

Université de Montréal

**Étude de la formation d'une communauté épistémique :
L'influence réciproque d'acteurs sociaux et de conceptualisations en santé
publique**

par
Chanel Boucher

Département de sociologie
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en sociologie

Août, 2011

© Chanel Boucher, 2011.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

**Étude de la formation d'une communauté épistémique :
L'influence réciproque d'acteurs sociaux et de conceptualisations en santé
publique**

présenté par:

Chanel Boucher

a été évalué par un jury composé des personnes suivantes:

Stéphane Moulin,	président-rapporteur
Andrée Demers,	directeur de recherche
Alain Degenne,	codirecteur
Maurice Lévesque,	membre du jury

Mémoire accepté le: 25 octobre 2011

RÉSUMÉ

L'étude de la formation d'une communauté épistémique québécoise en santé publique en ayant recours à l'interactionnisme-structural permet d'appréhender ce phénomène social sous l'angle d'une influence réciproque entre d'une part des acteurs sociaux interagissant entre-eux et d'autre part, des conceptualisations variées des objets de santé publique ; ces éléments sociaux et sémantiques subissent des transformations simultanées.

Notre étude démontre et illustre qu'au même moment où un réseau social de chercheurs prend forme, une thématique nouvelle prend place et rallie ces mêmes chercheurs, non pas seulement dans leurs relations sociales, mais dans les idées qu'ils manipulent lors de leur travail de production et de diffusion de connaissances ; les identités sociales se lient, mais pas au hasard, parce qu'elles partagent des éléments sémantiques communs.

C'est notamment en explorant 20 ans de collaborations scientifiques à l'aide d'outils d'analyse de réseaux, d'analyse en composantes discrètes et par l'exploration de treillis de Galois que cette étude a été menée. Notre approche est principalement exploratoire et une attention toute particulière est portée sur les aspects méthodologiques et théoriques du travail de recherche scientifique.

Mots clés: Analyse de réseaux sociaux, copublication, communauté épistémique, analyse sociosémantique.

ABSTRACT

The study of the formation of an epistemic community in Quebec's public health by using structural interactionism allows to understand this social phenomenon in terms of an interaction between, on the one hand, social actors interacting together and, on the other hand, various conceptualizations of the objects of public health. These social and semantic elements undergo simultaneous transformations.

Our study demonstrates that the formation of a network of researchers leads to the emergence of a new thematic at the center of interest of those researchers. Researchers are linked, not only in social networks term, but also at a conceptual level : social identities are linked, but not randomly, they are linked by semantics elements.

It is by exploring 20 years of scientific collaboration by using tools from social networks analysis, discrete components analysis and the exploration of concept lattice in particular that this study was conducted. Our approach is more exploratory and a special attention is put on methodological aspects of scientific work.

Keywords: Social networks analysis, copublication. epistemic community, sociosemantic analysis.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES	ix
LISTE DES ANNEXES	x
LISTE DES SIGLES	xi
NOTATION	xii
DÉDICACE	xiii
REMERCIEMENTS	xiv
INTRODUCTION	1
CHAPITRE 1: LA SANTÉ PUBLIQUE COMME COMMUNAUTÉ ÉPISTÉMIQUE	5
1.1 Émergence et évolution en santé publique	5
1.1.1 Émergence de la santé publique en tant qu'institution sociale au Québec du XIX ^e siècle	6
1.1.2 Évolution de la santé publique	9
1.1.3 La santé publique contemporaine	12
1.2 Comment le savoir se construit ?	15
1.2.1 Construction de la connaissance en santé publique	15
1.2.2 Les communautés épistémiques	17

1.2.3	Questions de recherche	20
CHAPITRE 2: MÉTHODOLOGIE ET OUTILS D'ANALYSE		21
2.1	Le corpus de données	21
2.2	Collecte des données	23
2.3	La classification des chercheurs et des thématiques	28
2.3.1	La classification des chercheurs selon leur champ disciplinaire de formation	28
2.3.2	La classification des thématiques de recherche : le <i>topics modeling</i>	31
2.4	Les analyses	37
2.4.1	L'analyse de réseaux	37
CHAPITRE 3: LA CLASSIFICATION DES THÉMATIQUES DE RECHER- CHE : LE TOPICS MODELING		42
3.1	Les résultats des analyses sémantiques	42
3.1.1	Illustration des résultats obtenus par <i>topics modeling</i>	46
3.2	Les grandes thématiques retenues	50
3.3	Un portrait descriptif de l'usage des thèmes	51
3.4	Conclusion	52
CHAPITRE 4: UN RÉSEAU SOCIAL PREND FORME...		54
4.1	1990 à 1994	55
4.2	1994 à 1999	56
4.3	2000 à 2004	58
4.4	2005 à 2009	60
4.4.1	Conclusion	63
CHAPITRE 5: ...AUTOUR D'UN MÊME THÈME : LA « NOUVELLE SANTÉ PUBLIQUE »		64
5.1	Thématiques de recherche centrales et périphériques	64
5.2	Évolution des thématiques de 1990 à 2009	67

5.3	Convergence vers la « nouvelle santé publique »	70
5.4	Conclusion	77
CHAPITRE 6: DISCUSSION		79
6.1	Le développement d'un réseau de collaboration interdisciplinaire	80
6.2	Vers une communauté épistémique	82
6.3	Limites de cette étude	85
6.4	Conclusion	86
BIBLIOGRAPHIE		89

LISTE DES TABLEAUX

2.I	Synthèse du nombre de chercheurs et d'articles pour chaque période . . .	27
3.I	Présentation des 10 premiers articles expliqués par le <i>Topic 67</i> : « Déterminants sociaux de la santé »	49
5.I	Centralité des thématiques, selon les périodes d'analyse	67
I.I	Liste des chercheurs et des disciplines en date du 1 ^{er} septembre 2009 . . .	xv
IV.I	Liste complète des fusions de <i>topics</i>	xli

LISTE DES FIGURES

1.1	Relations entre santé et facteurs sociaux et environnementaux : Schéma de Evans et al.	13
1.2	Structure du programme national de santé publique	14
2.1	Histogramme de la distribution du nombre d'auteurs par article	27
2.2	Au cours du temps, distribution des chercheurs selon leurs disciplines de formation	30
3.1	Perplexité du modèle, en fonction du nombre de <i>topics</i>	43
3.2	Congruence du modèle, selon le nombre d'itérations	44
3.3	Le résumé d'un article, traité par <i>topic modeling</i>	47
3.4	Histogramme de la distribution du nombre de <i>topics</i> par article	52
4.1	Collaborations entre futurs chercheurs de l'IRSPUM entre 1990 à 1994 (T-3)	55
4.2	Collaborations entre futurs chercheurs de l'IRSPUM entre 1994 à 1999 (T-2)	56
4.3	Collaborations entre futurs chercheurs de l'IRSPUM entre 2000 à 2004 (T-1)	59
4.4	Collaborations entre futurs chercheurs de l'IRSPUM en 2005 à 2009 (T0)	61
5.1	Réseau sémantique complet, en T0	65
5.2	Tronçon axé sur l'usage des déterminants sociaux de la santé en T-3	73
5.3	Tronçon axé sur l'usage des déterminants sociaux de la santé en T-2	74
5.4	Tronçon axé sur l'usage des déterminants sociaux de la santé en T-1	75
5.5	Tronçon axé sur l'usage des déterminants sociaux de la santé en T0	77
V.1	MouliNet en action...	xlvi

LISTE DES ANNEXES

Annexe I:	Liste des chercheurs	xv
Annexe II:	Liste des <i>topics</i>	xxxiv
Annexe III:	Légendes des graphes	xl
Annexe IV:	Liste complète des thèmes fusionnés	xli
Annexe V:	<i>MouliNet</i>	xlvi

LISTE DES SIGLES

admin	Administration et gestion
ammén	Amménagement
bio	biologie et chimie
diet	Nutrition
droit	Droit
edu	Sciences de l'éducation
env	Sciences environnementales
epid	Épidémiologie et biostatistiques
infir	Sciences infirmières
med	Médecine et sciences biomédicales
pharm-éco	Pharmacoéconomie
philo	Philosophie
phar	Pharmacologie
psycho	Psychologie
rel	Relations industrielles
santé	Santé publique
ss	Sciences sociales
TM	Topics modeling
vet	Médecine vétérinaire et zoologie

NOTATION

P Probabilité

D Densité

C Centralité

Pour Maël

REMERCIEMENTS

« You need to believe in things that aren't true. How else can they become ? »

Terry Pratchett(1996) dans « Hogfather » p.409

Que je sois parvenue aussi loin dans mon cheminement académique n'a rien de trivial ; c'est grâce au soutien de nombre des membres du corps professoral et administratif du département de sociologie de l'Université de Montréal. J'aimerais notamment remercier M. Paul Bernard qui était convaincu que parfois les gens n'avaient besoin que d'un coup de main ; il a changé ma vie de façon drastique et me fit rencontrer Mme Andrée Demers, qui deviendra d'abord mon employeur, puis la directrice de mes recherches pour ce présent mémoire.

Mme Demers a su me soutenir dans mes épreuves personnelles aussi bien que dans mes projets de recherche. Malgré les risques que je représente, elle a accepté de me diriger et elle a réalisé cette tâche avec brio. Je n'ai rien d'une étudiante docile et je ne peux être qu'infiniment reconnaissante envers Mme Demers d'avoir accepté le défi !

Je tiens aussi à remercier mon mentor, M. Alain Degenne ainsi que son épouse Anic pour leur soutien. M. Degenne a, sans relâche, veillé à me faire réfléchir en guidant mes lectures, corrigeant mes travaux et m'inondant de ses si constructifs et encourageants commentaires. Son assiduité rassurante ainsi que sa confiance paisible m'ont été d'un grand secours, tout au long de mon apprentissage. Je lui dois de m'avoir fait découvrir la sociologie whittienne qui occupe maintenant mon principal centre d'intérêt en recherche. J'estime avoir eu une chance incommensurable d'avoir pu travailler avec lui.

Sans le soutien de mon conjoint Damien Rilly, qui m'a aidée dans la programmation des divers scripts dont j'ai eu besoin, je n'aurais pu accomplir autant. Merci aussi à mes amis qui ont cru en moi davantage que je ne pouvais moi-même le faire !

C'était mon plus grand rêve que d'étudier à l'université et vous tous, vous m'avez permis de le réaliser et d'aller encore plus loin que je ne pouvais imaginer, alors je vous dis encore une fois: MERCI !

INTRODUCTION

« Les groupes dont l'individu fait partie constituent en quelque sorte un système de coordonnées tel que chaque coordonnée nouvelle qui vient s'y ajouter le détermine de façon plus exacte et plus nette. »

Georg Simmel (2010[1908]) dans « Sociologie: Études sur les formes de la socialisation » p.416

La santé publique est omniprésente dans notre vie quotidienne. Dans nos sociétés contemporaines, nous attendons de la santé publique qu'elle nous instruisse et nous informe des risques pour notre santé liés à notre environnement, à notre alimentation ou à nos habitudes de vie. Nous attendons de la santé publique qu'elle réduise les risques pour la santé et ainsi la prévalence des « maladies » et autres troubles nuisant à la « santé », qu'elle réduise les inégalités sociales de santé et nous permette de vivre plus longtemps et en meilleure santé. L'emphase sur la santé se perçoit dans de nombreux marchés : l'industrie agro-alimentaire tente de faire valoir les aspects nutritifs et « bons pour la santé » de ses produits ; l'industrie pharmacologique tente de répondre à nos petits malaises en tout genre et à corriger, grâce à des médicaments, nos mauvaises habitudes de vie ; l'industrie de l'ameublement se concentre sur la promotion de meubles sécuritaires pour les bébés. Il en va de même en politique ; les « politiques saines » ne cessent de croître. (Pommier et Grimaud, 2007) La santé publique rejoint les préoccupations sociales contemporaines et son discours y trouve donc aisément un public intéressé.

De nos jours, les étiquettes sont nombreuses pour illustrer comment la connaissance de type scientifique joue un rôle déterminant dans la compréhension de la contemporanéité : « société de l'information », « société du savoir », « société de la technologie », etc. (Cetina, 2000). Un point commun ressort de ces appellations : l'apparente ampleur qu'occupe la connaissance dans les diverses sphères de la vie sociale contemporaine. La santé publique n'échappe pas à la scientification de l'action sociale qui semble s'opérer sur l'ensemble de la vie sociale contemporaine. La place que prend la connaissance en santé publique dans les sociétés contemporaines est toutefois particulière parce qu'elle touche des préoccupations humaines profondes dont la souffrance et la mort. Un des *leit-*

motiv des acteurs en santé publique, tant chez les intervenants que chez les décideurs, est de fonder l'action sur des données probantes. À cet égard, la recherche en santé publique joue un rôle majeur.

Depuis plus d'un siècle, la santé publique s'est profondément transformée. D'abord préoccupée de l'hygiène publique et de la vaccination pour combattre les maladies infectieuses, la santé publique s'est élargie aux habitudes de vie avec le développement des maladies chroniques, puis aux facteurs structureaux et aux phénomènes sociaux qui affectent la santé des populations. Aujourd'hui, la santé publique se définit largement comme « l'action sur les déterminants de la santé et du bien-être, au niveau de la population et des systèmes qui la régissent cette action étant circonscrite par les fonctions essentielles et de soutien de santé publique. » (Laurendeau et al., 2007, p. 7) Les fonctions essentielles englobent la surveillance de l'état de santé de la population, la promotion de la santé et du bien-être, la prévention des maladies, des problèmes psychosociaux et des traumatismes, alors que les fonctions de soutien concernent les politiques publiques ayant un effet sur la santé, la recherche et le développement des compétences en santé publique (MSSS, 2003). La recherche en santé publique est ainsi appelée à couvrir un champ large. Laurendeau et al. (2007) la définissent comme : « l'ensemble des activités de recherche reliées à la santé et au bien-être de la population et à leurs déterminants, qui visent la production, l'intégration, la diffusion et l'application de connaissances scientifiques, valides et pertinentes à l'exercice des fonctions de santé publique. Les déterminants réfèrent ici à un ensemble de caractéristiques tant individuelles (ex. : démographiques, socio-économiques, génétiques, perceptuelles, comportementales) que collectives (ex.: structure sociale) et contextuelles (ex. : organisation des services de santé et services sociaux, milieu de vie, environnement physique et social, culturel, économique et politique) qui ont une influence directe ou indirecte sur la santé. » (Laurendeau et al., 2007, p. 7). Ainsi définie, la recherche interpelle des expertises disciplinaires très diversifiées.

En 2009, l'Université de Montréal créait l'Institut de recherche en santé publique de l'Université de Montréal (IRSPUM), avec comme objectif de regrouper sous une même entité l'ensemble de la recherche en santé publique à l'Université de Montréal, afin de

favoriser l'émergence de nouvelles synergies entre les chercheurs. Sont ainsi regroupés les chercheurs appartenant à diverses entités (principalement le Groupe de recherche interdisciplinaire en santé (GRIS), le Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), le Centre Léa Roback et l'axe de recherche en santé des population du Centre de recherche du Centre hospitalier universitaire de l'Université de Montréal (CRCHUM)) autour de la vision élargie de la recherche en santé publique.

Le but de ce mémoire est de comprendre comment se structurent les liens entre les chercheurs de l'IRSPUM ainsi que le développement de ces liens à travers le temps, en portant une attention particulière à la place qu'y occupe la recherche sur les déterminants sociaux de la santé. Le présent mémoire abordera le regroupement des chercheurs au sein de l'IRSPUM en tant que « communauté épistémique », c'est-à-dire un ensemble de chercheurs œuvrant sur un thème commun (Roth, 2006, 2008). Une communauté épistémique se constitue à la fois d'un réseau social, c'est-à-dire des liens effectifs qui s'établissent entre les chercheurs, et d'un réseau sémantique, c'est-à-dire des thématiques de recherche autour desquelles ils se regroupent (Roth, 2006, 2008). Or, comme nous l'avons souligné précédemment, la recherche en santé publique est un champ de recherche vaste et éclaté qui manipule une diversité d'objets et qui interpelle un large spectre de disciplines scientifiques. Ce simple constat soulève de nombreuses questions auxquelles nous tenterons de répondre dans ce mémoire. Les chercheurs de l'IRSPUM forment-ils une seule communauté épistémique ou si l'IRSPUM regroupe plusieurs communautés épistémiques n'ayant pas ou peu de liens entre elles, communautés qui pourraient se constituer soit autour de thématiques de recherche soit autour de l'appartenance disciplinaire ? Quelle place y occupe la thématique des déterminants sociaux qui est au cœur de la nouvelle santé publique ? Constitue-t-elle un pôle rassembleur de l'ensemble des chercheurs, peu importe leur appartenance disciplinaire, ou regroupe-t-elle principalement les chercheurs des sciences sociales ? Comment se sont développés les liens entre les chercheurs ? Préexistaient-ils à la création de l'IRSPUM ? Les chercheurs des différentes entités qui ont été intégrées au sein de l'IRSPUM collaborent-ils entre eux ou si ces entités demeurent autant de communautés épistémiques au sein de l'IRSPUM ?

Pour répondre à ces questions, nous analyserons les interactions entre les chercheurs et les thèmes qu'ils manipulent par le biais de l'étude de leurs publications. Cinq chapitres composent ce mémoire. Au chapitre premier, nous ferons d'abord un bref rappel de l'évolution de la santé publique, de ses objets et des différents acteurs qu'elle mobilise, en portant une attention particulière à l'émergence de la « nouvelle santé publique », pour ensuite traiter de la production des connaissances scientifiques en santé publique. Ce dernier volet nous permettra de porter une attention particulière sur l'interdépendance des idées et des identités sociales. Le second chapitre est méthodologique. Dans un premier temps, nous présenterons la construction du corpus de données pour ensuite expliquer les méthodes d'analyse que nous avons utilisées. Nous y présentons principalement l'analyse en composantes discrètes et l'analyse de réseaux sociaux et sociosémantiques. Le troisième chapitre présente les analyses de classification en composantes discrète des mots contenus dans le corpus soumis aux analyses. Ce chapitre permettra de comprendre comment ont été construites les données utilisées pour les analyses ultérieures et d'en évaluer la validité. Les résultats de notre étude se retrouvent dans les deux derniers chapitres. Dans un premier temps, au chapitre 4, nous examinerons les collaborations entre les chercheurs de l'IRSPUM en fonction de leur appartenance disciplinaire au moment de la création de l'IRSPUM ainsi que l'émergence de ces collaborations au cours des deux décennies précédentes. Dans un second temps, au chapitre 5, nous examinerons la convergence des intérêts de recherche des membres de l'IRSPUM autour de grandes thématiques en fonction de leur appartenance disciplinaire ainsi que l'émergence de la thématique des déterminants sociaux. Les résultats seront discutés au dernier chapitre.

CHAPITRE 1

LA SANTÉ PUBLIQUE COMME COMMUNAUTÉ ÉPISTÉMIQUE

Le présent chapitre se divise en deux sections. La première section a comme objectif de situer la santé publique contemporaine. Nous y examinerons les divers courants de pensée qui l'ont traversée depuis la fin du XIXe siècle, les divers acteurs sociaux et les différentes disciplines qui s'en trouvent interpellées ainsi que l'évolution de la recherche en ce domaine. Cette démarche permettra d'appréhender les grandes thématiques de recherche en santé publique et de mieux saisir leur émergence. Comme le rappelle Keating, la santé publique renvoie non seulement à l'histoire des institutions et des pratiques scientifiques mais aussi à celle des représentations et attitudes collectives devant la maladie, la souffrance et la mort. » (Keel et Keating, 1995, p. 10) Puis, nous explorerons le concept de « communauté épistémique » afin de le définir et l'articuler dans notre démarche de recherche. Les deux sections de ce premier chapitre sont liées : la santé publique est à la fois une culture épistémique et une institution sociale. Une communauté épistémique axée sur ce domaine de la recherche est aussi, à la fois, ces deux facettes, sous une forme plus abstraite. L'organisation sociale co-évolue en interaction avec la connaissance épistémique en santé publique.

1.1 Émergence et évolution en santé publique

Cette première partie abordera la santé publique principalement en tant qu'institution sociale, c'est-à-dire en se focalisant sur les identités sociales. Dans quel contexte est-elle apparue ? À quels besoins sociaux répond-elle ? Qui sont les acteurs principaux qui lui ont permis de voir le jour ? Nous chercherons à comprendre les grandes lignes de son évolution afin de saisir les changements influant sur les visées de la santé publique. À la fin de cette démarche, nous pourrons définir la santé publique contemporaine et mettre en lumière les thèmes et approches actuels dans ce domaine.

1.1.1 Émergence de la santé publique en tant qu'institution sociale au Québec du XIX^e siècle

Bien que les préoccupations thérapeutiques et préventives en matière de bien-être n'aient rien de nouveau, dès les débuts de l'humanité les humains ont développé des techniques de soins thérapeutiques (Merbs, 1971), l'émergence de la santé publique ne peut s'expliquer sans prendre en compte le développement de la médecine moderne (Keel et Keating, 1995). L'avènement de la science moderne, qui causa selon Weber le « désenchantement du monde » (Weber, 2003), et l'industrialisation sont les cadres socio-historiques dans lesquels on peut voir émerger la santé publique (Kunitz, 2007).

Avec l'industrialisation et l'urbanisation croissante qu'ont connues de nombreuses sociétés occidentales au cours du XIX^e siècle, de « nouveaux » débats sont apparus concernant les conditions de vie des populations. (Evans et al., 2004, Keel et Keating, 1995, Kunitz, 2007) Ces préoccupations, d'ordre principalement mercantile, concernaient plus particulièrement la « dégénérescence » des classes ouvrières et l'impact d'un tel processus sur l'économie. (Keel et Keating, 1995) Au même moment une rupture épistémologique majeure se produisait en sciences et plus particulièrement en médecine via l'adhésion à la théorie des germes, l'explication de l'apparition des maladies n'était plus perçue comme étant due aux miasmes mais à l'explication, plus scientifique, de la présence de micro-organismes causant les maladies. (Goulet, 1994) Parallèlement, le développement des statistiques sur les populations et plus spécifiquement de celles concernant la distribution de la mortalité et des naissances a permis de développer « l'épidémiologie ». (Omran, 2005) L'épidémiologie consiste sommairement en l'étude des facteurs influençant la mortalité, la morbidité et la fécondité des populations. (Omran, 2005) L'impératif social conjugué aux nouvelles connaissances scientifiques favorisera le développement de la santé publique.

La transformation dans l'attribution causale des maladies et la production croissante des connaissances en bactériologie et en épidémiologie, vont fournir à certains acteurs sociaux matière à discours. En particulier, des « médecins progressistes » vont produire un discours particulier afin de tenter de s'ancrer en tant qu'acteurs principaux de la santé

publique et d'ancrer la médecine en tant que l'institution sociale principale dont relève la santé publique. Dans leur démarche, ils interpellèrent diverses autorités publiques afin de concrétiser via des politiques publiques l'action de la santé publique. L'étude « *Autour du Journal de médecine de Québec/Quebec medical journal (1826-1827) : programme scientifique et programme de médicalisation* de Keel et Keating (1995) résume le processus d'ancrage de l'identité de médecin et de la santé publique dans l'organisation sociale au Québec :

« Tout d'abord il s'agit de connaître et de faire connaître la situation sanitaire dans les différentes régions du pays (quelles sont les maladies régnantes : endémiques et épidémiques), ce qui devrait permettre de prendre les mesures prophylactiques qui s'imposent. Il s'agit ensuite de mettre sur pied une meilleure organisation du service médical dans la société et d'implanter la fonction médicale dans les régions ou populations où elle ne l'est pas et donc de combattre les charlatans et les torts qu'ils causent à la santé publique. Il faut combattre aussi les habitudes et les préjugés populaires, soit les pratiques ou modes traditionnelles de prise en charge de la maladie, et cela en répandant des notions modernes d'hygiène domestique. Enfin il s'agit de réaliser un programme de mesures sanitaires (vaccination, etc.), d'obtenir l'appui des autorités à cet effet et de mettre en place pour cela un certain nombre de nouvelles institutions sanitaires et de nouvelles instances administratives du domaine de la santé publique. » Keel et Keating (1995, p. 52-53)

Bien que le cas des discours promus par le *Quebec Medical Journal* soit très spécifique et mériterait d'être contextualisé davantage, un peu partout ailleurs en Occident, l'institution de la santé publique a été portée entre autres par des « élites médicales » qui ont dû se tourner vers les autorités publiques pour mettre en œuvre leurs divers plans d'intervention sur la population. C'est une chose pour les hygiénistes de la fin du XIX^e siècle de croire en l'utilité publique de l'hygiène publique et de la vaccination, mais c'en est une autre de mettre ces mesures en place à l'échelle d'une population (Keel et Keating, 1995). L'appui des autorités est nécessaire, tant d'un point de vue logistique qu'économique. L'arrimage de la médecine et de l'État se met ainsi en place pour développer la santé publique .

Le raisonnement de Foucault amène à saisir l'importance de la santé publique pour les États modernes, via son rôle d'expert du bien-être et de la santé d'une population en permettant de légitimer une forme typiquement contemporaine de contrôle social¹.

1. Voir entre autre à ce sujet : « La naissance de la médecine sociale » dans Foucault (2001)

D'autres chercheurs dont Fassin (2000), vont extrapoler la notion de bio-pouvoir chez Foucault et mettre en évidence la primauté contemporaine du corps sur l'intégrité de la personne : « L'hypothèse que l'on peut avancer est que, dans ces opérations de mise en priorité, une reconnaissance plus grande est accordée à l'intégrité du corps qu'à l'intégrité de la personne ou, pour le dire autrement, que la reconnaissance de la personne passe par la reconnaissance du corps altéré ou souffrant » (Fassin, 2000, p. 22)

D'un point de vue économique, les médecins trouvent un certain public pour leurs discours sur les craintes concernant la santé des masses ouvrières, tout en promettant de favoriser une force de travail en bonne santé : « Or, pour les théoriciens du mercantilisme, et pour les médecins attachés à leurs idées, la santé des populations s'impose comme norme par excellence de la croissances démographique. De là l'importance de la médecine pour la prospérité de l'État » (Keel et Keating, 1995, p. 58). D'autre part, économiquement parlant, la santé publique a permis l'émergence de nombreux marchés : outillages et technologies médicales, industries des produits d'entretien domestiques, industries pharmaceutiques, construction de batiments pour pourvoir aux soins de santé, etc. (Keel et Keating, 1995) De nombreuses identités sociales différentes interagissent de façon tout autant différentes autour de la santé publique, ce qui renforcera son ancrage dans l'organisation sociale.

