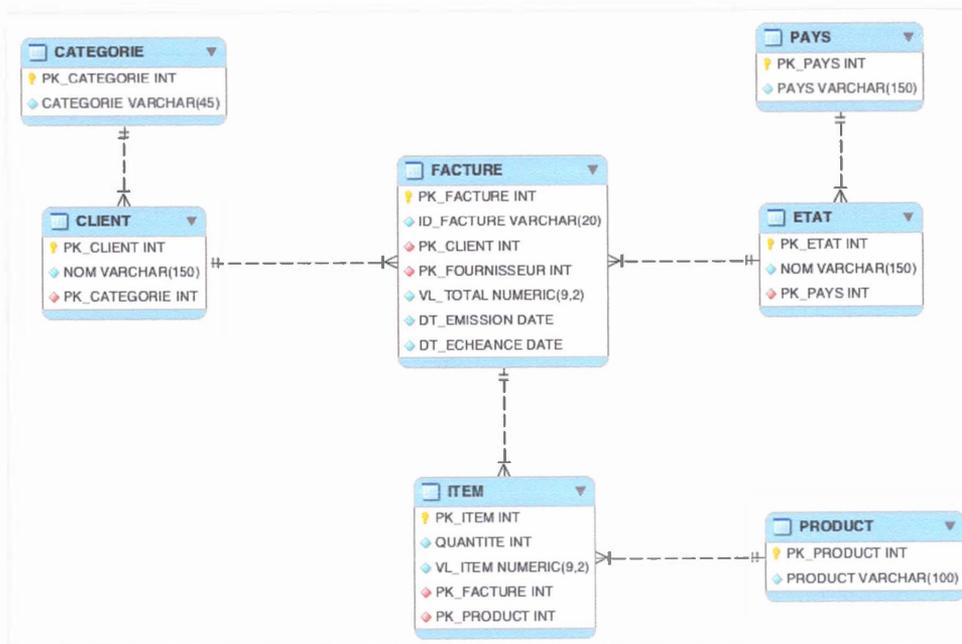


Figure 4.2 – Exemple d'un modèle relationnel.

Pour passer du modèle relationnel au multidimensionnel, il faut trouver des faits existants dans le modèle précédent. En les cherchant, c'est possible d'identifier un champ qui représente la valeur totale de la facture (champ « VI_Total » dans la table Facture). Ce champ peut être considéré un fait, car il est une métrique pour mesurer des factures. Un autre possible fait est le champ qui contient la valeur de chaque élément d'une facture (le champ « VI_Item » de la table Item).

Il y a une caractéristique des tables de faits qui s'appelle granularité. L'ensemble des dimensions liées à une table de faits détermine son niveau de granularité. Dans l'exemple proposé, les faits de la facture et de l'élément ont de différentes granularités. Un élément est plus détaillé qu'une facture, une fois que la valeur de la facture peut être calculée par la somme des valeurs des éléments. Autrement dit, le grain est la grandeur d'un fait. Alors, à cause des deux différents niveaux de granularités, deux tables de fait seront créées. C'est important de dire qu'une table de faits peut recueillir plusieurs faits s'ils sont valides pour le même ensemble de dimensions.

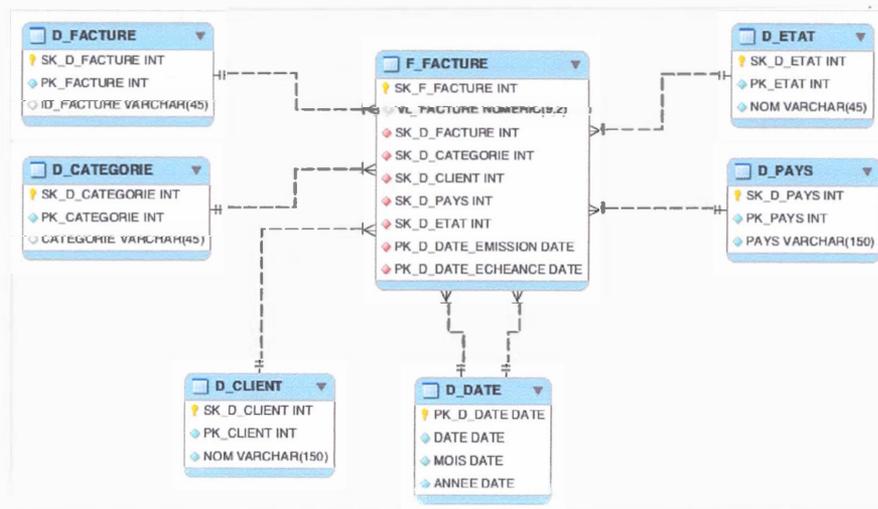


Figure 4.3 – Le schéma en étoile de la table de faits F_Facture.

La Figure 4.3 illustre un Schéma en Étoile de la table de faits F_Facture, basé sur la modélisation multidimensionnelle. La table Facture contient la clé primaire, un seul fait représenté par le champ de la valeur de la facture et des clés étrangères pour chaque dimension. Les dimensions qui gardent des relations avec le fait sont facture, catégorie, client, état, pays et date. Donc, ils permettent d'analyser le fait selon ces perspectives. Il y a, encore, une dimension et une table de fait avec le suffixe facture. Cela signifie dire que la modélisation permet d'identifier la facture qui a généré ce fait. Dans ce cas, la plus petite partie du fait, le grain, est détaillé au niveau de la facture.

En regardant la Figure 4.4, la table de faits F_Item présente une granularité au niveau des éléments des factures. Le schéma adopté est en flocon de neige pour illustrer les deux types principaux d'architecture. Cette table de faits, différemment de F_Facture, contient seulement des clés étrangères pour les dimensions facture, état, client et date. Les dimensions pays et catégorie ne sont pas liées directement à la table de fait, car dans le modèle relationnel elles sont liées aux tables état et client, respectivement. C'est la principale différence entre les deux types de schémas : dans le schéma en étoile, toutes les dimensions doivent toucher la table de faits.

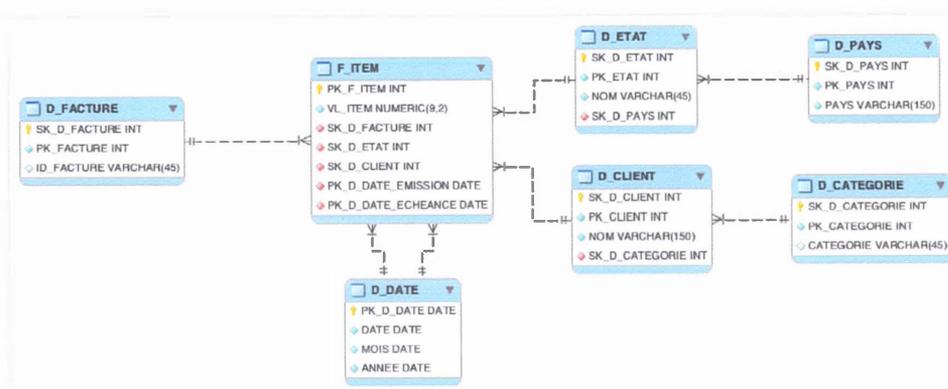


Figure 4.4 – Le Schéma en flocon de neige de la table de faits Item.

La table Facture du modèle relationnel contient deux dates : une date d'émission et une date d'échéance de la facture. En passant au modèle multidimensionnel, il y a une particularité concernant les champs de temps. Au-delà de favoriser l'intégration, un Entrepôt de Données est fortement basé sur une perspective temporelle et sur l'historique des faits. C'est pourquoi une dimension Date a été créée. Elle contient toutes les dates (jours, mois, années) et, dans certains cas, des agrégations comme les trimestres ou les semestres. Les deux schémas multidimensionnels présentent la dimension de date et les deux clés étrangères (date d'émission et d'échéance) dans les tables de faits. Si les deux tables de faits sont maintenues, il n'y a pas besoin d'ajouter les champs de date à la table de faits F_Item parce que cette information serait disponible dans la table F_Facture.

Au-delà des impacts de performance et flexibilité, avec ces exemples, il est possible de se rendre compte que le modèle en étoile élimine la liaison entre un pays et ses états dans les cas où il ne se trouve pas des faits. De manière analogue, il y a aussi des pertes d'information entre des clients et catégories en cas d'absence de faits d'un client. Bien que la perte des données des associations ne soit pas acceptable dans les systèmes d'information transactionnels, elle n'est pas assez grave dans un modèle multidimensionnel parce que tout est basé sur les faits.

Ainsi, la structure proposée est capable de recevoir les données des sources internes et externes et les intégrer dans un seul endroit. La nouvelle organisation avec la modélisation multidimensionnelle permet la croissance du volume des données sans compromettre la performance et la flexibilité d'évolution. En considérant toutes les données qui peuvent faire partie des prises de décisions, cette intégration représente une étape critique pour le succès de l'application de la soutenabilité.

4.4.4 Transporter les données

Après la création d'une base unique pour héberger les données des sources internes et externes, il faut concevoir et mettre en œuvre des mécanismes qui transporteront les données entre les origines et la base d'intégration.

En utilisant des techniques très connues dans les projets d'intelligence d'affaires, l'ETL comprend des processus de transporter les données à partir des sources jusqu'aux cibles. Cet acronyme est très connu en anglais comme « Extraction, Transformation and Load » qui signifie Extraction, Transformation et Charge en français. L'extraction à partir des sources est un processus que vise à garantir que les données sont capturées sans pertes ou changements à partir de leurs sources originales. Transformer est une activité d'adapter les données extraites selon la structure cible, les règles d'affaires ou les formats d'entrée. Par exemple, le passage d'un modèle relationnel à un modèle multidimensionnel peut exiger par lui-même des adaptations à cause des différents niveaux de normalisation. Il faut toujours garantir que cette transformation soit cohérente avec la sémantique de l'information. Enfin, l'étape du chargement est responsable pour enregistrer les données, transformées ou non, dans la base ou le format cible.

Avant de faire arriver les données à la base d'intégration, quelques tables peuvent être utilisées pour faire des stockages temporaires. Cette technique est souvent utilisée avec les sources volumineuses ou qui méritent un traitement plus complexe. Elle décompose les processus ETL dans petites parties pour faciliter les tâches de transfert des données. Alors, les transformations ne sont pas effectuées en même temps de l'extraction. Au lieu de cela, les données sont stockées dans une ou plusieurs tables qui seront traitées pour appliquer les changements nécessaires et finalement enregistrées dans la base de destination. Cette base intermédiaire est, en général, différente de la base de l'Entrepôt de Données et elle est communément appelée « Staging ».

Les clés étrangères entre ces tables temporaires ne sont pas utilisées parce que l'objectif est de stocker temporairement les données des sources sans faire aucun traitement. Alors, ces tables sont préparées pour recevoir toutes les données des sources sans appliquer des validations ou des règles. Un avantage de cette pratique est la réduction du temps que la source sera occupée pendant l'extraction. Un autre point fort est qu'un échec au moment des transformations ou enregistrement au destin, n'oblige pas de refaire l'extraction.

La gamme de différents formats des données, soit à l'origine ou à la base d'intégration, rend les tâches de transport et d'intégration plus difficiles à accomplir. Les diverses technologies de stockage et d'accès aux données doivent aussi être considérées. Mais, les outils disponibles sur le marché aident à gérer plusieurs petits programmes d'ETL avec des interfaces pour accéder aux différents types de bases de données, fichiers divers, Web services, etc.

La fréquence des mises à jour est déterminée, initialement, par la perspective temporelle des décisions à prendre. Cela veut dire qu'une décision qui est prise chaque mois a besoin de mettre à jour les données impliquées avec une fréquence au moins mensuelle. Mais, si la source fournit les données chaque semaine au lieu de à chaque mois sans garder l'historique, il faut la collecter à une fréquence plus courte. Alors, la disponibilité des données par la source est un facteur qui peut influencer à la fréquence de mise à jour.

Le volume de données est un sujet très important à cause de son impact sur la fréquence des mises à jour qui peut rendre non viable le processus d'ETL. Selon les ressources disponibles pour exécuter l'ETL, la quantité de données dans la source, les manipulations et le chargement peuvent être vraiment coûteux au niveau de budget et de traitement. La granularité des données est aussi un facteur qui peut mener à une augmentation importante du volume de données.

Les processus de transport de données peuvent devenir vraiment lourds à cause du volume et de la fréquence de mises à jour. Ils peuvent influencer le succès d'un projet en intelligence d'affaires. Alors, l'automatisation de l'ETL devient un outil puissant pour éliminer les procédures manuelles et garantir l'indépendance par rapport l'alimentation de données. Mais, l'ETL doit présenter une performance appropriée à la fenêtre de temps destinée à l'exécution du processus.

Supposez l'exemple d'un transport journalier des données financières à partir des sources vers la base d'intégration d'une grosse entreprise, qui prend un temps important pour exécuter. Normalement, les processus d'ETL sont démarrés après les heures régulières de travail pour ne pas rivaliser avec les systèmes d'information transactionnels. En considérant que le processus est déclenché à partir de 23 heures et qu'il prend au moins de 10 à 11 heures pour terminer, peut-être que les informations ne seront pas disponibles dans la base d'intégration avant 9h. Cette situation est critique, car les informations intégrées ne sont pas disponibles dans les premières heures de travail. Dans ce cas, une optimisation au niveau de l'ETL est requise pour rendre le projet viable.

Mais, même avec une automatisation bien faite et optimisée, il faut vérifier si les données transportées sont bien arrivées à la base d'intégration. Ces tests sont faits initialement pour attester la consistance des données avant de les utiliser. Après obtenir la confiance nécessaire, les données sont considérées valides et disponibles pour l'utilisation. Mais, c'est intéressant de, parfois, répéter les tests de vérification de la qualité des données.

Enfin, il y a des conditions légales qui peuvent empêcher l'extraction ou la manipulation des données à partir des sources. Normalement, les sources externes sont plus soumises à ce type de restriction. Pour éviter des problèmes judiciaires, il faut toujours consulter la propriété et confidentialité des données.

4.4.5 Traiter les inconsistances

Chaque cycle d'exécution vise à insérer plus de données dans la base d'intégration et la laisser à jour (ou presque) selon la fréquence déterminée. En général, un Entrepôt de Données intègre plusieurs sources dans une base unique. Cette intégration souvent expose des incohérences entre les données des différentes bases, mais qui devraient être consistantes entre eux. Par exemple, un client peut avoir le même code, mais avec deux différents noms ou une facture présente une valeur différente pour un élément. Ou, encore, une donnée qui devrait être toujours remplie arrive à l'entrepôt sans une valeur (la valeur nulle).

Les raisons pour ces types d'inconsistances sont diverses : des défaillances des règles des sources de données, mises à jour manuelles directement à la source, problèmes d'exécution des processus ETL qui cause une mise à jour partielle, etc. Un exemple d'une inconsistance est un client qui détient un code de la catégorie à laquelle il appartient, mais la catégorie n'existe pas dans la base d'intégration. Si la donnée est obligatoire, probablement une erreur empêchera l'insertion. Cette méthode recommande que toutes les données soient insérées dans la base d'intégration, même s'ils ne suivent pas les règles d'affaires. Une technique pour résoudre ce type de problème est de créer des records de contrôle dans les dimensions comme un code -1 pour la catégorie « Non Définie ». Au moment qu'arrive un client avec une catégorie vide, le record sera inséré, avec une catégorie générique.

Alors, l'Entrepôt de Données est capable d'identifier les inconsistances et d'alerter les fournisseurs des données (les sources) pour les corriger. Pour le traitement, les sources doivent être adaptées parce que les données stockées dans la base d'intégration ne peuvent pas être changées. Après les corrections dans les sources, les données seront naturellement insérées dans l'entrepôt dans le cas dont l'ETL est automatisé.

Enfin, un projet d'intégration basé sur les approches d'intelligence d'Affaires sert aussi pour valider et améliorer la qualité des sources comme des systèmes d'information, par exemple.

4.5 Module 3 : découverte de la connaissance

Après l'exécution des Modules 1 e 2, il est prévu que les indicateurs de soutenabilité sont identifiés et les données nécessaires pour les gérer sont déjà enregistrées dans la base d'intégration, après l'accomplissement des cycles d'ETL. Le Module 3 vise à combiner tous les éléments de l'Entrepôt de Données et les utiliser pour extraire des informations sous la forme de rapports, cubes, tableaux de bord ou, encore en utilisant les outils d'exploitation de données pour rechercher des modèles de comportement des données.

4.5.1 Comprendre les liaisons entre les faits et dimensions

Avant d'utiliser les données intégrées dans les schémas multidimensionnels, il faut comprendre les faits et les dimensions sous lesquelles ils seront analysés. Un fait est une mesure qui permet de classer un concept et l'analyser par rapport les autres.