Les discours des médecins ont aussi trouvé un « public » réceptif en la bourgeoisie qui tentait de son côté d'imposer des normes de bienséance et de raffinement qui lui sont propres aux populations pauvres et/ou immigrantes :

« The importance of protecting and preserving food, of using clean water, of bathing, of washing one's hand after using toilet, of keeping the home clean and airy, of screening windows against flies,² and not spitting were all subject of health education activities whose legitimacy was enhanced by germ theory. Until about 1910, it was reformers and public servant who were largely responsible for persuading the masses of the benefits of keeping clean. » Kunitz (2007, p.15)

Ainsi, en quelque sorte, les discours et actions de diverses « élites » se sont rencontrés et ont rendu possible l'émergence de la santé publique. En tant que mode de régulation, l'institution de la santé publique touche à de nombreux domaines de la vie sociale. Il ne

2. Les auteurs citent ici le texte de : N. Rogers, « *Germ with legs : flies, disease, and the new public health* » . Bulletin of the history of medicine 63 : 599-617, 1989

faut pas croire que l'émergence de la santé publique a été planifiée, malgré les efforts de plusieurs types d'acteurs pour s'ancrer et ancrer la santé publique dans la réalité sociale. Les contingences décrites dans cette sous-section permettent de saisir le contexte structurel qui a permis l'émergence de la santé publique mais les interactions sociales ont lieu dans un chaos d'interactions. Car, comme le rappelle White, les identités sociales sont prises dans un chaos d'interactions qui les traversent. (White, 2008) À l'instar de l'analogie au jeu d'échec proposée par Elias (1985), les acteurs sociaux ne peuvent individuellement et sciemment faire évoluer la vie sociale dans une direction spécifique. Dans la prochaine sous-section, nous décrirons de façon succincte les transformations sémantiques majeures qui se sont produites en santé publique.

1.1.2 Évolution de la santé publique

Durant la première moitié du XX^e siècle, l'ancrage des acteurs sociaux et de la santé publique dans l'organisation sociale s'est poursuivi, non sans heurt. Les premières écoles de santé publique, apparues aux États-Unis, ont tôt fait de déployer des efforts de contrôle afin de se distinguer des écoles de médecine classique :

« Ce qui n'était au départ conçu que comme une cohabitation sans heurt allait prendre une tournure plus conflictuelle à chaque fois qu'une intersection des champs de pratique mettait en jeu le statut des professionnels de la santé publique. Ayant des programmes, des écoles et des professionnels distincts, médecine préventive et médecine curative étaient vouées à se développer de manière isolée, la première confinée à la santé publique et la seconde associée à la pratique privée et lucrative. » Keel et Keating (1995, p. 162)

Les hygiénistes se sont distingués de la médecine, plus orthodoxe, en cherchant à modifier les conditions de vie des populations dans le but affirmé d'agir sur leur santé. Leur champ d'action ne se limitait pas au pathologique ou même à la promotion de l'hygiène, mais s'ouvrait aussi sur des réalités sociales complexes non couvertes par la médecine : « L'hygiène, loin de se limiter à l'action sanitaire proprement dite, aura pour mandat de contribuer à régler les problèmes de maladies, d'arriération mentale, de pauvreté, de criminalité, d'alcoolisme, d'instruction publique, ainsi que des problèmes industriels. » (Keel et Keating, 1995, p. 163) La santé publique a ainsi étendu son action sur de plus en plus de domaines de la vie sociale.

Les actions en santé publique focalisaient dans un premier temps sur la diffusion « d'informations sanitaires ». Les hygiénistes vont œuvrer pendant des décennies à produire des brochures, des affiches et des messages publicitaires afin d'intervenir sur la santé des populations via l'éducation : « Ces pratiques étaient fondées sur la croyance incontestée et profondément enracinée qu'il était intrinsèquement bon d'éduquer la population ; on espérait alors que la santé allait s'améliorer grâce à la science et à des moyens plus systématiques de faire l'éducation à la santé. » (O'Neill et al., 2006b, p. 6) Ce type d'action en santé publique qui consiste à diffuser des informations sanitaires n'a pas disparu.

La science à laquelle se réfère la santé publique du début du XX^e siècle est essentiellement l'épidémiologie qui elle se retrouve entre les mains des médecins pour qui la théorie des germes prend une importance particulière. La tendance à la pathologisation de nombreuses réalités sociales par la santé publique du début du XX^e siècle croîtra jusqu'à la fin du XX^e siècle. (O'Neill et al., 2006b) La gestion des maladies infectieuses va, dans la même veine, occuper principalement les hygiénistes qui vont œuvrer activement entre autres pour la vaccination obligatoire des populations. (Keel et Keating, 1995)

De nombreuses institutions relatives à la prise en charge de la santé des populations verront le jour durant le début du XX^e siècle. Dans leur démarche d'assainissement des milieux de vie de la population et d'ancrage de la médecine en tant que domaine dont relève la santé, les médecins hygiénistes vont peu à peu investir les hôpitaux et s'identifier comme « experts » des questions de santé. La médecine semble, pour l'époque, offrir la voie la plus lucide afin d'améliorer la santé des populations.

La santé publique connaîtra un nouvel essor dans la seconde moitié du XX^e siècle. En 1946, l'Organisation Mondiale de la Santé (OMS) adoptera sa célèbre définition de la santé qui la définit comme étant non seulement l'absence de maladie et d'infirmité, mais aussi un état complet de bien-être physique, mental et social. (O'Neill et al., 2006a) D'abord centrée sur l'éducation sanitaire visant à favoriser une bonne utilisation des services de santé, notamment dans le contexte de mise en place des systèmes publics de santé, l'importance épidémiologique grandissante des maladies chroniques

et de leur facteurs de risque comportementaux amènera un déplacement du focus de la santé publique vers les habitudes de vie et mobilisera les spécialistes des sciences sociales, notamment les psychologues et les spécialistes de la communication, pour mieux comprendre et modifier les comportements. Au Canada, le Rapport Lalonde en 1974, bien qu'ouvrant la voie vers l'inclusion des facteurs environnementaux et les conditions de vie, mettra davantage l'accent sur la responsabilisation personnelle des gens face à leur état de santé et sur les habitudes de vie.

La Charte d'Ottawa, produite à la suite du Rapport Black (1980) et en collaboration avec l'OMS, constituera un point tournant pour la santé publique. Le gouvernement canadien et l'association canadienne pour la santé publique, annonceront le début d'une « nouvelle ère en santé publique »: la « promotion de la santé ». La Charte d'Ottawa décrit la santé comme étant une ressource de la vie quotidienne ayant un impact majeur sur le bien-être des gens. (OMS, 1986) Elle définit aussi les conditions préalables à la santé : « la paix, un abri, de la nourriture et un revenu » (OMS, 1986) En ce sens, les facteurs influençant la santé prennent une allure davantage déterminée par les conditions d'existence des populations que par simplement des facteurs bio-médicaux ou comportementaux. Il apparaît alors de plus en plus important de créer des milieux favorables à la santé et donc d'agir sur les facteurs sociologiques.

Vers la fin du XX^e siècle, la santé publique s'est tournée davantage vers la « promotion de la santé » et la prise en compte de facteurs sociologiques sur la santé. (O'Neill et Pederson, 1994) La promotion de la santé désigne un processus visant à permettre aux gens d'exercer un certain contrôle sur leur santé ainsi qu'à l'améliorer.³ Ainsi, la promotion de la santé est à la fois une approche en santé publique et une politique sociale. (Bunton et MacDonald, 1992) Tel qu'il est inscrit en première page de la Charte d'Ottawa : « la promotion de la santé va bien au delà des soins. Elle inscrit la santé à l'ordre du jour des responsables politiques des divers secteurs en les éclairant sur les conséquences que leurs décisions peuvent avoir sur la santé, en leur faisant admettre leur responsabilité à cet égard. » (OMS, 1986, p. 1) L'intervention en promotion de la santé diffère de « l'ancienne santé publique » notamment par sa prise en compte de facteurs

3. Cette phrase s'inspire fortement de Bunton et MacDonald (1992, p. 1)

sociaux et environnementaux en tant que déterminants de la santé « The health promotion was based on a perspective that health is determined in important way by social and environmental factors (Kickbusch, 1986 ; World Health Organisation, 1984) » (O'Neill et Pederson, 1994, p. 41). Il est possible de voir en quoi la « nouvelle santé publique » tout en s'en distinguant, a renoué avec ses racines hygiénistes.

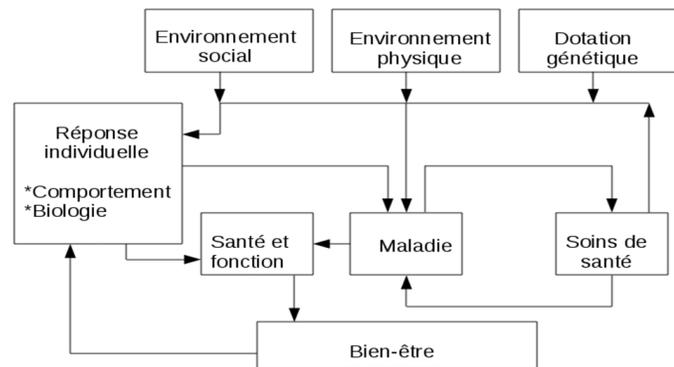
1.1.3 La santé publique contemporaine

Dans la foulée du rapport Black (Working Group on Inequalities in Health, 1980) et de l'étude de Whitehall (Marmot, 1994a, b) en Angleterre, le développement le plus marquant des dernières décennies, tant au niveau de l'action que de la recherche en santé publique, est l'intérêt pour les déterminants sociaux de la santé et les inégalités sociales de santé. Au Canada et au Québec, cet intérêt a particulièrement marqué le développement de la recherche en santé publique au cours des deux dernières décennies, notamment suite à la publication de « *Why are some people healthy and others not* » en 1994 (Evans et al., 2004). L'importance de ce thème de recherche a été reconnue lors de la création des Instituts de recherche en santé du Canada (IRSC) en 2000, qui ont reçu comme mandat de développer la recherche en santé dans cinq grands domaines dont celui des dimensions sociales et culturelles de la santé. Émerge ainsi ce qui sera nommé la « nouvelle santé publique ». Pour souligner l'importance grandissante de cette perspective, notons qu'en 2005 l'Organisation mondiale de la Santé (OMS) a établi la « Commission des déterminants sociaux de la santé » afin de réduire les inégalités de santé liés aux conditions de vie des populations (Marmot et al., 2009). La prise en compte des inégalités sociales en matière de santé s'est accentuée encore au courant des années 1990, démontrant encore davantage à quel point les écarts de santé se corrélaient au statut socio-économique. L'évolution de la santé publique l'amène aujourd'hui à accorder une place plus importante aux aspects sociologiques liés à la santé des populations. La « promotion de la santé » est une approche de la santé axée sur la prise en compte de divers types de déterminants de la santé, dont les « déterminants sociaux de la santé » et sur des interventions visant à réduire les inégalités sociales de santé.

Le modèle de santé des populations (voir la figure 1.1 en page 13) proposé par Evans

et Stoddart, propose une reconceptualisation des déterminants de la santé des populations incluant parmi ceux-ci, l'environnement physique et social et le bagage génétique. Les comportements y sont largement déterminés par l'environnement physique et social. Quant aux soins de santé, ils sont vus davantage comme une réponse à la maladie que comme un déterminant de la santé des populations.

Figure 1.1 – Relations entre santé et facteurs sociaux et environnementaux : Schéma de Evans et al.



SOURCES Issue de Evans et al. (2004, p.46)

Ce modèle a été fortement critiqué pour le peu d'importance accordée aux soins de santé comme déterminants de la santé (Poland et al., 1998). Les travaux récents tendent à démontrer que si les soins de santé ne sont pas le seul ou principal déterminant de la santé des populations, ils n'en sont pas moins un déterminant important. Toutefois, l'attention que ce modèle a reçu, notamment de la part des décideurs et des acteurs de la santé publique, a grandement favorisé le développement de la recherche sur les déterminants sociaux de la santé et de l'action publique sur ceux-ci.

La compréhension nouvelle de la santé comme étant un phénomène social et non pas strictement biologique demeure néanmoins une transformation majeure de la santé publique. (McQueen et al., 2007). Malgré ce tournant majeur, pour nombre de chercheurs en sciences sociales, la santé publique ne porte pas encore assez d'attention aux théories sociales ; « Nevertheless, public health has continued to concentrate on changing individual behaviours and exposures to pathology rather than targeting the conditions influencing those behaviours and exposures. » (Bauman et al., 1998, p. 68)

Comme nous l'avons montré dans cette section, la santé publique évolue par cumul plutôt que par substitution : de nouvelles fonctions et de nouveaux objets apparaissent sans que les fonctions et les objets antérieurs disparaissent. Par conséquent, le champ de la santé publique s'élargit constamment. Au Québec, actuellement, la santé publique couvre un large spectre, comme l'illustre la figure 1.2. Et, suivant de près le modèle de santé des populations de Evans et al., sont considérés comme déterminants de la santé « les prédispositions biologiques des personnes, les comportements qu'elles adoptent, les relations familiales et sociales et, plus largement, les conditions de vie, l'environnement physique ainsi que la qualité et l'accessibilité aux services sociaux et de santé ». (Cardinal et al., 2003, p. 15)

Figure 1.2 – Structure du programme national de santé publique



SOURCES : Issue de Cardinal et al. (2003, p.7)

La recherche en santé publique se déploie de manière à couvrir l'ensemble du champ de la santé publique. Dans la prochaine section, nous examinerons comment se construit la connaissance en santé publique

1.2 Comment le savoir se construit ?

1.2.1 Construction de la connaissance en santé publique

La santé publique n'est pas qu'institution sociale, elle représente aussi une discipline académique. Par exemple, l'École de santé publique de l'Université de Montréal (ESPUM) offre depuis quelques années un programme de Ph.D. en Santé publique. C'est une discipline académique qui se veut interdisciplinaire et regroupe des chercheurs des sciences humaines et sociales, de médecine, des sciences naturelles et d'autres horizons encore. Il n'est cependant pas nécessaire de détenir un doctorat en santé publique pour faire de la recherche en santé publique. L'important ici est de comprendre que la santé publique regroupe de nombreuses disciplines académiques en interaction et qui participent toutes à la construction des connaissances en santé publique.

Mullins (1975) s'est intéressé à la question de l'origine du changement au cours du développement des disciplines scientifiques. Ce chercheur confronte deux visions du changement : l'approche interne du changement, idéaliste, sous-tend que ce sont les chercheurs eux-même qui développent les idées dans leur esprit ; tandis que l'approche externe, matérialiste, sous-tend que les changements sont plutôt produits par les structures sociales. (Mullins, 1975) Dans une approche interactionniste, il démontre que les changements sont à la fois d'origine interne et externe : il s'agit de prendre en compte simultanément les structures externes aux disciplines (structure de financement, de communication, etc.) et celles qui y sont internes (positions et interactions des chercheurs dans les réseaux épistémiques, formation des étudiants, etc.). Le cas de la santé publique ne nous semble pas échapper à cette dynamique. Il y a, à la fois des réseaux de chercheurs réunis par des idées communes contribuant au développement de cette discipline et des conditions externes à ces communautés épistémiques : motivations économiques, politiques, institutions sociales, valorisation des intérêts des classes dirigeantes... Pour notre part, nous focaliserons essentiellement sur l'analyse des aspects internes influençant le développement de la santé publique, c'est-à-dire sur les relations entre les chercheurs et thèmes de recherches.

Selon Karin Knorr Cetina, la culture épistémique explique « comment nous savons

ce que nous savons »⁴. Plus spécifiquement, elle démontre comment à la fois, les cultures épistémiques pénètrent les sociétés actuelles, « les sociétés du savoir » et comment ces cultures épistémiques sont aussi imprégnées des « sociétés du savoir. » (Cetina, 2000) Le domaine d'expertise que constitue la santé publique a lui aussi ses propres cultures épistémiques. Dozon et Fassin (2001) parlent de la diversité culturelle de la santé publique qui, loin d'être stable et homogène, dépend en partie des contextes nationaux et de circonstances historiques. (Dozon et Fassin, 2001).

Les diverses cultures épistémiques en santé publique sont en interaction. Pour White (2008), les formes que peuvent prendre ces interactions se résument à trois « disciplines »: Soit on se rallie (*council*), soit on négocie (*arena*), soit on entreprend une action ayant un objectif extérieur (*interface*). Une discipline est une forme « d'effort de contrôle » typique et situationnel qui régule les interactions sociales. (Grossetti et Godart, 2007) En d'autres mots, dans une situation donnée, les interactions sociales vont prendre une des trois formes typiques. Avant de décrire davantage ces disciplines, il nous semble important de préciser que les « efforts de contrôle » que déploient les « identités sociales » sont au cœur de l'explication whitienne de l'action sociale. (Tilly, 1993) L'approche whitienne sous-tend que les diverses formations sociales que nous pouvons observer, par exemple un centre de recherche en santé publique, sont le produit d'efforts de contrôle que les identités sociales déploient afin de réduire l'incertitude endémique dans laquelle elles sont plongées : « Contrôle ici ne signifie pas le pouvoir, mais plutôt une recherche d'appuis ou d'ancrages qui peuvent apporter une certaine stabilité à une identité et une tentative de maîtriser un environnement turbulent. » (Grossetti et Godart, 2007, p. 10)

Pour en revenir aux disciplines, illustrons nos propos à l'aide d'interactions en santé publique. Lorsque les chercheurs en santé publique confrontent leurs diverses approches de la santé, ils sont dans une discipline de type *arena*, où l'échelle de valeur est la « pureté ». La sélection s'opère alors sur des critères de pureté tels (la valeur scientifique, l'expertise apportée, etc.) Le processus en cours est celui de la sélection, c'est-à-dire qu'ici, il importe de « trier le bon grain de l'ivraie » en réfutant ou confirmant

4. Traduction libre de : "how we know what we know" dans Cetina (2000)

les résultats présentés par les pairs. Lorsque ceux-ci exhortent les autorités publiques à mettre en place des programmes d'intervention en santé publique ou encore les autres chercheurs à s'intéresser à tel sujet de santé jugé des plus primordiaux, ils sont dans une discipline de type *council*. L'échelle de valeur associée à cette discipline est le prestige, au sens où il s'agit d'influencer l'action collective, de rallier. Le processus qui s'opère dans une interaction de type *council* est celui de la médiation. Finalement, les chercheurs en santé publique sont aussi engagés dans une dynamique d'échange entre producteurs de connaissance en santé publique et demandeurs de telles connaissances. Cette forme d'interaction est dite l'*interface* car son rôle est davantage d'ordre logistique. L'échelle de valeur qui lui est associée est celle de la qualité et le processus est celui de l'engagement. Ce processus mène les acteurs à poursuivre un objectif extérieur commun, celui de produire de la connaissance en santé publique.

Nous avons abordé les disciplines chez White car ces formes typiques que peuvent prendre les interactions ont aussi, à notre avis, une influence sur la construction de la connaissance en santé publique. La connaissance en santé publique vise à la fois, qualité, pureté et prestige et s'opère tant par sélection, médiation, qu'engagement. Comme de nombreux auteurs l'ont démontré auparavant, la connaissance est le produit du social. D'autre part, à l'instar de l'idée soutenue par la notion de « modernité réflexive » (Beck, 2008, Giddens, 1990), nous contemporains, sommes imprégnés des apports continus de la connaissance (scientifique). La connaissance et l'expertise transforment les représentations et actions sociales, tout autant que les actions et représentations sociales influent sur la connaissance et l'expertise. La santé publique n'y échappe pas.

1.2.2 Les communautés épistémiques

Il n'est pas rare d'entendre parler de la « communauté » des chercheurs en santé publique au sens de l'ensemble des chercheurs œuvrant en santé publique. Cette définition des plus simples laisse néanmoins percevoir que ces chercheurs sont liés par un thème de recherche commun, celui de la santé publique. Elle permet aussi de mettre en évidence l'importance d'analyser non seulement des acteurs sociaux, mais de prendre simultanément en compte ce qui les unit, i.e. les thèmes de recherche qu'ils partagent.

Une communauté épistémique se constitue à la fois d'un réseau social et d'un réseau sémantique. (Barabasi et al., 2002, Newman, 2001, Roth et al., 2009) Selon Roth et al. (2009), c'est la co-évolution entre concepts et agents qui explique la dynamique et la morphogenèse des communautés épistémiques.

Les communautés épistémiques sont donc ici conceptualisées comme étant des réseaux de chercheurs œuvrant sur des thèmes communs. Il existe dans la sociologie whitienne une notion pour exprimer ce type de relations entre sémantique et social : « La notion de « *netdom* » (pour *network-domain*) rend compte de cette dualité entre les domaines sémantiques, les réseaux de sens et les réseaux sociaux. » (Grossetti et Godart, 2007, p. 7) Cependant, les *netdoms* ne sont pas des choses, mais des amalgames de relations et de thèmes (White, 2008). Le thème de la santé publique par exemple est un large *netdom* : un réseau de gens œuvrant autour d'un thème commun. Dans un *netdom*, la relation est spécialisée, c'est-à-dire qu'elle est directement liée au thème unissant des gens en réseaux.

Les gens dans les *netdoms* sont considérés selon l'approche whitienne comme constituant un « public »: « Des identités en relation, spécialisées, liées à un « *netdom* » constituent un « public ». (Grossetti et Godart, 2007, p. 17) En santé publique, nous avons vu que de nombreuses identités interagissaient autour de la santé publique : corps médical, industriels, ingénieurs, autorités publiques, citoyens, bourgeoisie, malades, pour ne nommer que ceux-là, il s'agit du public de la santé publique.

Dans l'étude d'une communauté épistémique, ces concepts deviennent utiles pour appréhender la question des interactions entre divers *netdoms*. Lorsqu'une identité passe d'un *netdom* à un autre, elle s'aventure dans un différent espace de sens, et selon White, ce sont ces commutations qui amènent des idées nouvelles. (White, 2008) Peut-on l'observer en santé publique ? Ces concepts sont centraux dans la théorie whitienne car ils permettent de rendre compte à la fois de l'action sociale et des transformations sociales : « Thus, the world comes from identities attempting control within their relations to other identities. In their search for control, identities switch from *netdom* to *netdom*, and each switching is at once a decoupling from somewhere and embedding into somewhere. » (White, 2008, p. 2)

L'action sociale, provient des efforts de contrôle qu'émettent les identités lors d'interactions. Lors des interactions, les identités produisent des histoires qui peuvent permettre l'émergence de liens ou encore de relations. (White, 2008) Un médecin soignant des malades, la production d'une publicité gouvernementale sur les bienfaits de l'exercice physique ou encore un dépressif prenant un anxiolytique, aucun de ces gestes n'a de sens sorti de son contexte et sans les histoires qui contribuent à constituer les liens entre ces identités. L'importance des histoires dans la constitution de réseaux sociaux s'entrevoit dans notre prise en compte des thèmes de recherche manipulés par les chercheurs de la santé publique, ce qui les lie n'est pas seulement « social » mais aussi sémantique, un réseau de sens : « Les histoires sont donc un médium essentiel pour la constitution, non seulement des « réseaux sociaux », mais aussi de toutes les configurations sociales plus complexes qui en dérivent. » (Grossetti et Godart, 2007, p. 15)

C'est via les concepts d'encastrement et de découplage que la sociologie whitienne permet d'appréhender la reproduction sociale et le changement :

« L'encastrement et le découplage permettent de penser la dynamique d'émergence ou de dissolution d'entités agissantes ou de formes sociales, qu'il soit question d'identités individuelles, d'organisations, ou de marchés. Toute entité est à la fois encadrée dans les liens qu'elle a tissés avec d'autres identités, et découplée, c'est-à-dire disposant d'une marge de manoeuvre spécifique. Tout découplage débouche sur un nouvel encastrement (?). » Grossetti et Godart (2007, p. 38)

Pour prendre un exemple, les chercheurs en santé publique s'encastrent dans un univers de sens et de pratiques spécifiques qui tend à déterminer dans un sens faible leurs actions en tant que chercheurs. Ils se découplent aussi en rejoignant diverses *netdoms*, par exemple le thème de la « santé mentale » où ils vont s'encastrent à nouveau. Les chercheurs ont ainsi la possibilité de sortir d'un certain cadre pour en rejoindre un autre. Le découplage permet l'émergence de nouvelles formes sociales et de nouveaux *netdoms*. L'approche whitienne nous semble des plus utiles pour l'étude des communautés épistémiques en santé publique, des relations entre thèmes et chercheurs, et pour saisir le développement de ce domaine de la vie sociale, .

1.2.3 Questions de recherche

En nous inspirant de la sociologie whittienne, notre problématique se centre sur une approche interactionniste-structurale des communautés de savoir et a pour but d'étudier les transformations simultanées, entre réseaux sociaux et réseaux de sens afin de mieux appréhender la gestion de telles structures. Nous cherchons à étudier de la façon la plus formelle et rigoureuse possible l'émergence d'une communauté épistémique spécifique, l'IRSPUM, et de voir comment les éléments sociaux et sémantiques interagissent entre eux. Plus spécifiquement, ce que nous voulons savoir est :

Est-ce que l'IRSPUM représente effectivement une communauté épistémique au sens où les acteurs sociaux la composant constituent un réseau de collaboration et convergent tous vers une problématisation commune de la santé publique ? Ou si, au contraire, l'IRSPUM est constitué de plusieurs communautés épistémiques, plus ou moins isolées les unes des autres, se constituant soit sur la base de leur appartenance disciplinaire, soit autour des thématiques manipulées par les chercheurs. Est-ce que la « nouvelle santé publique » se retrouve dans le paysage sémantique du groupement épistémique que nous avons choisi d'étudier ? Est-elle pleinement intégrée dans le champ de la santé publique ou si elle constitue un champ à part ?

Nous utiliserons l'analyse de réseau pour répondre à ces questions. Nous examinerons d'abord les liens de collaborations entre les chercheurs puis leur regroupement autour de thématiques de recherche, toujours en tenant compte de l'appartenance disciplinaire des chercheurs. Le prochain chapitre présente la méthodologie utilisée dans cette étude.

CHAPITRE 2

MÉTHODOLOGIE ET OUTILS D'ANALYSE

L'objectif de ce chapitre est d'explicitier la méthodologie et les outils d'analyse qui ont servi le double objectif de ce mémoire : d'une part à étudier la formation d'une communauté épistémique particulière en examinant les collaborations entre chercheurs et leur regroupement autour de thématiques ; d'autre part, à examiner la place spécifique qu'y occupent les déterminants sociaux de la santé. Comme mentionné au chapitre 1, la conceptualisation d'une communauté épistémique se fonde ici sur l'idée que les chercheurs sont liés entre eux, non seulement à travers leurs interactions sociales, mais aussi (et surtout) à travers les univers de sens dans lesquels ils œuvrent. Pour atteindre nos objectifs de recherche, nous avons opté pour une approche en terme de réseaux.

Dans ce chapitre, nous présenterons d'abord comment nous avons construit le corpus de données, puis comment nous avons procédé pour classifier les champs disciplinaires des chercheurs et surtout les thématiques des articles, pour ensuite présenter les méthodes d'analyse utilisées, soit l'analyse de réseaux et les treillis de Galois.

2.1 Le corpus de données

Le corpus des données est constitué des publications des chercheurs membres réguliers ou associés de l'IRSPUM en 2009, lors de sa création. L'IRSPUM regroupait en septembre 2009 226 chercheurs en santé publique. Il se décrit comme étant un centre de recherche interdisciplinaire qui s'organise autour des « grands enjeux contemporains en matière d'amélioration de la santé des populations et de gouvernance des soins de santé ». ¹ Rappelons que l'IRSPUM est constitué essentiellement des chercheurs issus de trois centres de recherche lui préexistant : Le groupe de recherche sur les aspects sociaux de la santé et de prévention (GRASP), le Groupe de recherche interdisciplinaire en santé (GRIS) et le centre Léa-Roback. Le premier s'axait sur les déterminants sociaux

1. Sources : <http://www.gris.umontreal.ca/actualite.asp>

de la santé, le second sur l'administration et la promotion des soins de santé ainsi que sur l'environnement, et le dernier sur les inégalités sociales de santé.