Dans l'exemple de la table de faits F_Facture (Figure 4.3), la valeur de la facture est une métrique qui peut indiquer la relevance de cette facture parmi les autres. La comparaison de cette facture avec les autres va la classer dans un univers corporatif spécifique. Cette situation est basée sur la perspective de la facture, mais en considérant la métrique qui indique la somme des valeurs des factures de chaque pays, la perspective aura un niveau moins détaillé. En revanche, la table de faits F_Item contient la valeur de chaque élément d'une facture. Ainsi, c'est correct affirmer que la granularité de cette table est plus petite que celle de la table F_Facture.

Bien que les schémas soient indépendants, les tables de faits F_Facture et F_Item gardent une relation d'agrégation : la valeur d'une facture est composée par la somme de ses éléments. Mais, supposez une extension de l'exemple où un autre département de l'entreprise ou même une autre entreprise veuille collecter les réclamations de ses consommateurs. Soit une source interne ou externe, une possible table de faits F_Réclamation aurait un champ pour enregistrer la « quantité de réclamations » selon les dimensions client, entreprise, pays et date. Dans ce cas, aucune potentielle relation d'agrégation n'est vue entre F_Réclamation et les autres tables de F_Facture ou F_Item.

Par conséquent, plusieurs tables de fait sont créées avec de différents faits et perspectives, conformément les décisions à prendre et la stratégie d'affaires. L'ensemble de dimensions d'un schéma (en étoile ou en flocon de neige) définit la granularité d'une table de faits. Mais, les faits de différentes tables ou schémas peuvent être liés entre eux, même si les granularités sont différentes et s'il n'a pas des relations d'agrégation. Dans cette situation, il faut exister au moins une dimension conforme, autrement dit, une dimension commune entre les schémas qui seront intégrés et qui touchent les tables de faits. Par exemple, les dimensions « date, client et pays » sont communes aux dimensions F_Facture et F_Réclamation et, alors, c'est possible d'intégrer les faits de ces deux tables à travers ces dimensions.

La dimension de temps (ou Date) est un bon exemple et elle est présente dans tous les schémas. Toutes les données du type date dans une solution d'intelligence d'affaires ne sont plus un champ de date dans la table de faits. Ils seront remplacés par les clés primaires de par une table de dimension qui contient toutes les dates possibles d'exister dans les sources. L'intervalle de valeurs idéal varie selon la nature des champs. Par exemple, un champ de date de naissance peut contenir un intervalle des dates avant l'année de 1900 jusqu'à aujourd'hui. Mais, les données de date de fin du contrat

peuvent contenir des dates futures. Cette technique de la conception d'un Entrepôt de Données concernant cette dimension est largement utilisée.

Alors, cette étape du Module 3 est responsable de l'identification des faits qui peuvent être intégrés à travers des dimensions conformes. Cette particularité est la base de la capacité intégrative d'un projet en intelligence d'affaires.

4.5.2 Analyser les données

Après avoir compris comment les faits et les dimensions sont reliés, cette étape vise à construire des artefacts comme des rapports à partir de l'Entrepôt de Données. Les outils du type OLAP, « On-Line Analytical Processing », sont adaptés pour extraire de l'information à partir des bases de données ou d'autres types de sources. Ils offrent des ressources pour faciliter l'analyse de données sous la forme de rapports avec des tables ou des graphiques, des cubes, des tableaux de bord interactifs, des statistiques diverses, des outils de découverte de patrons, etc. Mais, il faut comprendre comment ces outils peuvent aider à la prise de décisions. Il ne suffit pas d'avoir de divers instruments sans les utiliser correctement.

Les rapports avec des tables des valeurs gèrent bien les données plus détaillées et en quantités plus importantes. Ils sont des documents souvent utilisés au niveau opérationnel pour bien montrer l'information transactionnelle plus nombreuse. Mais, ils sont aussi bien utilisés avec les rapports plus condensés.

Les ressources graphiques servent aussi pour illustrer des relations entre les données. Un diagramme à barres, par exemple, peut montrer les relations entre des faits selon une dimension comme le pays ou le temps. Un diagramme linéaire présente des évolutions par rapport le temps et des diagrammes circulaires (ou pizza) sont d'autres formes d'illustrer les distributions des données dans les dimensions. Le choix dépend

de quel type d'information sera focalisée. Si les valeurs en détail doivent être visibles à la prise de décisions, c'est intéressant d'utiliser les tables. Mais, si la comparaison entre les valeurs est plus importante pour la prise de décision, c'est recommandé de mettre des graphiques pour mieux illustrer ces relations.

Un tableau de bord (en anglais, « Dashboard ») est une ressource de présentation d'informations qui peut rassembler plusieurs types différents d'information. En général, un tableau de bord est dynamique et son interactivité permet aux utilisateurs d'analyser les données sous de différentes perspectives.

Les outils d'Exploration de Données (en anglais, « Data Mining ») utilisent les techniques de statistique pour comprendre les relations entre les données et, le cas échéant, trouver des modèles entre eux. En général, ces types d'approches ne considèrent pas les règles d'affaires, mais seulement les fréquences de combinaisons similaires des données. Par exemple, les données peuvent montrer que 38% des clients d'une catégorie A ont un montant annuel moyen de \$23,000 en factures. Ou, encore, un type d'élément X est présent dans les 62% des factures qui contiennent un élément Y. Il n'y a pas des règles régissant ces cas, mais ils sont des modèles de comportement qui ont été trouvés par l'analyse statistique des données.

En parlant d'intégration, l'objectif est de rassembler plusieurs données des trois dimensions de soutenabilité : économique, sociale et environnementale. Comme la base d'intégration est une place unique des toutes les données pertinentes aux prises de décisions, en utilisant des options de présentation ci-dessus, c'est possible de combiner les informations des différentes dimensions dans un même rapport, cube ou tableau de bord. En outre, cette base de données peut être un objet de recherche des outils d'exploration de données dans une tentative de trouver des liaisons cachées entre les données.

Cette étape de la méthode ne vise pas seulement à présenter les données, mais en fait, l'objectif est d'attirer l'attention des utilisateurs aux informations qui ressortent et méritent des actions à cause de son comportement particulier. En d'autres termes, le défi est d'ajouter de la valeur aux entreprises et leurs clients.

4.5.3 Supporter les prises de décisions

À ce moment, une dernière étape est mise en évidence : réviser les décisions à prendre et évaluer la capacité de la solution en intelligence d'affaires à les supporter. Souvent, le lien entre la motivation principale d'un BI et ses résultats est perdu tout au long du chemin.

La Module 1 de cette méthode s'occupait d'identifier les décisions et les indicateurs à chercher. Soit par la longue durée ou par sa complexité, un projet en intelligence d'affaires peut changer les objectifs et, ainsi, manquer des réponses attendues par les décisions à prendre. C'est pourquoi il faut appliquer la solution construite au milieu pratique.

Pendant cette vérification, quelques ajustements peuvent être identifiés pour que la solution ait plus d'adhérence que celle proposée initialement. Mais, c'est possible aussi de trouver d'autres types de décisions qu'il pourrait supporter. Cette étape vise à créer des racines plus profondes de la solution BI et de soutenabilité au sein des entreprises.

4.6 Les résultats attendus

Le principal objectif de cette recherche est la proposition d'une méthode qui facilite l'application de la soutenabilité au sein des entreprises donnant lieu à des connaissances stratégiques, faisant usage de techniques d'intelligence d'affaires pour intégrer les indicateurs de natures économiques, sociales et environnementales.

L'application de la méthode exposée dans ce chapitre apporte des bénéfices aux entreprises d'une façon à l'ouvrir les horizons décisionnels à partir de la compréhension et de l'utilisation de nouveaux éléments d'autres perspectives de connaissance. La dimension économique continue à jouer un rôle principal dans les processus décisionnels, mais ses relations avec le social et l'environnement sont prises en compte aussi.

Donc, un premier résultat attendu est l'existence d'indicateurs dans les processus décisionnels qui ne sont pas liés directement à la production ou à la prestation de services des entreprises. L'idée est d'ouvrir une porte à la compréhension des influences causées et souffertes par les facteurs à l'extérieur de l'entreprise.

Différemment de l'application des rapports de soutenabilité, la méthode vise sélectionner un nombre réduit d'indicateurs de soutenabilité selon le besoin des décisions à prendre. Ainsi, les premiers résultats sont rapidement obtenus et les évolutions sont effectuées dans les cycles courts.

Une base d'intégration a été créée pour recevoir des données de diverses sources et elle est préparée pour les évolutions, même si elles exigent des données volumineuses. Un processus d'ETL est responsable de l'alimentation de cette base et, en général, il est automatisé. Les inconsistances sont identifiées par les tests et traitées selon le besoin.

Finalement, la possibilité d'extraire des informations intégrées devient réelle et la découverte de nouvelles formes de prises de décisions est expérimentée. La qualité de ces résultats varie selon les définitions des décisions, la qualité de données des sources et la capacité de comprendre les informations et les outils d'analyse de données.

CHAPITRE V

PREUVE DE CONCEPT

Ce chapitre vise à réunir les informations liées à une preuve de concept pour illustrer l'application de la méthode proposée dans le Chapitre 4 dans un contexte pédagogique universitaire. La structure de ce chapitre est composée par une description de ce contexte, suivi par des renseignements sur chaque étape exécutée, distribuées dans les trois modules prévus. À la fin, un bref des résultats obtenus sera exposé.

5.1 Le contexte du cas

Le cas proposé aborde un contexte décisionnel hypothétique au milieu de l'éducation de niveau universitaire. Bien que la soutenabilité soit un thème très important et actuel pour la préservation de notre planète, Lozano (2011) explique que le nombre d'institutions d'éducation et le niveau de leurs rapports de soutenabilité suggèrent que les universités soient encore à ses débuts en comparaison avec le milieu corporatif. Selon Fassin (2015), la soutenabilité est arrivée au milieu universitaire avec 20 ans de retard par rapport aux entreprises. Ce terrain peu exploré représente un défi pour la preuve de concept et, pour cette raison, il a été choisi.

En général, les universités offrent des programmes de formation professionnelle dans plusieurs domaines de l'enseignement : arts, communication, langues, santé, sciences, etc. Ces sujets sont, en général, regroupés dans les facultés qui comprennent plusieurs

départements et écoles, par exemple une faculté des arts qui peut être composée par un département de musique et une école de théâtre. Cette preuve de concept est limitée au niveau de la faculté, sans détailler les informations des départements ou des écoles.

Chaque faculté propose des programmes de formation selon l'expertise de ses professeurs. Le système d'éducation universitaire québécoise est basé sur trois cycles. Le premier cycle contient des options de baccalauréat, des certificats, des mineures et des majeurs. Le baccalauréat est une formation complète qui vise à donner des connaissances sur un sujet spécifique. Les crédits sont une façon de mesurer le nombre d'heures de la formation. Un programme de baccalauréat contient 90 crédits ou 30 cours, car chaque cours a normalement 3 crédits. Il y a aussi des programmes de 120 crédits, qui peuvent exiger un stage complémentaire. Les certificats et mineures sont des programmes plus courts, de 30 crédits et les majeures exigent 60 crédits.

Le deuxième cycle est couvert à travers des diplômes d'études supérieures spécialisées (DESS) ou des programmes de maîtrise. Une maîtrise vise à développer la capacité de la recherche pour comprendre les phénomènes ou pour proposer des améliorations. Les programmes du troisième cycle sont les doctorats qui visent à former des chercheurs et ils sont fortement basés sur les contributions au domaine d'étude. Chaque type de programme mène à un diplôme spécifique qui atteste la réussite de l'étudiant.

La durée des programmes varie selon les régimes. Le temps complet est un régime où l'étudiant suit sa formation comme l'activité principale. Le temps partiel est moins exigeant en termes de quantité de crédit par session. Il y a trois sessions principales : Automne, Hiver et Été. Alors, la durée est influencée par le régime et les cours qui sont offerts dans chaque session. Un baccalauréat de 90 crédits a une durée moyenne de 3 ans et un programme de doctorat au moins 4 ans.

Physiquement, chaque université peut avoir un ou plusieurs campus qui sont distribués stratégiquement, selon les objectifs et les besoins d'infrastructure et de ressources des

programmes. La gestion des programmes, cours, campus de tous les cycles et types de régimes peut représenter un grand défi aux gestionnaires des universités.

Selon les estimations de la Statistique Canada (2016), le Canada a 1.3 million d'étudiants inscrits aux programmes universitaires, distribués dans « l'environnement de 110 universités ou établissements qui confèrent des grades au Canada » (Statistique Canada, 2008). Les programmes couvrent tous les domaines d'études. Ces chiffres soulèvent un autre point important : la responsabilité avec l'environnement et les questions sociales. Selon les objectifs de réductions de la pollution et de la destruction de l'environnement, fixés par les gouvernements, les actions des entreprises deviennent la cible des organismes de régulation et du grand public. Alors, les universités sont aussi responsables que les autres types d'entreprises.

En revenant à la problématique, la difficulté d'envisager les trois dimensions de soutenabilité dans les prises de décisions stratégiques est illustrée dans cette preuve de concept en utilisant deux scénarios décisionnels. Chaque scénario pourrait considérer simplement la dimension économique pour les analyses, mais la méthode proposée au Chapitre 4 vise à fournir d'autres éléments pour enrichir les prises de décisions.

Face à ce contexte, les prises de décisions devraient considérer les aspects de différentes dimensions : économique, sociale et environnementale. Les indicateurs de soutenabilité peuvent intégrer des analyses et aider à la création d'une gestion plus responsable et efficace. Enfin, c'est grâce à la pertinence du sujet, complexité et à la grande quantité de variables que cette preuve de concept se concentrera sur ce contexte et essaiera d'appliquer la soutenabilité avec la méthode proposée.

La décision d'utiliser un cas fictif est basée sur la difficulté d'aller chercher sur le terrain des données de nature sensible et stratégique d'une université, comme les adresses des étudiants, leurs programmes, les coûts, etc. Ces données pourraient résider dans différents départements de l'université et cela rendrait la collecte des données une

tâche difficile à accomplir. D'un point de vue technique, l'effort de l'institution pour rendre les informations anonymes avant de les mettre à la disposition du chercheur, pourrait être un obstacle à la construction de cette preuve de concept.

Bien que quelques données de l'Université du Québec à Montréal (UQÀM) soient utilisées dans cette preuve de concept, rien de spécifique à cette université ne sera considéré et, ainsi, son applicabilité dans un contexte plus large est possible. Ces données sont disponibles dans le site de l'UQÀM. La même stratégie de collecte de données ouvertes a été utilisée pour les informations de l'Emploi Québec, Statistiques Canada et Google Maps. La section 5.3.1 abordera plus en détail les sources de données.

5.2 Module 1 : sélection des indicateurs

Selon les recommandations du premier module de la méthode et le contexte exposé, les objectifs de cette partie sont d'identifier deux potentielles décisions à prendre par les gestionnaires d'une université, puis choisir les indicateurs à utiliser, les approches de soutenabilité et définir une stratégie d'exécution.

5.2.1 Définir les décisions à prendre

Une université a une envergure qui peut générer plusieurs défis aux gestionnaires au niveau des prises de difficiles décisions qui doivent considérer plusieurs variables et situations. Une mauvaise décision peut conduire à de nombreux problèmes. C'est pourquoi le choix des indicateurs les plus pertinents est une tâche essentielle pour une analyse plus précise.

Cette preuve de concept propose deux scénarios décisionnels : une décision sur les ajustements des programmes de formation et une décision sur un possible partenariat

entre l'université et des entreprises de transports collectifs comme, par exemple, la Société de Transport de Montréal (STM). C'est important de souligner que ces décisions sont hypothétiques, mais elles sont bien utiles et réalistes face au contexte proposé. Chaque scénario sera détaillé dans les prochaines sections.

5.2.1.1 Scénario 1

Le premier processus décisionnel est basé sur l'intention d'une université à l'analyse de ses programmes de formation de n'importe quel cycle ou type de diplôme pour les ajuster selon le marché du travail et l'attractivité de ces programmes. Ces ajustements peuvent causer des fusions entre des programmes, des divisions en plusieurs autres et la suppression ou la création des programmes.

Si les gestionnaires prennent des décisions stratégiques comme celle-là en considérant seulement la dimension économique, les indicateurs seraient basés sur la performance économique des programmes de formation : des revenus, coûts, dépenses, factures, résultats, etc. C'est bien de dire que cette analyse est logique et cohérente dans un contexte purement économique. Cependant, l'objectif de cette recherche est de donner aux gestionnaires d'autres éléments pour analyser aussi les dimensions sociale et environnementale dans la prise de décisions.