C'est à travers les publications des chercheurs de l'IRSPUM que nous avons décidé d'explorer le réseau social et sémantique de cette communauté épistémique. Avec qui et sur quoi ont-ils écrit dans le cadre de leur travail de chercheur ? Comme il était trop ardu de collecter et d'analyser le contenu sémantique de tous les types de publications possibles en recherche, nous avons choisi de ne collecter que les articles publiés. Nous avons retenu tous les articles publiés par les membres de l'IRSPUM au cours de leur carrière.

Concernant le choix du type de publication à retenir, nous avons opté pour l'article car c'est l'une des formes contemporaines les plus fréquentes de la publication scientifique (Beaver et Rosen, 1978). En outre, les articles peuvent être perçus comme des artefacts ou des traces que les chercheurs ont laissés derrière eux au cours de leur carrière (Roth, 2006, 2008). Les articles que les chercheurs ont publiés contiennent de nombreuses informations concernant l'évolution de la recherche : Quand l'article a-t-il été publié ? Qui l'a écrit ? De quoi traite l'article ? Qui a collaboré avec qui et sur quel(s) thème(s) ?

L'accès aux articles complets étant souvent restreint, nous avons opté, comme d'autres chercheurs (Barabasi et al., 2002, Newman, 2001, Roth, 2008), pour ne collecter que les noms des auteurs, l'année de publication, le titre et le résumé des articles publiés.² Ces données nous informent sur les collaborations épistémiques, les collaborations et les thèmes de recherche traités, au fil du temps par un groupe de chercheurs qui se retrouveront tous membres de l'IRSPUM lors de sa création en 2009.

Notre base de données brute est un fichier contenant l'ensemble des articles publiés par les chercheurs de l'IRSPUM au cours de leurs carrières de chercheur, sous forme de bibliographies indexées dans un logiciel de gestion bibliographique.³ Dans la sec-

2. Nous n'avons cependant tenu compte que des publications comprises dans les 20 dernières années dans nos analyses puisque le nombre de publications collectées datant d'avant 1990 était trop faible pour en permettre une analyse cohérente

3. Nous avons utilisé le logiciel *JabRef* version 2.6. Voir : JabRef Development Team (2009) « JabRef Version 2.6 » pour plus d'information sur ce logiciel.

tion suivante nous expliciterons davantage la méthodologie entourant la collecte de nos données.

2.2 Collecte des données

Cette section vise dans un premier temps à présenter comment s'est effectuée la collecte de nos données. Comme mentionné au préalable, la principale source de données dont nous disposons correspond à une liste bibliographique d'articles publiés par les chercheurs de l'IRSPUM au cours de leurs carrières et indexée dans un logiciel de gestion de données bibliographiques. Son format brut est totalement numérisé. Par la suite, nous veillerons à décrire le contenu de notre base de données.

Notre matériau primaire est l'ensemble des notices bibliographiques des articles publiés des chercheurs membres de l'IRSPUM au cours de leur carrière. Bien que, grâce aux outils contemporains et aux bibliothèques en ligne les articles puissent sembler simples à collecter, la réalité en est tout autre ! Surtout, lorsqu'il est question de collecter des publications de chercheurs en particulier, plutôt que de collecter tous les articles publiés à un moment donné sur un thème donné, sans se soucier de collecter des articles sur des auteurs spécifiques. Nous y avons porté une sérieuse minutie lors de la collecte de nos données puisque celles-ci constituent le socle empirique sur lequel s'appuient nos analyses.

Deux considérations préliminaires distinguant ce travail d'autres similaires sont à prendre d'emblée en compte. En premier lieu, précisons que la « scientificité » des articles collectés n'a pas été prise en compte, bien qu'en général les chercheurs en analyse de réseaux épistémiques excluent les articles « non scientifiques » de leur corpus empirique. La santé publique a fondamentalement une composante épistémique (au sens d'entité productrice de connaissances), mais nous ne nous aventurerons pas à trancher sur la scientificité des connaissances produites par les chercheurs de l'IRSPUM. Outre la difficulté technique consistant à repérer des articles qui seraient jugés comme n'étant pas « scientifiques » selon une définition précise, ce type d'exclusion semble apporter moins d'avantages que d'inconvénients pour notre étude.

La seconde considération à retenir concerne notre volonté de ne pas exclure d'article sur des critères linguistiques. Bien que la recherche en santé publique, comme en général, se publie majoritairement en anglais, l'IRSPUM est situé dans un milieu francophone et nombre des chercheurs contenus dans notre base de données publient aussi en français. La langue de publication ne nous a pas semblé problématique compte-tenu de la possibilité de traduire les titres et résumés d'articles afin d'obtenir un corpus uniforme du point de vue linguistique. Nous expliciterons ces considérations ainsi que les autres éléments méthodologiques mobilisés afin d'uniformiser le corpus plus tard dans la section concernant les traitements des données et les méthodes d'analyse utilisées.

Tout comme les autres chercheurs employant les publications d'articles en analyse de réseaux épistémiques, nous avons employé pour cette tâche de recherche et de collecte de données brutes, des moteurs de recherche de données bibliographiques en ligne. Compte-tenu de notre découpage structurel centré sur l'IRSPUM, le nom de chacun des chercheurs de l'IRSPUM a été recherché sur trois bases de données bibliographiques différentes : *PubMed*⁴, *ISI Web of Knowledge*⁵ et *Érudit*.^{6,7} La spécialisation de *PubMed* dans les articles biomédicaux et concernant la santé rend son usage des plus pertinents pour la collecte de publications de chercheurs œuvrant en santé publique. Quand à *ISI*, il recense davantage de revues francophones que *PubMed* et concerne un champ plus large de la recherche. Enfin, compte tenu de la localisation géographique de l'IRSPUM, le recours à *Érudit* se justifie puisque cette base de données se concentre sur les publications canadiennes et s'axe davantage sur les sciences sociales que les deux premières bases de données mentionnées, son usage s'avère pertinent. Le recours à ces bases de données bibliographiques fait en sorte que le format des données collectées soit entièrement

4. <http://ncbi.nlm.nih.gov/PubMed>

5. <http://isiwebknowledge.com>

6. <http://www.erudit.com>

7. Afin d'obtenir un maximum de publications, la recherche des publications pour chaque chercheur s'est effectuée en entrant dans chacun des moteurs de recherche utilisés, d'abord le nom complet d'un chercheur donné (nom et prénom) puis en ne recherchant que par le nom de famille. Cette démarche fastidieuse nous a permis de collecter un maximum de publications, puisque les formes d'étiquetage des noms d'auteurs sur les articles sont très variées : parfois il n'y a que les initiales du prénom et le nom de famille, parfois le nom est complet, mais sans les seconds prénoms, etc. Bref, il n'y a pas de standard effectif en ce domaine.

numérique.

La collecte des données s'est effectuée du 1^{er} septembre 2009 au 1^{er} Mars 2010. Les articles retenus sont ceux publiés au plus tard en décembre 2009 : les articles de 2010 sont exclus. Un tableau contenant notamment la liste des chercheurs de l'IRSPUM en date du 1^{er} septembre 2009⁸ a servi de guide afin de collecter spécifiquement les articles rédigés par ces chercheurs. Cette liste de chercheurs contient aussi des informations concernant les disciplines dans lesquelles les chercheurs ont été formés ainsi que sur leur statut au sein de l'IRSPUM. Elle se trouve en annexe I, à la page xv.

Une fois les articles collectés, nous avons par la suite exclu des analyses les articles n'ayant pas de résumé parce qu'ils ne permettent pas l'analyse des thèmes de recherche, le titre n'étant pas toujours assez précis pour classer ces articles. Nous avons aussi exclu de notre corpus analysé, les articles d'avant 1990, bien que les plus anciens articles collectés datent de 1969. La raison en est que trop peu d'articles appartiennent à la période précédant 1990. Sur 10 ans, entre 1969 et 1979 inclusivement, il y a eu 54 articles publiés par des chercheurs actuellement à l'IRSPUM. Par la suite, entre 1980 et 1989 inclusivement, il y a eu 318 articles publiés, ceux-ci concernant moins de 30 chercheurs. Nous n'avons donc retenu que les articles publiés entre le 1^{er} janvier 1990 et le 31 décembre 2009, inclusivement. Au total 3165 articles se retrouvent dans notre corpus.⁹

Il importe de mentionner maintenant que comme toute autre recherche s'appuyant sur des sources documentaires, une importante phase de construction et de nettoyage des données a dû être effectuée. En outre, un effort considérable a été déployé afin de normaliser les noms des chercheurs. Les noms étant inscrits sous plusieurs formes dans les bibliographies, il a fallu détecter tous les homonymes des chercheurs de l'IRSPUM qui auraient pu se glisser dans nos données et mettre toutes les variantes possibles de noms sous une même forme, afin de bien lier chaque article à son ou ses véritable(s) auteurs et de standardiser la façon d'écrire les noms des auteurs. L'analyse de réseaux nous a permis par exemple de détecter si un « A. Demers » était bien « Andrée Demers »

8. Voir l'annexe I en page xv

9. la liste des publications est accessible à cette adresse : <http://idealites.info/IRSPUM.html>

ou pas, en analysant les co-auteurs des articles visés par cette ambiguïté. Les divers scripts de normalisation des noms ont été mis en œuvre dans l'outil *MouliNet* que nous avons créé pour les fins de notre recherche.¹⁰

Afin de rendre possible une analyse longitudinale de notre corpus empirique, nous avons découpé nos données en quatre périodes couvrant chacune cinq années. Comme mentionné plus tôt, nos premiers essais nous ont indiqué que les périodes qui précèdent 1990 concernaient trop peu de cas pour nous permettre d'en tenir compte dans nos analyses. Pour rendre compte le plus justement possible de ce que nos données représentent, notre découpage s'inspire d'un compte à rebours : puisque la mise sur pied de l'IRSPUM survient dans la dernière année couverte par notre corpus, la période qui couvre son apparition fait référence au « temps zéro » (T0).

Pour décrire notre découpage temporel en ordre chronologique, la période allant du 1^{er} janvier 1990 au 31 décembre 1994 inclusivement (T-3) représente la période la plus ancienne couverte par notre corpus ; c'était il y a 15 à 20 ans. Des 226 chercheurs membres de l'IRSPUM en 2009 lors de notre enquête, nous avons retrouvé des articles publiés pour 65 d'entre eux. Ensuite, vient le « temps -2 » (T-2) qui s'étend du 1^{er} janvier 1995 au 31 décembre 1999 inclusivement. Nous avons collecté des articles publiés au T-2 pour 82 chercheurs. Notons ici qu'il ne s'agit pas d'une simple addition : des 65 chercheurs pour lesquels nous avons collecté des données au T-3 certains ne sont pas représentés en T-2. Les détails de ces mouvements seront exposés dans le chapitre portant sur les résultats de nos analyses de réseaux. La période qui précède celle qui verra l'IRSPUM se concrétiser est appelée « temps -1 » (T-1) et concerne les articles publiés par les chercheurs qui seront membres de l'IRSPUM en 2009 entre le 1^{er} janvier 2000 au 31 décembre 2004 inclusivement. À cette période, nous retrouvons dans nos données 186 chercheurs sur les 226 qui seront membres de l'IRSPUM. Au T0, soit du 1^{er} janvier 2005 au 31 décembre 2009 inclusivement, nos données concernent 206 chercheurs ; c'est-à-dire que nous n'avons pas été en mesure de repêcher au moins un article publié durant le T0 pour 20 des 226 chercheurs embauchés par l'IRSPUM lors de notre collecte de données.¹¹ Le

10. Pour plus amples informations concernant cet outil : voir l'annexe V.1 à la page xlvi

11. Nous avons porté une attention particulière aux chercheurs dont nous n'avons pu retrouver de publications, afin de nous assurer d'avoir un corpus des plus exhaustifs.

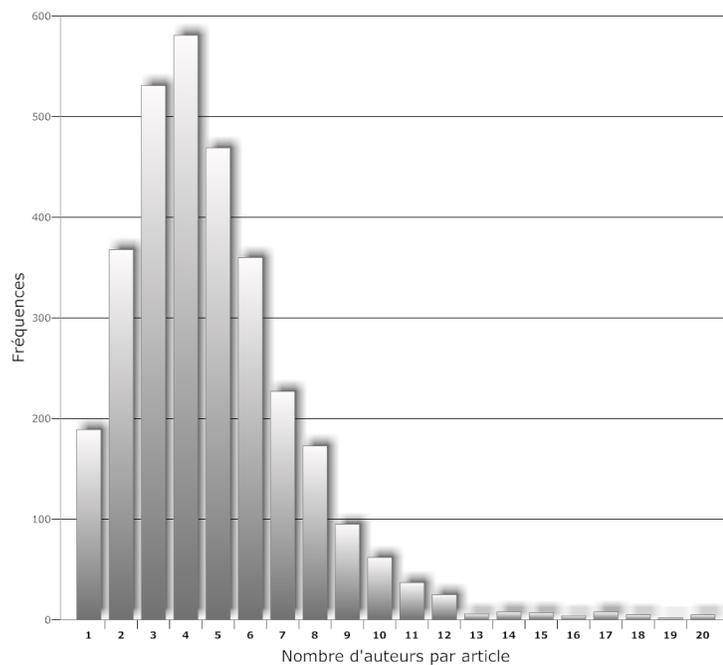
tableau 2.I de la page 27 présente une synthèse du nombre de chercheurs et d'articles pour chacune des périodes.

Tableau 2.I – Synthèse du nombre de chercheurs et d'articles pour chaque période

Temps:	Descriptif	N. de chercheurs	N. d'articles
T-3	(1 ^{er} janvier 1990 au 31 décembre 1994)	65	345
T-2	(1 ^{er} janvier 1995 au 31 décembre 1999)	82	525
T-1	(1 ^{er} janvier 2000 au 31 décembre 2004)	186	964
T0	(1 ^{er} janvier 2005 au 31 décembre 2009)	206	1331

Il est important de rappeler que cette base de données est constituée des articles publiés au cours de chacune de ces périodes par les chercheurs qui ont formé l'IRSPUM en 2009. Elle ne représente donc pas l'ensemble des publications en santé publique et ne peut être utilisée pour étudier l'évolution de la recherche en santé publique à l'Université de Montréal depuis 1990.

Figure 2.1 – Histogramme de la distribution du nombre d'auteurs par article



SOURCES : Les données sont obtenues à partir de l'analyse de la base de données COLLABORATIONS à l'aide du logiciel SPSS

La figure 2.1 de la page 27 montre le nombre d'auteurs moyen par article. Le nombre d'auteurs moyen par article est de quatre chercheurs, ce qui correspond aux résultats obtenus par nos collègues (Newman, 2001, Roth, 2008). Nous avons exclu de ces analyses trois cas extrêmes d'articles signés par plus de 20 collaborateurs, ce qui porte le nombre d'articles soumis à cette analyse à 3162. Seulement 189 (6%) des articles ont été publiés par un seul auteur et 107 (4%) par plus de 10 chercheurs. Une collaboration de deux à six chercheurs sur un même article se présente dans 2309 articles (ce qui concerne 73% des articles, c'est-à-dire près des trois quarts des publications). Le mode est de quatre auteurs par article. La collaboration semble ainsi la norme en recherche à l'IRSPUM, comme c'est le cas généralement dans la recherche contemporaine.

2.3 La classification des chercheurs et des thématiques

2.3.1 La classification des chercheurs selon leur champ disciplinaire de formation

Pour déterminer le champ disciplinaire des chercheurs, nous avons utilisé la liste des chercheurs qui nous a été fournie par la direction de l'IRSPUM.¹² Cette liste recense plus d'une cinquantaine de domaines d'expertise distincts. Lorsque deux formations étaient inscrites, nous n'avons pris que la première. Il s'agit principalement du cas des médecins. Ils ont été classés sous « Médecine/Sciences biomédicales ». Nous avons aussi porté un intérêt sur les domaines dans lesquels ont été formés les chercheurs de l'IRSPUM, afin d'étudier la relation entre domaines de formation et thèmes abordés en recherche ainsi que dans le but d'explorer les aspects possibles de la collaboration interdisciplinaire entre ces chercheurs. Afin d'alléger l'analyse, nous avons réduit le nombre de ces disciplines à neuf (9), en fusionnant les domaines apparentés. Nous avons par exemple, fusionné « psychologie sociale », « psychologie », « psychologie clinique » « psychoéducation » et « psychométrie » dans une seule catégorie : « Psychologie ». Voici la liste de ces disciplines fusionnées:

1. « Santé publique et communautaire »
2. « Sciences humaines et sociales »

12. Cette liste est consultable en annexe I en page xv

3. « Médecine et sciences médicales »
4. « Épidémiologie et biostatistiques »
5. « Psychologie »
6. « Biologie et chimie »
7. « Sciences vétérinaires et zoologie »
8. « Sciences environnementales »
9. « Administration »

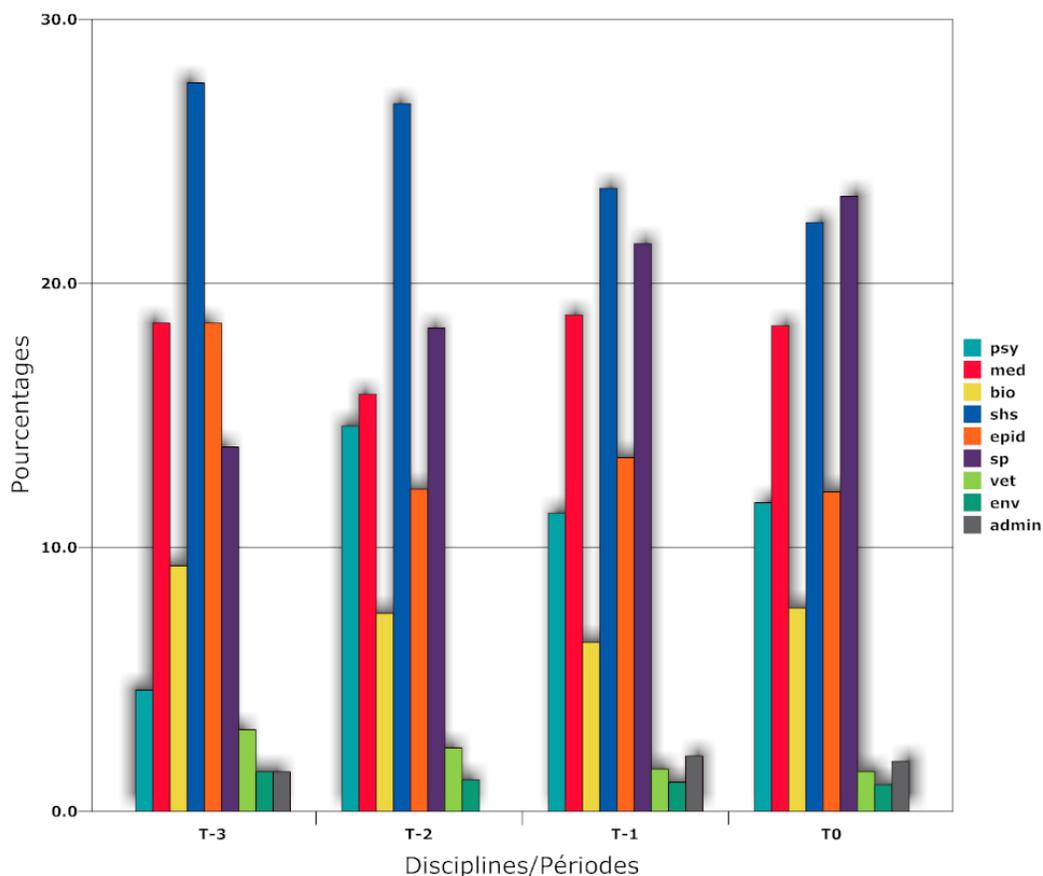
Bien que la plupart des fusions soient en apparence faciles à déduire, certaines demandent à être détaillées. C'est le cas des « Sciences humaines et sociales », qui regroupe l'aménagement, le droit, l'économie, l'histoire, l'anthropologie, la sociologie, les sciences politiques, la philosophie, la pharmacoéconomie, la démographie, les sciences humaines appliquées, l'éducation et les relations industrielles. C'est aussi celui de la catégorie « Biologie et chimie », qui outre contenir les disciplines de la biologie, de la biochimie, de la chimie inorganique, pour ne nommer que ces disciplines, contient aussi celles de la pharmacologie et de la toxicologie qui sont des disciplines chapeautées par la chimie et la biologie. L'ensemble des fusions est donné dans l'annexe III en page xl. Pour ces analyses descriptives, la totalité des chercheurs de l'IRSPUM (en date du 1^{er} septembre 2009) a été prise en compte, soit 226 cas.

Si les neuf grandes disciplines sont toutes présentes au T0¹³, ce n'est pas le cas lors des périodes précédant celle de la création de l'IRSPUM. Grâce à la figure 2.2 en page 30, il nous est possible d'explorer la distribution des disciplines dans lesquelles ont été formés les chercheurs de l'IRSPUM.

Les neuf grandes disciplines dans lesquelles les chercheurs de l'IRSPUM ont été formés ainsi que les périodes couvertes par notre étude sont représentées à la figure 2.2. Le pourcentage de chercheurs issus des sciences humaines et sociales (shs) diminue progressivement au cours des 20 dernières années, tandis que celui des chercheurs en santé publique, pour sa part augmente. En T0, 23% des chercheurs provenaient de santé publique et 22% des shs, en T-3 ils étaient 14% et 28%, respectivement. Si l'écart était important en T-2 il a presque disparu en T0. La proportion des chercheurs issus

13. Période allant du 1^{er} janvier 2005 au 31 décembre 2009 inclusivement

Figure 2.2 – Au cours du temps, distribution des chercheurs selon leurs disciplines de formation



SOURCES : Les données sont obtenues à partir de l'analyse de la base de données TOPICS à l'aide du logiciel SPSS. La liste des abréviations est fournie à la page xi

d'épidémiologie ou des biostatistiques décroît en T-2 pour demeurer à peu près stable par la suite (à environ 12%). La proportion des chercheurs formés en psychologie pour sa part triple de T-3 à T-2, pour se stabiliser à 2,5 fois la représentativité du T-3 (soit aussi à près de 12%). La proportion des chercheurs formés en médecine et sciences médicales varie peu (de 3 points de pourcentage) ; elle se situe à 18% en T0. Il en va de même pour les chercheurs ayant été formés en sciences biologiques et/ou chimiques, qui varie entre 6% et 9% selon les périodes, pour se situer à 8% en T0. La proportion des chercheurs en sciences environnementales elle aussi demeure stable (autour de 1,5%), tandis que celle des chercheurs en santé animale diminue légèrement, pour se stabiliser

en T-1 à 1%. Aucun chercheur provenant des sciences administratives n'est présent en T-2, mais pour les autres périodes, cette discipline est représentée à 2%.

Ces résultats, somme toute descriptifs permettent de mettre en lumière l'importance de la collaboration dans la recherche faite par les chercheurs de l'IRSPUM ; la plupart des articles publiés par ces chercheurs sont l'œuvre de collaborations. Nous pouvons constater que l'épidémiologie, plus ancienne discipline associée à la santé publique reste présente, mais semble de moins en moins représentée au fil des périodes d'analyse. Ils rendent compte aussi de la distribution du nombre de thèmes traités dans chaque article. Ces résultats s'accordent avec ceux obtenus par Roth (2008) et Newman (2001).

2.3.2 La classification des thématiques de recherche : le *topics modeling*

Le problème qui nous occupe ici est de détecter quels thèmes de recherche ont été manipulés à travers les articles contenus dans notre corpus empirique et ce, dans des délais qui prennent en compte le cadre spécifique d'un mémoire de maîtrise. Le temps ne nous permettant pas de lire chaque résumé d'article et de les classer qualitativement en un minimum de thèmes, nous avons dû songer à une autre solution. La plus évidente semble *a priori* de recourir aux mots clés apposés aux articles par les éditeurs ou encore les auteurs eux-même. L'usage des mots clés n'est cependant pas approprié, à notre avis, pour plusieurs raisons : Les mots clés sont parfois absents, parfois trop nombreux (nous avons observé parfois plus de 30 mots clés pour un seul article) ; ils correspondent souvent soit à des choix éditoriaux arbitraires, soit à la représentation des mots les plus occurrents, sans tenir compte de la signification et de la pertinence de ces dit mots clés. Bref, ils n'apparaissent pas permettre de rendre compte du contenu sémantique de notre corpus empirique :

« Trivialement on pourrait par exemple prendre en compte les mots-clés que les auteurs attribuent généralement à leurs articles. En pratique, il est cependant problématique d'adopter cette approche car d'une part, ces champs ne sont pas systématiquement renseignés et en outre, ce serait risquer de s'appuyer sur l'appréciation subjective des auteurs à des concepts distincts pouvant être étiquetés par des mots clés identiques, d'autres concepts, uniques, pouvant être étiquetés par des mots-clés différents. »Roth (2008) p.528

Roth (2008), tout comme d'autres chercheurs, a opté pour une approche par fréquences d'occurrences de termes jugés pertinents *a priori*, puis par regroupements, afin de constituer sa liste des thèmes traités dans les articles. Ces chercheurs travaillent sur des corpus beaucoup plus homogènes que le nôtre, ce qui leur permet de procéder ainsi. Notre corpus est hétérogène car il contient des articles provenant de nombreuses disciplines scientifiques différentes et tant des sciences pures, que des sciences sociales ou encore de l'administration. Dans notre corpus, un même mot peut représenter un concept important dans une discipline et ne rien signifier de particulièrement important dans une autre. Nous avons donc dû trouver une autre solution et c'est ainsi que nous avons découvert un outil issu des champs de la linguistique et de l'intelligence artificielle permettant d'isoler la structure sémantique de large corpus textuels (et même visuels).

Afin d'extraire de nos données les thèmes que les chercheurs ont manipulés dans leurs publications, nous avons opté pour une technique automatisée d'analyse de discours fournissant une structure sémantique du contenu de notre corpus. Ce type d'analyse appartient à la catégorie des analyses en composantes discrètes (Buntine et Jakulin, 2006) et sont plus particulièrement appelées « *topics modeling* » (Blei et al., 2003, Buntine et Jakulin, 2006, Griffiths et Steyvers, 2004).

Le *TM* nous apparaît comme étant un outil privilégié en sciences sociales afin de trouver des thèmes de recherche puisqu'il vise à classer l'ensemble du vocabulaire présent dans un corpus en divers « *topics* » (Blei et Lafferty, 2009, Griffiths et Steyvers, 2004). Le *TM* est une technique d'analyse sémantique du discours, en grande partie automatique et quantitative. Le rôle du chercheur et les aspects qualitatifs de la méthode ne sont cependant pas négligeables : celui-ci choisit les mots à inclure dans l'analyse et ceux à exclure et il décide de donner ou non un sens aux partitions (appelées « *topics* ») résultant de l'analyse par *TM*.

Un *topic* est une partition (« *cluster* ») d'une distribution probabiliste de mots qui occurrent fréquemment ensemble. Le classement s'opère en fonction de la probabilité de chaque mot d'appartenir à un *topic* donné. Autrement dit, une même variable latente « explique » la distribution dans le *topic*. Tous les mots sont classés sous chaque *topic* en fonction de la probabilité relative qu'ils ont d'y appartenir, du plus probable au moins

probable (Blei et al., 2003). En ce sens, chaque mot du corpus appartient à chacun des *topics*, mais sous une distribution différente. Cette approche combinatoire n'oblige donc pas un mot à ne prendre qu'un seul sens ; la classification se fait en fonction de l'utilisation des termes dans les articles et de leur agencement en fonction des co-occurrences. Cette approche permet de prendre en compte la polysémie des mots en les classant en fonction des contextes dans lesquels ils sont employés. Pour illustrer ce propos, prenons l'exemple du terme « *drug* » : selon le contexte dans lequel il est utilisé dans un résumé d'article, il peut signifier différentes choses.¹⁴ En prenant en compte la co-occurrence de ce terme avec les autres termes contenus dans les articles dans lesquels il se retrouve, au moins deux différents *topics* apparaissent : celui où ce terme occure principalement avec « *addiction* », « *injection* » et « *illicit* », et celui où il occure principalement avec « *Burkina* », « *access* » et « *fees* ». Ainsi le mot « *drug* » a une probabilité élevée de se retrouver dans le *topic* concernant les « Soins de santé dans les pays en développement » ainsi que dans celui concernant les « Usagers de drogues par injection intraveineuse ». C'est l'amalgame (la mixture) des mots qui détermine le sens que prend un mot dans un *topic* et donc le sens que prend le *topic* (Blei et Lafferty, 2009, Blei et al., 2003).

Le *TM* se distingue d'autres approches statistiques de l'analyse de discours car il est basé sur l'inférence bayésienne, c'est-à-dire que son estimation est conditionnelle à certains énoncés probabilistes. (Blei et Lafferty, 2009, Blei et al., 2003, Buntine et Jakulin, 2006) Plus précisément, la classification par *TM* se fait en fonction de distributions spécifiques (un prior) puis la distribution s'affine au fil des itérations. Pour réaliser nos analyses, nous avons opté pour le logiciel *DCA version 0.202* développé et distribué par Buntine et Jakulin (2006). Afin de faciliter les passages entre les divers types de format de données, *DCA* a été mis en œuvre dans *MouliNet*. Nous avons testé d'autres logiciels de *TM* (*GibbsLDA++* et *Stanford Topic Modeling Toolbox (TMT)*) et obtenu des classifications similaires à celle obtenue à l'aide de *DCA*, mais ce dernier offre davantage d'outils d'analyses intégrés permettant l'inférence et il est, à notre avis, d'une plus grande flexibilité. Tout ces outils sont basés sur le même principe selon lequel les mots sont analysés sans tenir compte de leur ordre dans la phrase (modèle dit « *bag of*

14. Cet exemple est tiré des résultats de nos analyses et n'est pas fictif.

words »); chaque document est présenté comme un sac de mots pèle-mèle qui lui est propre. (Blei et al., 2003) Le modèle généré dans *DCA* peut être testé afin entre autres d'en vérifier sa vraisemblance, nommée « perplexité » dans le jargon de la linguistique et de l'intelligence artificielle. Le logiciel *DCA* permet de choisir entre deux modèles génératifs : le premier utilisant comme hyperparamètre un prior de Dirichlet et le second utilisant plutôt une distribution de Poisson, afin dans les deux cas d'adoucir la classification en posant des conditions de distribution des mots à travers les topics et à travers la distribution des topics dans le corpus. (Buntine et Jakulin, 2006) Bien que nous ayons testé les deux modèles et qu'ils fournissent des résultats similaires, nous avons choisi d'utiliser le modèle basé sur le prior de Dirichlet puisqu'il est davantage couvert par la littérature.

Pour réaliser une analyse par *TM*, il faut d'abord entraîner le modèle à classer le plus justement possible au fil des itérations. La phase d'apprentissage permet d'estimer *a posteriori* la valeur des hyperparamètres d'ajustement du modèle ainsi que le nombre optimal de *topics* à demander en fonction de ces hyperparamètres. Le chercheur doit ainsi décider du nombre de topics idéal pour rendre compte de son corpus.

Afin d'uniformiser le corpus, nous avons traduit de façon semi-automatisée les résumés et titres non anglophones en anglais, en nous inspirant de la démarche de van Meter et de Saint Leger (2009) dans « *German and french contemporary sociology compared : Text analysis of congress abstract* »¹⁵. Les biais occasionnés par la traduction automatique demeurent à notre avis moins importants que ceux occasionnés par l'exclusion des articles en langue non-anglophone. De surcroît, dans la mesure où les analyses par *TM* se basent sur la co-occurrence de mots, mot non pas compris par le programme dans leur aspect sémantique mais plutôt en tant que forme spécifique, la traduction nuit peu aux résultats des analyses. La traduction s'est effectuée en deux phases: d'abord en laissant l'outil *Google Translate* détecter automatiquement la langue de publication des articles puis traduire en anglais si nécessaire. Finalement, en recherchant à travers le corpus et à l'aide d'outils de recherche textuelle, les mots n'ayant pas été traduits afin d'en

15. Nous avons implémenté *Google translate API version 2* dans *MouliNet*. *Google translate API version 2* est disponible à l'adresse suivante: <http://code.google.com/intl/fr/apis/language/translate/overview.html>

effectuer la traduction à l'aide de *Google Translate*. Peu d'articles ont été concernés par ce second traitement (68 articles).

Un traitement linguistique de base a été ensuite appliqué sur le corpus avant d'être soumis aux analyses par TM. Les nombres, les symboles, les mots les moins discriminants ou trop communs et les mots n'apparaissant qu'une seule fois dans tout le corpus ont été retirés. En ce qui concerne les mots non-discriminants, les symboles et les nombres, nous avons établi une liste des éléments à exclure du corpus avant l'analyse. Cette liste comprend des mots tels que *of, the, this, why...* mais aussi des verbes et des adjectifs jugés peu discriminants tels que *having, describing, know...* Les pourcentages et autres symboles ainsi que les chiffres ont ensuite été ajoutés à la liste des éléments à exclure. La création de cette liste (appelée « *stop-words* ») a nécessité de prendre en compte tout le vocabulaire de chacune des disciplines auxquelles appartiennent les chercheurs de l'IR-SPUM afin de ne pas désavantager certains domaines de recherches. Par exemple, le mot « *capital* » peut n'être qu'un adjectif dans certains domaines et se retrouver notion dans d'autres, comme dans la notion de « *social capital* ». Enlever trop de mots rend l'analyse moins efficace, tout comme ne pas en retirer assez ; il faut trouver un équilibre en fonction des besoins spécifiques que nous avons de recourir à cette méthode. Il ne nous sert à rien d'obtenir un nombre important de catégories n'ayant aucun lien avec la santé publique et se rapportant à, par exemple, l'« intervalle de confiance » (CI) en statistiques.

Ensuite, nous retirons les mots n'occurant qu'une seule fois dans tout le corpus car ils ne permettent pas d'être classés efficacement en fonction des cooccurrences. Pour cela, il suffit de détenir la distribution des mots dans le corpus et d'ajouter à la liste des mots à exclure, ceux apparaissant moins de deux fois dans le corpus. Lorsque le corpus est plus homogène, il vaut mieux fixer plus haut le seuil d'acceptabilité des mots en fonction de leur occurrence (Buntine et Jakulin, 2006), mais dans notre cas, l'hétérogénéité des thèmes et des disciplines sollicités nous oblige à fixer au plus bas ce seuil. Nous obtenons de meilleurs résultats en acceptant des termes de faible occurrence.

Au total, le corpus soumis aux analyses comprend près de 700 000 mots dont 21 060 mots uniques. Les mots occurent donc entre une fois dans tout le corpus à 3621 fois (dans le cas du mot « *health* »). Après avoir exclu la liste prédéfinie des mots, symboles

et chiffres, il reste 8871 mots uniques dans le corpus. Ce traitement a vidé certains articles de leur contenu. En comptant aussi les articles exclus car ils ne possédaient pas de résumé, nous obtenons 3165 articles qui seront soumis aux analyses par *TM*.

Comme nous l'avons mentionné plus tôt, c'est après avoir entraîné le modèle qu'il est possible de fixer certains paramètres essentiels à l'analyse par *TM*. Afin de fixer le nombre d'itérations nécessaires, en fonction des spécificités de notre modèle, nous avons pris en compte *le log probability* calculable dans *DCA*. (Buntine et Jakulin, 2006) Cette approximation nous permet d'analyser la congruence du modèle au fil des itérations afin d'observer à quel moment le modèle se stabilise, c'est-à-dire que les itérations suivantes n'améliorent plus le score de perplexité du modèle. L'ajustement de l'échantillonnage par les hyperparamètres a un effet sur la granulosité du modèle, c'est-à-dire que la classification sera plus ou moins grossière selon sa valeur. Les valeurs optimales de cet ajustement se font en fonction du nombre de partitions demandées. Afin d'aider le chercheur dans ses choix, la quantification de la qualité du modèle est une mesure utile et c'est ce à quoi sert le score de perplexité. (Buntine, 2009) C'est une mesure approximative de vraisemblance. La perplexité ne permet pas de trancher aveuglément en faveur d'un modèle obtenant le meilleur score ; elle est davantage indicative que prescriptive: « However, be aware that models with better perplexity scores don't always produce more interpretable topics or topics better suited to a particular task. Perplexity score can be used as stable measures for picking among alternatives, for lack of better option. »¹⁶

La classification automatique par *TM* nous a permis de détecter des thèmes de recherche représentatifs du contenu sémantique de notre corpus empirique. Sans le recours à ce type d'analyse, il nous semblait irréaliste de songer à analyser la part sémantique contenue dans les milliers d'articles (3165 précisément) qui constituent notre base de données. Malgré les lacunes des analyses automatiques, notamment les risques de mauvaises classifications, cette méthode nous a permis d'obtenir des informations précieuses concernant les sujets manipulés par les chercheurs de l'IRSPUM dans les articles qu'ils ont publiés au cours des 20 dernières années. Les résultats de ces analyses sont présentés au prochain chapitre et ils seront utilisés dans les analyses principales présentées aux

16. SOURCES: <http://nlp.stanford.edu/software/tmt/tmt-0.3/>

chapitres 4 et 5.

2.4 Les analyses

2.4.1 L'analyse de réseaux

Les méthodes d'analyse que nous avons mobilisées pour étudier l'émergence de l'IRPSUM relèvent principalement de l'analyse de réseaux et la forme des données collectées s'y prête. Nous avons opté pour cette approche car elle nous permet de rendre compte dans un premier temps des aspects formels de la copublication entre les chercheurs de l'IRSPUM et des relations entre les chercheurs et les thèmes qu'ils manipulent dans leurs travaux. Dans un second temps, beaucoup plus significatif car en lien avec l'approche théorique à laquelle nous faisons référence, c'est un type de méthode permettant d'explorer le social sous l'angle des interactions entre identités sociales et éléments sémantiques donnant forme à des régularités observables. Ce type de méthode semble des plus adaptés pour s'accorder avec l'approche de l'interactionnisme-structural. (Degenne et Forsé, 2004, White, 2008) C'est dans cette perspective que nous y avons recours.

Selon la théorie whitienne, un réseau social n'est pas simplement la prise en compte d'interactions entre acteurs sociaux ; ils doivent aussi communément partager des « histoires » pour que cela soit considéré comme constituant un réseau (Grossetti et Godart, 2007). Lorsque des identités sociales sont liées entre elles par le partage d'éléments sémantiques, celles-ci sont considérées comme appartenant à un « *netdom* » : « Netdom is a suitable descriptor : "dom" from domain of topics and "net" from network relations. » (White, 2008, p. 7) La communauté épistémique de l'IRSPUM peut représenter un de ces *netdoms*, au sens où il s'agit effectivement d'un réseau de relations sociales qui se retrouvent liées par un domaine épistémique spécifique, la santé publique, et c'est ce que cette méthode peut nous permettre d'étudier.

Du point de vue méthodologique, plusieurs choix s'offrent au chercheur qui veut explorer les communautés épistémiques à l'aide de l'analyse de réseaux. D'abord, il lui faut choisir quels types de liens seront retenus. Dans notre cas, nous avons opté pour

l'étude des copublications, bien que la plupart des études en analyse de réseaux dans ce domaine aient été réalisées à partir de réseaux de cocitations¹⁷ et encore plus rarement en prenant en compte la part sémantique des réseaux sociaux:

« En dehors des nombreux travaux consacrés aux processus locaux de collaboration scientifique (Latour et Woolgar, 1998 ; Matz et Martin, 1997), les études structurelles à grande échelle sur ce type de communautés se sont essentiellement penchées sur les liens d'autorité plutôt que d'interaction, c'est-à-dire, sur les réseaux de citations (comme Solla Price, 1965, 1976) plutôt que de collaborations (Newman, 2004), et/ou ont généralement pris en compte la seule structure sociale sans s'intéresser aux interrelations avec des propriétés sémantiques » (Roth, 2008, p. 524)

Puisque notre intérêt porte principalement sur l'étude de l'émergence de l'IRSPUM et non pas sur les phénomènes d'autorité, nous avons suivi l'inspiration de Roth (2006) et choisi d'étudier les collaborations. Le formalisme qu'il est possible de donner aux réseaux permet d'observer certaines propriétés dont nous traiterons sous peu, mais aussi de visualiser et de manipuler de tels réseaux (Degenne et Forsé, 2004, Wasserman et Faust, 1994, White, 2008). Les auteurs sont liés entre eux s'ils ont publié ensemble un article mais aussi via les thèmes qu'ils ont manipulés dans leurs articles. Nous avons construit deux types de matrices pour les fins de nos analyses de réseaux : des matrices d'adjacence (chercheurs x chercheurs) et des matrices bipartites (chercheurs x thèmes). Nous allons maintenant expliciter la construction de nos matrices.

Comme mentionné un peu plus tôt, notre corpus empirique a été découpé en quatre périodes de cinq ans. Nous avons ainsi une matrice d'adjacence par période correspondante. Les matrices d'adjacence sont des tableaux symétriques où chaque chercheur ayant publié au temps donné se retrouve à la fois en ligne et en colonne. La variable se situe à l'intersection entre deux chercheurs distincts.¹⁸ Nos matrices sont binaires, ce qui signifie que lorsqu'un chercheur a publié avec un autre un article commun (ou plus) dans une période donnée, un « 1 » se trouve à l'intersection de ces deux chercheurs pour indiquer la présence d'un lien. S'il n'y a eu aucune publication entre deux chercheurs dis-

17. Deux articles cités ensemble dans un troisième article représente une cocitation ; on suppose alors que les deux articles cités sont nécessaires au troisième, surtout si la cocitations se produit dans plusieurs articles.

18. Si la transposition des lignes en colonnes est exactement la même, la diagonale n'est pas prise en compte, car c'est un non sens de considérer que l'auteurX a publié avec lui-même

tincts, un « 0 » indique l'absence de lien. Mentionnons tout de suite que deux chercheurs liés directement l'un à l'autre représentent ce que l'on nomme une dyade en théorie des graphes.¹⁹

Les matrices bipartites, aussi appelées matrices « acteurs-événements » (Wasserman et Faust, 1994), diffèrent des matrices d'adjacence car elles sont asymétriques ; nous retrouvons en ligne tous les chercheurs présents à un temps donné et en colonne tous les thèmes manipulés au même temps donné par ces chercheurs. Une dyade représente alors dans ce cas-ci une relation entre un chercheur et un thème. Encore une fois, il s'agit d'un tableau où un « 1 » inscrit à l'intersection d'un chercheur et d'un thème indique la présence d'un lien. L'absence de lien est indiquée par un « 0 ». Nous avons construit une matrice bipartite par période analysée, soit quatre matrices bipartites. C'est à partir des résultats par *topics modeling* que nous avons construit la part sémantique de ces matrices sociosémantiques.

Sur les matrices d'adjacence, nous avons exploré comment la collaboration entre des chercheurs qui se retrouveront tous membres de l'IRSPUM en 2009 a évolué au fil des périodes, pour former un réseau social de chercheurs, et ce avant l'institution de l'IRSPUM. En partant des 65 chercheurs ayant publié au T-3, nous avons porté notamment notre attention sur les collaborations durables, c'est-à-dire qui perdurent pendant plus de deux périodes consécutives. Pour ce faire nous avons noté chacune des dyades de chaque période pour vérifier lesquelles se maintenaient au temps suivant et durant combien de périodes consécutives elles continuaient d'exister.

La visualisation nous permet aussi d'apprécier certaines des propriétés des réseaux. Dans l'étude ci-présentée, les chercheurs et les thèmes correspondent à ce que l'on nomme des sommets en théorie des graphes. Les liens sont appelés arêtes. Des points (les sommets) sont liés par des lignes (des arêtes). Les relations entre les chercheurs et les thèmes qu'ils manipulent peuvent se représenter par des graphes. La visualisation permet d'apprécier des propriétés qualitatives des réseaux. Notamment, certaines structures peuvent avoir la même densité et pourtant se différencier par leurs formes. (Degenne et Forsé, 2004) Si un lien entre deux chercheurs est essentiel au maintien de la structure

19. Plus précisément, une dyade représente une paire de sommets liés (par une arête).

(sans ce lien, le réseau se scinderait), nous pouvons en déduire que leur collaboration assure une liaison entre sous réseaux. Ils forment un « pont »: un lien qui unit deux parties d'un réseau ensemble, c'est-à-dire que sa suppression entraînerait la scission du réseau.

Nous avons aussi cherché à connaître la centralité des sommets dans les réseaux, afin de mieux visualiser la popularité des thèmes et l'activité des chercheurs. Nous avons ainsi calculé la centralité de degré sur les réseaux complets. La centralité de degré consiste à mesurer le nombre de liens qui lient un chercheur ou un thème aux autres sommets du réseau.²⁰ La possibilité de connaître précisément la valeur de centralité qu'obtient chaque thème aide à notre interprétation de leur usage au cours du temps.

Le logiciel employé pour réaliser les analyses de réseaux est Visone version 2.6.2.²¹

2.4.1.1 Treillis de Galois

Bien que nous ayons aussi exploré les matrices bipartites à l'aide de graphes, ce sont les treillis de Galois qui nous ont fourni les résultats les plus intéressants du point de vue des possibilités d'analyser simultanément les liens entre les thèmes et les chercheurs. Cette approche mathématique permet de prendre en compte les multiples chevauchements d'appartenances des individus à des groupes différents. (Freeman, 1992) Les treillis de Galois sont surtout l'une des rares façons de traiter des données de réseaux bipartites (avec l'analyse des correspondances).(Wasserman et Faust, 1994)

Un treillis de Galois regroupe l'ensemble des combinaisons « acteurs-événements » existantes dans une matrice bipartite pour en extraire la structure hiérarchique ; il permet de visualiser simultanément quels chercheurs travaillent sur quels thèmes et avec qui dans une logique permettant de classer les relations chercheur-thèmes des plus spécifiques aux plus générales(Roth, 2006, 2008).

L'interprétation d'une visualisation en « diamant » d'un treillis de Galois, telle que celle qui sera ici présentée, exige de s'y familiariser un minimum puisqu'il est peu courant d'en rencontrer dans les recherches en sociologie. Chaque sommet d'un treil-

20. Cette phrase s'inspire de Degenne et Forsé (2004) p. 156

21. Pour en savoir plus sur ce logiciel, il est possible de consulter ce document:Brandes et Wagner (2004) « Visone - Analysis and Visualization of Social Networks » In Graph Drawing Software

lis comporte deux étiquettes : une correspondant aux acteurs et l'autre aux thèmes car le treillis souligne l'avènement des deux éléments de façon simultanée. Un sommet représente donc la co-occurrence de chercheurs et de thèmes. Les visualisations présentées dans ce mémoire peuvent se lire dans deux sens. Les thèmes les plus élevés sont les thèmes manipulés par le plus grand nombre de chercheurs ; plus on descend dans le treillis, moins le thème se retrouve manipulé par les chercheurs. Les chercheurs les plus élevés sont ceux manipulant le moins de thèmes et les plus bas ceux manipulant le plus de thèmes.

En somme, notre méthodologie requiert l'usage de méthodes d'analyse variées et nous avons souvent dû trouver des solutions inusitées afin de répondre à nos questions de recherche qui consistent, rappelons-le à vérifier si la nouvelle santé publique se reflète dans l'univers sémantique d'un groupe de chercheurs dont nous tâcherons de déterminer s'ils forment communauté. L'usage d'autant de méthodes d'analyse ne doit pas occulter la part de réflexion et de travail des chercheurs qui a été investie dans ce travail.

CHAPITRE 3

LA CLASSIFICATION DES THÉMATIQUES DE RECHERCHE : LE *TOPICS MODELING*

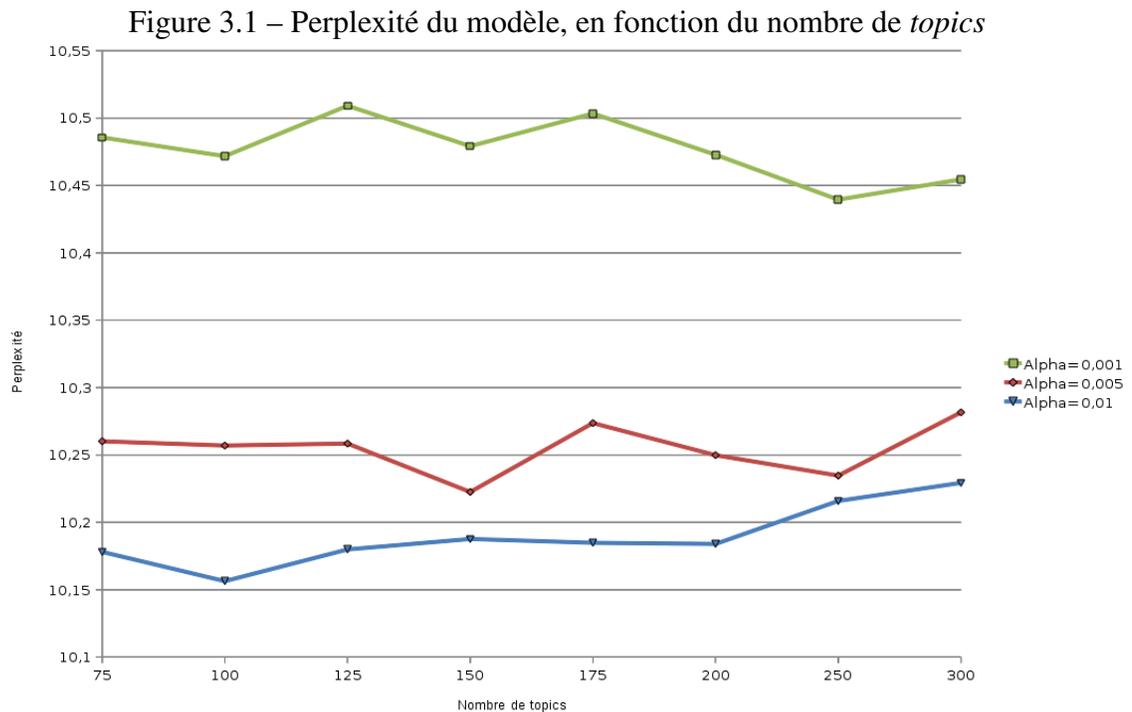
Ultimement, ce que la recherche ci-présentée vise à approfondir se trouve dans l'étude de la co-évolution entre les chercheurs et les thèmes de recherche qu'ils ont manipulés dans leurs publications. Nous avons mentionné au chapitre précédent que nous avons eu recours à une approche par *topics modeling(TM)* afin d'isoler les thèmes manipulés dans les articles rédigés, au cours des 20 dernières années, par les chercheurs qui sont actuellement membres de l'IRSPUM. Ce troisième chapitre a pour objectif de présenter les résultats des analyses par *TM*. Ces analyses nous permettent de vérifier la validité apparente de notre classification. Cette classification sera utilisée au chapitre 4 où nous examinerons les liens entre chercheurs et les thématiques de recherche qu'ils manipulent.¹

3.1 Les résultats des analyses sémantiques

Comme nous l'avons mentionné précédemment, les analyses par *TM* exigent d'entraîner le modèle à la classification. Une fois le modèle « entraîné », il est possible d'observer les divers scores de perplexité qu'obtiennent les différents modèles testés. Nous avons testé de nombreux modèles variant en nombre de catégories et selon différentes valeurs d'ajustement de l'échantillonnage. Seront présentés les trois différents modèles dont les résultats étaient les plus satisfaisants pour les fins de notre étude.

La figure 3.1 illustre la variabilité du score de perplexité en fonction du nombre de partitions demandées et de la valeur du paramètre d'ajustement de la granulosité (du *alpha*) (en fonction de trois valeurs de *alpha*(α) : $\alpha = 0,001$; $\alpha = 0,005$ et $\alpha =$

1. Comme il en sera bientôt question, nous avons demandé une classification en 150 partitions; il ne sera pas ici question de présenter ces 150 *topics*, mais les mesures prises pour vérifier la validité des résultats et nous présenterons seulement quelques thèmes choisis, afin de rendre compte des résultats obtenus. Le output complet de l'analyse par *TM* est fourni à l'adresse suivante : <http://idealites.info/maitrise/150T/biblio.150d.html>

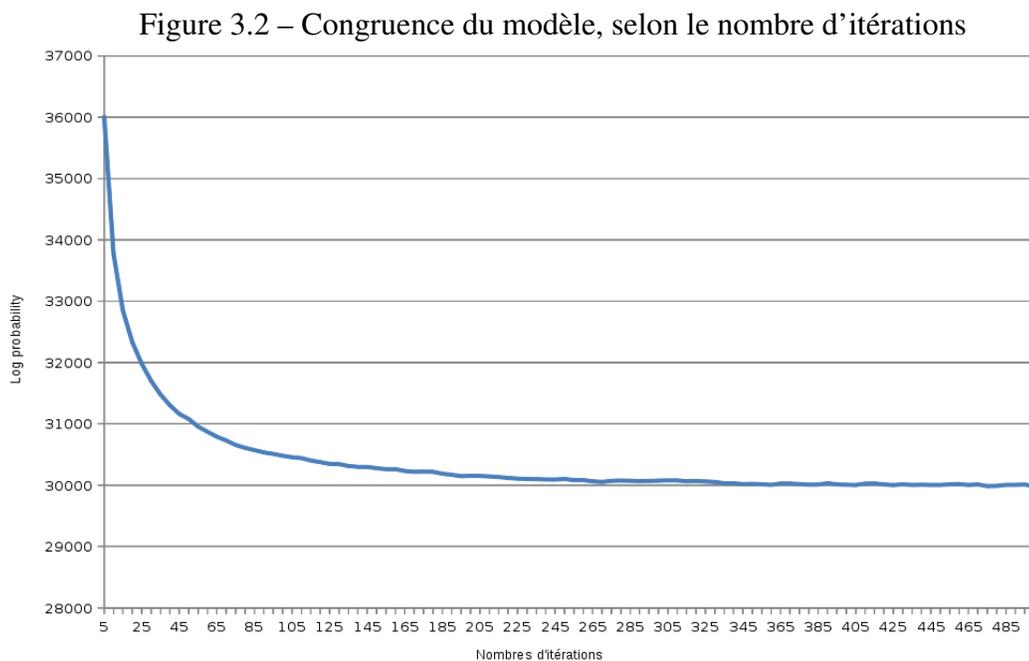


SOURCES : Les données sont obtenues à partir de l'analyse de la base de données CORPUS_IRSPUM, à l'aide du logiciel *DCA*. Le graphique est obtenu à partir du logiciel *Excel*. $\alpha = 0,005$

0,01), pour des classifications allant de 75 partitions à 300 partitions. En dessous de 75 partitions, les *topics* semblaient trop généraux et indistincts, indépendamment de la granulosité (de la valeur de α). Les modèles au-delà de 300 partitions contiennent de nombreux topics non significatifs car ils ne sont reliés aux articles qu'avec une très faible probabilité. Ce sont des catégories « *junk* ». En fait, d'un point de vue strictement qualitatif, les modèles de moins de 100 *topics* et de plus de 250 *topics* n'apparaissent pas rendre compte de la façon la plus optimale du corpus soumis aux analyses. Le modèle que nous avons retenu est celui avec un *alpha* de 0,005 et 150 partitions. Son score de perplexité est de 10,22 ($\alpha = 0,005$) et n'est pas le plus bas que nous avons obtenu, mais ce score correspond à celui du modèle qui nous a semblé le plus satisfaisant parce qu'il permet le plus justement de rendre compte des thématiques de recherche manipulées dans les articles constituant notre corpus.