Par exemple, si un programme de formation présente une réduction de la performance économique, naturellement la pensée est de le supprimer ou de faire une fusion avec un autre programme pour améliorer sa performance. Mais, analyser tout simplement cet indicateur n'est pas assez précis pour la bonne compréhension de ce contexte. En ajoutant la dimension sociale, par exemple, c'est possible de considérer les perspectives des carrières comme un facteur fiable pour aider à interpréter le comportement des données par rapport les programmes. Donc, un programme qui présente une baisse de performance économique, mais les perspectives des carrières sont favorables, peut-être

que l'université ne doive pas le supprimer, mais seulement optimiser les coûts liés à ce programme ou améliorer sa qualité pour attirer plus d'étudiants.

D'un point de vue environnemental, les universités et les entreprises en général veulent réduire l'impact de ses activités sur l'environnement. C'est pourquoi elles ont des bureaux qui visent la gestion de ses activités et la surveillance de leurs impacts par rapport l'environnement. Fonseca et al. (2011) ont constaté que les rapports de soutenabilité des 25 plus grandes universités du Canada sont faits par ces types de bureaux internes, mais ils ne sont pas soumis à une évaluation externe indépendante. Et pourtant, seulement un de ces rapports présentait la lettre du président de l'université, un fait qui met en évidence une faible intégration avec la gestion et la représentativité questionable de ces informations. Enfin, les auteurs avertissent que ces rapports ont une valeur limitée et qu'ils ne sont pas fiables pour orienter les décisions orientées vers la soutenabilité.

Face à ce contexte, ce scénario offre une possibilité d'approcher la gestion et la soutenabilité. Il propose l'analyse des impacts sur la qualité de l'air, causée par la pollution générée par les déplacements des étudiants pour aller à l'université. C'est un indicateur qui aide à établir un type simple de performance environnementale et qui peut mener à l'identification des programmes qui ont des impacts les plus sévères sur la nature. Ainsi, l'analyse des programmes de formation pourrait considérer aussi la dimension environnementale.

En bref, de cette manière, les trois dimensions de la soutenabilité seraient considérées dans la prise de décisions pour le réarrangement des programmes de formation.

5.2.1.2 Scénario 2

Une deuxième situation décisionnelle est basée sur une stratégie où une université vise à établir un accord avec une entreprise de transports collectifs pour créer des lignes d'autobus électriques ou hybrides pour réduire l'impact sur l'environnement. Les autobus seraient moins polluants et, alors, la performance environnementale de l'université pourra dépasser ses frontières, au-delà des actions internes de réduction de la consommation d'énergie et d'eau, par exemple. Du côté de l'entreprise de transport, au-delà des avantages financiers, il y a aussi la possibilité d'offrir un service segmenté qui peut améliorer son image face aux clients et d'expérimenter une autre façon de mieux connaître son public.

Bien que ce scénario ne soit pas un cas réel, il est très pertinent. Selon Drewes et Michael (2006), la distance entre l'institution et la résidence de l'étudiant et les frais supplémentaires reliés aux études sont les critères les plus influents sur quelle université de la Province de l'Ontario, au Canada, les étudiants choisissent. Petruzzellis et Romanazzi (2010) expliquent que le manque de soins et d'attention individualisée pour les étudiants représente un problème critique que la gouvernance universitaire est confronté à la concurrence moderne et mondiale. Alors, la difficulté de déplacement pour aller à l'université et la compréhension des besoins des étudiants sont de remarquables critères pour choisir une université.

Le contexte décisionnel de ce scénario propose une analyse plus profonde dans les trois dimensions de la soutenabilité. L'analyse de la dimension économique sera basée sur une possible augmentation du nombre d'inscriptions grâce à la stratégie d'investissement sur l'optimisation des déplacements de l'étudiant et d'une image plus responsable de l'université auprès de l'environnement. La dimension sociale peut considérer un point d'équilibre entre le temps et la facilité de prendre un transport plus optimisé pour les étudiants qui se dirigent vers le campus. Enfin, l'impact sur la qualité

de l'air est proposé comme un indicateur intéressant pour la performance environnementale.

Il faut aussi considérer le côté de l'entreprise de transport collectif, car elle fait partie de l'accord et elle a ses objectifs particuliers à suivre. Quelques actions comme le réarrangement des lignes existantes ou la création des nouveaux types de tarification de transport, subventionnées par l'université seraient des types de décisions à prendre. Cette stratégie peut être appliquée dans plusieurs universités ou entreprises pour comprendre les comportements des étudiants et des travailleurs. Mais, en raison de la portée de cette preuve de concept, les décisions du côté de l'entreprise de transport ne seront pas abordées. Ainsi, d'autres éléments devraient participer de cette analyse parce que les objectifs ne sont pas les mêmes.

5.2.2 Choisir les indicateurs

Selon la méthode proposée, la tâche de choisir les indicateurs les plus pertinents peut causer des confusions, car il y a une grande quantité d'indicateurs disponibles dans les approches de soutenabilité. Dans le cas où l'objectif est la création des rapports de soutenabilité, de nombreux indicateurs sont utilisés et une seule approche est recommandée pour assurer la cohérence entre eux.

Mais, la preuve de concept ne vise pas la création des rapports et, au lieu de collecter plusieurs indicateurs, les choix seront guidés par les besoins des décisions à prendre. Chaque scénario sera analysé pour identifier les indicateurs pertinents déjà existants dans les approches de soutenabilité. Le cas échéant, d'autres indicateurs seront ajoutés.

5.2.2.1 Scénario 1

Ce premier cas aborde l'évaluation des programmes de formation en utilisant les dimensions économique, sociale et environnementale pour les optimiser selon le marché du travail. Les indicateurs majeurs de la dimension économique seront les coûts, les chiffres d'affaires et les résultats de chaque programme. Ils sont calculés par la somme des indicateurs plus détaillés.

Au niveau des coûts, il faut calculer les salaires des enseignants et du personnel des fonctions administratives, l'énergie dépensée, l'espace physique, les meubles, l'eau et plusieurs d'autres coûts, soit variables ou fixes.

Les chiffres d'affaires sont obtenus à partir des revenus créés par les programmes de formation. Chaque inscription comprend des frais de scolarité et d'autres frais additionnels. Une partie des frais de scolarité des étudiants canadiens ou des résidents permanents est subventionnée par les organismes gouvernementaux. Les frais de scolarité des étudiants étrangers sont payés en totalité par eux-mêmes, car les subventions ne sont pas applicables à ce public. Enfin, la relation entre les coûts et les chiffres d'affaires devient, en général, l'indice basique pour les prises de décisions économiques. Cette relation est favorable à l'université si les chiffres d'affaires sont plus grands que les coûts.

Un facteur social qui influence les chiffres d'affaires est la perspective des carrières professionnelles des étudiants qui sont formés par ces programmes. Ce facteur indique si une carrière a des tendances positives ou négatives selon le marché. La quantité d'emplois disponibles, le taux de chômage, et la valeur moyenne des salaires sont aussi importantes pour une analyse de la dimension sociale. Ces indicateurs permettent aux gestionnaires de faire une analyse des tendances des carrières professionnelles pour comprendre, par exemple, si une performance économique faible d'un programme peut être influencée par un ralentissement des perspectives des carrières liées.

L'environnement sera aussi considéré en ajoutant des données des temps et distance liées aux déplacements des étudiants pour participer aux cours présentiels de l'université. Ces déplacements ont des impacts sur la qualité de l'air soit en utilisant une voiture, des transports en commun. Il y a des études qui disent que même l'utilisation des vélos, de façon indirecte, cause un impact sur l'environnement, car l'énergie dépensée par les gens est plus grande en comparaison à ceux qui utilisent des voitures ou des transports publics et, alors, et ils génèrent plus de dioxyde de carbone et ils ont besoin de plus d'aliments. De cette façon, une analyse de l'impact des déplacements des étudiants est importante pour comprendre la dynamique de la performance environnementale.

Selon l'organisme Statistiques Canada (2006), le transport est une source majeure d'émission des gaz de GES, les Gaz d'Effet de Serre (en anglais, GHG or « Greenhouse Gas »). Au Canada en 2004, le transport était responsable par 26% des émissions totales de GES au pays. La combustion des carburants des moyens de transport est responsable pour contribuer à la pollution de l'air à cause des émissions des gaz comme le dioxyde de carbone (CO₂), le méthane (CH₄) et l'oxyde nitreux (N₂O). Chaque mesure de volume de gaz sera un indicateur et utilisé dans cette preuve de concept.

5.2.2.1 Scénario 2

Ce cas propose un partenariat entre l'université et des entreprises de transports collectifs pour l'utilisation des autobus électriques selon les comportements de déplacements des étudiants.

Différemment du premier scénario, les indicateurs économiques seront basés sur l'évolution du nombre d'inscriptions par programme. Les indicateurs de coûts et revenus ne seront pas considérés, car il n'y a pas des distinctions entre les étudiants de

différents cycles ou de statuts. Une granularité minimale qui sera utilisée par cet indicateur sera au niveau du programme.

La dimension sociale sera basée sur les indicateurs de distance et de durée pour le déplacement de chaque étudiant. C'est important de considérer la distribution par chaque type de transport utilisé par les étudiants pour comprendre les exigences d'un possible nouveau transport. La région ou la ville d'où vient l'étudiant seront aussi considérées.

La dimension environnementale analysera aussi la qualité de l'air comme un indicateur, par le contrôle des émissions des chaque gaz du GES émis par les moyens de transport comme le train, le métro, les autobus, les voitures de différents modèles et les transports alternatifs. Dans ce cas, la granularité sera au niveau de chaque déplacement.

5.2.3 Choisir les approches de soutenabilité

Cette preuve de concept vise à démontrer l'intégration entre les trois dimensions de la soutenabilité : économique, sociale et environnementale. L'approche G4 Guidelines (GRI, 2014a et 2014b) a été choisie comme un guide de référence pour cette preuve de concept grâce à ses concepts didactiques par rapport aux dimensions de soutenabilité. Ce guide est composé par des catégories, aspects et indicateurs. Les catégories sont équivalentes aux dimensions économique, sociale et environnementale. Les aspects sont des groupes d'indicateurs qui sont reliés aux mêmes sujets comme, par exemple, les indicateurs de l'aspect Émissions.

Les indicateurs économiques choisis par ces deux scénarios sont trouvables dans ce guide de soutenabilité. Les mesures des revenus et coûts sont situés dans l'aspect de la performance économique. Mais, pour les calculer, des indicateurs particuliers au

scénario seront collectés : le nombre d'inscriptions et les frais de scolarité. Ces indicateurs ne font pas partie des approches de soutenabilité à cause de sa spécificité, mais ils sont indirectement liés aux indicateurs économiques principaux.

Au niveau social, la quantité d'emplois disponibles et le taux de chômage sont des indicateurs qui permettent aux entreprises d'évaluer leurs capacités d'attirer les nouveaux employés ou de comprendre que les risques de chaque profession par rapport l'entreprise. Ces deux indicateurs sont abordés par l'aspect Emplois du guide G4 Guidelines. C'est important de remarquer que la preuve de concept utilise ces indicateurs pour évaluer les perspectives des professions liées aux programmes de formation.

La catégorie de l'environnement du même guide contient des aspects émissions et transport, qui abordent les indicateurs de mesure des émissions de gaz à effet de serre (ou GHG) en général et reliées au transport de gens. Les indicateurs des émissions du guide comprennent tous les gaz qui endommagent la couche d'ozone, mais dans la preuve de concept seulement ceux concernant le transport seront considérés.

En bref, cette preuve de concept aborde les indicateurs des trois dimensions de soutenabilité, en utilisant le guide G4 Guidelines comme une base de référence. C'est important de souligner que d'autres approches pourraient être choisies en fonction de leurs pertinences aux scénarios.

5.2.4 Créer une stratégie d'exécution

Un projet en intelligence d'affaires est utilisé par cette preuve de concept pour gérer les indicateurs et supporter les décisions à prendre. La gestion de ce projet sera basée sur des itérations successives et incrémentales. Une version très simple avec seulement

deux décisions à prendre sera utilisée comme un point de départ. Un cycle au complet en passant par les trois modules sera fait avec cette portée initiale.

Pendant ce cycle, la gestion agile n'a pas été mise en pratique parce que le chercheur a travaillé tout seul pendant la construction. Mais, dans un projet réel, il peut profiter des pratiques agiles pour augmenter la productivité et paralléliser les activités pour optimiser les professionnels engagés dans le développement de la solution BI.

C'est important de souligner que le choix de commencer avec un petit groupe de décisions est une action que vise à créer une espèce de première version sur un cas réel. De cette manière, c'est possible de comprendre les objectifs de chaque étape et d'expérimenter les premiers résultats concrets au début du travail.

Mais, même pour seulement ces deux décisions, certains indicateurs seront priorisés sur les autres dans le contexte de la preuve de concept. Supposez le cas hypothétique où, après une recherche sur les sources des données financières du Scénario 1, il est constaté que l'université a plusieurs systèmes non intégrés dans chaque faculté pour gérer les coûts à cause de la nature spécifique de chaque programme. Alors, même avec des systèmes internes, la difficulté d'intégrer toutes les sources pourrait impacter les délais pour la première version de cette solution en intelligence d'affaires. Alors, pour prendre une décision basée sur des indicateurs, un gestionnaire peut commencer par la compréhension de l'évolution des inscriptions pour vérifier quelles sont les tendances d'un programme et si ces perspectives sont favorables pour l'université. Ainsi, la stratégie d'exécution décide de travailler seulement les données liées à la quantité d'inscriptions dans le premier cycle et les données de coûts et revenus, seront traitées dans une deuxième itération.

Du côté social, les données seront toutes collectées et considérées dans le premier cycle. Par contre, la dimension environnementale vise à considérer les types de transport utilisés par les étudiants. Mais, supposez que l'université n'enregistre pas

dans ses systèmes d'information internes quelles voitures les étudiants utilisent ou même quels types de transport ils utilisent pour aller à l'université. Dans ce cas, les statistiques seront utilisées dans un premier moment et elles deviendront plus précises après l'adaptation des systèmes d'information et la collecte de ces données.

Le Scénario 2 sera plus simple au niveau de la dimension économique, car il aborde seulement les nombres d'inscriptions par programme, qui seront disponibles à la fin du premier cycle. Les statistiques seront aussi utilisées dans un premier moment par les indicateurs sociaux et environnementaux.

Ce plan d'exécution par cycles n'est pas rigide et peut être changé à n'importe quel moment pour adapter le projet aux nouvelles conditions.

5.3 Module 2 : adaptation et intégration

Après avoir défini ce qu'il faut faire, c'est le moment de découvrir comment le faire. Alors, ce module s'occupera d'identifier les sources de chaque indicateur et les données qui seront nécessaires pour supporter les décisions.

5.3.1 Identifier les sources des données

Basés sur le contexte de l'éducation au niveau universitaire, les scénarios veulent démontrer comment fournir des informations des trois dimensions de soutenabilité pour supporter deux processus de décision simples. La plupart des données utilisées dans la preuve de concept viennent des sources ouvertes, disponibles dans plusieurs sites Web. Il y a trois catégories de données : réelles, approximatives et fictives.

Les données réelles utilisées sont classées en tant que ouvertes et sont trouvables dans les sites Web de plusieurs organismes ou entreprises. Ce type d'information est précise

et n'utilise pas des artifices statistiques dans leurs constructions. D'autre part, les données approximatives sont considérées comme officielles, mais elles sont calculées en utilisant des formules statistiques. Alors, c'est possible d'avoir un aperçu moins précis de la réalité, mais avec un degré acceptable de fiabilité. Pour assurer la qualité des données, les organismes officiels de mesure comme Statistiques Canada ou Québec ont été préférés. Ces deux types de données ne sont pas modifiés, mais seulement filtrés selon le besoin de cette recherche.

Quelques données fictives ont été générées pour remplacer celles qui ne pouvaient pas être collectées. Elles sont normalement stockées dans les sources internes aux entreprises ou organismes et sont des informations sensibles qui demandent des autorisations de manipulation. Alors, pour contourner ce problème, elles ont été aléatoirement générées par des routines de l'ETL, par exemple les informations des régions où chaque étudiant habite pour calculer ses déplacements.

Le Tableau 5.1 présente les principales sources des données qui seront manipulées par les deux scénarios et distribuées dans les dimensions de soutenabilité. La granularité informe le niveau de détail des données offertes par les sources. La colonne de source contient des références pour le document ou la page Web qui détient les données. La nature dit si les données sont réelles, approximatives ou fictives. Enfin, l'itération définit dans quel cycle les données seront disponibles.