La congruence du modèle est assurée par une mesure de *log probability* permettant de vérifier combien d'itérations sont nécessaires afin que le modèle se stabilise, que les

classes soient fixées et qu'elles ne changent plus au fil des itérations. Le graphique de congruence du modèle présenté à la figure 3.2 en page 44 montre qu'après 350 itérations, le modèle atteint la stabilité, c'est-à-dire que les itérations suivantes n'améliorent plus le score de perplexité du modèle. La classification ne subit plus de modifications, elle s'est fixée. Cette information permet de s'assurer que le modèle retenu est bel et bien stable.



SOURCES : Les données sont obtenues à partir de l'analyse de la base de données CORPUS_IRSPUM, à l'aide du logiciel *DCA*. Le graphique est obtenu à partir du logiciel *Excel*.

Puisque les mesures de la qualité des modèles par *topics modeling* ne sont que les meilleures estimations du moment, nous avons veillé à vérifier la qualité du modèle retenu d'un point de vue plus qualitatif. Pour ce faire, nous avons d'abord apposé une « étiquette » à chacune des partitions obtenues.² Cette étiquette se veut descriptive du thème qui correspondrait à chacune des partitions obtenues. Précisons d'emblée que 10 *topics* sur 150 nous sont apparus comme indéfinissables, c'est-à-dire que nous n'avons pas pu distinguer un thème de recherche spécifique. De ces 10 *topics* trop généraux, la moitié n'est liée à aucun article. Cela se produit car l'analyse par *topics modeling* est sémantique et vise à regrouper sous une même partition des mots qui co-occurrent.

2. La liste des 150 *topics* étiquetés est fournie à l'annexeII en page xxxiv

Certains mots co-occurrent dans un contexte particulier, même si le sens que prennent ces *topics* n'explique rien d'aucun article. Il s'agit ainsi de classifications n'expliquant aucun des articles et qui peuvent être interprétées comme étant des « *junk cluster* ». L'autre moitié est liée de façon très improbable à des articles ($P \leq 10\%$). Il s'agit alors de *topics* « limpides », mais abordés dans un seul article ou dans une très faible proportion d'articles.

Nous avons retiré ces 10 *topics* peu discernables des analyses ultérieures que nous présenterons. Il en va de même de 12 articles qui n'étaient liés à aucun *topic*. Plusieurs raisons expliquent que des articles se retrouvent sans *topic* : dans certains cas il s'agit d'articles dont les résumés semblent trop courts pour que le *topics modeling* ait pu les classer ; dans d'autres cas encore, de sujets difficilement distinguables car trop rares dans l'ensemble du corpus. Il ne faut pas négliger ces exclusions qui soulignent que le *topics modeling* a des difficultés à traiter des corpus aussi hétéroclites et petits que celui que nous avons soumis aux analyses.

Au cours des analyses tests par *topics modeling*, nous avons remarqué une très importante stabilité des classifications obtenues entre chaque test. Indépendamment des logiciels de *topics modeling* utilisés, nous obtenions des partitions similaires. Cependant, nous avons aussi remarqué que certains thèmes apparaissaient parfois, puis disparaissaient. Ces instabilités se sont produites principalement autour des thèmes concernant les recherches sur les populations autochtones ainsi que celles traitant de la profession de sage-femme. La classification retenue ne laisse pas percevoir ces deux thèmes, qui pourtant se retrouvent effectivement dans le corpus.³ Nous ne pouvons ainsi pas prétendre que la classification obtenue soit parfaitement représentative du corpus.

Afin d'augmenter le niveau de confiance envers le modèle retenu, nous avons sélectionné aléatoirement un échantillon de 50 articles parmi la totalité des articles soumis au *topics modeling*.⁴ Pour chaque article de l'échantillon, nous avons vérifié si l'é-

3. Nous avons tenté de repêcher manuellement les articles concernés par les thématiques instables, mais la tâche fut plus complexe que prévue : le temps alloué à l'élaboration d'une seconde classification, cette fois-ci qualitative, par-dessus la classification automatique, a dépassé de loin nos délais prévus et nous avons dû mettre de côté cette tentative chronophage qui n'aboutissait à aucun résultat satisfaisant.

4. L'échantillon est disponible à l'adresse web suivante : <http://idealites.info/networks-irspum/outputs/LDA/irspum.sample.html>

tiquette lui étant apposée permettait effectivement, à notre avis, de rendre compte ou pas du sujet dont traite l'article. Sur les 50 articles que contient l'échantillon, 48 (96%) nous apparaissent correctement décrits par leurs étiquettes. En considérant cet échantillon aléatoire, les *topics* ont une forte probabilité d'expliquer effectivement les articles auxquels ils sont liés.

En résumé, nous avons retenu le modèle à 150 partitions ($P = 0,005$) puisqu'il représente celui nous permettant le mieux de rendre compte des thèmes qu'ont traités les chercheurs de l'IRSPUM au cours de leur carrière. Malgré l'imperfection de ce type d'analyse, le *topics modeling* nous a permis de mettre à jour la structure sémantique d'un important corpus d'articles variés. Une tâche qui aurait demandé une patience et un temps exceptionnellement importants si elle avait été réalisée sans le soutien de l'analyse automatique de contenu qu'est l'analyse en composantes discrètes.

3.1.1 Illustration des résultats obtenus par *topics modeling*

Le *topics modeling* fournit des partitions dans lesquelles se trouvent des mots. Ces mots co-occurrent, c'est-à-dire qu'ils se retrouvent ensemble dans un même contexte. Ils sont classés selon leur probabilité d'appartenance aux *topics*. Avant de présenter la classification que nous avons retenue, il sera question d'illustrer ce que représentent concrètement ces *topics* lorsqu'ils se retrouvent rattachés aux articles. Ensuite, nous présenterons quelques partitions afin d'illustrer comment nous avons effectué l'étiquetage.

Comme mentionné dans le chapitre précédent, nous appliquons le *topics modeling* à un corpus d'articles écrits par les chercheurs de l'IRSPUM au cours de leurs carrières respectives. Voici un exemple choisi au hasard afin d'illustrer en quoi le *topics modeling* permet de mettre à jour les thèmes manipulés par les chercheurs dans leurs articles. Le résumé d'un article se présente ainsi :

Le titre de cet article de 1997 est "*Diet and premenopausal bilateral breast cancer : a case-control study*". Il est lié au *topic* 64 (avec une probabilité de 51% d'expliquer l'article) et au *topic* 96 (avec une probabilité à 49% d'expliquer l'article). Les couleurs dans le titre et le résumé de l'article désignent les *topics* d'appartenance des mots. Par exemple, la couleur bleue se réfère au *topic* 96 et en rouge nous retrouvons les mots

Figure 3.3 – Le résumé d’un article, traité par *topic modeling*

Article #3131

Topics distribution : 64^(0.507463) 96^(0.492537)

Diet⁽⁶⁴⁾ and **premenopausal**⁽⁹⁶⁾ **bilateral**⁽⁹⁶⁾ **breast**⁽⁹⁶⁾ **cancer**⁽⁴²⁾ : a case - control study . :
 We investigated associations⁽⁴²⁾ between **diet**⁽⁶⁴⁾ and **premenopausal**⁽⁹⁶⁾ **bilateral**⁽⁹⁶⁾ **breast**⁽⁹⁶⁾ **cancer**⁽⁴²⁾
 in a familial⁽¹⁴⁸⁾ matched case - control study . We studied 140 cases from population - based⁽¹⁴⁹⁾
 registries⁽¹³⁰⁾ in Los⁽¹²²⁾ Angeles⁽³⁷⁾ County⁽²¹⁾ (California)⁽¹⁴⁸⁾ and Connecticut,⁽²²⁾ and from the major
 hospitals⁽⁸⁸⁾ in the southern⁽¹²³⁾ parts of the Province⁽¹⁴⁷⁾ of Quebec⁽⁴⁾ . Unaffected sisters⁽¹⁵⁾ of the
 cases served as matched controls (222 total) . **Dietary**⁽⁶⁴⁾ **intake**⁽⁶⁴⁾ were assessed⁽¹¹⁴⁾ with a **food**⁽⁶⁴⁾
 frequency questionnaire . Total **fat**⁽⁶⁴⁾ monounsaturated⁽¹⁸⁾ **fat**⁽⁶⁴⁾ polyunsaturated⁽¹⁸⁾ **fat**⁽⁶⁴⁾ **oleic**⁽³⁵⁾
acid⁽⁶³⁾ and linoleic⁽¹¹³⁾ **acid**⁽⁶³⁾ **intake**⁽⁶⁴⁾ was inversely associated with **premenopausal**⁽⁹⁶⁾ **bilateral**⁽⁹⁶⁾
breast⁽⁹⁶⁾ **cancer**⁽⁴²⁾ risk⁽⁴²⁾ . Consumption⁽⁸⁹⁾ of **carbohydrates**⁽³⁵⁾ (and sweetened beverages)⁽⁸⁹⁾ was
 associated with an increased risk⁽⁴²⁾ . We observed no associations⁽⁴²⁾ for **dietary**⁽⁶⁴⁾ **fiber**⁽¹⁰⁸⁾
 antioxidants, or major **food**⁽⁶⁴⁾ groupings, but we did observe inverse associations⁽⁴²⁾ for **intake**⁽⁶⁴⁾ of
 low⁽¹⁴²⁾ **fat**⁽⁶⁴⁾ **dairy**⁽¹³⁸⁾ **products**⁽⁷²⁾ and **tofu** . These findings suggest that monounsaturated⁽¹⁸⁾ and
 polyunsaturated⁽¹⁸⁾ **fats**⁽⁸⁹⁾ as well as soy⁽⁸³⁾ **foods**⁽⁶⁴⁾ might reduce the risk⁽⁴²⁾ of **premenopausal**⁽⁹⁶⁾
bilateral⁽⁹⁶⁾ **breast**⁽⁹⁶⁾ **cancer**⁽⁴²⁾ .

SOURCES : Les données sont obtenues à partir de l’analyse de la base de données - CORPUS_IRSPUM, à l’aide du logiciel DCA. Pour visualiser tous les résultats du *topics modeling* : <http://idealites.info/networks-irspum/outputs/LDA/biblio.articles.0.html> L’image est obtenue à partir d’un script implémenté dans *MouliNet*.

liés au *topic* 64. L’intensité de la couleur varie selon la probabilité du mot d’appartenir au *topic* en question : plus la couleur est intense, plus la probabilité est élevée, plus la couleur est faible, moins la probabilité est élevée. Dans le résumé et le titre de l’article, accolé au mot et entre parenthèses, nous retrouvons le numéro correspondant au *topic* affecté à ce dit mot, dans le contexte de cet article. Les mots en gris sont ceux exclus de l’analyse par la *stop-words list*.

Mais que peuvent bien représenter les *topics* 64 et 96 ? C’est ce que nous allons maintenant regarder. En listant les cinq premiers mots les plus probables d’appartenir à ces deux *topics*⁵, nous serons en mesure de tenter de comprendre le sens que peuvent prendre ces *topics*, c’est-à-dire de les étiqueter. Les cinq thèmes les plus probables d’appartenir au *topics* 64 sont : *food, dietary, traditional, intake, vitamin*. Les cinq thèmes les plus probables d’appartenir au *topics* 96 sont : *breast, cancer, women, risk, mammogra-*

5. Ils sont classés en ordre décroissant : du terme le plus probable au moins probable

phy.

L'interprétation qualitative des partitions 64 et 96 permet au chercheur qui en voit la nécessité de « nommer » les thèmes traités dans les articles auxquels ces *topics* sont associés. Il s'agit ici de trouver à quoi pourrait correspondre la variable latente expliquant que ces termes se retrouvent ensemble et dans cet ordre de probabilité, sous ces partitions. Tel que mentionné plus tôt, nous avons préalablement étiqueté les *topics* : le *topic* 64 a été vu comme traitant des « régimes alimentaires » et le *topic* 96 du « cancer du sein ». Si l'on retourne au résumé de l'article présenté précédemment, les thèmes « régimes alimentaires » et « cancer du sein » semblent effectivement permettre de rendre compte des sujets traités dans cet article. Ainsi le *topic modeling* permet de découvrir les thèmes abordés dans les articles par les chercheurs.

De même, on peut examiner les articles qui sont expliqués par un *topic*. Nous prendrons ici l'exemple du *topic* 67 que nous avons nommé « Déterminants sociaux de la santé ». Bien que d'autres *topics* ont aussi traité de déterminants sociaux de la santé, comme par exemple le *topic* 142 « statuts socioéconomiques », le *topic* 67 traite, selon notre interprétation, du thème des déterminants sociaux de la santé de façon générale. Le tableau 3.1 présente les titres des 10 premiers articles expliqués par le *topic* 67 et la probabilités de ces articles d'être associés à ce *topic*.

C'est à la fois en observant les termes liés à un *topic* donné et les articles auxquels ce *topic* est lié que nous parvenons à interpréter quel pourrait être la variable latente sous-tendue par le *topic* en question. Concernant le cas du *topic* 67 : « Déterminants sociaux de la santé », les 10 premiers articles liés à ce *topic* le sont dans une très grande probabilité.⁶ Il nous a fallu prendre en compte un très important nombre d'articles (500) pour voir ce qui pouvait les réunir sous ce *topic*, puisque dans les dix premiers, il est davantage question de préoccupations d'ordre social. Pour tous les *topics*, nous avons pris en compte les 500 premiers articles (ou moins, selon le cas) de chaque *topic* avant d'apposer une étiquette. L'analyse des 500 premiers articles nous permet de comprendre que ce qui unit ces articles sous le *topic* 67 est la prise en compte de déterminants sociaux de la santé. D'après ce qui est possible d'observer dans les 10 articles listés, on y aborde des

6. la plus faible présentée étant à 95%. Voir les chiffres entre parenthèses suivant les titres des articles.

Tableau 3.I – Présentation des 10 premiers articles expliqués par le *Topic 67* : « Déterminants sociaux de la santé »

Titres des articles	Probabilité
<i>"The role of NGOs in global health research for development."</i>	1.000000
<i>"The concept of social capital and its uses"</i>	1.000000
<i>"Social inequalities in health. A partnership of research and education"</i>	1.000000
<i>"Social inequality of health - Partnering research and training"</i>	1.000000
<i>"Can models be built on the basis of the relationship between social contexts and health ?"</i>	1.000000
<i>"Social support and mental health : concepts, measures, recent research and implications for clinicians"</i>	1.000000
<i>"Knowledge translation research in population health : establishing a collaborative research agenda."</i>	0.990991
<i>"Growth of world population and environment issues"</i>	0.976744
<i>"Social Policy : a public health issue ?"</i>	0.964286
<i>"Transcending the known in public health practice : the inequality paradox : the population approach and vulnerable populations."</i>	0.953488

SOURCES : Les données sont obtenues à partir de l'analyse de la base de données CORPUS_IRSPUM, à l'aide du logiciel *DCA*.

notions telles que « capital social », « inégalités sociales de santé », « soutien social » et « politiques publiques ». Ce *topic* nous a semblé dans son ensemble traiter de facteurs sociologiques influant sur la santé des populations. C'est pour cette raison que nous avons qualifié ce *topic* ainsi. L'ensemble des 150 *topics*, la liste (en ordre décroissant) des 25 premiers mots les plus probables de leur appartenir ainsi que les articles et leurs probabilités d'être expliqué par ces *topics* (jusqu'à un maximum de 500 articles) est disponible à l'adresse suivante : <http://idealites.info/networks-irspum/outputs/LDA/biblio.150d.html>.

Certes, les *topics* ne rendent pas compte des articles dans toute leur finesse et leurs nuances, mais les résultats nous semblent suffisants pour les fins de ce mémoire.

3.2 Les grandes thématiques retenues

C'est dans un objectif très spécifique que nous avons réalisé des analyses par *topics modeling*. Il s'agissait d'obtenir les grands thèmes principaux qu'ont manipulés les chercheurs de l'IRSPUM à travers les articles qu'ils ont publiés. Bien que l'analyse en 150 partitions nous ait fourni un modèle des plus représentatifs de notre corpus, il est apparu à la fois utile et nécessaire d'effectuer certains regroupements, à la fois pour réduire le nombre de thèmes et pour éviter de se retrouver avec des thèmes parfois très similaires.

Un de ces cas a été mentionné plus tôt ; nous avons entre-autres choses fusionné les *topics* 67 (« déterminants sociaux de la santé ») et 142 (« statut socioéconomique »). Pourquoi ? Parce que le statut socioéconomique est généralement considéré en tant que déterminant social de la santé. Ainsi, tous les thèmes faisant référence à des déterminants sociaux de la santé se sont vus fusionner dans le *topic* correspondant aux « déterminants sociaux de la santé ».

Un autre exemple de ces fusions concerne les *topics* 49 et 140 : tous les deux traitent de pneumoconioses. L'un posant l'emphase sur l'amiantose mais en traitant aussi d'autres pneumoconioses et l'autre en traitant des pneumoconioses en général. Notre modèle contenait aussi plusieurs partitions traitant d'expositions à des substances toxiques spécifiques.⁷. Tous les *topics* abordant un thème relevant de l'exposition à des contaminants et leurs composites dans l'environnement ont été fusionnés pour ne former qu'un seul thème : « Exposition et contamination par l'environnement ». Voici la liste des thèmes fusionnés ; les détails concernant les fusions sont présentés à l'annexe IV de la page xli :

Th1 : Santé dentaire

Th2 : Médecine préventive

Th3 : Soins de santé

Th4 : Développement des enfants et adolescents

7. Tel que le manganèse, le trichloréthylène, le pyrène, les composés organochlorés, pour ne nommer que ceux-là

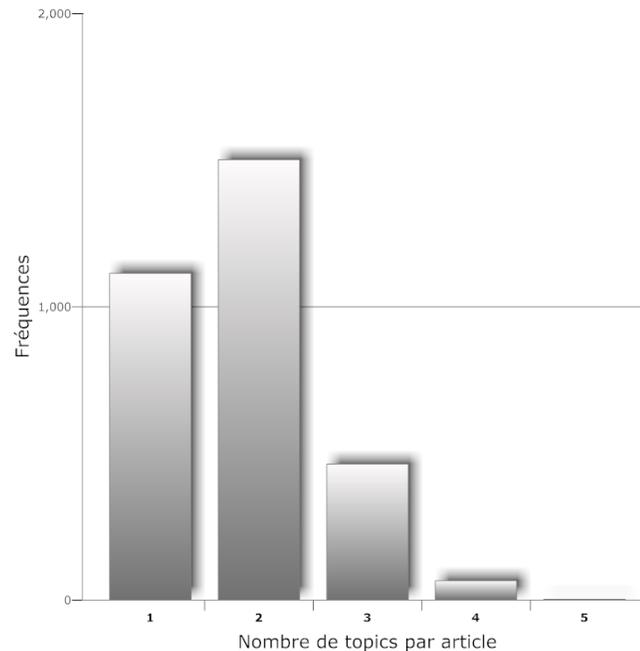
- Th5** : Santé des femmes et des enfants
- Th6** : Santé internationale
- Th7** : Soins des personnes âgées
- Th8** : Autres troubles médicaux
- Th9** : Maladies chroniques
- Th10** : Maladies infectieuses
- Th11** : Lignes directrices cliniques
- Th12** : Médicaments
- Th13** : Immunologie
- Th14** : Génétique et protéomique
- Th15** : Zoonoses et parasites
- Th16** : Maladies professionnelles
- Th17** : Santé environnementale
- Th18** : Santé mentale
- Th19** : Déterminants sociaux de la santé
- Th20** : Comportements sociaux
- Th21** : Politiques publiques
- Th22** : Éthique en recherche et en sciences
- Th23** : Promotion de la santé
- Th24** : Évaluation des technologies de la santé
- Th25** : Gestion des services de soins de santé

3.3 Un portrait descriptif de l'usage des thèmes

Nous avons entre autres voulu savoir combien de *topics* en moyenne faut-il pour rendre compte d'un article afin de vérifier si nos résultats avaient une certaine validité apparente. Si les articles traitent en moyenne d'un nombre important de thèmes, nous

devrons nous interroger sur la pertinence de la classification utilisée puisqu'en principe les articles sont spécifiques à un nombre réduit de thèmes.

Figure 3.4 – Histogramme de la distribution du nombre de *topics* par article



SOURCES : Les données sont obtenues à partir de l'analyse de la base de données TOPICS à l'aide du logiciel SPSS

Pour les résultats d'analyses présentés à la figure 3.4 aucun cas n'a été retiré : les 3165 articles retenus y sont représentés. La distribution du nombre de *topics* par article y est représentée. Pour près de la moitié des articles (1509, soit 48%), il y a deux *topics* traités. Il est plus rare de voir plus de deux *topics* traités ; 538 articles (17%) traitent de trois thèmes ou plus. Un seul *topic* est traité dans 35% des cas (soit 1118 articles). Du point de vue de la validité apparente, ces résultats semblent pertinents puisque les articles tendent à se centrer sur un nombre limité de thèmes.

3.4 Conclusion

Le but de ces analyses par *TM* était de mettre à jour des thèmes de recherche qu'ont manipulés des chercheurs afin d'être en mesure de construire un réseau épistémique.

c'est-à-dire un réseau social de chercheurs autour de thèmes communs, et non pas d'explorer en profondeur le contenu sémantique d'articles scientifiques. C'est pour remplir notre objectif que nous avons apposé des étiquettes aux *topics*. Cette pratique n'a pas cours en linguistique où les *topics* ne sont pas qualifiés.

En terminant, c'est en utilisant la classification obtenue par *TM* pour nos analyses ultérieures que nous serons le plus en mesure d'évaluer la qualité de la classification obtenue, parce que les thèmes apparaîtront en fonction de périodes et d'intensités particulières.

CHAPITRE 4

UN RÉSEAU SOCIAL PREND FORME...

Ce mémoire porte sur l'étude de l'émergence d'une communauté épistémique spécifique. Notre conceptualisation de ce qu'une communauté épistémique représente nous exhorte à prendre en compte tant les aspects sociaux que sémantiques permettant à ce type de formation sociale d'apparaître. Le présent chapitre porte sur l'exploration des réseaux sociaux entre les chercheurs de l'IRSPUM. C'est dans un second temps, au chapitre prochain, que nous présenterons l'évolution sémantique coïncidant avec celle du réseau social. Il faut comprendre ce chapitre et le suivant comme étant interdépendants ; les deux faces, d'une même pièce, qui se complètent.

L'objectif de ce chapitre est de montrer à quel point la cohésion entre les chercheurs actuels de l'IRSPUM s'est accrue avant même que l'IRSPUM soit créé et comment un réseau social s'est mis en place. Nous verrons d'abord comment la collaboration croît, au fil du temps, entre des chercheurs en particulier et les réseaux se développent autour de quelques chercheurs.

Nous l'avons vu au chapitre 2, la population choisie pour les fins de l'étude ci-présentée est constituée des chercheurs actuels d'une jeune institution de recherche créée en 2009, l'IRSPUM, mais dont les assises reposent sur la fusion de Centres de recherche bien établis. Les données prises en compte dans notre étude présentent exclusivement des chercheurs qui seront (dans le futur) des chercheurs de l'IRSPUM.

Bien que l'étude de la transitivité des liens sociaux soit des plus pertinentes et intéressantes, le cadre d'un mémoire nous oblige à ne pas en tenir compte. Seules les collaborations directes à un article commun entre chercheurs sont prises en compte. Nous avons ainsi exclu les collaborations entre futurs chercheurs de l'IRSPUM et les chercheurs qui ne seront pas membres de l'IRSPUM en 2009. En examinant l'évolution à partir de janvier 1990 (du T-3) des collaborations entre les chercheurs qui se retrouveront membres de l'IRSPUM en 2009, il sera possible de comprendre en partie comment s'est constituée la communauté des chercheurs de l'IRSPUM. Une série de graphes sera maintenant

présentée afin d'illustrer l'évolution de la collaboration au sein des futurs chercheurs de l'IRSPUM. Pour tous les graphes qui seront dorénavant présentés, les légendes sont disponibles à l'annexe III en page xl.

4.1 1990 à 1994

Figure 4.1 – Collaborations entre futurs chercheurs de l'IRSPUM entre 1990 à 1994 (T-3)



SOURCES : Les résultats sont obtenus à partir de l'analyse de la base de données COLLABO_T-3 à l'aide du logiciel *Visone*

En T-3, 65 (32 associés et 33 réguliers) des 226 chercheurs qui constituent actuellement l'IRSPUM ont publié au moins un article. Comme nous pouvons le constater à la figure 4.1, des 65 chercheurs présents, huit collaborent avec un autre futur membre de l'IRSPUM ; en T-3 ces 65 chercheurs collaborent peu entre eux puisque nous n'observons que quatre liens de collaboration. Deux de ces liens impliquent des chercheurs ayant actuellement le statut de chercheurs réguliers de l'IRSPUM (les deux autres impliquent des chercheurs associés, distingués par le pointillé rouge). Cinq des huit chercheurs présents sur ce graphe sont maintenant chercheurs réguliers à l'IRSPUM. Le pourcentage de chercheurs ne copubliant pas avec au moins un autre chercheur de l'IRSPUM est de 88% ; ce qui représente 57 des 65 chercheurs présents en T-3. Bref, cette période se caractérise par la faiblesse des liens entre chercheurs.

L'homophilie dans les collaborations apparaît : les chercheurs en sciences humaines et sociales (SHS) (en bleu) s'unissent spécifiquement avec des chercheurs en santé publique (en mauve), les médecins et chercheurs en sciences médicales (en rouge) demeurent entre eux tandis que les chercheurs en épidémiologie/biostatistiques (en orange) s'unissent aux chercheurs en sciences biologiques et chimiques (en jaune).

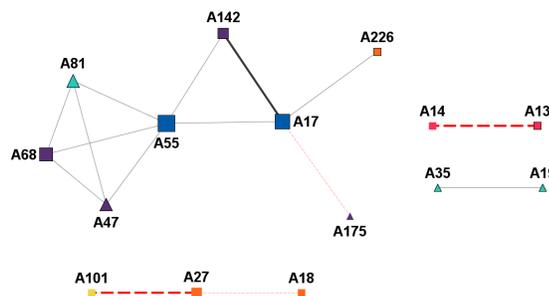
En T-3, les futurs chercheurs de l'IRSPUM sont loin de représenter une communauté,

c'est-à-dire un groupement uni et partageant des éléments sociaux et sémantiques. La vaste majorité des chercheurs n'est liée à aucun autre futur chercheur de l'IRSPUM. Il n'y a pas de cohésion significative au T-3 au sein de la population étudiée.

4.2 1994 à 1999

En T-2, 82 (39 associés et 43 réguliers) des 226 chercheurs qui constituent actuellement l'IRSPUM ont publié au moins un article. Quarante-neuf (49) des 65 chercheurs en T-3 ont publié en T-2 et 33 nouveaux chercheurs font leur apparition dans nos données. Pour l'ensemble des chercheurs ayant publié en T-2, nous retrouvons 15 chercheurs ayant collaboré à la rédaction conjointe d'un article (ou plus) avec (au moins) un autre futur chercheur de l'IRSPUM et aucune collaboration pour 67 chercheurs.

Figure 4.2 – Collaborations entre futurs chercheurs de l'IRSPUM entre 1994 à 1999 (T-2)



SOURCES : Les résultats sont obtenus à partir de l'analyse de la base de données COLLABO_T-2 à l'aide du logiciel *Visone*

À la figure 4.2, nous observons 15 liens impliquant 15 chercheurs. Trois des quatre dyades présentes en T-3 : (A14, A13), (A101, A27) et (A17, A142), se poursuivent en T-2 (liens plus gras). Des huit chercheurs présents à la figure 4.1 un seul ne se retrouve pas sur la figure 4.2. : A38 ; sept sur huit chercheurs demeurent présents pendant 2 périodes consécutives. Le collaborateur d'A38, A55, s'est quant-à lui allié à une dyade déjà existante (A17, A142).

Fait particulièrement intéressant, un petit réseau commence à se développer. A55

représente l'acteur social le plus central sur ce graphe¹ en étant celui qui a le plus grand nombre de liens de collaboration (5). Il fait également office d'articulateur entre deux cliques (A55, A17, A142) et (A55, A81, A68, A47). Les deux chercheurs en SHS (A55 et A17) ont des positions particulières, leur suppression scinderait le graphe. D'ailleurs deux des nouveaux entrants s'associent à A17. Une triade se forme autour de A27 qui maintient son lien avec A101 et développe une nouvelle collaboration avec A18. Enfin une nouvelle alliance se crée, entre deux chercheurs en psychologie (A35 et A19). Il est à noter que les nouveaux entrants présents à la figure 4.2 proviennent de psychologie (3) et de santé publique (2).

La plus grosse composante visible en figure 4.2 de la page 56 est formée principalement de chercheurs ayant actuellement le statut de chercheurs réguliers à l'IRSPUM (à l'exception de A175), la majorité (7/8) étant à l'époque déjà membres réguliers du GRIS. Il semble que ce noyau de chercheurs soit en partie celui à partir duquel s'est constitué l'IRSPUM et c'est ce que nous tâcherons de vérifier. La triade où A27 se trouve l'articulateur est formée uniquement de chercheurs ayant actuellement le statut de chercheur associé à l'IRSPUM. L'exclusion des chercheurs associés désintégrerait entièrement une des 4 composantes observables de la figure 4.2 et laisserait dans l'isolement la dyade (A14, A13) qui perdure depuis T-3.

Globalement, en T-2 la collaboration entre futurs chercheurs de l'IRSPUM demeure faible, quoi qu'elle soit passée à près de 20 %. Encore 80% des chercheurs ne collaborent pas avec d'autres futurs membres de l'IRSPUM. Nous notons que les collaborations illustrées à la figure 4.2 témoignent d'une appartenance commune à une institution de recherche préexistante à celle de l'IRSPUM, le GRIS.

Des quatre composantes présentes en T-2 à la figure 4.2 seulement deux se maintiendraient si l'on supprimait les liens vers les chercheurs associés, dont une de façon intégrale. Un seul sommet se détacherait de la composante la plus grosse. Aussi ne resterait-il qu'un lien perdurant depuis T-3. Le même chercheur régulier qu'en T-3 (A13) se retrouverait encore une fois isolé par la suppression des liens vers les chercheurs associés. Quatre liens seraient supprimés. La place des chercheurs associés semble complexe

1. Ce qui est illustré par la surface plus grande du sommet le représentant

et certains ont des positions clés dans le réseau. Passons au T-1.

4.3 2000 à 2004

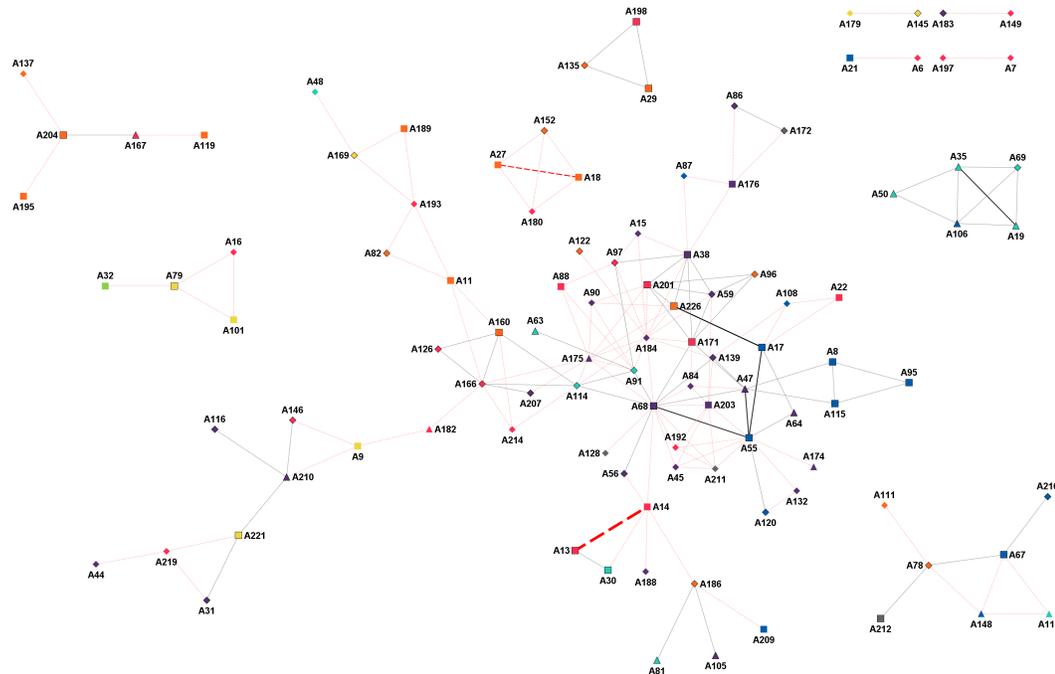
La période 2000 à 2004 (T-1) apparaît comme une période charnière pour la création de l'IRSPUM puisque c'est à ce moment qu'un réseau de chercheurs prend forme. Durant cette période, 186 des 226 chercheurs de l'IRSPUM ont publié au moins un article ; pour 99 de ceux-ci nous n'avons pas trouvé d'article au cours des périodes précédentes. De ces 186 chercheurs (dont 91 sont actuellement chercheurs réguliers à l'IRSPUM), 105 copublient (dont 58 chercheurs réguliers) avec (au moins) un autre chercheur de l'IRSPUM.

Au plan des origines disciplinaires, les identités se mélangent peu. La période se caractérise par une grande homophilie disciplinaire. Les spécialistes en sciences sociales collaborent avec les spécialistes en santé communautaires, les épidémiologistes entre eux ou avec des bio-statisticiens, et les psychologues collaborent entre eux.

La figure 4.3 montre le graphe des collaborations entre futurs chercheurs de l'IRSPUM en T-1. Nous observons 173 liens impliquant 105 chercheurs. La figure 4.3 montre 11 composantes dont quatre paires de chercheurs (A179, A145), (A183, A149), (A21, A6) et (A197, A7), une triade dont tous les membres sont liés (une clique) (A135, A29, A198), deux composantes à quatre sommets dont une clique (A27, A152, A180, A18) et (A32, A79, A16, A101), deux composantes de cinq sommets (A137, A204, A195, A167, A119) et (A50, A35, A106, A69, A19), une composante formée par sept sommets (A12, A78, A11, A67, A148, A110, A216) et la plus grosse composante au centre, comportant 69 sommets et 136 liens.

Malgré le développement de liens, plusieurs petites composantes ou dyades demeurent isolées. C'est le cas notamment de la composante formée des chercheurs en sciences sociales, de celle des chercheurs en psychologie. En fait les identités se mélangent encore peu. Les médecins collaborent surtout avec les épidémiologistes et les bio-statisticiens, les psychologues collaborent presque exclusivement entre eux et les spécialistes des sciences sociales collaborent surtout avec les chercheurs en santé commu-

Figure 4.3 – Collaborations entre futurs chercheurs de l'IRSPUM entre 2000 à 2004 (T-1)



SOURCES : Les résultats sont obtenus à partir de l'analyse de la base de données COLLABO_T-1 à l'aide du logiciel *Visone*

nautaire. Les seuls qui semblent transgresser largement les frontières disciplinaires sont ceux formés en santé communautaire, une spécialité par définition interdisciplinaire.

La figure 4.3 de la page 59 montre une relation perdurant depuis T-3 (A14, A13) et six relations se maintenant depuis T-2 (A55, A68), (A55, A47), (A55, A17), (A17, A226), (A35, A19) et (A27, A18). A68 est le chercheur le plus central, avec 16 liens vers d'autres membres de l'IRSPUM. Il est suivi par A184 qui possède 13 liens vers d'autres chercheurs et A55 qui en possède 11. Il est à noter que A184 fait son entrée en T-1.

La présence de certains chercheurs associés est essentielle au maintien de la grosse composante ; leur exclusion entraînerait la scission du graphe en plusieurs parties. C'est notamment le cas de A11, de A9 et A182 (dont l'arête crée un pont) et de A14, pour ne nommer que les principaux. La suppression des liens vers les chercheurs associés

créerait la disparition de deux composantes et isolerait A79. Elle ferait aussi disparaître la relation entre A13 et A14 qui dure depuis trois périodes consécutives. La présence des chercheurs associés importe dans le maintien de la structure présentée en figure 4.3, bien qu'un noyau perdurerait. Les chercheurs associés permettent de faire des ponts entre les diverses composantes du réseau du T-1.

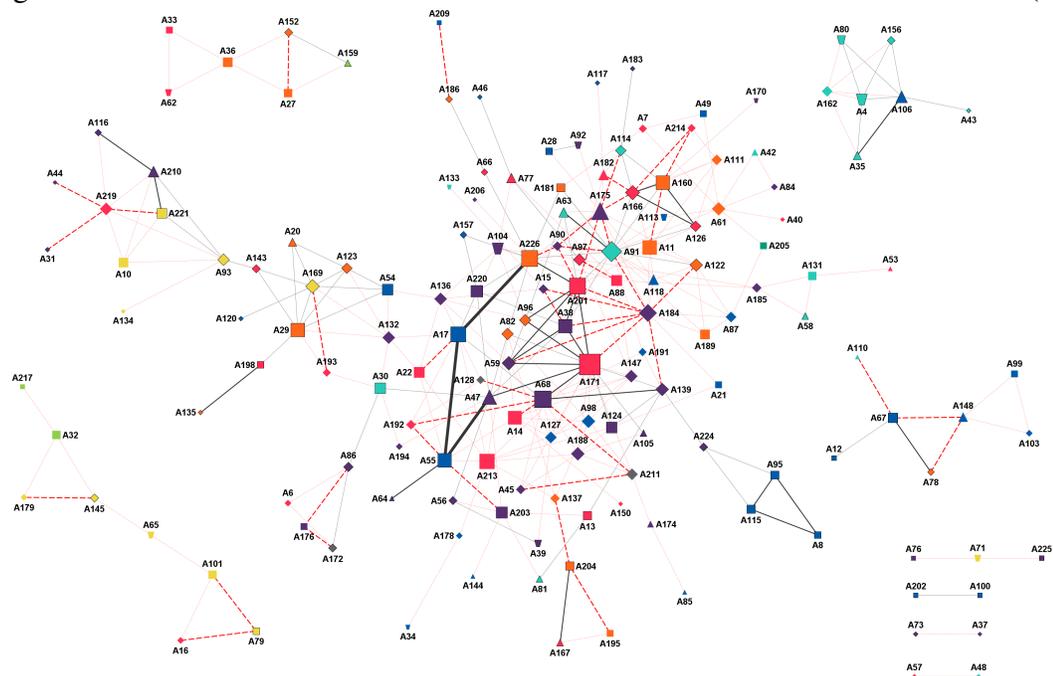
Par ailleurs, il est aussi important de noter que 44% des chercheurs ne copublient avec aucun autre des futurs chercheurs de l'IRSPUM, ce qui signifie qu'en T-1 c'est maintenant la majorité des chercheurs qui collaborent avec au moins un de leurs futurs collègues de l'IRSPUM. Rappelons qu'en T-3 seulement 12% des futurs chercheurs de l'IRSPUM collaboraient entre eux, en T-2, ils étaient rendus à 20% et que pour le T-1, ce pourcentage passe à 66%. Nous notons donc une croissance des relations entre futurs chercheurs de l'IRSPUM ; un réseau social se met en place, une communauté commence à émerger. La cohésion sociale s'est accentuée au fil du temps. Passons au T0.

4.4 2005 à 2009

En T0, la période précédant immédiatement la création de l'IRSPUM, 20 chercheurs font leur entrée pour la première fois dans nos données. Des 206 chercheurs présents en T0, 60 sont présents depuis T-3, 32 depuis T-2 et 94 depuis T-1 (la forme des sommets indique la première période d'apparition du chercheur). Des 206 chercheurs actuels de l'IRSPUM ayant publié au T0, 99 ont le statut de chercheurs réguliers. Nous pouvons noter ici que cela signifie que sur les 226 chercheurs de l'IRSPUM que comporte notre population d'analyse, il a ÃtÃ impossible de trouver d'article en T0. Rappelons que ceci ne signifie pas qu'ils n'ont aucunement publié durant cette période, puisque notre étude prend en compte uniquement les articles et pas les autres types de publication de recherche.

La figure 4.4 représente la période précédent la création de l'IRSPUM. Nous observons 319 liens impliquant 159 chercheurs (dont 79 sont des chercheurs réguliers). Les auteurs qui publient avec le plus de (futurs) collègues de l'IRSPUM sont des auteurs déjà présents dès T-3 : A171, A226, A17, A55, A213, hormis A91, qui provient de T-2. Cet

Figure 4.4 – Collaborations entre futurs chercheurs de l'IRSPUM en 2005 à 2009 (T0)



SOURCES : Les résultats sont obtenus à partir de l'analyse de la base de données COLLABO_T0 à l'aide du logiciel *Visone*

élément nous permet de penser que ces chercheurs sont des « piliers » autour desquels le réseau social des collaborations au sein des membres de l'IRSPUM s'articule.

Nous retrouvons neuf composantes à la figure 4.4. La plus grosse composante est formée de 122 sommets et 276 liens. Cette composante se caractérise par son interdisciplinarité ainsi que par la présence de *leaders* dans les principales disciplines de la santé publique (A171 et A201), en médecine, en sciences humaines et sociales (A17 et A55), en santé communautaire (A68, A175 et A184) et en épidémiologie (A29 et A226).

Pour ce qui est des huit autres composantes, nous retrouvons trois paires de chercheurs (A202, A100), (A73, A37) et (A57, A48), une triade (A76, A71, A226), une composante de six sommets (A33, A62, A36, A152, A27, A159), deux composantes de sept sommets (A80, A156, A162, A4, A106, A35, A43) et (A110, A12, A67, A148, A99, A103) et une composante de huit sommets (A217, A32, A179, A145, A65, A101, A16, A79). Comme on peut le constater, ces composantes demeurent en marge de l'IRSPUM. Il s'agit large-

ment de greffons qui ont été fait au GRIS pour constituer l'IRSPUM : Les spécialistes en santé animale et médecine vétérinaire (A33, A62, A36, A15, A159 ainsi que A217, A32, A179, A145, A65, A101, A79 et A16) et des spécialistes en sciences humaines et sociales (A67, A110, A148, A12, A78) ainsi qu'en psychologie (A80, A162, A156, A35, A106, A43)

Trois relations perdurent depuis T-2, c'est-à-dire depuis plus de 10 ans : (A55, A47), (A55, A17) et (A17, A226). Le chercheur A171 maintient six de ses liens avec d'autres futurs chercheurs de l'IRSPUM depuis T-1. Il est important de souligner que la densité s'accroît, ce qui signifie que la communauté devient plus cohésive.

Certains associés apparaissent vitaux au maintien de la structure observée à la figure 4.4 de la page 61, notamment A132. L'exclusion de cet associé scinderait la plus grosse composante. En fait, l'exclusion des chercheurs associés ne laisserait intègre qu'une seule des neuf composantes présentes dans le graphe de la figure 4.4 : la dyade (A202, A2100). Toutes les autres composantes se transformeraient par l'exclusion des chercheurs associés ; une des composantes se désintégrerait entièrement, ne laissant que des chercheurs réguliers isolés. La grosse composante se scinderait en cinq parts dont deux dyades, une composante de 13 sommets ainsi qu'une plus grosse de 46 sommets et 69 liens. Six chercheurs réguliers se retrouveraient isolés par la suppression des liens vers les associés : A58, A79, A117, A225, A31 et A 73. Des collaborations à long terme se perdraient aussi. La plupart demeureraient néanmoins, dont les trois s'étalant sur trois périodes consécutives. Il y aurait suppression de 219 liens sur 319, dont 46 liens qui existaient depuis T-1. Les chercheurs associés apparaissent essentiels au maintien de la structure observable en figure 4.4.

Pendant la période qui verra l'IRSPUM émerger, seulement un peu moins de 23% des chercheurs ne copublient pas avec un autre chercheur de l'IRSPUM. La proportion des chercheurs ayant copublié avec au moins un (futur) collègue de l'IRSPUM s'est inversée entre le T-3 et T0. Rappelons que 12% des chercheurs copubliaient avec un autre futur membre de l'IRSPUM en T-3, qu'ils étaient 20% en T-2, 66% en T-1 et qu'ils sont 77% en T0.

4.4.1 Conclusion

Ce qui ressort principalement des résultats d'analyse présentés ci-haut, c'est qu'au fil des périodes analysées, un réseau social émerge : de plus en plus les chercheurs se mettent à collaborer entre eux, à s'allier à la rédaction conjointe d'articles. La cohésion sociale s'accroît ; l'intensité des liens augmente et se perpétue dans le temps et les collaborations deviennent de plus en plus interdisciplinaires. La collaboration entre chercheurs se construit de façon graduelle au fil des périodes analysées. L'homophilie est plus forte entre membres d'une même composante qu'entre les composantes.

Pour clore ce chapitre, rappelons que les résultats de nos analyses suggèrent l'émergence d'un réseau social de chercheurs ; de quasi inexistante, la collaboration entre futurs chercheurs de l'IRPSUM s'accroît jusqu'à toucher la majorité de la population étudiée. Le prochain chapitre vise à prendre en compte ce qui fait en sorte que ces chercheurs se soient mis à collaborer entre eux ; qu'est ce qui peut bien les réunir et faire en sorte que nous puissions parler de l'émergence d'une communauté épistémique autour de la « nouvelle santé publique » ?

CHAPITRE 5

...AUTOUR D'UN MÊME THÈME : LA « NOUVELLE SANTÉ PUBLIQUE »

Au chapitre précédent nous avons montré comment les chercheurs membres de l'IRSPUM ont, au fil du temps et ce avant la création de cet institut de recherche, formé un réseau social de plus en plus cohésif. Cette cohésion seule pourrait permettre à certains de conclure qu'ils sont en présence d'une communauté épistémique. Mais, selon notre perspective des réseaux sociaux, les aspects sémantiques sont d'une importance capitale puisqu'ils permettent d'aborder ce qui peut bien lier entre-eux des gens. En d'autres mots, les gens sont liés certes, mais parce qu'ils partagent quelque chose.

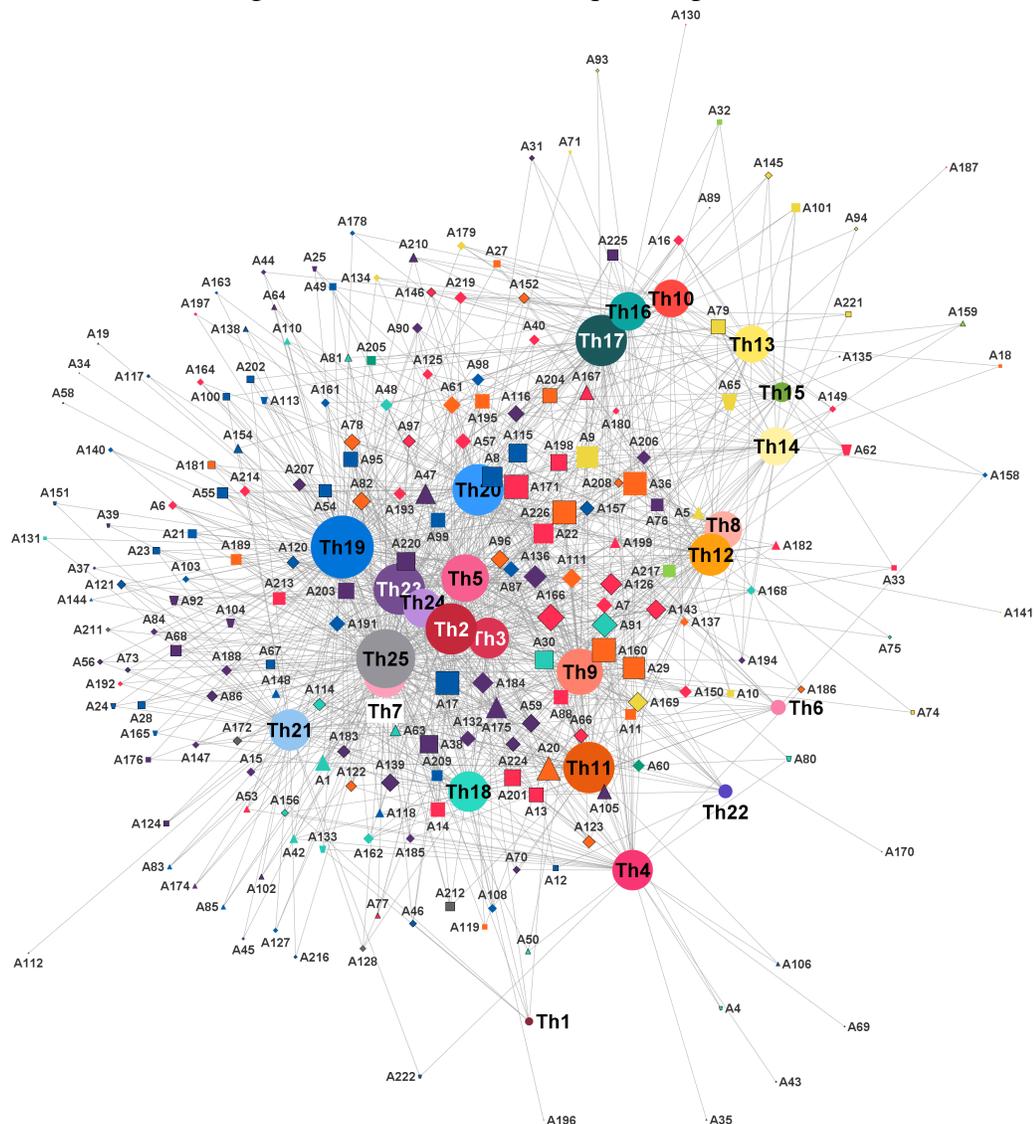
C'est grâce à nos analyses préliminaires que nous avons été en mesure de prendre en compte les aspects sémantiques liant entre-eux les chercheurs de l'IRSPUM. Celles-ci nous ont permis, rappelons-le, de détecter des thèmes de recherche manipulés par la population étudiée dans les articles qu'ils ont copubliés.

Ce dernier chapitre vise à répondre à une ultime question : quelle place occupe la « nouvelle santé publique » dans le paysage sociosémantique de la communauté épistémique de l'IRSPUM ? Est-ce possible de déceler une évolution de la cohésion sémantique au sein de la communauté épistémique dont nous étudions l'émergence ? Il sera d'abord question de présenter la place qu'occupe la « nouvelle santé publique » au sein de l'IRSPUM, en se référant spécifiquement au T0. Ensuite, nous démontrerons comment les chercheurs se sont non seulement mis à collaborer entre eux, mais aussi, se sont mis à partager de plus en plus le thème des déterminants sociaux de la santé.

5.1 Thématiques de recherche centrales et périphériques

C'est en présentant le graphe sociosémantique complet du T0 que nous comptons dans un premier temps démontrer l'importance occupée par la nouvelle santé publique au sein de l'IRSPUM. Rappelons que pour tous les graphes présentés, les légendes sont disponibles à l'annexe III en page xl.

Figure 5.1 – Réseau sémantique complet, en T0



SOURCES : Les résultats sont obtenus à partir de l'analyse de la base de données SOCIOSEM_T0 à l'aide du logiciel *Visone*

À la figure 5.1 (voir aussi le tableau 5.I à la page 67), le thème le plus souvent manipulé par les chercheurs est celui des déterminants sociaux de la santé (Th19) ($C = 5,799$), suivi de près par celui de la gestion des services de soins de santé (Th25) ($C = 5,198$). En troisième position se trouve le thème des comportements sociaux (Th20) ($C = 3,308$). Les thèmes les moins manipulés sont respectivement la santé dentaire (Th1)

($C = 0,301$) et la santé internationale (Th22) ($C = 0,344$) ; ils font partie des rares thèmes ayant une centralité de degré inférieure à celle de certains chercheurs ; c'est-à-dire que certains chercheurs publient sur plus de thèmes qu'il y a de chercheurs en tout publiant sur ces thèmes peu populaires. Entre les thèmes centraux se trouvent entre autres la promotion de la santé (Th23), la médecine préventive (Th2), et la santé des femmes et des enfants (Th5). Il est à noter que lorsque des sommets se retrouvent sous d'autres sommets en figure 5.1, c'est que leur centralité de degré est moindre que celle des sommets les occultant.

La figure 5.1 de la page 65 nous montre aussi à quel point certains thèmes sont centraux, mais nous retrouvons en périphérie de nombreux thèmes, dont un regroupement de thèmes axés davantage sur des thématiques touchant la biologie et la chimie (autour de Th14, Th13 et Th10). Autour de la santé environnementale (Th17), nous retrouvons les maladies professionnelles (Th16) et les maladies infectieuses (Th10). Les études sur les jeunes (Th4) se retrouvent aussi, un peu plus à part. Nous ne prendrons malheureusement pas le temps de commenter la totalité de ce qui peut s'observer en figure 5.1, puisque ce graphe est d'une complexité apparente ; nous l'avons présenté essentiellement pour montrer la place occupée par les déterminants sociaux de la santé à l'IRSPUM.