Tableau 5.1
Sources de données identifiées

Dimension Économique				
Données	Granularité	Source	Nature	Itération
Liste de programmes	Programme	Site de l'UQAM	Réelle	1
Campus, facultés, cycles, diplômes, régimes et sessions	Programme	Site de l'UQAM	Réelle	1
Nombre d'inscriptions	Inscription	Site de l'UQAM	Réelle	1
Coûts fixes et variables	Faculté	Systèmes d'Information	Fictive	2
Frais de scolarité selon le cycle et le régime d'études	Programme	Systèmes d'Information	Approximative	2
Revenus par subventions	Faculté	Systèmes d'Information	Approximative	2
Dimension Sociale				
Données	Granularité	Source	Nature	Itération
Classification Nationale des Professions	Profession	Statistiques Canada	Réelle	1
Salaires des carrières	Profession	Site de l'Emploi Québec	Approximative	1
Taux de chômage, perspectives, nombre des postes.	Profession	Site de l'Emploi Québec	Approximative	1
Programmes de formation des carrières	Profession	Site de l'Emploi Québec	Réelle	1
Dimension Environnementale				
Données	Granularité	Source	Nature	Itération
Régions de la grande Montréal	Région	Statistiques Canada	Réelles	1
Émissions par type de transport	Émission	The Guardian	Approximative	1
Adresses et transport des étudiants	Étudiant	Générées	Fictives	1
Distance et durée des déplacements	Déplacement	Google Maps API	Fictives	1
Impacts des transports sur la qualité de l'air.	Transport et occupation	Générées	Fictive	1
Autres Types de Données				
Données	Granularité	Source	Nature	Itération
Dates	Jour	Générées	Réelle	1
Université	Université	Saisie Manuelle	Réelle	1

La première section du Tableau 5.1 a défini les données liées aux inscriptions sous la perspective des programmes qui seront traitées dans une itération initiale. Ces données ouvertes ont été collectées dans le site Web de l'Université du Québec à Montréal – UQÀM. Elles se présentent sous le format d'un document PDF ou HTML publiés chaque année depuis 1994. Les autres données liées aux coûts et revenus seraient obtenues à partir des potentiels systèmes d'information qui seraient responsables de la

gestion de ce type d'information. Pendant l'itération 2, l'objectif serait d'utiliser des statistiques pour répartir les valeurs de coûts et revenus de l'université par le nombre d'inscriptions et, ainsi, créer un nouvel indicateur économique. D'autres itérations seraient ajoutées pour augmenter la précision des données à travers l'extraction à partir des systèmes d'information au niveau des facultés ou des programmes.

Le traitement des informations de la dimension sociale sera fait dans le premier cycle, en considérant la liste de professions selon la Classification Nationale des Professions (CNP), et les salaires, le taux de chômage, les demandes et les perspectives des carrières. Enfin, pour créer un lien entre professions et programmes, les informations de la formation seront enregistrées dans la base d'intégration.

Une liste des régions de la grande Montréal sera utilisée par la dimension environnementale comme une base pour la création d'informations fictives de déplacement. Les données liées aux étudiants, par exemple l'adresse et le type de voiture utilisé, seront générées aléatoirement, mais en respectant des statistiques de distribution de la population et le niveau de scolarité par région. Comme l'objectif est de démontrer que la méthode est capable de mener les entreprises à la connaissance de soutenabilité, ces données fictives n'affectent pas les résultats finaux de la preuve de concept.

Au-delà des données de régions et types de transport, les déplacements des étudiants seront générés à partir des API du Google Maps en utilisant la région où ils habitent et le campus de destination. Chaque déplacement sera enregistré avec les données de distance et de durée. Ainsi, les statistiques d'émission des gaz par type de transport seront bien pertinentes pour calculer les impacts sur la qualité de l'air à travers des déplacements.

C'est important de souligner qu'une liste des dates sera générée parce qu'elles composent une dimension très importante pour la conception de solution en

intelligence d'affaires. Cette dimension de temps recueille toutes les dates existantes dans les tables de la base de données pour appliquer une ligne de temps unique. Enfin, les données liées à l'université seront saisies manuellement.

Cette preuve de concept propose l'intégration entre les données internes à l'entreprise et les données externes pour enrichir la recherche et profiter des statistiques officielles des organismes pertinents. Cette stratégie de collecte de données vise à souligner aussi la possibilité d'appliquer les concepts de soutenabilité sans avoir besoin des efforts majeurs au niveau des changements des systèmes d'information. C'est possible, alors, d'utiliser directement la base de données d'intégration pour stocker ces données sans impacter les systèmes d'information.

C'est important de définir la granularité des données qui sont utilisées. La granularité est la plus petite partie de l'information. Par exemple, le nombre d'inscriptions peut être enregistré à un niveau plus haut comme par faculté ou plus détaillé comme par programme. En général, une granularité plus détaillée permet d'agréger par de différentes perspectives.

5.3.2 Adapter les systèmes d'information

La dernière section a identifié les sources de données qui alimenteront la base d'intégration. La méthode indique que quand il y a des données qui n'existent pas dans les sources externes et internes, il faut créer une place pour les enregistrer.

Les données de la dimension économique sont, de façon générale, contrôlées par les systèmes d'information corporatifs ou des progiciels intégrés. Après l'introduction des dimensions sociale et environnementale dans l'entreprise, autres systèmes peuvent être impliqués dans la solution.

Tenant compte la dimension environnementale, les adresses des étudiants sont normalement enregistrées dans le logiciel qui gère les inscriptions. Mais, le type de transport que l'étudiant utilise pour aller à l'université n'est pas souvent présent dans ces systèmes d'information. Donc, il faut adapter le logiciel d'inscription pour inclure de nouvelles informations de l'étudiant. La modification se traduira dans l'inclusion d'une table avec les types de transport (« T_Transport »). Et, aussi, un nouveau champ (une clé étrangère) avec le type de transport (« ID_Transport ») sera créé dans la table qui contient les étudiants (« T_Etudiant ») pour enregistrer cette donnée. Une autre adaptation devra être faite au niveau de l'interface du système d'information qui permettra aux utilisateurs d'enregistrer le type de transport de chaque étudiant.

L'adaptation présentée figure comme une modification simple à faire. Mais, dans certains cas, les délais sont longs et ne furent pas à attendre. Parfois, c'est plus rapide de créer une nouvelle application et dupliquer les données des étudiants, par exemple, pour enregistrer de nouvelles informations. Et, aussi, l'introduction d'autres données peut causer l'apparition des doublons. Supposez le cas où il existe déjà une table avec quelques types de transport dans un système de contrôle de stationnement. Alors, un même concept sera répliqué dans deux différentes places.

Au moment d'intégrer les deux systèmes d'information, il peut avoir des inconsistances entre les données. Pour éviter ce problème, toutes les données sont intégrées dans une solution BI qui permettra l'extraction de la connaissance des trois dimensions de soutenabilité sans ambiguïtés.

5.3.3 Créer une base d'intégration de données

Le point d'intégration sera une base de données multidimensionnelle gérée par un serveur MySQL. Le choix pour ce type de 'Système de Gestion de Base de Données' (DBMS) c'était grâce à sa licence libre et « open source », mais d'autres DBMS

pourraient être choisis. Cette base recevra les données à partir des sources qui ont été identifiées à la dernière section.

Trois schémas en étoile sont proposés pour recevoir des données identifiées : Revenu (« F_Revenu »), Déplacement (« F_Deplacement ») et Perspectives (« F_Perspectives »). Elles sont basées sur les dimensions économique, environnementale et sociales, respectivement. Bien que dans cette preuve de concept les trois schémas sont correspondants aux dimensions de soutenabilité, il n'existe pas une règle de modélisation et, donc, une dimension peut avoir plusieurs schémas comme celui-là ou un schéma peut même toucher de différentes dimensions.

5.3.3.1 Schéma 'Revenu'

La dimension économique est représentée par le schéma en étoile Revenu de la Figure 5.1. La table de faits « F_Revenu » recevra les faits de nombre d'inscriptions. Chaque fait peut être représenté dans plusieurs dimensions : université, campus, faculté, programme, cycle, diplôme, régime, genre, session et temps. La granularité de cette table de faits est au niveau de l'inscription. Cela signifie qu'un enregistrement (nombre d'inscriptions égal à 1) dans la table de faits représente une inscription d'une personne de tel genre, dans telle session, de tel programme, sous tel régime, de tel cycle, pour tel diplôme, dans telle faculté, de tel campus et, finalement, de telle université.

Chaque table de dimension contiendra les valeurs possibles. Par exemple, la dimension régime (« D_Regime ») contient les valeurs des régimes existants : temps complet, temps partiel et rédaction de thèse.

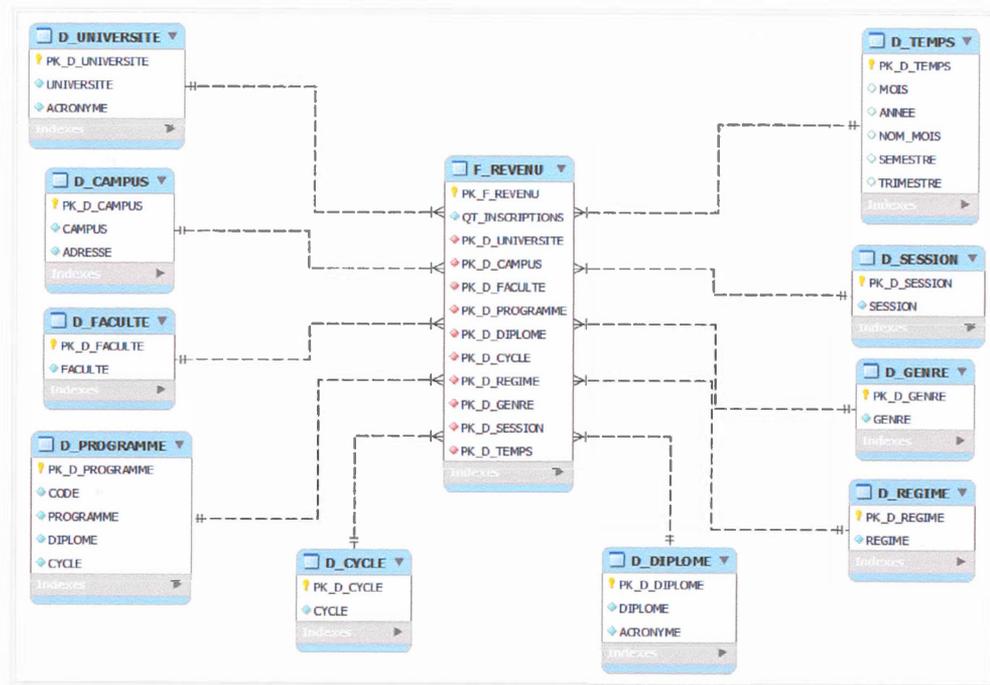


Figure 5.1 - Schéma en étoile des inscriptions.

Bien que le nom du schéma Revenu ne contienne pas les données de coûts ni de revenus, les prochaines itérations visent à collecter ces données.

5.3.3.2 Schéma 'Perspective'

Chaque carrière aura ces renseignements par rapport ses perspectives enregistrés dans ce schéma. La table de faits contient deux indicateurs : la quantité d'emplois (« QT_Emplois ») et la valeur estimée du revenu (« VL_Revenu »). Les perspectives d'analyse sont : programme, la profession, CNP 1 à 3 (les trois niveaux des professions selon la Classification Nationale de Professions), la demande, le taux de chômage et les perspectives de chaque profession.

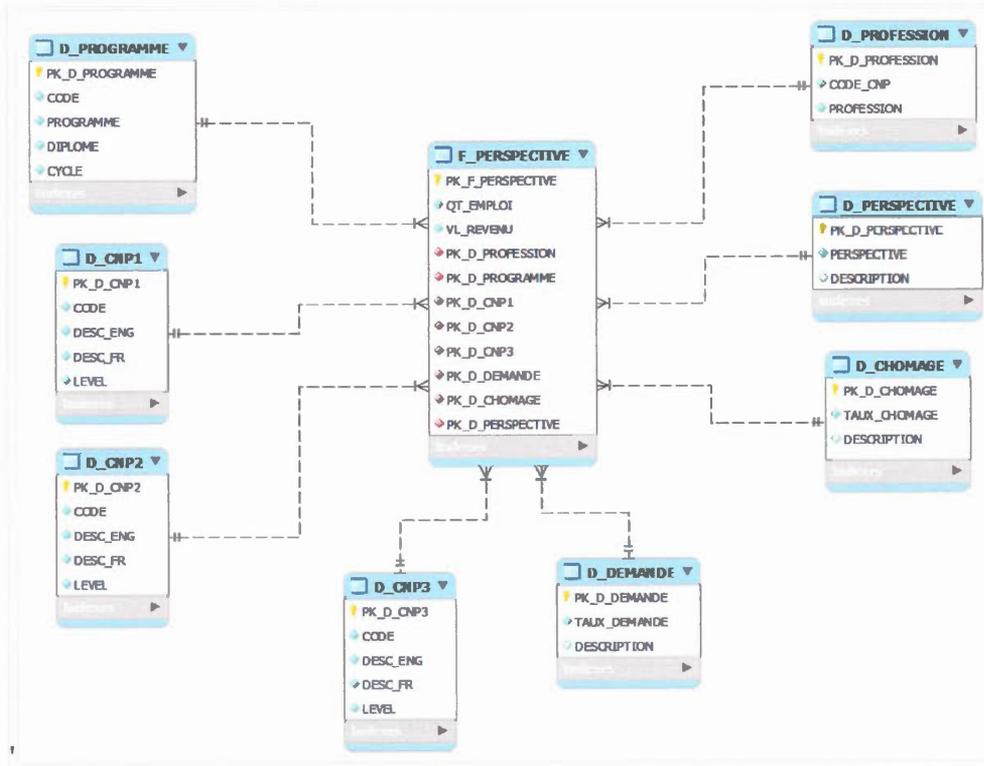


Figure 5.2 - Schéma en étoile de perspectives des professions.

5.3.3.3 Schéma 'Déplacement'

Chaque déplacement sera enregistré dans ce schéma en utilisant les tables de faits et de dimensions. Un déplacement est compris comme le parcours fait entre un point de départ et la destination finale. Les indicateurs sont la durée de temps, la distance parcourue et plusieurs champs liés à la qualité de l'air.

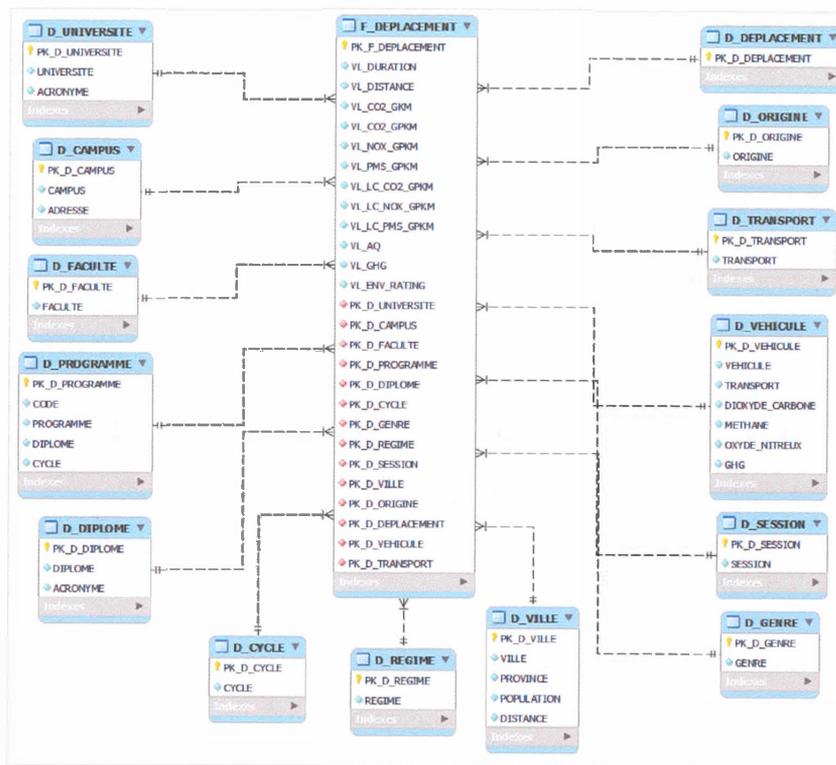


Figure 5.3 - Schéma en étoile des déplacements.

Chaque schéma sera responsable pour stocker les données des plusieurs sources et assurer l'intégration entre eux.

5.3.4 Transporter les données

Cette section présentera le processus ETL (« Extraction, Transformation and Load ») de la preuve de concept. Il a été créé avec l'objectif de transférer les données à partir des sources identifiées vers la base d'intégration multidimensionnelle. Comme cette preuve de concept ne vise pas à gérer un environnement de production, le processus ETL est concentré sur les tâches d'amener les données vers la destination sans s'occuper des détails d'exécution automatique ou même des optimisations. Ce

processus ETL est composé par les trois étapes majeures : l'extraction, les transformations et la charge.

L'outil 'Pentaho Data Integration' (PDI) a été utilisé pour concevoir les routines ETL et les exécuter. Cet outil a été choisi grâce à sa réputation dans le marché, à ses caractéristiques de logiciel libre et à ses fonctions qui facilitent l'ETL. En utilisant une interface visuelle, plusieurs composants sont disponibles pour extraire des données à partir d'une gamme de types de sources (fichiers, bases de données, Web services, etc.), les manipuler à travers des fonctions de transformation et les enregistrer avec plusieurs options de destinations.

Les explications sur le processus ETL seront faites dans trois parties, en utilisant des figures correspondantes à chaque dimension de soutenabilité.

5.3.4.1 Dimension économique

Après l'identification des sources et la création de la base d'intégration selon les principes d'un Entrepôt de Données, la première étape du processus de transport de données peut commencer. L'extraction consiste à collecter les données des diverses sources qui satisfont aux exigences des filtres de recherche requis. Selon les sources de la première itération identifiées dans le Tableau 5.1, les données de la dimension économique ont été extraites manuellement à partir des informations publiées dans le site Web de l'Université du Québec à Montréal.

La liste de programmes (UQÀM, 2016b) a été collectée dans une page Web (www.etudier.uqam.ca/programmes) de l'université et enregistré le fichier. Une routine d'ETL (Figure 5.4) était mise en œuvre avec PDI pour extraire de ce fichier les programmes, les facultés, les campus et les types de diplômes et les enregistrer dans leurs respectives tables de dimension. Quelques transformations de nettoyage des

champs et de suppression de doublons ont été faites pour garantir une bonne qualité des données.

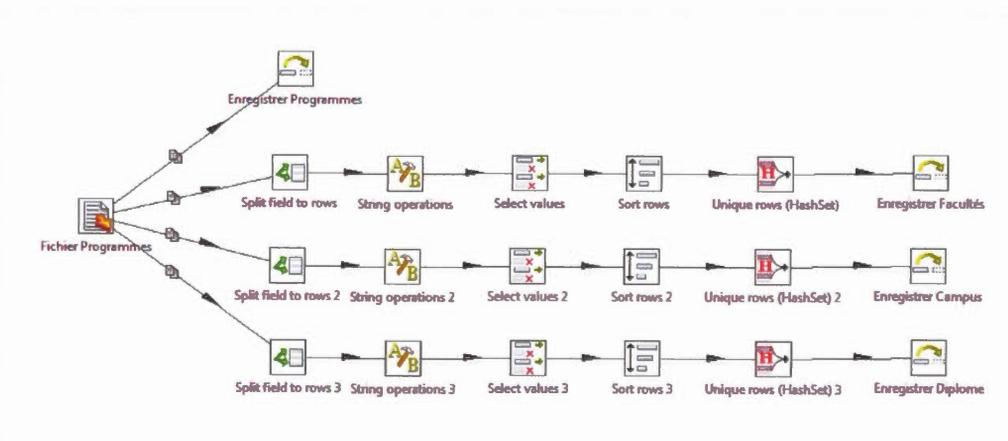


Figure 5.4 - ETL des programmes, facultés, campus et diplômes.

De façon similaire, les données relatives aux inscriptions sont trouvable dans les publications « La population étudiante de l'UQAM – Statistiques d'inscription » des périodes de 2014-2015 (UQÀM, 2015) et 2015-2016 (UQÀM, 2016a) disponibles sur le site Web de l'UQÀM. Ces documents en format PDF contiennent des renseignements par rapport les inscriptions faites par campus, facultés, programmes, cycles, diplômes, genres, régimes d'études et sessions entre les années 2014 et 2016.

La routine d'ETL des inscriptions est plus complexe que celle utilisée pour les dimensions. Quelques transformations ont été faites pour adapter les données des sources à la nouvelle structure de l'Entrepôt de Données (ou Data Warehouse – DW). Des petits changements ont été faits pour adapter les données extraites : adaptations de normalisation, modifications de formats, scission de champs, la suppression de caractères, nettoyage de champs, etc.

Dans ce cas, un « job » a été utilisé à la place d'une seule transformation. Un « job » est un ensemble de transformations qui sont exécutées selon un ordre de priorité. Selon

la Figure 5.5, l'activité « Créer la liste de fichiers » est responsable lire les données d'une liste de fichiers à traiter et passer ses données à la prochaine étape. Comme il y a deux fichiers à lire, ce processus sera exécuté au complet deux fois.

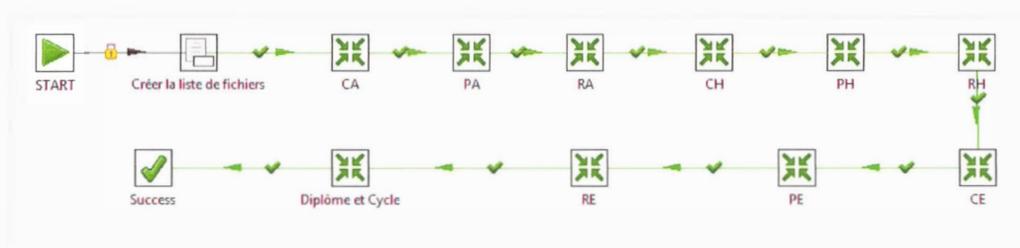


Figure 5.5 - ETL des inscriptions, composé par un "job".

Les étapes suivantes vont exécuter la même procédure plusieurs fois. La différence est que chaque transformation « CA, PA, RA, etc. » sera basée sur de différents paramètres d'entrée. Le nom de la transformation « CA » est un acronyme pour le régime et la session. Par exemple, « CA » est une transformation qui va extraire les données des inscriptions qui sont à temps complet (« C ») pendant la session d'Automne (« A »). La liste des acronymes au complet est : « Temps complet » (C), « Temps partiel » (P), « Rédaction de Thèse » (R) et pour « Automne » (A), « Hiver » (H) et « Été » (E). De cette façon, chaque étape (CA, PA, RA, etc.) représente l'exécution paramétrée de la transformation illustrée dans la Figure 5.6.

Cette transformation est responsable pour alimenter la table de faits « F_Revenu ». Après recevoir le fichier à traiter et les paramètres, elle va nettoyer les données pour les identifier et supprimer des caractères inappropriés. Ensuite, les dimensions seront consultées en utilisant les descriptions des champs pour trouver leurs clés primaires. Cette clé sera enregistrée dans la table de faits. À la fin, un ensemble de records seront enregistrés dans la table « F_Revenu ». C'est important de souligner que cette transformation sera exécutée plusieurs fois selon les paramètres utilisés.

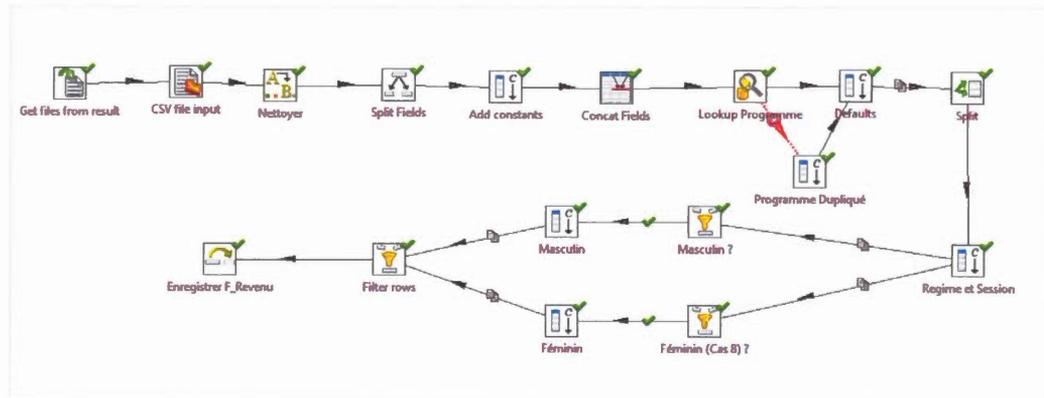


Figure 5.6 - ETL paramétré des inscriptions.

La dernière étape « Diplôme et Cycle » est une transformation (Figure 5.7) responsable pour découvrir le diplôme et le cycle pour les enregistrer dans la table de faits « F_Revenu ». Les documents des statistiques d'inscriptions ne contiennent pas ces informations. La découverte est faite à partir des programmes qui sont associés à un cycle et à un diplôme.

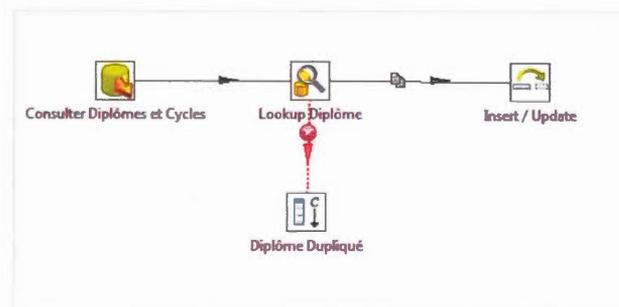


Figure 5.7 - ETL de diplômes et cycles.

Après les transformations, les données ont été stockées dans le schéma Revenu de la base d'intégration multidimensionnelle. Une tâche très importante dans la phase de charge (« Load ») est la gestion de création ou de mises à jour des données. En considérant plusieurs exécutions de l'ETL, si une donnée arrive, est-ce qu'elle devrait être créée ou mise à jour ? Comme l'objectif de la preuve de concept est de démontrer

l'application de la méthode, ceci n'impose pas un problème, car l'exécution se passe une seule fois. Mais, dans un cas réel, l'automatisation est faite selon la fréquence pertinente.

5.3.4.2 Dimension sociale

Les informations principales de cette dimension sont les indicateurs qui reflètent les perspectives des professions au marché du travail. Le premier pas est l'extraction de la Classification Nationale de Professions (Statistiques Canada, 2011) – CNP. À partir du site Web de la Statistiques Canada, c'est possible de télécharger un fichier sous le format CSV. La Figure 5.8 présente le processus ETL pour transporter les données et enregistrer les trois niveaux de la CNP sous la forme de dimensions et aussi la dimension Profession. Ces niveaux sont des catégories qui rassemblent des professions similaires.

Par exemple, « Arts, culture, sports et loisirs » est une catégorie de niveau 1 qui contient un sous-niveau 2 avec le « Personnel professionnel des arts et de la culture ». Ce dernier niveau est composé de plusieurs catégories (niveau 3) comme la « Professionnels/professionnelles des bibliothèques, des archives, des musées et des galeries d'art » et « Professionnels/professionnelles des arts plastiques et des arts de la scène ». Des professions comme « Musiciens/musiciennes et chanteurs/chanteuses » et « Danseurs/danseuses » sont quelques exemples de professions dans la dernière catégorie de niveau 3.

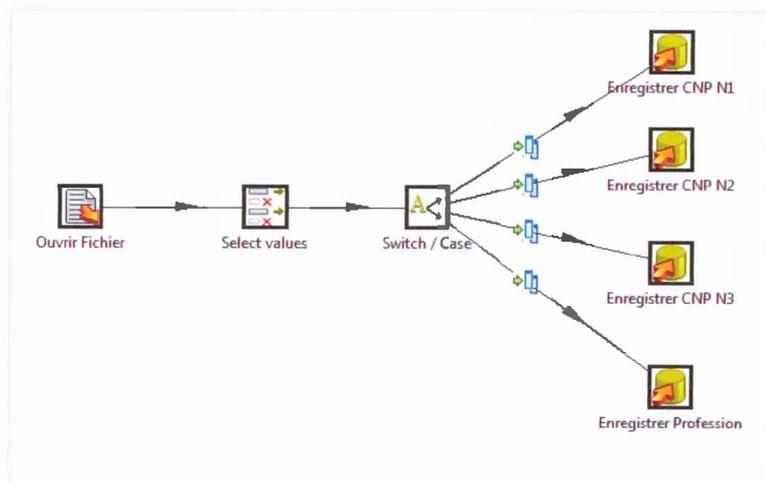


Figure 5.8 - ETL des professions.

Ensuite, après avoir enregistré les renseignements de la CNP et les professions, il manquait la création d'une liaison entre les professions et les programmes de formation. Le site Web de l'Emploi Québec contient les données manquantes pour relier les professions et les programmes de formation de plusieurs universités. Mais, à cause de quelques difficultés d'automatiser l'extraction, la collecte a été faite manuellement à travers la création d'un fichier texte CSV. La page Web de l'Emploi Québec utilisée est obtenue après à partir de la page d'accueil en cliquant sur le lien « Explorer un métier ou une profession » (Emploi Québec, 2016). La recherche a été faite en utilisant les paramètres avancés et après la sélection d'une catégorie professionnelle. Chaque profession est présentée avec les options de formation par université.

La routine d'ETL de la Figure 5.9 illustre la transformation utilisée pour traiter et enregistrer dans une table temporaire (« t_profession_programme » stockée dans une base de « Staging ») les correspondances entre les professions et les programmes de formation de l'UQÀM. En bref, les clés primaires des professions et des programmes ont été trouvées à partir des descriptions contenues dans ce fichier. Un ajustement a été

fait pour distinguer quelques noms identiques de programmes de différents cycles trouvés dans le site de l'UQÀM. Par exemple, le programme « Chimie » maintient le même nom pour les cycles de baccalauréat, maîtrise et doctorat.

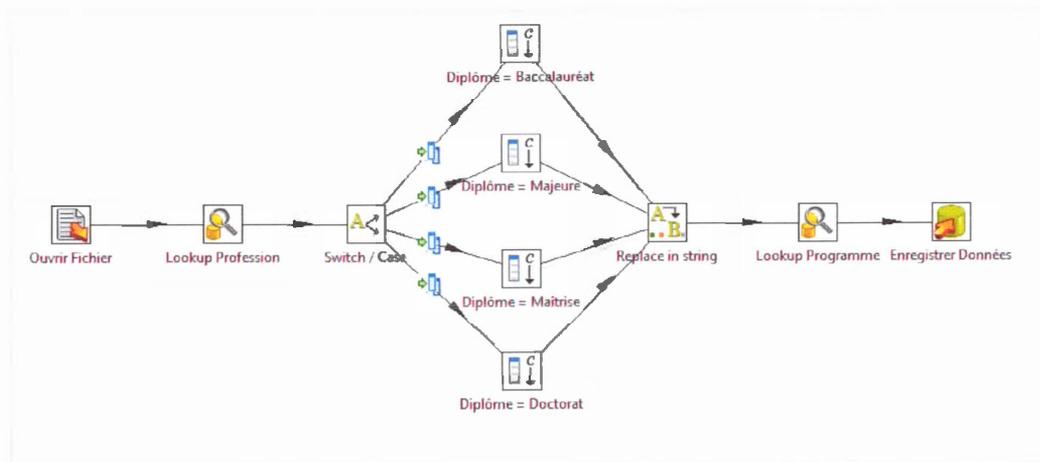


Figure 5.9 - ETL de la table temporaire de professions et formations.

D'autres données comme les possibles valeurs des perspectives (favorables, acceptables, restreintes et non publiées), des taux de chômage (faible, modéré, élevé) et de demande de main-d'œuvre (nul/négatif, faible, modéré, élevé), ont été manuellement insérées dans les dimensions correspondantes. Cette décision est basée sur le nombre d'options et la stabilité de ces valeurs.

Une dernière étape de l'ETL est l'alimentation de la table de faits des perspectives des professions à partir du document « Perspectives d'emploi par profession 2013-2017 » (Emploi Québec, 2014). Après la conversion de ce document PDF en CSV et quelques ajustements, l'ETL a traité et importé les données, selon la Figure 5.10. Les transformations étaient simples à faire, seulement la recherche des clés primaires des tables de dimensions et quelques ajustements des champs.