Notons aussi que l'on observe à la figure 5.1 de la page 65 que dans le voisinage des grandes thématiques se retrouvent un certain nombre de thème de centralité moindre. Ces thèmes sont ceux qui sont, la plupart du temps, manipulés conjointement à l'une des grandes thématiques. Comme nous pouvons l'observer, le thème des soins de santé aux personnes âgées se retrouve en bonne partie éclipsé sous le thème de la gestion des services de soins de santé. La proximité des points dans le graphe signifie une proximité dans l'usage par de mêmes chercheurs, de ces thèmes. Ces thèmes voisins viennent apporter une nuance sur le recours aux grandes thématiques.

Les thèmes les plus centraux se retrouvant dans les articles publiés par les chercheurs maintenant membres de l'IRSPUM reflètent ce qui est désigné par la « nouvelle santé publique », c'est-à-dire, un intérêt marqué pour l'étude des déterminants sociaux de la santé.

5.2 Évolution des thématiques de 1990 à 2009

Nous ne présenterons pas les graphes sociosémantiques complets des autres périodes, quoi qu'on y constate la part croissante de l'intérêt pour les déterminants sociaux ; ces graphes sont complexes, comme il est possible de le constater à la figure 5.1 de la page 65. Un tableau indiquant le degré de centralité de chaque thème, selon chaque période permet d'explorer davantage l'évolution dans l'usage des différents thèmes abordés à travers les publications analysées.

Tableau 5.I: Centralité des thématiques par période, en fonction du nombre d'auteurs les manipulant

Thèmes	T-3	T-2	T-1	T0
(Th1) Santé dentaire	0,43 (2)	0,98 (6)	0,51 (8)	0,30 (7)
(Th2) Médecine préventive	1,29 (6)	1,44 (7)	1,40 (22)	1,93 (45)
(Th3) Soins de santé	1,72 (8)	2,12 (13)	2,54 (40)	1,89 (44)
(Th4) Développement des jeunes et adolescents	1,50 (7)	2,29 (14)	1,97 (31)	1,72 (40)
(Th5) Santé des femmes et des enfants	3,44 (16)	2,12 (13)	2,86 (45)	2,75 (64)
(Th6) Santé internationale	NA (0)	0,16 (1)	0,70 (11)	0,56 (13)
(Th7) Soins aux personnes âgées	1,72 (8)	2,79 (17)	1,59 (25)	2,36 (55)
(Th8) Autres troubles médicaux	0,64 (3)	0,33 (2)	0,76 (12)	1,03 (24)
(Th9) Maladies chroniques	1,93 (9)	2,45 (15)	2,22 (35)	2,56 (60)
... Suite à la prochaine page				

Thèmes	T-3	T-2	T-1	T0
(Th10) Maladies infectieuses	3,00 (14)	2,61 (16)	1,84 (29)	1,12 (26)
(Th11) Lignes directrices en santé	1,72 (8)	1,31 (8)	1,40 (22)	1,93 (45)
(Th12) Médicaments	2,58 (12)	2,61 (16)	1,78 (28)	2,23 (52)
(Th13) Immunologie	1,72 (8)	1,96 (12)	1,27 (20)	1,38 (32)
(Th14) Recherche génétique et protéomique	1,29 (6)	1,44 (7)	1,33 (21)	1,42 (33)
(Th15) Zoonoses	1,93 (9)	1,31 (8)	0,83 (13)	0,73 (17)
(Th16) Maladies professionnelles	2,58 (12)	2,45 (15)	2,29 (36)	1,55 (36)
(Th17) Santé environnementale	3,86 (18)	2,79 (17)	2,22 (35)	2,19 (51)
(Th18) Santé mentale	1,72 (8)	1,31 (8)	1,46 (23)	1,72 (40)
(Th19) Déterminants sociaux de la santé	4,72 (22)	5,56 (34)	6,42 (101)	5,80 (135)
(Th20) Comportements sociaux	2,36 (11)	3,10 (19)	3,88 (61)	3,31 (77)
(Th21) Politiques publiques	2,36 (11)	1,80 (11)	1,84 (29)	2,10 (47)
(Th22) Éthiques en recherche génétique	0,43 (2)	0,16 (1)	0,57 (9)	0,52 (12)
(Th23) Promotion de la santé	1,50 (7)	1,80 (11)	1,97 (31)	2,19 (51)
(Th24) Évaluation des technologies de la santé	0,86 (4)	0,82 (5)	1,02 (16)	1,59 (37)
... Suite à la prochaine page				

Thèmes	T-3	T-2	T-1	T0
(Th25) Gestion des services de santé	4,72 (22)	4,90 (30)	5,34 (84)	5,20 (121)
SOURCES : Résultats obtenus par l'analyse des bases de données SOCIOSEM_T-3 à T0, avec <i>Visone</i>				

Le tableau 5.1 présente le degré de centralité de chaque thématique par période d'analyse et ce, en fonction du nombre d'auteurs ayant manipulé la thématique donnée pendant la période donnée. Entre les parenthèses se trouve le nombre de chercheurs ayant manipulé le thème. Par exemple, il est possible de constater que la thématique de la santé dentaire (Th1) est restée marginale et ne s'est pas significativement accrue, malgré la croissance de notre population d'analyse, tandis que celle des déterminants sociaux de la santé (Th19) a eu une forte attraction particulièrement au début des années 2002, le nombre de chercheurs manipulant ce thème est passé de 34 pour la période de 1995-1999 à 101 pour la période 2000-2004. L'attraction pour ce thème va se poursuivre à la période 2005-2009 avec 135 chercheurs ayant publié sur ce thème. Il est possible de constater que c'est la thématique ayant attiré la plus grande proportion de chercheurs

La thématique des déterminants sociaux de la santé (Th19) était déjà centrale en 1990-1994 (T-3), soit la plus ancienne période prise en compte dans nos analyses. Cette thématique est suivie de près par celle de la gestion des services de soins de santé (Th25) et celle des comportements sociaux (Th20). La thématique des comportements sociaux a aussi attiré une vague de chercheurs en T-1, en passant de 19 chercheurs en T-2 à 61 en T-1.

Le tableau 5.1 nous permet aussi de constater le déclin de la manipulation de certaines thématiques. C'est le cas des zoonoses (Th15), qui occupent de moins en moins de chercheurs depuis T-3. Il en va de même pour les maladies infectieuses (Th10). Certaines thématiques demeurent marginales. C'est le cas de la santé internationale (Th6), de l'éthique en recherche génétique (Th22) et de la santé dentaire (Th1).

Si l'on regarde la manipulation des thématiques par période, l'on constate qu'en T-3, les thématiques les plus centrales sont la santé des femmes et des enfants (Th5), les maladies infectieuses (Th10), la santé environnementale (Th17), les déterminants

sociaux de la santé (Th19) et la gestion des services de soins de santé (Th25). En T-2, les thématiques Th19 et Th25 sont encore les plus centrales ; les autres thématiques n'ont pas une centralité élevée. Les comportements sociaux (Th20) apparaissent aux côtés de Th19 et Th25 en T-1. Ces thématiques demeurent les plus centrales en T0. Les thématiques centrales ont changé au fil du temps ; les deux dernières périodes s'axent davantage sur une perspective sociale de la santé publique.

Bref, les deux grandes thématiques majeures de l'IRSPUM sont les déterminants sociaux de la santé et la gestion des services de soins de santé et c'est largement autour de ces thèmes que s'est constituée l'IRSPUM. Ce constat laisse penser que l'IRSPUM s'est développé en regroupant la recherche en santé publique à l'Université de Montréal et en recrutant de nouveaux chercheurs, que les intérêts des chercheurs ont de plus en plus convergé mais que la place relative de ces deux thématiques a peu évolué. Pour mieux comprendre la convergence vers la nouvelle santé publique, il nous apparaît pertinent de mettre en relation les chercheurs et les thématiques dont ils traitent. C'est ce que nous ferons dans la prochaine section.

5.3 Convergence vers la « nouvelle santé publique »

C'est en analysant les données sociales et sémantiques de façon simultanée que nous serons en mesure d'aller puiser dans la richesse des données collectées. Tel que mentionné au chapitre 2, le treillis de Galois nous permet d'ordonner et hiérarchiser les paires de chercheurs-thèmes, tout en respectant les chevauchements qui peuvent exister. Ainsi, il devient possible de savoir « qui travaille sur quoi et avec qui », ce que les graphes ne nous permettent pas de faire. Comme le rappelle Wasserman et Faust (1994, p. 326), les treillis de Galois possèdent deux caractéristiques importantes : l'idée de dualité et celle de sous-ensemble. Chaque sommet d'un treillis de Galois représente un sous-ensemble de chercheurs liés à une thématique commune, mais aussi à un sous-ensemble de thématiques que les chercheurs manipulent. Chaque sommet possède ainsi deux types d'étiquettes : une identifiant les chercheurs, et l'autre identifiant les thématiques. Autrement dit, un sommet représente la co-occurrence d'au moins un chercheur

et un thème de recherche. L'idée de dualité se réfère à ces deux perspectives complémentaires : d'un côté, l'on retrouve les relations entre les chercheurs manipulant des thématiques communes, et de l'autre, l'on retrouve les relations entre les thématiques qui attirent les chercheurs. Les graphes, comme celui de la figure 5.1 ne nous indiquent que les liens entre les chercheurs et les thèmes ; les treillis permettent un classement hiérarchique des relations d'affiliations.

Si les graphes nous permettent de savoir qui manipule quelle thématique, les treillis de Galois, quant à eux peuvent nous renseigner sur l'avènement simultané des chercheurs et des thèmes : qui manipulent quoi, avec qui et quoi ? La classification hiérarchique que produit le treillis de Galois permet de découvrir des sous-ensembles de chercheurs et de thématiques, allant des plus généraux aux plus spécifiques. Malgré leurs forces, les treillis de Galois sont vites difficiles à visualiser. Les treillis de Galois obtenus à partir de nos données sont extrêmement volumineux et complexes et il n'est pas possible de les présenter sans émonder ces treillis. Nous avons choisi de présenter des tronçons provenant des treillis complets.

L'exploration de ces treillis pourrait se poursuivre longuement, mais afin de répondre à notre question de recherche, il nous a paru pertinent d'analyser la manipulation du thème le plus central que nous avons observé (les déterminants sociaux de la santé Th19) par les chercheurs autour desquels s'est formée la communauté que nous avons observée. Ces chercheurs sont les sept chercheurs réguliers ayant le degré de centralité le plus élevé en T0, ¹.

Les sept chercheurs réguliers les plus centraux et présents tout au cours des périodes analysées proviennent de 5 grandes disciplines différentes : (A68) est en santé publique, (A55) et (A17) sont en sciences humaines et sociales, (A171) et (A13) sont en médecine, (A226) est en épidémiologie et (A79) est en biologie/chimie. Ces chercheurs représentent aussi la diversité des formations dont relèvent les chercheurs de l'IRSPUM.

Pour lire un treillis de Galois², il faut d'abord comprendre qu'il peut s'interpréter de haut en bas et/ou de bas en haut. Le sommet le plus élevé représente l'ensemble de

1. En fait, ils sont les plus centraux, pour la majorité, depuis le T-3

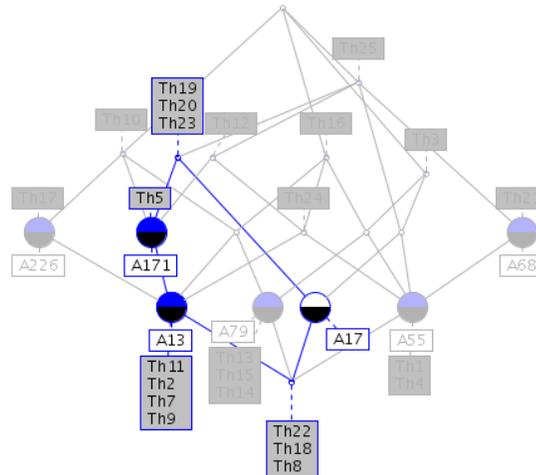
2. Il s'agit ici d'une forme de visualisation (en « diamant ») des treillis de Galois

tous les chercheurs. Le sommet le plus bas est l'ensemble de tous les thèmes. Parfois, il y a des thèmes liés à ce sommet ; ce sont les thèmes qui ne sont manipulés par aucun des sept chercheurs sélectionnés. Dans un treillis de Galois, il faut suivre les « chemins » pour saisir les hiérarchies. Comme mentionné plus tôt, les thèmes inutilisés par les sept chercheurs se retrouvent tout en bas. Si l'on monte dans le treillis, nous trouvons enfin des sommets correspondant à la co-occurrence de chercheurs et de thèmes. Plus un chercheur est bas, plus il a manipulé de thèmes, mais plus un thème est bas, moins il est manipulé par les chercheurs. La base est donc constituée de chercheurs et de leurs thèmes spécifiques, leur « niche ». Plus un chercheur apparaît haut, moins il a manipulé de thèmes, mais plus un thème est haut, plus il attire de chercheurs.

Afin d'attirer l'attention sur le phénomène qui nous intéresse, nous avons choisi de montrer les chemins qui mènent vers les déterminants sociaux de la santé (Th19) ainsi que ses dépendances. Les chemins sont tracés en bleu. Hormis les deux ensembles aux extrémités, chaque sommet représente la rencontre de chercheurs et de thèmes, mais pour faciliter la visualisation, seule la première apparition (d'un chercheur ou d'un thème) est identifiée. Si on part de la première apparition d'un chercheur (en bas), on peut donc suivre les lignes (chemins) vers les autres thèmes qu'il a manipulés. En partant de la première apparition d'un thème (en haut), on peut, en suivant les lignes qui s'y rattachent, trouver tous les auteurs qui l'ont manipulé. Prenons un exemple concret :

La figure 5.2 de la page 73 montre qu'en T-3 trois chercheurs sur les sept présents manipulent le thème des déterminants sociaux de la santé (Th19) (A13, A17 et A171). A13 est le seul à manipuler les thèmes des lignes directrices en santé (Th11), de la médecine préventive (Th2), des soins aux personnes âgées (Th7) et celui des maladies chroniques (Th9). Il manipule aussi celui de la santé des femmes et des enfants (Th5) (tout comme A171) et des déterminants sociaux de la santé (Th19), des comportements sociaux (Th20) et celui de la promotion de la santé (Th23). A17 manipule exclusivement le thème des déterminants sociaux de la santé (Th19), des comportements sociaux (Th20) et de la promotion de la santé (Th23), tandis que A171 manipule ces trois thèmes ainsi que Th5 qu'il partage avec A13). Trois thèmes sont manipulés par aucun des sept chercheurs : l'éthique en recherche génétique (Th22), les autres troubles médicaux (Th18) et la santé

Figure 5.2 – Tronçon axé sur l’usage des déterminants sociaux de la santé en T-3



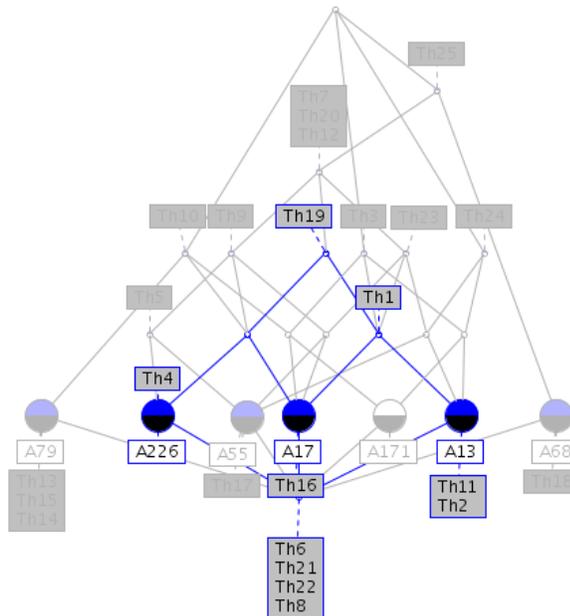
SOURCES : Les résultats sont obtenus à partir de l’analyse de la base de données SOCIOSEM_T-3 à l’aide du logiciel *Concept Explorer*

mentale (Th8).

En T-3 ce sont les deux médecins (A171 et A13) et un chercheur en sciences humaines et sociales (A17) qui manipulent le thème des déterminants sociaux de la santé dans leurs articles publiés. Les déterminants sociaux de la santé sont hauts dans la hiérarchie, mais au même niveau hiérarchique que le thème des médicaments (Th12) et plus bas que les maladies infectieuses (Th10) et que la gestion des services de soins de santé (Th25).

En T-2, la figure 5.3 de la page 74 montre qu’il y a encore seulement trois chercheurs sur sept manipulant le thème Th19, comme en T-3 : A13 et A17 ont continué de manipuler ce thème dans leurs articles, s’ajoute à eux A226. A13 demeure le seul à manipuler le thème des lignes directrices en santé (Th11) et de la médecine préventive (Th2). Il manipule aussi la santé dentaire (Th1) et les déterminants sociaux de la santé (Th19). A17 est le seul à manipuler le thème de la santé environnementale (Th16), mais partage avec A13 les thèmes de la santé dentaire (Th1) et celui des déterminants sociaux de la santé (Th19). A226 est le seul à manipuler le thème du développement des jeunes

Figure 5.3 – Tronçon axé sur l’usage des déterminants sociaux de la santé en T-2



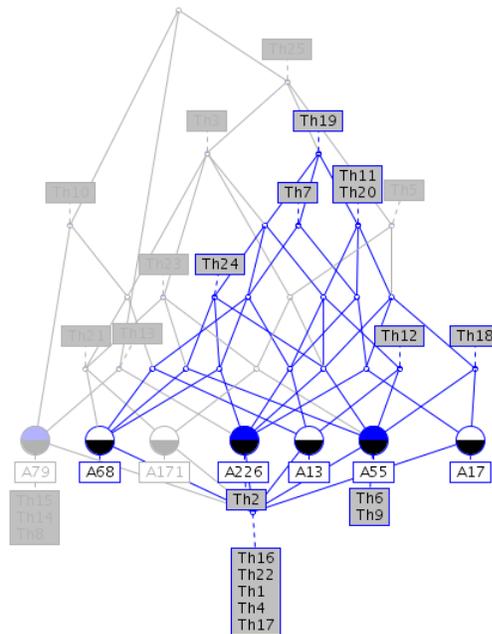
SOURCES : Les résultats sont obtenus à partir de l’analyse de la base de données SOCIOSEM_T-2 à l’aide du logiciel *Concept Explorer*

et adolescents (Th4) et il manipule Th19. Quatre thématiques n’ont pas été manipulées pendant cette période par aucun des sept chercheurs : la santé internationale (Th6) les politiques publiques (Th21) l’éthique en recherche génétique(Th22) et la santé mentale (Th8)

En T-2, les trois chercheurs manipulant le thème des déterminants sociaux de la santé proviennent de trois différentes disciplines : A226 provient d’épidémiologie, A17, déjà présent en T-3, quant à lui provient des SHS et A13 aussi présent en T-3 provient de médecine ; ce thème de recherche ne semble pas l’apanage exclusif des chercheurs en sciences sociales. Les déterminants sociaux de la santé se retrouve au même niveau que de nombreux thèmes : maladies infectieuses (Th10), maladies chroniques (Th9), soins de santé (Th3), promotion de la santé (Th23), évaluation des soins de santé (Th24), mais il attire moins de chercheurs qu les thématiques de la médecine préventive (Th2), des

comportements sociaux (Th20), des médicaments (Th12) et la gestion des services de soins de santé (Th25).

Figure 5.4 – Tronçon axé sur l’usage des déterminants sociaux de la santé en T-1



SOURCES : Les résultats sont obtenus à partir de l’analyse de la base de données SOCIOSEM_T-1 à l’aide du logiciel *Concept Explorer*

C’est en T-1 que la convergence vers le thème des déterminants sociaux de la santé se fait la plus marquée comme il est possible de le constater à la figure 5.4 de la page 75. La forme du treillis s’est allongée, puisque de plus en plus les chercheurs manipulent des thèmes communs. Les déterminants sociaux de la santé interpellent maintenant cinq chercheurs sur sept (A68, A226, A13, A55 et A17). Notons que A17 et A13 manipulent ce thème depuis T-3 et A226 depuis T-2. Il y a maintenant neuf thèmes sur 25 se trouvant manipulés conjointement à Th19. Le thème de la gestion des soins de santé (Th25) reste manipulé par le plus grand nombre de chercheurs. Cinq thématiques ne sont pas abordées par ces chercheurs : les maladies professionnelles (Th16), l’éthique en recherche génétique (Th22), la santé dentaire (Th1), le développement des jeunes et des

adolescents (Th4) et la santé environnementale (Th17).

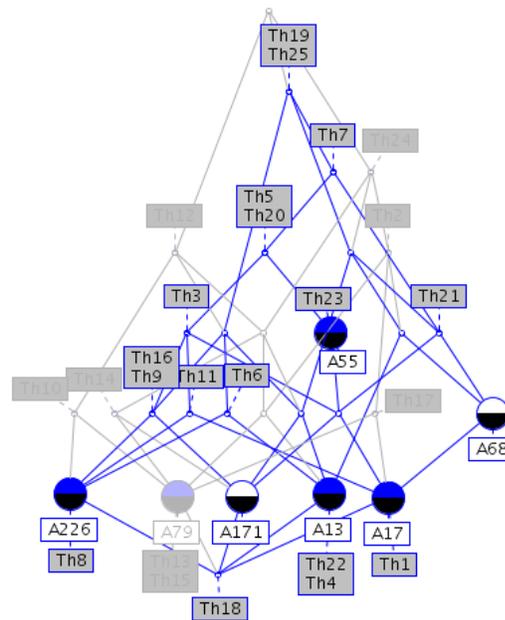
En T-1 s'ajoute à A226, A13 et A17, déjà présentés, A55, issu des SHS et A68, issue de la santé publique. La gestion des soins de santé (Th25) continue de rallier un nombre plus important de chercheurs, mais il semble que ces chercheurs convergent vers des thématiques communes ; tous les chercheurs ayant manipulé Th19 ont aussi manipulé Th25. Allons voir ce qui se produit en T0. Du point de vue hiérarchique, seul Th25 est plus manipulé que Th19. Le thème des soins de santé (Th3) est autant manipulé que celui des déterminants sociaux.

On peut aussi remarquer que les chercheurs ont de plus en plus de thèmes de recherche communs. A226 est le seul à manipuler la médecine préventive (Th2), tandis que A55 est le seul à manipuler les thèmes de la santé internationale (Th6) et des maladies chroniques (Th9). La plupart des chercheurs (4/5) qui partagent Th19, partagent aussi les thématiques des soins aux personnes âgées (Th7), des médicaments (Th11) et des comportements sociaux (Th20). Allons voir ce qui se produit en T0.

En T0, c'est six des sept chercheurs sélectionnés qui manipulent le thème des déterminants sociaux de la santé (Th19), comme l'illustre la figure 5.5 de la page 77. Seul A79 ne manipule pas ce thème. Le nombre des thèmes conjointement manipulés à celui des déterminants sociaux de la santé passe à 14 sur 25, pendant cette période. Le thème des déterminants sociaux de la santé est maintenant le plus populaire, tout comme celui de la gestion des services de soins de santé (Th25). Tous ceux ayant manipulé Th25 ont aussi manipulé Th19. Le thème des déterminants sociaux de la santé devient de plus en plus structurant de la recherche de ces chercheurs et s'intègre maintenant à presque toutes les autres thématiques.

En T0, les sept chercheurs manipulent Th19 hormis A79 qui provient de la biologie/chimie. Les thèmes de la biologie/chimie demeurent, comme nous l'avons montré en figure 5.1 de la page 65, en périphérie des thèmes centraux. Seule la santé mentale (Th18) demeure non abordée par ces sept chercheurs. Dans la hiérarchie des thématiques, celle des soins aux personnes âgées (Th7) se retrouve élevée ; juste sous Th19 et Th25. On retrouve ensuite les thèmes de la santé des femmes et des enfants (Th5) et des comportements sociaux (Th20).

Figure 5.5 – Tronçon axé sur l’usage des déterminants sociaux de la santé en T0



SOURCES : Les résultats sont obtenus à partir de l’analyse de la base de données SOCIOSEM_T0 à l’aide du logiciel *Concept Explorer*

5.4 Conclusion

Pour conclure ce dernier chapitre, faisons maintenant le pont entre le chapitre précédent et celui-ci : notre démarche visait à démontrer que l’émergence et l’évolution des communautés épistémiques est à la fois liée à la formation d’un réseau d’acteurs sociaux et à la formation d’un réseau de sens. Nous avons vu dans le chapitre 4 qu’un réseau social s’est mis en place par la collaboration entre chercheurs à la publication d’articles conjoints. Dans le chapitre 5, nous avons voulu montrer que ce réseau social de collaboration s’est constitué en parallèle à un intérêt grandissant pour des thématiques communes de recherche.

Les sept chercheurs inclus dans ces analyses représentent ceux autour desquels s’est constituée la communauté épistémique de l’IRSPUM et ce que nous pouvons constater

avec les résultats présentés dans ce chapitre c'est qu'ils sont non seulement centraux, mais que leurs centres d'intérêt en recherche ont convergé au fil du temps pour s'orienter principalement vers les déterminants sociaux de la santé et ce malgré leur formations différentes.

Ce qui ressort des résultats présentés dans ce chapitre c'est la convergence croissante des intérêts de recherche de ces chercheurs vers la thématique des déterminants sociaux de la santé (Th19) et son arrimage aux côtés du thème plus ancien de la gestion des services de soins de santé (Th25). Ce phénomène n'est pas spécifique à ce sous-groupe de sept chercheurs puisque la centralité de ce thème croît tout au fil du temps, dans notre population totale.

CHAPITRE 6

DISCUSSION

Nous avons porté notre intérêt tout au cours de ce mémoire sur le développement de l'IRSPUM comme communauté épistémique, ici conceptualisée comme étant un réseau social s'organisant autour d'une thématique de recherche commune. Pour qu'émerge une communauté épistémique telle que celle de l'IRSPUM, il faut davantage que la mise sur pied d'un centre de recherche et la mise à disposition d'une infrastructure. Il faut, selon Roth et al. 2009, qu'au fil du temps des liens de collaboration se tissent entre les chercheurs et que, simultanément, ils interagissent autour d'une même idée, ce que Harrison White (2008) qualifie de *netdom*.

Le défi pour l'IRSPUM est de regrouper sous une même entité l'ensemble des chercheurs en santé publique de l'Université de Montréal, visant en cela à incarner la nouvelle santé publique où l'ensemble des grandes thématiques de santé publique sont présentes. Ces chercheurs proviennent de différentes entités de recherche, principalement le Groupe de recherche interdisciplinaire en santé (GRIS), le Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), le centre Léa Roback et l'axe de recherche en santé des populations du Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM). Ces chercheurs proviennent aussi de diverses disciplines et sont rattachés comme professeurs à diverses facultés. Le défi est donc pour l'IRSPUM de constituer une communauté épistémique à travers cette diversité pour favoriser de nouvelles synergies et l'émergence de nouvelles idées. Compte tenu de la création récente de l'IRSPUM (2009), notre recherche porte sur la communauté de chercheurs qui s'y regroupent au moment de sa création afin de prendre en compte le développement des collaborations et l'évolution des thématiques de recherche mais aussi, ce faisant, le chemin à parcourir pour que l'IRSPUM rencontre ses objectifs de regroupement interdisciplinaire et interfacultaire.