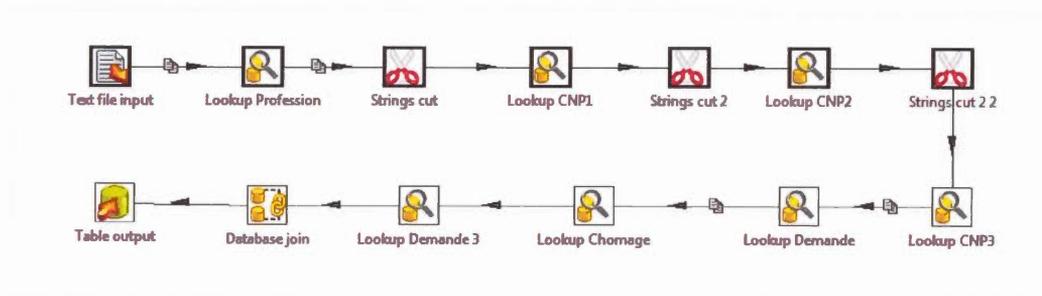


Figure 5.10 - ETL des perspectives de professions.

Alors, après ces routines d'ETL, les données de la dimension sociale ont été transportées des sources vers le schéma Perspective.

5.3.4.3 Dimension environnementale

La stratégie adoptée pour représenter la dimension environnementale est de comprendre les émissions causées par les déplacements des étudiants pour aller à l'université. Une phase initiale avec une base de « staging » a été utilisée pour stocker temporairement les données de régions, type de transports et leurs statistiques par rapport les émissions de gaz.

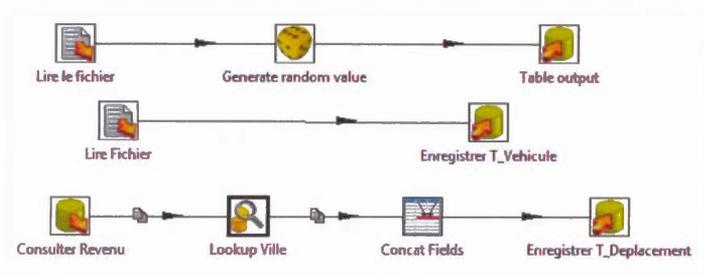


Figure 5.11 - ETL des régions, véhicules et déplacements ("staging").

La Figure 5.11 présente les trois routines d'ETL qui s'occupent de ces tâches. Les données des régions (Statistiques Canada, 2011b) ont été collectées pour générer les adresses aléatoires des étudiants. Les types de transport et leurs émissions ont été trouvés dans le document « Emissions by mode of transport » (en anglais) publié par le journal The Guardian dans leur site Web. Dans cette publication, une feuille de calcul avec les émissions a été collectée et importée par la routine d'ETL.

La table temporaire de déplacements a été générée à partir de la table de faits Revenu, en considérant le nombre d'inscriptions de chaque programme et session. Alors, cette transformation génère aléatoirement une région d'origine pour chaque étudiant. Ces données sont utilisées par une autre transformation que s'occupe de calculer les chemins et la durée de ces déplacements.

Les tables créées dans la base temporaire ont été utilisées par autres transformations pour alimenter les dimensions origine, déplacement, transport et véhicule. La Figure 5.12 rassemble les quatre routines indépendantes d'ETL. Elles sont simples et servent seulement à éliminer les doublons des tables temporaires. La dimension déplacement garde un numéro d'identification de chaque déplacement équivalent à un trajet fait par un étudiant vers l'université, par exemple.

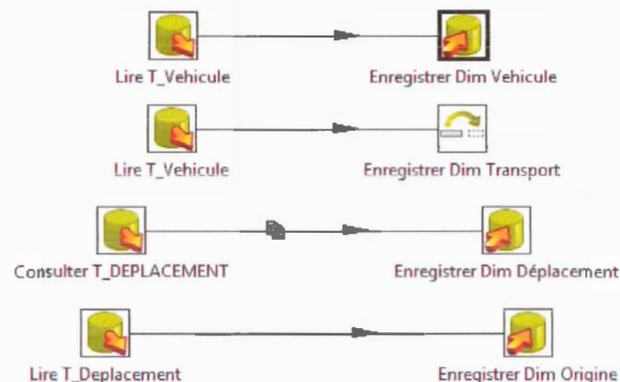


Figure 5.12 - ETLs de dimensions de véhicule, transport, déplacement et origine.

Mais la transformation la plus complexe de cette preuve de concept est responsable d'alimenter la table de faits de déplacements et calculer les indicateurs des émissions de gaz, représentée par la Figure 5.13. À partir des données hypothétiques des déplacements des étudiants, temporairement stockées dans la table « t_déplacement », les données approximatives des chemins à suivre pour arriver à l'université sont créées. Un type de déplacement est attribué par hasard à chaque déplacement, selon une règle de distribution fictive : 50% des déplacements en transports communs, 30% en voitures, 10% sur le vélo et 10% à pied.

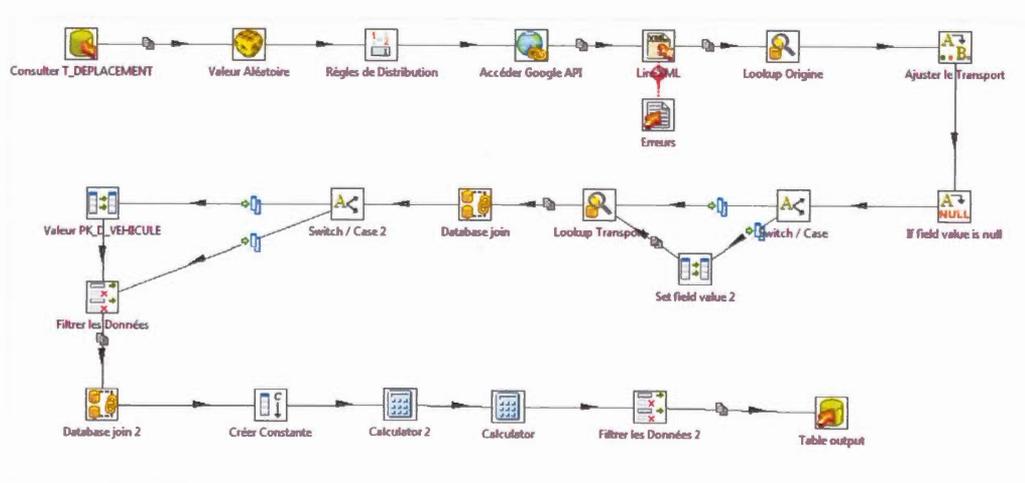


Figure 5.13 - ETL de déplacements et émissions de gaz.

Ensuite, la routine d'ETL se connecte avec les API du Google Maps pour obtenir le chemin en détail de chaque déplacement en format XML. Après avoir traité les réponses XML, les données de distance et durée sont collectées. Les clés primaires des dimensions sont identifiées et quelques calculs sont faits pour alimenter les indicateurs d'émission de gaz selon le type de transport de chaque déplacement. À la fin de la transformation, les données sont enregistrées dans la table de faits de déplacements.

5.3.5 Traiter les inconsistances

Cette preuve de concept n'a pas trouvé des inconsistances grâce à des petits ensembles de données et aux traitements qui ont été faits dans les documents extraits des sources. Mais l'avancement du projet probablement trouvera des inconsistances, surtout au moment d'intégrer des données potentiellement similaires.

Par exemple, le deuxième cycle du projet ira considérer les données de coûts et revenus des facultés et programmes qui sont distribuées dans différents systèmes d'information. Les inconsistances au niveau des indicateurs peuvent aussi apparaître au moment de comparer des subventions et les confronter avec les données approximatives calculées à partir des nombres d'inscriptions.

5.4 Module 3 : découverte de la connaissance

Le Module 1 a défini deux scénarios décisionnels de cette preuve de concept et les indicateurs ont été identifiés. Le premier scénario vise à supporter les ajustements des programmes de formation et le deuxième, de donner les éléments nécessaires pour une prise de décision d'un possible partenariat avec la STM pour l'utilisation d'un autobus moins polluant par les étudiants.

5.4.1 Comprendre les liaisons entre les faits et dimensions

Cette étape de la méthode vise à identifier les possibles relations entre plusieurs faits, en utilisant des dimensions conformes qui sont les dimensions qui font partie de deux ou plusieurs schémas. La Section 5.3.3 a montré les trois schémas de cette preuve de concept, F_Revenu, F_Déplacement et F_Perspective, qui seront responsables pour

stocker les données sous un format multidimensionnel dans la base d'intégration. Alors, l'objectif est de comprendre les points d'intersection entre les schémas et de discuter les possibilités de regroupement des données stockées dans la base d'intégration.

Tableau 5.2

Liste des dimensions conformes

Dimension	F_Revenu	F_Perspective	F_Déplacement
Temps	X	X	X
Session	X	-	X
Université	X	-	X
Campus	X	-	X
Faculté	X	-	X
Programme	X	X	X
Diplôme	X	-	X
Cycle	X	-	X
Régime	X	-	X
Genre	X	-	X
Ville	X	-	X

Le Tableau 5.2 rassemble une liste des dimensions conformes qui sont présentes au moins dans deux parmi les trois schémas multidimensionnels. La dimension Temps fait partie, en général, de tous les schémas d'une solution en intelligence d'affaires. Ainsi, tous les champs du type date font référence à la dimension Temps et ils peuvent être alignés dans une même perspective temporelle. De la même façon, la dimension programme est importante grâce à sa présence dans tous les schémas.

Plusieurs dimensions conformes sont trouvées dans les schémas F_Revenu et F_Déplacement. Cela signifie que leurs faits peuvent être intégrés quand ils sont basés sur ces dimensions. Par exemple, si un gestionnaire veut analyser la quantité d'inscriptions d'une faculté, il peut aussi croiser ces données avec le volume d'émissions ou la durée totale liée aux déplacements des étudiants de la même faculté.

L'explication de cette particularité de modélisation est que les déplacements sont liés aux étudiants qui sont inscrits aux programmes. Alors, au moment de la conception de la solution d'intelligence d'affaires, les déplacements ont été créés sans enregistrer les étudiants, mais en utilisant des dimensions de genre, régime d'études et la ville de l'étudiant. De manière similaire, les références à l'université, le campus, la faculté, le type de diplôme et le cycle des programmes de chaque étudiant ont été copiés dans la dimension F_Déplacement.

Cependant, le schéma F_Perspective contient seulement deux dimensions conformes avec les autres schémas : les dimensions Temps et Programme. Alors, ces deux dimensions sont les points d'intégration entre F_Perspective et les autres schémas. C'est possible encore de croiser ce schéma avec les autres en utilisant les dimensions comme cycle ou diplôme.

La preuve de concept a présenté ces cas pour illustrer deux types de modélisation différents. Les avantages de l'approche de réplique des dimensions, comme vues dans les schémas F_Revenu et F_Déplacement, sont une meilleure performance et aussi la capacité de garder l'historique par rapport les changements des relations entre les dimensions. Par contre, la modélisation de F_Perspective est plus simple à maintenir.

Enfin, pour chercher les faits d'un schéma par ses dimensions, l'utilisateur doit simplement mettre le fait et les dimensions dans un rapport ou d'autre type de présentation de l'information. Mais, s'il veut évaluer un fait selon une dimension d'un

autre schéma, il faut comprendre la sémantique de la relation entre leurs dimensions. Par exemple, si un gestionnaire vise à savoir la quantité d'emplois de la profession de danseur dans la ville de Beauharnois, il doit croiser le fait de F_Perspective avec le schéma F_Revenu ou F_Déplacement. Pour relier ces deux schémas par la dimension conforme de programme, la quantité d'emplois sera mal calculée, car un programme peut être lié à plusieurs villes des étudiants. Par contre, le calcul sera correct s'il cherche la quantité d'emplois par faculté parce qu'un programme fait partie d'une seule faculté.

5.4.2 Analyser les données

La preuve de concept exposera chaque scénario et quelques modèles de présentation pour aider aux processus décisionnels. La qualité et la précision sont dépendantes des limites imposées par les sources de données.

5.4.2.1 Scénario 1

Le premier cas est un scénario décisionnel où les gestionnaires veulent utiliser des informations pour effectuer des ajustements par rapport les programmes de formation d'une université. L'idée est de créer un ensemble d'artefacts pour présenter les informations selon l'intérêt des utilisateurs.

Un tableau de bord interactif est proposé dans la Figure 5.14 pour guider les gestionnaires dans l'analyse des programmes. L'objectif est de fournir un mécanisme d'évaluation capable de mettre en évidence les programmes qui sont plus susceptibles d'avoir une mauvaise performance par rapport les autres programmes.



Figure 5.14 - Tableau de bord des programmes.

Cet artefact se concentre à montrer les comparaisons entre les programmes au lieu de fournir des détails sur chaque programme, car l'objectif est d'identifier dans un premier temps les programmes moins performants qui seront analysés en profondeur.

Le tableau de bord est composé par sept sections. Quelques titres sont écrits sans les accents à cause des limitations de l'outil. La section « Programmes » se trouve sur le coin supérieur gauche et elle amène la liste de programmes de la Faculté d'Arts et le type de diplôme à obtenir. L'interactivité se passe au moment de choisir un type de diplôme ou un programme et toutes les autres sections sont filtrés. Dans ce cas, le choix était le diplôme du type baccalauréat de la Faculté d'Arts. Ainsi, les données présentées par les autres sections sont limitées à ces paramètres.

La section « Évolution des Inscriptions », placée en haut dans la colonne centrale, présente un diagramme à barres avec le nombre total d'inscriptions de chaque session des programmes sélectionnés dans la section « Programmes ». Ce diagramme montre que tous les programmes de cette faculté ont présenté des réductions par rapport aux

valeurs dans chaque session en comparaison avec l'année précédente de cet indicateur total.

Par rapport la comparaison entre les programmes, ils sont classés dans la section « Nombre d'Inscriptions » selon l'indicateur de quantité d'inscriptions pendant la période de 2014 à 2015. La capacité de créer un bon revenu commence avec la quantité d'étudiants. Le programme d'Art Dramatique est celui qui a le moins d'inscriptions. Cette constatation met en évidence ce programme qui se pose comme un candidat pour un possible ajustement.

Mais, la section « Emplois par Programmes » montre que les professions liées à la formation en Art Dramatique sont à la deuxième place en termes de nombre d'emplois requis par le marché. Cet indicateur est incompatible avec le nombre d'inscriptions et peut mener à réviser le programme pour évaluer la qualité et pour comprendre ce comportement des inscriptions.

Le « Revenu par Programme » aborde les salaires moyens des professionnels qui ont des formations dans chaque programme. La comparaison entre la performance de cet indicateur pour chaque programme présente l'Art Dramatique bien placée avec un bon salaire moyen.

Le taux d'augmentation des nombres d'inscriptions de 2015 par rapport la période précédente est illustré par la section « Comparaison des Inscriptions par Année ». C'est possible de remarquer que la plupart des programmes ont enregistré des taux de réduction allant jusqu'à 65% du nombre d'inscriptions. L'Art Dramatique a baissé en 47% ses inscriptions entre les périodes comparées.

Enfin, la pollution générée par les programmes est mesurée dans la section « Pollution par Programme ». L'indicateur utilisé est le volume d'émissions de monoxyde de carbone dans l'atmosphère, calculé par les déplacements des étudiants avec des types différents de transport. Les performances des cours de Design Graphique et Design de

l'Environnement sont les plus nocives pour la qualité de l'air. L'Art Dramatique figure à la quatrième place en termes de cet indicateur avec 12% de l'impact total des programmes de baccalauréat de la Faculté d'Arts.

Avec ces informations dans tableau de bord assez simple, les gestionnaires peuvent faire des inférences par rapport les programmes de formation. En analysant la dimension économique du programme d'Art Dramatique, il y a une impression que le programme devrait être abandonné ou fusionné avec un autre pour améliorer sa performance économique. Mais, les perspectives des carrières liées à ce programme montrent que le nombre d'emplois et le revenu moyen est bien placé par rapport les autres programmes. Cela souligne le besoin d'intervenir dans ce programme pour l'améliorer et, ainsi, faire monter le nombre d'inscriptions.

PROGRAMME	DIPLOME	SESSION					
		Autonne		Hiver		Été	
		2014	2015	2015	2016	2014	2015
		QT INSCRIPTIONS					
Art dramatique	Baccalauréat	66.00	54.00	67.00	30.00	28.00	1.00
Arts visuels et médiatiques	Baccalauréat	293.00	156.00	245.00	59.00	48.00	6.00
	Maîtrise	53.00	43.00	39.00	18.00	41.00	11.00
Danse	Maîtrise	39.00	11.00	19.00	8.00	30.00	3.00
	Baccalauréat	113.00	183.00	105.00	70.00	1.00	8.00
Design de l'environnement	Maîtrise	19.00	13.00	17.00	7.00	15.00	3.00
	Baccalauréat	235.00	152.00	216.00	76.00	79.00	10.00
Enseignement des arts	Maîtrise	37.00	12.00	24.00	4.00	25.00	2.00
	Baccalauréat	134.00	61.00	125.00	42.00	46.00	1.00
Histoire de l'art	Doctorat	31.00	12.00	21.00	6.00	26.00	4.00
	Maîtrise	71.00	25.00	47.00	19.00	50.00	15.00
Musique	Baccalauréat	104.00	75.00	95.00	46.00	7.00	16.00
Muséologie	Maîtrise	38.00	16.00	33.00	2.00	15.00	1.00
Muséologie, médiation et patrimoine	Doctorat	1.00	1.00	NA	1.00	3.00	1.00
	Sémiologie	42.00	21.00	41.00	22.00	33.00	20.00
Théâtre	Maîtrise	63.00	33.00	42.00	22.00	36.00	11.00
	Doctorat	100.00	49.00	65.00	38.00	97.00	27.00
Études et pratiques des arts	Baccalauréat	470.00	238.00	426.00	149.00	114.00	23.00
	Doctorat	53.00	21.00	36.00	20.00	48.00	16.00
Études littéraires	Maîtrise	178.00	120.00	111.00	94.00	137.00	53.00
	Total	2,140.00	1,296.00	1,774.00	733.00	879.00	232.00

Figure 5.15 - Un rapport des inscriptions.

Une autre constatation importante est la réduction de presque toutes les inscriptions dans les programmes de cette faculté entre les deux périodes considérées. Mais, pour bien comprendre le scénario, les gestionnaires peuvent demander plus d'informations. Par exemple, un rapport avec les inscriptions de chaque programme par session et année. Ce rapport est illustré par la Figure 5.15. Avec ces données, les décideurs peuvent constater qu'il y a eu des réductions majeures pendant l'Été. Par exemple, le programme d'Art Dramatique dans l'Été 2014 avait 28 nouvelles inscriptions. Mais, pendant l'Été de 2015, seulement une inscription a été faite : une baisse de 96% du nombre d'inscriptions. À partir de cette nouvelle information, les prises de décisions seraient plus précises.

C'est important de souligner que cette preuve de concept se détient à un petit échantillon de données en comparaison avec les approches réelles d'intelligence d'affaires. Mais, même avec cette limitation, quelques découvertes ont été effectuées et une nouvelle façon de gérer les trois dimensions de la soutenabilité a été proposée.

5.4.2.2 Scénario 2

Le processus décisionnel du deuxième scénario est basé sur l'intention de l'université de signer des accords avec des entreprises de transport collectif qui gèrent des lignes des villes des banlieues vers Montréal. L'objet de cet accord est la création des lignes avec les autobus électriques pour optimiser les temps de déplacements des étudiants et réduire les émissions de gaz polluants. Le but majeur de ce scénario décisionnel est d'identifier les villes où l'impact de ce projet serait plus pertinent selon les dimensions de soutenabilité.

La Figure 5.16 représente un tableau de bord qui sert à classer les villes où les étudiants habitent selon le nombre d'inscriptions, la durée et l'impact sur la qualité de l'air des déplacements des étudiants. Ce tableau de bord interactif, comme celui présenté dans

le scénario 1, est composé par six sections et à chaque sélection, les contenus des autres composants sont filtrés. Il y a trois sessions ci-dessus qui s’occupent de la comparaison entre les villes. Les autres trois sessions qui se posent au-dessous des celles premières, visent à montrer les détails de chaque ville sélectionnée.

La section « Inscriptions par Ville » utilise un diagramme à barres pour classer les douze villes (Top 12) qui ont plus d’étudiants inscrits à cette université. Dans ce cas, les villes de Boisbriand, de Beauharnois et de l’Épiphanie sont les leaders en termes d’inscriptions. Cette liste de villes sera l’entrée utilisée par les sections de durée et pollution par ville. Cette stratégie vise à atteindre un nombre considérable d’étudiants, mais les parties sociale et environnementale apporteront d’autres éléments pour l’analyse.

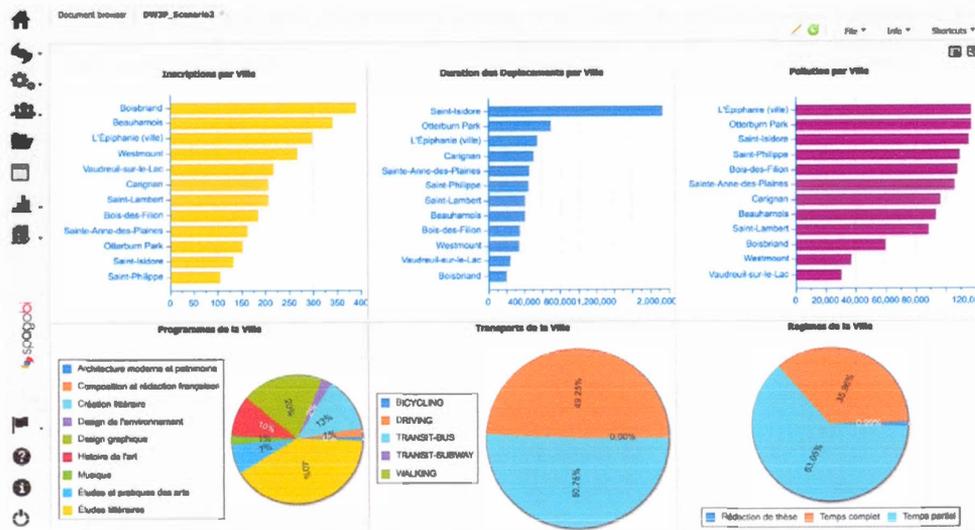


Figure 5.16 - Tableau de bord des villes.

La prochaine section, la « Durée des Déplacements par Ville », utilise le même principe de la première, mais dans ce cas l’indicateur évalué sera la durée de temps pour l’étudiant sortir de sa maison et arriver à l’université. La classification des villes

par la durée des déplacements présente les villes Saint Isidore, Otterburn Park et l'Épiphanie. Mais, le nombre d'inscriptions modeste peut suggérer que la distance a influencé la performance de Saint Isidore et Otterburn Park. Dans le cas de l'Épiphanie elle figure bien aussi dans le nombre d'inscriptions.

La section « Pollution par Ville » est responsable pour calculer les émissions du monoxyde de carbone (CO₂) résultantes des déplacements. Bien que plusieurs indicateurs soient disponibles dans le schéma F_Déplacement, seulement le monoxyde de carbone est considéré dans cette situation. Les villes bien classées par l'indicateur de durée sont aussi dans les mêmes trois premières places.

Ces trois premières sections sont responsables pour trouver les meilleures options pour démarrer un projet de test. L'Épiphanie est un bon choix, car elle détient beaucoup d'inscriptions, la durée des déplacements est élevée et, par conséquent, la pollution est considérable. Mais, une autre ville est Carignan qui figure dans les positions médianes et qu'elle pourrait être choisie par les gestionnaires.

Après choisir une ville dans les trois premières sections, les sections ci-dessous sont responsables pour montrer la distribution des programmes, types de transport et régimes d'études de la ville. C'est un ensemble d'indicateurs plus détaillés, au niveau d'une ville choisie.

La section « Programmes de la Ville » fait une association entre la ville et les programmes de formation des étudiants de cette ville. Les types de transport utilisés sont présentés par la section « Transport de la Ville ». Enfin, la section « Régimes de la Ville » montre la distribution des types de régimes (temps complet, partiel ou rédaction de thèse).

Dans le cas de Carignan (Figure 5.16), une bonne diversité de programmes de formation de la Faculté d'Arts peut être trouvée. Au-delà les critères de nombre d'inscriptions, la diversité des programmes est un avantage, car le projet serait plus

connu et l'image de l'université serait diffusée plus rapidement dans les programmes. Au niveau des déplacements, les types de transport plus fréquents sont la moitié des autobus et l'autre moitié des voitures. Les statistiques d'utilisation de 50% des voitures peuvent être un avantage en visant la captation des nouveaux clients qui utiliseront les services spécialisés des autobus électriques. À la fin, la pollution est considérable et, ainsi, le projet aura un impact important et positif sur l'environnement. Alors, ces trois facteurs rendent faisable un possible projet de test dans les villes de l'Épiphanie et Carignan.

5.4.3 Supporter les prises de décisions

Les deux scénarios décisionnels ont trouvé de bons résultats dans ce premier cycle. Le scénario 1 a utilisé les indicateurs de nombre d'inscriptions, quantité d'emplois, nombre d'emplois disponibles, revenus et pollution pour évaluer les programmes et guider vers les décisions des ajustements des programmes. Le scénario 2, en utilisant des indicateurs de nombre d'inscriptions, durée des déplacements et la qualité de l'air pour supporter la décision de définir les villes qui feront partie d'un projet pilote pour créer des lignes spéciales avec les autobus électriques. Les prochains cycles seront responsables pour améliorer la précision et les options d'informations disponibles. D'autres types de rapports, tableaux de bord et cubes devraient être créés pour supporter aussi l'accès aux données plus détaillées.

Au niveau technique, c'est intéressant de motiver une discussion avant le prochain cycle sur la modélisation du schéma F_Perspectives pour ajouter des dimensions conformes de F_Revenu et F_Déplacement. Cette action permettra de relier plus facilement les faits et dimensions.

5.5 Les résultats obtenus

Cette preuve de concept visait appliquer la méthode proposée dans le Chapitre 4 pour comprendre les influences liées aux deux scénarios proposés. Après la définition de chaque scénario, les indicateurs ont été identifiés selon les approches de soutenabilité. Quelques indicateurs spécifiques ont été créés.

La stratégie adoptée a choisi un petit ensemble d'indicateurs dans les dimensions économique, sociale et environnementale. Les évolutions ont été programmées par cycles en espérant les résultats plus rapidement.

Une base d'intégration a été créée avec trois schémas multidimensionnels pour recevoir les données provenant des sources internes et externes. Un processus d'Extraction, Transformation et Charge (ETL) a été créé pour automatiser les mises à jour des données, mais une partie a été faite manuellement. Les données ont été alimentées dans la base d'intégration. Les inconsistances n'ont pas été trouvées, mais les tests étaient simples et pour une bonne qualité, il faut les approfondir.

Quelques tableaux de bords et rapports ont été créés avec des indicateurs de trois dimensions pour comprendre les relations entre eux et pour analyser les informations. À la fin, les prises de décisions des deux scénarios étaient plus complètes et basées sur l'information.

Par conséquent, l'application de la méthode proposée dans cette recherche et testée par cette preuve de concept peut amener les entreprises à considérer les trois dimensions de soutenabilité aux processus décisionnels.

5.6 Renseignements techniques

Les logiciels utilisés dans la preuve de concept sont tous libres et avec des codes sources ouverts (« Open Source »). La décision d'utiliser ce type de logiciels est pour

montrer qu'une solution d'application de soutenabilité en utilisant l'intelligence d'affaires peut être créée sans dépenser beaucoup d'argent avec l'achat des outils propriétaires. C'est sûr qu'il y a des outils qui sont plus puissants et qui offrent plus de fonctionnalités, mais une première version de la solution peut être créée avec des outils libres et, le cas échéant, une migration pourra être appliquée au futur.

Le SGBD ou le Système de Gestion de Base de Données (en anglais, « Database Management System – DBMS ») utilisé a été le MySQL Server 5.6, un logiciel facile à utiliser et très connu parmi les options du marché. L'outil MySQL Workbench 6.3 a été utilisé comme un client pour modéliser les bases de données.

Les routines ETL ont été créées dans le Pentaho Data Integration 6.1, un outil très complet et stable pour faire la conception et exécution de l'ETL. Et, finalement, l'outil utilisé pour la création des rapports et des tableaux de bord est le Spago BI 5.2.

CHAPITRE VI

DISCUSSIONS ET CONCLUSION

Ce chapitre revient sur chaque section de cette recherche pour motiver une discussion sur les concepts exposés dans ce mémoire, ainsi qu'une conclusion sur les impressions du chercheur par rapport le sujet développé.

6.1 Discussions

La soutenabilité est un sujet très pertinent à cause de l'importance des actions des entreprises et leurs impacts sur la planète. Mais, la soutenabilité ne se préoccupe pas seulement de l'environnement. Elle vise établir un équilibre entre l'économie, le social et l'environnement. Alors, ce terme est actuellement interprété dans ces trois dimensions.

La problématique souligne des difficultés de choisir les approches et les indicateurs selon le domaine d'une entreprise à cause de la variété d'options disponibles. Comme les approches sont souvent utilisées pour créer des rapports de soutenabilité, elles contiennent de nombreux indicateurs des trois dimensions. Mais les recommandations standards des rapports ne sont pas suffisantes pour expliquer l'intégration entre les indicateurs et pour interpréter leurs possibles relations cachées. L'intelligence d'affaires est une bonne technique pour résoudre ces problèmes, mais son application est encore peu connue à la fois des affaires et dans la littérature.

La question principale de cette recherche est « En quoi les approches en intelligence d'affaires sont-elles utiles et pertinentes pour générer de la connaissance stratégique de soutenabilité au-delà de l'identification et de l'intégration des indicateurs économiques, sociaux et environnementaux ? ». Et pour répondre, l'objectif majeur de cette recherche est de proposer une méthode d'application de la soutenabilité au sein des entreprises en utilisant des approches d'intelligence d'affaires, permettant de sélectionner et d'intégrer les indicateurs économiques, sociaux et environnementaux pertinents à une entreprise pour générer et utiliser des connaissances stratégiques.

Pour bien comprendre les sujets impliqués, la Revue de la Littérature a exposé les concepts de soutenabilité et les approches d'application plus célèbres, qui sont utilisées principalement pour générer les rapports de soutenabilité. Ces rapports visent à améliorer la réputation des entreprises face aux clients et investisseurs, mais ils n'abordent pas les relations entre les indicateurs.

La technologie de l'information est utilisée pour automatiser, informer et transformer (Seidel et al., 2015) les entreprises et les systèmes d'information sont des outils très bien adaptés pour ces objectifs. L'intégration des données en utilisant l'intelligence d'affaires est une stratégie commune dans la littérature. Bien que les deux sujets soient souvent abordés séparément, mais il n'y a pas beaucoup d'articles qui discutent les deux sujets ensemble.

La Science du Design (SD) a été choisie comme la stratégie pour présenter la méthode d'application de la soutenabilité, proposée par cette recherche. La SD est très souvent utilisée dans les recherches liées aux systèmes d'information. L'étape de conception et développement a été effectuée avec une liste d'étapes qui décrivent la méthode proposée. Ensuite, une preuve de concept a démontré l'applicabilité de la méthode dans le but de faire émerger la connaissance des indicateurs de soutenabilité.

La méthode est divisée en trois modules. Le premier module définit ce qu'il faut faire pour appliquer la soutenabilité. La définition des décisions à prendre est, selon la méthode, le pas principal pour guider à une bonne sélection des indicateurs et des approches. Ensuite, la stratégie d'exécution va définir les cycles itératifs pour l'accélération d'obtention des résultats et des évolutions futures. Ce module peut répondre à l'une des questions secondaires de recherche pour découvrir les approches et indicateurs les plus pertinents d'une entreprise. À partir des décisions à prendre et la considération des trois dimensions de soutenabilité, les indicateurs seront toujours alignés aux objectifs stratégiques de l'entreprise.

Le Module 2 liste les étapes d'adaptation de systèmes d'information et d'intégration des données. L'identification des sources internes ou externes permet de savoir où sont les données qui supportent les indicateurs. Les données qui n'existent pas déjà dans les sources seront stockées après les adaptations dans les systèmes d'information. Une base d'intégration multidimensionnelle, modélisée selon les principes des Entrepôts de Données (un « Data Warehouse »), a été proposée visant la convergence de toutes les données dans les schémas en étoile ou en flocon de neige. Le transfert des données est exécuté par les routines de l'ETL (Extraction, Transformation et Charge) qui copient les données des origines vers la base d'intégration. Les inconsistances sont des conséquences naturelles des processus d'intégration et doivent être traitées pour augmenter la précision et la qualité des informations. Une autre question secondaire sur comment les indicateurs sociaux et environnementaux peuvent être intégrés aux indicateurs corporatifs est répondu par cette partie de la méthode.

Le dernier module est basé sur la compréhension des connexions entre les faits et dimensions et sur l'analyse de données pour promouvoir l'extraction de la connaissance à partir des indicateurs de soutenabilité. Des ressources et des techniques d'intelligence d'affaires permettent d'analyser les données et les transformer en

information. Mais, il faut toujours vérifier si les informations obtenues sont encore pertinentes pour les prises de décisions définies dans le premier module.

À ce moment, la question principale de cette recherche peut être répondue : les approches d'intelligence d'affaires sont parfaitement capables de générer de la connaissance stratégique par rapport la soutenabilité et ses trois dimensions. Pour attester les réponses aux questions de recherche, une preuve de concept a été développée avec deux scénarios décisionnels hypothétiques.

Dans le Scénario 1, les gestionnaires voulaient ajuster les programmes de formation universitaire en considérant les trois dimensions de soutenabilité. Le Scénario 2 se concentrait sur un possible partenariat entre l'université et des entreprises de transport collectif pour créer des lignes avec des autobus électriques qui seraient moins polluants que les véhicules conventionnels.

L'application du Module 1 de la méthode proposée a bien identifié les décisions à prendre pour les deux scénarios, aussi bien que les indicateurs. L'approche de soutenabilité utilisée était le G4 Guidelines qui est complète, simple et très didactique. Mais, d'autres approches sont aussi applicables. Le plus important est de toujours aligner les processus décisionnels avec les indicateurs pour assurer que les efforts soient efficaces. À la fin, la stratégie d'exécution a permis de créer des cycles itératifs d'une façon à accélérer l'obtention des résultats qui pourront changer les prochains cycles et même les décisions à prendre.

Le Module 2 a guidé les tâches d'identification des sources fiables des données pour alimenter les indicateurs. Les adaptations sur les systèmes d'information pour recevoir de nouvelles données non considérées ont montré que des petits changements peuvent élargir les horizons de l'analyse de données. Mais, cet élargissement a été canalisé vers un Entrepôt de Données pour intégrer tous les variables et concepts. De cette façon, les données avant pulvérisées ont été connectées selon leurs significations. Encore, les

routines d'ETL ont permis le transfert des données entre les sources et cette base d'intégration.

À ce stage, le Module 3 a été introduit pour transformer les données en information valable à l'aide des outils d'analyse de données. En utilisant des tableaux de bord interactifs, la découverte de la connaissance a émergé à partir des données brutes sous la forme des graphiques interconnectés. Une expérience plus dynamique a été présentée pour les deux scénarios et les premières inférences ont été réalisées. Alors, beaucoup d'intelligence peut être découverte même avec une petite analyse des indicateurs de soutenabilité.

6.2 Difficultés rencontrées et leçons apprises

Normalement, les gens pensent à la soutenabilité seulement sous la perspective de l'environnement. Mais, la première leçon apprise avec cette recherche est que la soutenabilité considère aussi les dimensions sociale et économique. La principale question de cette recherche a été confirmée, quand l'utilisation des approches d'intelligence d'affaires pour appliquer la soutenabilité au sein des entreprises est vraiment efficace. C'était une leçon apprise qui renforce l'importance du sujet de la soutenabilité dans les processus décisionnels. Et aussi la découverte de la connaissance peut mener aux prises des décisions plus responsables et basées sur une vision plus large du business et du monde.

Les données statistiques sont disponibles et très utiles pour comprendre les concepts de soutenabilité. Pendant la preuve de concept, la recherche de ces données ouvertes et la découverte d'autres informations ont créé de nouvelles idées pour incrémenter des scénarios, mais pour maintenir la portée, ces informations n'ont pas été utilisées.

Au niveau de l'expérience pratique obtenue avec cette recherche, une leçon apprise a été la compréhension de la soutenabilité pour nos vies et la préservation de notre planète. La façon aveugle de conduire les affaires par rapport les autres dimensions en considérant seulement la dimension économique ne permet pas d'analyser nos actions sur l'environnement et sur la vie des gens.

Quelques difficultés ont été rencontrées, comme de trouver des données sous la même perspective temporelle. Parfois, une partie des données étaient disponibles sur une période de temps différente de l'autre. Alors, quelques adaptations et limites ont été imposées. Les formats des données à partir des origines ont été une autre difficulté à surmonter. Quelques pages Web ou des feuilles de calcul ont été traitées manuellement. Ces types de problèmes dans un environnement de production peuvent causer des lenteurs de traitement de l'information. La recherche des sources dans les formats plus faciles à traiter peut apporter des avantages.

Même avec un peu d'expérience du chercher dans le domaine de l'intelligence d'affaires, quelques défauts dans la création de l'Entrepôt de Données ont causé des retards dans la construction de la preuve de concept.

Un autre problème a été observé à cause des quelques défauts avec les logiciels libres utilisés dans la preuve de concept. Soit par incompatibilités avec les systèmes d'exploitation ou des erreurs d'opérations, quelques actions pour les corriger ou même pour ne pas utiliser des fonctions ont été effectuées. Mais, en général, tous les objectifs de la recherche ont été atteints.

6.3 Limites

La stratégie de cette recherche d'utiliser des données ouvertes, principalement pour les données sensibles (les adresses des étudiants, les indicateurs économiques de

l'université, etc.) par rapport la preuve de concept, a limité ce travail. Mais ces types de données des entreprises sont toujours difficiles à collecter, car ils sont très stratégiques. Par contre, l'applicabilité de la méthode n'était pas affectée par cette limitation. Bien que la preuve de concept soit aussi limitée au contexte de l'éducation universitaire, la méthode peut être utilisée dans n'importe quel domaine.

6.4 Contribution

L'utilisation de la méthode a démontré que l'intelligence d'affaires peut faciliter l'application de la soutenabilité dans un contexte d'éducation universitaire. C'est une contribution importante dans un scénario peu exploré par la littérature. Mais, la contribution majeure de cette recherche a été d'approximer la soutenabilité et les approches d'intelligence d'affaires dans le but d'extraire de la connaissance à partir des indicateurs de dimensions économique, sociale et environnementale.

6.5 Recherches futures

Une gamme de possibles évolutions peuvent être proposées à partir de cette recherche. Au niveau de la problématique, une analyse des difficultés avec l'utilisation des approches d'intelligence d'affaires pourrait être ajoutée à la liste de problèmes proposés.

Sur la méthode, l'information concernant les rôles et les profils de chaque type de professionnel impliqué dans les étapes pourraient incrémenter la méthode et faciliter son application selon un domaine spécifique. Par contre, l'applicabilité serait moins diverse.

Au niveau de la preuve de concept, plusieurs recherches peuvent se produire. Un exemple est de considérer aussi les inscriptions de plusieurs années ou encore d'autres

universités pour comprendre un scénario plus large. Faire une étude de cas avec les données réelles d'une ou plusieurs universités. Calculer les subventions fournies par les gouvernements aux organismes universitaires. Peut-être, rechercher les relations entre la qualité de l'air à l'intérieur d'une entreprise avec l'absentéisme des employés. Ou, enfin, la création d'une application de Data Mining serait très intéressante pour illustrer, par exemple, le point optimal entre la distance parcourue et le nombre d'inscriptions dans la décision proposée pour le Scénario 2.

BIBLIOGRAPHIE

1. Adams, C. A., & Frost, G. R. (2008). Integrating sustainability reporting into management practices. In *Accounting Forum* (Vol. 32, No. 4, pp. 288-302). Elsevier.
2. Agiu, D., Mateescu, V., & Muntean, I. (2014). Business Intelligence overview. *Database Systems Journal*, 5(3), 23-36.
3. Ahmed, M. D., & Sundaram, D. (2012). Sustainability modelling and reporting: From roadmap to implementation. *Decision Support Systems*, 53(3), 611-624.
4. Aras, G., & Crowther, D. (2009). Corporate sustainability reporting: a study in disingenuity?. *Journal of business ethics*, 87(1), 279-288.
5. Bara, A., Botha, I., Diaconita, V., Lungu, I., Velicanu, A., & Velicanu, M. (2009). A model for business intelligence systems' development. *Informatica Economica*, 13(4), 99.
6. Brundtland, G. H. (1987). Report of the world commission on environment and development: our common future. United Nations.
7. Bucher, T., Gericke, A., & Sigg, S. (2009). Process-centric business intelligence. *Business Process Management Journal*, 15(3), 408-429. doi:10.1108 / 14637150910960648
8. Curkovic, S., & Sroufe, R. (2011). Using ISO 14001 to promote a sustainable supply chain strategy. *Business Strategy and the Environment*, 20(2), 71-93.

9. Drewes, T., & Michael, C. (2006). How do students choose a university?: an analysis of applications to universities in Ontario, Canada. *Research in Higher Education*, 47(7), 781-800.
10. Emploi Québec (2016). Explorer un métier ou une profession. Retrieved from http://imt.emploiquebec.gouv.qc.ca/mtg/inter/noncache/contenu/asp/mtg121_rech_reprofs_01.asp?lang=FRAN&Porte=1
11. Emploi Québec (2014). Le marché du travail au Québec – Perspectives professionnelles 2013-2017. Retrieved from http://www.emploiquebec.gouv.qc.ca/publications/pdf/00_imt_perspectives-emploi-profession.pdf
12. Fassin, Y. (2015). The Status of CSR and Sustainability Reporting at Universities in Europe: a Survey. In *Proceedings of the International Association for Business and Society* (Vol. 26, pp. 179-191).
13. Figge, F., Hahn, T., Schaltegger, S., & Wagner, M. (2002). The sustainability balanced scorecard—linking sustainability management to business strategy. *Business strategy and the Environment*, 11(5), 269-284.
14. Fonseca, A., Macdonald, A., Dandy, E., & Valenti, P. (2011). The state of sustainability reporting at Canadian universities. *International Journal of Sustainability in Higher Education*, 12(1), 22-40.
15. Gatto, M. (1995). “Sustainability: is it a well defined concept?”. 1181-1183.
16. GRI (2015a). Sustainability Reporting Guidelines. Available: <https://www.globalreporting.org/resourcelibrary/GRIG4-Part1-Reporting-Principles-and-Standard-Disclosures.pdf>

17. GRI (2015b). G4 Sector Disclosures – Food Processing. Available: <https://www.globalreporting.org/resourcelibrary/GRI-G4-Food-Processing-Sector-Disclosures.pdf>
18. GRI (2013). G4 Sustainability Reporting Guidelines. Retrieved from <https://www.globalreporting.org/reporting/g4/Pages/default.aspx>
19. GRI (2014a). G4 Guidelines–Reporting principles and standard disclosures.
20. GRI (2014b). G4 Guidelines–Implementation Manual.
21. Hansson, S. O. (2010). Technology and the notion of sustainability. *Technology in Society*, 32(4), 274-279.
22. Haupt, R., Scholtz, B., & Calitz, A. (2015). Using Business intelligence to Support Strategic Sustainability Information Management. In *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists* (p. 20). ACM.
23. Hilpert, H., Kranz, J., & Schumann, M. (2014). An Information system design theory for green information systems for sustainability reporting-integrating theory with evidence from multiple case studies.
24. International Organization for Standardization (2015a), ISO 14001 Key benefits, Retrieved October 11, 2015, from http://www.iso.org/iso/iso_14001_key_benefits.pdf.
25. International Organization for Standardization (2015b), ISO 14001:2015 – Environmental management systems — Requirements with guidance for use, Retrieved Mars 20, 2016, from <https://www.iso.org/obp/ui/#iso:std:60857:en>.

26. Kimball, R.: *Data Warehouse Toolkit – Practical Techniques for Building Dimensional Data Warehouses*, Wiley & Sons, New York, 1996.
27. Kimball, R.: *A Dimensional Modeling Manifesto*, in: *DBMS Online*, 1997, <http://www.dbmsmag.com/9708d15.html>.
28. Kimball, R.; Reeves, L.; Ross, M.; Thornthwaite, W.: *The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing and Deploying Data Warehouses*, Wiley & Sons, New York, 1998.
29. Ko, R. K. (2009). A computer scientist's introductory guide to business process management (BPM). *Crossroads*, 15(4), 4.
30. Khomba, J. K., & Vermaak, F. N. (2012). Relevance of financial reporting systems: Single-bottom line or triple-bottom line. *African Journal of Business Management*, 6(9), 3519.
31. Kutay, N., & Tektüfekci, F. (2015) A New Era for Sustainable Development: A Comparison for Sustainability Indices.
32. Lejeune, A. (2017). *From gamestorming to mobile learning: a conceptual framework and a gaming proposition to explore the design of flourishing business models* (Doctoral dissertation, Université du Québec à Montréal).
33. Lozano, R. (2011). The state of sustainability reporting in universities. *International Journal of Sustainability in Higher Education*, 12(1), 67-78.
34. Malhotra, A., Melville, N. P., & Watson, R. T. (2013). Spurring impactful research on information systems for environmental sustainability. *MIS Quarterly*, 37(4), 1265-1274.

35. Marjanovic, O. (2010). The Importance of Process Thinking in Business intelligence. *International Journal of Business intelligence Research*, 1(4), 29-46. doi:10.4018/jbir.2010100102
36. McCool, S. F., & Stankey, G. H. (2004). Indicators of sustainability: challenges and opportunities at the interface of science and policy. *Environmental management*, 33(3), 294-305.
37. Mitchell, G. (2006). Problems and fundamentals of sustainable development indicators.
38. Morse, S., McNamara, N., Acholo, M., & Okwoli, B. (2001). Sustainability indicators: the problem of integration. *Sustainable development*, 9(1), 1-15.
39. Munajat, Q., & Kurnia, S. (2015). Towards Sustainability Framework Development: Identifying Organisational Capability For Sustainability Practices. In PACIS 2015 Proceedings. Paper 106.
40. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
41. Petrini, M., & Pozzebon, M. (2009). Méanaging sustainability with the support of business intelligence: Integrating socio-environmental indicators and organisational context. *The Journal of Strategic Information Systems*, 18(4), 178-191.
42. Petruzzellis, L., & Romanazzi, S. (2010). Educational value: how students choose university: Evidence from an Italian university. *International journal of educational management*, 24(2), 139-158.

43. Sacu, C., & Spruit, M. R. (2010). BIDM-The Business intelligence Development Model. In ICEIS (1) (pp. 288-293).
44. Samajdar, C. (2014). Triple bottom line strategy for sustainable steel industry. In Business and Information Management (ICBIM), 2014 2nd International Conference on (pp. 101-103). IEEE.
45. Searcy, C., & Buslovich, R. (2014). Corporate perspectives on the development and use of sustainability reports. *Journal of Business Ethics*, 121(2), 149-169.
46. Seidel, S., Székely, N., & vom Brocke, J. (2015). Green IS: Are We Still Thinking in Mere Economic Imperatives or Are We Striving for Eco-Effectiveness?.
47. SIGMA Project (2006). SIGMA Guidelines. Retrieved from <http://www.projectsigma.co.uk/Guidelines/>
48. S&P (2015). Dow Jones Sustainability Indices Methodology.
49. Slaper, T. F., PhD., & Hall, T. J. (2011). The triple bottom line: What is it and how does it work? *Indiana Business Review*, 86(1), 4-8.
50. Statistiques Canada (2006). Human Activity and the Environment: Annual Statistics 2006, Retrieved from <http://www.statcan.gc.ca/pub/16-201-x/2006000/9515-eng.htm>.
51. Statistiques Canada (2008). Enquête sur la commercialisation de la propriété intellectuelle dans le secteur de l'enseignement supérieur, Retrieved from http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&Id=58189
52. Statistiques Canada (2011). Classification Nationale des Professions (CNP) 2011 - Niveaux, Retrieved from <http://ouvert.canada.ca/data/fr/dataset/bc4d5e01-cc19-45f3-bebc-72969b6d4b42>.

53. Statistiques Canada (2011b). Série « Perspective géographique », Recensement de 2011. Retrieved from <http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-cma-fra.cfm?LANG=Fra&GK=CMA&GC=462>.
54. Statistiques Canada (2016). Effectifs postsecondaires selon le type d'institutions, le régime d'études, provinces et sexe. Retrieved from <http://www.statcan.gc.ca/tables-tableaux/sum-som/102/cst01/educ71a-fra.htm>
55. Tanzil, D., & Beloff, B. R. (2006). Assessing impacts: Overview on sustainability indicators and metrics. *Environmental Quality Management*, 15(4), 41-56.
56. UQÀM, Université du Québec à Montréal (2016a). La population étudiante de l'UQAM Statistiques d'inscription – 2015-2016. Retrieved from http://www.regis.uqam.ca/Pdf/Pop_etudiante/population_1516.pdf
57. UQÀM, Université du Québec à Montréal (2015). La population étudiante de l'UQAM – Statistiques d'inscription – 2014-2015. Retrieved from http://www.regis.uqam.ca/Pdf/Pop_etudiante/population_1415.pdf
58. UQÀM, Université du Québec à Montréal (2016b). Programmes d'Études. Retrieved from <http://www.etudier.uqam.ca/programmes>.
59. Yao, J., Wang, J., Chen, Q., & Xing, R. (2013). Core Methodologies in Data Warehouse Design and Development. *International Journal of Robotics Applications and Technologies*, 1(1), 57-66. doi:10.4018/ijrat.2013010104