Notre objectif de départ était d'examiner dans une perspective historique, soit au cours des 20 années qui ont précédé la création de l'IRSPUM, l'évolution des liens en-

tre les chercheurs qui constituent l'IRSPUM et l'évolution des thématiques de recherche afin de saisir si l'IRSPUM forme une communauté épistémique au sens où les chercheurs la constituant forment un réseau de collaboration gravitant autour d'une problématisation commune de la santé publique ou si, au contraire, l'IRSPUM est formée d'un regroupement de chercheurs ne présentant pas la propriété d'être liés par des éléments sémantiques communs. Dans le cadre de cette étude, nous avons centré notre attention sur les publications des chercheurs et c'est à travers celles-ci que nous avons examiné les collaborations entre chercheurs et les thématiques traitées.

Pour réaliser cette recherche, nous avons constitué une base de données des articles publiés entre 1990 et 2009 par les 226 chercheurs réguliers et associés constituant l'IRSPUM en 2009, soit au total 3165 articles. Les thématiques de recherche ont été déterminées en utilisant la méthode des topics modeling (Chapitre 3). Nous avons découpé ce corpus de données en quatre périodes, soit 1990-1994, 1995-1999, 2000-2004 et 2005-2009 afin de suivre le développement des collaborations au fil du temps et de l'arrivée de nouveaux chercheurs. Soixante-cinq des 226 chercheurs de l'IRSPUM étaient déjà actifs en recherche au début des années '90, mais plus de la moitié des chercheurs de l'IRSPUM sont de jeunes chercheurs ayant moins de 10 ans de carrière.

6.1 Le développement d'un réseau de collaboration interdisciplinaire

Les liens entre les chercheurs de l'IRSPUM se sont constitués graduellement au cours de ces vingt ans. Au départ, durant la période 1990-1994, même si 65 chercheurs étaient déjà actifs, ils collaboraient peu (nos résultats ont montré quatre dyades) et, pour ceux qui le faisaient, privilégiaient les collaborations avec des chercheurs de leur discipline ou proche de celle-ci. Bien que les collaborations demeurent peu nombreuses, un embryon de réseau commence à prendre forme entre 1995 et 1999, formé de chercheurs des sciences humaines et sociales et de santé publique, autour duquel se construira graduellement le noyau fort de l'IRSPUM. Cet embryon est constitué de chercheurs alors regroupés dans le GRIS. Les deux dernières périodes sont particulièrement intéressantes puisque c'est au cours de la dernière décennie que la recherche en santé publique se

développe fortement à l'Université de Montréal. En effet, la possibilité de recruter des chercheurs boursiers financés par le Fonds de recherche en santé du Québec et la reprise des embauches de professeurs à la fin des années '90 amènent un fort contingent de nouveaux chercheurs en santé publique (environ 100 chercheurs). De nouvelles collaborations se tissent, les collaborations se multiplient, mais l'homogamie disciplinaire demeure dominante : les médecins collaborent entre eux et avec les épidémiologistes et les bio-statisticiens ; les spécialistes en sciences humaines et sociales collaborent entre eux et avec les spécialistes de santé publique ; les psychologues collaborent entre eux. Bref, les identités disciplinaires se mélangent peu, à l'exception des spécialistes de santé publique dont la formation interdisciplinaire favorise vraisemblablement la collaboration avec un spectre élargi de disciplines. Toutefois, même si les collaborations se développent et que l'on voit apparaître une masse liant plus du tiers des chercheurs, plusieurs petits regroupements disciplinairement homogames demeurent isolés de celles-ci.

Ce développement des collaborations se poursuit au cours de la dernière période, les liens s'intensifient et surtout se diversifient disciplinairement au sein du noyau central formé très largement par les chercheurs du GRIS. Les chercheurs provenant d'autres regroupements demeurent sans ou avec peu de liens avec ce noyau et forment des cliques beaucoup plus homogènes disciplinairement. La création de l'IRSPUM par le regroupement de plus petites entités autour du GRIS ne s'appuie donc pas sur des collaborations effectives entre les chercheurs de ces différentes entités. Par ailleurs, les frontières entre les composantes ne sont pas disciplinaires puisque les collaborations au sein de la composante principale se font entre des chercheurs appartenant à tous les horizons disciplinaires. En fait, les cloisonnements semblent davantage facultaires. Les chercheurs regroupés au sein d'entités mineures sans lien avec la composante principale se caractérisent par leur appartenance à des facultés ou départements non identifiés à la santé publique, qu'il s'agisse de la médecine vétérinaire ou des sciences sociales.

6.2 Vers une communauté épistémique

Pour vérifier l'hypothèse selon laquelle une communauté épistémique représente un réseau social qui se développe autour d'un thème de recherche commun et montrer l'importance de la prise en compte des contextes dans l'étude des relations sociales, nous avons, comme présenté au chapitre 5, analysé les interactions entre les chercheurs et les thèmes qu'ils manipulent dans leurs articles. Parallèlement à l'accroissement des collaborations entre les chercheurs, nous avons également observé une convergence des intérêts de recherches. Deux thèmes sont dominants dès la première période d'analyse (1990-1994), celui des déterminants sociaux de la santé et celui de la gestion des services de soins de santé, et ces thèmes deviennent de plus en plus centraux au fil du temps. Par ailleurs, ces thèmes se développent dans un univers de sens. Ainsi, lorsque l'on regarde la dispersion spatiale des chercheurs autour des thématiques de recherche pour la période d'émergence de l'IRSPUM, soit 2005-2009, ces deux thèmes apparaissent bien centraux, ce qui indique qu'un grand nombre de chercheurs les manipulent, mais un ensemble de thématiques moins centrales gravitent autour de ces thèmes. Ainsi, les chercheurs qui travaillent sur la thématique des services de santé travaillent aussi sur les thématiques de soins aux personnes âgées, de soins de santé de médecine préventive et d'évaluation des technologies de la santé, ces thèmes constituant un univers de sens. De même, les chercheurs travaillant sur les déterminants sociaux de la santé s'intéressent également aux comportements de santé et à la promotion de la santé. Il faut toutefois souligner la proximité de ces deux grands univers de sens, ce qui indique un décloisonnement de ces thématiques, plusieurs chercheurs travaillant à la fois sur l'une et sur l'autre.

En périphérie de ces deux grands pôles de recherche, quatre autres pôles de moindre importance en termes de volume de recherche se distinguent : un premier autour des maladies professionnelles, des maladies infectieuses et la santé environnementale ; un deuxième autour de l'immunologie, de la génétique et des zoonoses ; un troisième autour du développement des jeunes et adolescents ; et un quatrième tourne autour des médicaments, des autres troubles médicaux (autre que strictement chroniques, ou strictement infectieux) et des maladies chroniques. Deux thèmes se distinguent par leur très faible

attraction : le thème de la santé dentaire et celui de la santé internationale. Concernant cette dernière thématique, mentionnons d'abord notre surprise puisqu'il s'agit officiellement d'une des thématiques identifiée comme majeure pour l'IRSPUM. Dans les faits, cette thématique n'apparaît aucunement « majeure » dans l'ensemble des travaux réalisés par les chercheurs de l'IRSPUM. Néanmoins, le centre d'attraction sémantique apparaît bel et bien.

La distribution des chercheurs selon les thèmes qu'ils manipulent s'est transformée au fil du temps. Au fil des périodes prises en compte dans nos analyses, nous avons remarqué la place grandissante prise par le thème des déterminants sociaux de la santé, par celui de la gestion des services de soins de santé et par celui des comportements sociaux. Certains thèmes prennent un recul significatif dans l'importance qui leur est accordée. C'est notamment le cas du thème des maladies infectieuses, des zoonoses, des maladies professionnelles et de la santé environnementale.

Ces transformations laissent des traces au niveau de la collaboration entre chercheurs: parallèlement à la convergence des intérêts de recherche des chercheurs, la collaboration entre les chercheurs s'accroît. Dit en d'autres mots, d'un groupement d'individus non liés ou à peine liés entre eux, nous avons vu émerger un **réseau social** (des chercheurs liés entre eux par des collaborations conjointes à des articles), simultanément à une convergence des thématiques abordées par ceux-ci. La théorie nous rappelle que les gens ne se lient pas aléatoirement, mais parce qu'ils partagent des éléments sémantiques. Parmi ces éléments, les thèmes de la nouvelle santé publique, notamment celui des déterminants sociaux de la santé, constitue un pôle majeur qui occupe une place grandissante non seulement parce qu'il rallie de plus en plus de chercheurs mais aussi, comme l'indiquent les treillis de Galois, mais aussi parce que hiérarchiquement il chapeaute de plus en plus de thèmes.

Paradoxalement, la place grandissante prise par la thématique des déterminants sociaux de la santé se produit en parallèle avec le recul relatif des chercheurs en sciences humaines et sociales principalement au profit des chercheurs formés en santé publique (ou communautaire). Nous avons aussi pu observer une certaine pluridisciplinarité dans la manipulation des thèmes ; l'appartenance disciplinaire ne peut à elle seule expliquer

l'intérêt qu'ont les chercheurs pour les thèmes qu'ils manipulent ; ils sont de toutes appartenances disciplinaires à manipuler le thème des déterminants sociaux de la santé. Bref, les frontières disciplinaires s'estompent tout particulièrement en ce qui concerne ce thème.

Les thèmes de recherche que nous avons identifié grâce aux *topics modeling* nous ont permis d'étudier le rapport entre la collaboration scientifique et les univers sémantiques de recherche. L'objectif de ces analyses était d'obtenir de grandes thématiques de recherche afin de voir si les chercheurs convergent ou non vers une problématique commune. Il aurait été intéressant d'étudier plus en profondeur les relations entre les thématiques elles-mêmes pour saisir davantage les nuances dans l'usage des grandes thématiques et la constitution des grands univers de sens, mais ce n'était pas l'objet de ce présent mémoire. Nous avons donc utilisé nos résultats des analyses par *topics modeling* afin d'étudier les relations entre les chercheurs et les idées qu'ils manipulent dans le but d'approfondir les connaissances concernant la formation des communautés épistémiques. Ainsi, la cohésion sociale s'est intensifiée, mais nous n'avons pas observé un seul réseau social (où tous les membres sont liés dans une seule même grande composante), mais plutôt plusieurs composantes, dont une très large (dont le thème le plus rassembleur est celui des déterminants sociaux de la santé). À certains endroits cette large composante n'est reliée que par des ponts structuraux, c'est-à-dire par un seul lien, sans lequel la composante se scinderait en deux composantes distinctes. Ce type de lien concerne majoritairement des couples de chercheurs issus de disciplines différentes ; les ponts sont source d'interdisciplinarité.

Nous voulions savoir si l'IRSPUM représentait une communauté épistémique au sens où tous les membres convergent vers une problématique commune de recherche. La population de l'IRSPUM ne forme pas une communauté épistémique, au sens strict de la définition ici utilisée puisqu'en périphérie du thème centralisateur se retrouvent d'autres *netdoms*. Cependant, il existe au sein de l'IRSPUM une large communauté épistémique axée sur l'étude des déterminants sociaux de la santé et la gestion des services de santé. Celle-ci ne regroupe pas tous les membres de l'IRSPUM, mais elle regroupe une proportion importante des chercheurs. Ainsi, en étant moins stricte, nous remarquons que

l'IRSPUM (en tant que centre de recherche) s'oriente principalement sur les thématiques de la nouvelle santé publique. Ces thèmes n'ont cependant pas éclipsé ceux de l'ancienne santé publique que nous retrouvons notamment par le biais des ensembles de thèmes autour de la santé environnementale (maladies professionnelles, maladies infectieuses et santé environnementale) et de l'immunologie.

La forme des réseaux nous informe aussi sur des aspects dont traite la théorie de l'information : la redondance de l'information. Plus un réseau est dense, plus l'information qui y circule risque d'être redondante. La redondance n'est pas mauvaise en soi et assure en fait une meilleure diffusion de l'information. Cependant, pour favoriser l'innovation, trop de redondance peut nuire ; il faut trouver un certain équilibre. Il demeure complexe de tirer des conclusions concernant le cas de l'IRSPUM. Au centre se trouvent des cliques qui perdurent. Ces noyaux durs permettent une structuration de la communauté épistémique de l'IRSPUM, du point de vue tant sémantique que social. Si tous les membres de notre population d'analyse étaient aussi fortement liés entre-eux, il risquerait par contre d'y avoir trop de redondance, ce qui nuirait à l'innovation qui pourrait émerger de cette communauté épistémique. Il ne faudrait pas, pour le dire en des termes simples, que tous les chercheurs soient liés entre eux. À ce sujet, mentionnons que nous n'avons pu tenir compte des liens externes ; c'est-à-dire, des liens que les membres de notre population d'analyse ont avec tous les autres chercheurs, externes à notre population, qui sont eux aussi potentiellement sources d'innovation.

6.3 Limites de cette étude

Malgré toute notre minutie, notre étude comporte plusieurs limites qu'il ne faut surtout pas négliger. Dans un premier temps, la classification par *topics modeling*, comme toutes techniques d'analyses sémantiques automatiques, ne fournit pas une classification sans faille du contenu de notre corpus. De plus, la polysémie linguistique nous oblige à envisager les erreurs quant à la capacité des *topics* pour rendre compte des articles analysés. De plus, nous avons nous même apposé des étiquettes à ces *topics*, en fonction de ce qui nous semblait réunir tous ces articles sous un même *topics*. Néan-

moins, nous avons vérifié qualitativement et quantitativement la qualité du modèle que nous avons retenu pour nous assurer que ces lacunes ne nuisent pas de façon significative à nos analyses ultérieures. De plus, les analyses que nous avons effectuées révèlent la proximité de certains des thèmes que nous avons distingués. Des analyses plus fines seraient nécessaires pour déterminer dans quelles mesures ces thèmes se distinguent.

Nous avons choisi de ne prendre en compte dans notre étude qu'un seul type de publication, l'article, et qu'un seul type de lien, la co-publication. Les chercheurs diffusent leurs recherches par de nombreux autres médiums : les conférences, les livres, les rapports de recherches, pour ne nommer que ceux-là. Ils peuvent aussi être liés via d'autres types de liens, par exemple comme collègues ou amis ainsi que par la participation à des événements scientifiques ou autres, et ces liens sont aussi source de cohésion et d'innovation. Un portrait différent pourrait ressortir en prenant en compte davantage d'éléments.

La plus grande limite de cette étude réside dans le fait que nous n'avons pas su expliquer avec autant de force que souhaitée, le lien déterminant qui existe entre d'une part, les aspects sémantiques (les thèmes de recherche) et d'autre part, les éléments sociaux (les chercheurs). Comment les *netdom* apparaissent et en quoi permettent-ils l'émergence de communautés épistémiques ? Nous pouvons affirmer que simultanément à l'émergence d'un réseau social donné, un « champ » se développe et que ce champ attire vers lui des éléments qui se lient entre eux, pour former une communauté épistémique. Il nous faudrait réaliser d'autres analyses pour étayer notre compréhension de ce phénomène social. Malgré tout, rappelons que la théorie nous laisse des pistes à explorer : ce qui unit les gens entre eux se situe dans les contextes de leurs interactions. Dans le cas qui nous préoccupe, ce contexte est celui de la nouvelle santé publique.

6.4 Conclusion

En somme, la communauté épistémique observée au sein de l'IRSPUM s'oriente effectivement, selon nos analyses, vers les thèmes de la nouvelle santé publique. Elle s'y retrouve pleinement intégrée. La connaissance en santé publique produite au sein de

l'IRSPUM reflète le mouvement qu'à connu la connaissance en santé publique, c'est-à-dire qu'elle s'est orientée vers l'étude des déterminants sociaux de la santé. Mais ce changement ne s'est pas produit en balayant l'ancienne santé publique et la thématique de la gestion des services de soins de santé demeure une problématique centrale.

Il n'y a pas de rupture profonde ni de mutation ; la communauté épistémique de l'IRSPUM se retrouve à la fois encastrée dans le cadre plus large de la santé publique contemporaine et s'en découple en représentant une entité unique possédant sa propre marge de manoeuvre. L'IRSPUM est à la fois une institution sociale dont la raison d'être est de réguler la production et la diffusion de connaissances en santé publique. Elle représente aussi une culture épistémique. Pour en rendre davantage compte, il faudrait comparer le contenu sémantique de notre corpus avec celui d'autres communautés épistémiques en santé publique à travers le monde. Les problématiques de santé publique ont une saveur sociale que nous n'avons malheureusement pas pu prendre en compte dans notre étude.

Nous avons surtout voulu approfondir les connaissances concernant les liens entre le social et le sémantique ; c'est-à-dire, cette idée selon laquelle il faut s'intéresser aux relations qui existent non seulement entre les gens, mais aussi et surtout, au fait que ces relations sociales sont directement liées à un élément (sémantique) qui les rallie dans un même univers socio-sémantique. Dans le cas de la communauté épistémique que nous avons observée, des chercheurs se sont mis de plus en plus à collaborer conjointement à la rédaction d'articles. Parallèlement au développement de ce réseau social, nous notons une convergence des intérêts des chercheurs vers le thème des déterminants sociaux de la santé. Il faudrait de plus amples recherches pour démontrer comment ces deux phénomènes sont liés, mais nous supposons pour le moment que les communautés épistémiques émergent dans les « *netdoms* ».

La richesse des données que nous avons collectées pour réaliser ce mémoire n'est pas rendue à sa juste valeur. Nous avons tout de même visé à explorer au maximum les relations sociales et socio-sémantiques entre les chercheurs et les connaissances qu'ils ont produites au fil des 20 dernières années. Cette exploration a permis de mettre en lumière, sans toutefois l'expliquer, une co-évolution entre réseau social et réseau de sens. La prise en compte des contextes des interactions sociales semble l'élément essentiel de

la recherche en sciences sociales. Ainsi, la cohésion sociale s'est intensifiée, mais nous n'avons pas observé un seul réseau social (où tous les membres sont liés dans une seule même grande composante), mais plutôt plusieurs composantes, dont une très large (dont le thème le plus rassembleur est celui des déterminants sociaux de la santé). À certains endroits cette large composante n'est reliée que par des ponts structuraux ; c'est-à-dire que sans ce seul lien, la composante se scinderait en deux composantes distinctes. Ce type de lien concerne majoritairement des couples de chercheurs issus de disciplines différentes ; les ponts apportent une certaine interdisciplinarité.

BIBLIOGRAPHIE

- A. L. BARABASI, H. JEONG, Z. NÉDA, E. RAVASZ, A. SCHUBERT ET T. VICSEK. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3-4):590–614, 2002.
- ZYGMUNT BAUMAN, GILLIAN BENDELOW, MICHAEL BURY, MIKE FEATHERSTONE, NICHOLAS FOX, MIKE HEPWORTH, PAUL HIGGS, RICHARD LEVINSON, MIKE SAKS, ANNETTE SCAMBLER, GRAHAM SCAMBLER ET SIMON WILLIAMS. *Modernity, medicine and health : Medical sociology toward 2000*. Routledge, 1998.
- D. BEAVER ET R. ROSEN. Studies in scientific collaboration : Part i. the professional origins of scientific co-authorship. *Scientometrics*, 1(1):65–84, 1978.
- URLICK BECK. *La société du risque : Sur la voie d'une autre modernité*. Flammarion, 2008.
- DAVID BLEI ET J. LAFFERTY. *Text Mining: Theory and Applications*, chapitre 1. Topic Models, pages 1–24. Taylor and Francis, 2009. URL <http://www.cs.princeton.edu/~blei/papers/BleiLafferty2009.pdf>.
- DAVID M. BLEI, Y. NG ANDREW ET I. JORDAN MICHAEL. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.
- ULRIK BRANDES ET DOROTHEA WAGNER. *Graph Drawing Software*, chapitre Visone - Analysis and Visualization of Social Networks., pages 321–340. Springer-Verlag, 2004. URL <http://www.inf.uni-konstanz.de/algo/publications/bw-vavsn-04.pdf>.
- WRAY BUNTINE. Estimating likelihoods for topic models. *Advances in Machine Learning*, 5828:51–64, 2009. Lecture Notes in Computer Science.
- WRAY BUNTINE ET ALEKS JAKULIN. Discrete component analysis. *Lecture notes in computer science*, 3940:1–33, 2006.

- ROBIN BUNTON ET GORDON MACDONALD. *Health Promotion: Disciplines and Diversity*. Routledge, 1992.
- LISE CARDINAL, DANIÈLE FRANCOEUR, MARTHE HAMEL ANS SYLVIE KIROUAC ET YVES THÉBERGE. Programme national de santé publique 2003-2013. Rapport technique, Ministère de la santé et des services sociaux, Gouvernement du Québec, 2003.
- KARIN KNORR CETINA. *Epistemic culture : How the sciences make knowledge*. Harvard university press, London, England, 2000.
- ALAIN DEGENNE ET MICHEL FORSÉ. *Les réseaux sociaux*. A. Colin, Paris, 2ème édition, 2004.
- JEAN-PIERRE DOZON ET DIDIER FASSIN. *Critique de la santé publique. Une approche anthropologique*. Coll. Voix et Regards. Éditions Balland, 2001.
- NORBERT ELIAS. *La société de cour*. Champs. Flammarion, 1985.
- ROBERT G. EVANS, MORRIS L. BARER ET THEODORE R. MARMOR. *Why Are Some People Healthy and Others Not ?*, chapitre Producing Health, Consuming Health Care, pages ??-?? Hawthorne, New York, 2004.
- DIDIER FASSIN. Entre politiques du vivant et politiques de la vie. pour une anthropologie de la santé. *Anthropologie et Sociétés*, 24(1):95–116, 2000.
- MICHEL FOUCAULT. *Dits et écrits 2 : 1976-1988*. Gallimard, 2001.
- LINTON FREEMAN. La resurreccion des cliques : application du treillis de galois. *Bulletin de Méthodologie Sociologique*, 37(1):3–24, 1992.
- ANTHONY GIDDENS. *The Consequences of Modernity*. Stanford University Press, 1990.
- DENIS GOULET. Entre les miasmes et les germes. l'impact de la bactériologie sur la pratique médicale en territoire canadien 1870-1930. *Manguinhos*, 1(2):20–38, 1994.

THOMAS L. GRIFFITHS ET MARK STEYVERS. Finding scientific topics. *PNAS*, 101 (1):5228–5235, 2004.

MICHEL GROSSETTI ET FRÉDÉRIC GODART. Harrisson white : des réseaux sociaux à une théorie structurale de l'action. *SociologieS*, ND:1–15, 2007. URL <http://sociologies.revues.org/index233.html>. [En ligne], Découvertes / Redécouvertes, Harrison White, mis en ligne le 17 octobre 2007, Consulté le 13 mai 2010.

JABREF DEVELOPMENT TEAM. *JabRef Version 2.6*. JabRef Development Team, 2009. URL <http://jabref.sourceforge.net>. GNU licence.

OTHMAR KEEL ET PETER KEATING. *Santé et société au Québec : XIXe - XXe siècle*. Eds. Peter Keating and Othmar Keel, chapitre Autour du Journal de médecine de Québec/Quebec Medical Journal (1826-1827): programme scientifique et programme de médicalisation., pages 35–59. Boréal, Montréal, 1995.

STEPHEN J. KUNITZ. *The health of populations: general theories and particular realities*. Oxford University Press, 2007.

MARC LALONDE. Nouvelle perspective de la santé des canadiens : un document de travail. Rapport technique, Ministre des Approvisionnements et Services Canada, 1981. URL http://www.hc-sc.gc.ca/hcs-sss/alt_formats/hpb-dgps/pdf/pubs/1974-lalonde/lalonde-fra.pdf. Cat : H31-1374.

MARIE-CLAIRE LAURENDEAU, MARTHE HAMEL, CHRISTINE COLIN ET MARIE-JEANNE DISANT. Bilan de la recherche en santé publique au québec (1999-2004). Rapport technique, Institut national de santé publique du Québec, 2007. URL <http://www.inspq.qc.ca/pdf/publications/625-BilanRechSantePubl1999-2004.pdf>.

MICHAEL MARMOT, SHARON FRIET, RUTH BELL, TANJA AJ HOUWELING ET SEBASTIAN TAYLOR. Comblent le fossé en une génération : Instaurer l'équité en santé

en agissant sur les déterminants sociaux. Rapport technique, Organisation mondiale de la Santé (OMS), 2009.

SIR MICHAEL G. MARMOT. Social differentials in health within and between populations. *Daedalus*, 123(4):197–216, 1994a. ISSN 00115266. URL <http://www.jstor.org/stable/20027276>.

SIR MICHAEL G. MARMOT. Work and other factors influencing coronary health and sickness absence. *Work and Stress*, 8(2):191–201, 1994b.

DAVID MCQUEEN, ILONA KILCKBUSCH, LOUISE POTVIN, JÜRGEN M. PELINKAN, LAURA BALBO ET THOMAS ABEL. *Health and modernity : The role of theory in health promotion*. Springer, 2007.

CHARLES F MERBS. Paleopathology: Diseases and injuries of prehistoric man. *American Anthropologist*, 73(6):1425–1426, 1971. ISSN 1548-1433. URL <http://dx.doi.org/10.1525/aa.1971.73.6.02a01040>.

MSSS. Programme national de santé publique 2003-2012. Rapport technique, Gouvernement du Québec (Ministère de la santé et des services sociaux), 2003. Publication no : 02-216-01F.

NICHOLAS C. MULLINS. Développement des disciplines scientifiques : origines internes et externes du changement. *Sociologie et sociétés*, 7(1):133–142, 1975.

M. E. J. NEWMAN. The structure of scientific collaboration networks. *PNAS*, 98(2): 404–09, 2001.

ABDEL R. OMRAN. The epidemiologic transition : A theory of the epidemiology of population change. *Milbank Quarterly*, 83(4):731 ?–757, 2005.

OMS. Charte d'ottawa pour la promotion de la santé. Rapport technique, OMS, Ottawa, 1986.

- M. O'NEILL ET A. PEDERSON. *Health Promotion in Canada*, chapitre Two analytic paths for understanding Canadian developments in health promotion, pages 40–55. W.B. Saunders, Toronto, 1994.
- MICHEL O'NEILL, SOPHIE DUPÉRÉ, ANN PEDERSON ET IRVING ROOTMAN. *Promotion de la santé au Canada et au Québec, perspectives critiques*. Collection sociétés, cultures et santé. PUL, Québec, Canada, 2006a.
- MICHEL O'NEILL, ANN PEDERSON, SOPHIE DUPÉRÉ ET IRVING ROOTMAN. *Promotion de la santé au Canada et au Québec, perspectives critiques.*, chapitre La promotion de la santé au Canada et à l'étranger : bilan et perspectives, pages ??–?? PUL, Québec, 2006b.
- B POLAND, D COBURN, A ROBERTSON ET J EAKIN. Wealth, equity and health care: a critique of a "population health" perspective on the determinants of health. critical social science group. *Social science medicine*, 46(7):785–798, 1998.
- JEANINE POMMIER ET OLIVIER GRIMAUD. Les fonctions essentielles de santé publique : histoire, définition et applications possibles. *Santé Publique*, 19(hs):9–14, 2007.
- TERRY PRATCHETT. *Hogfather*. Victor Gollancz, 1996.
- CAMILLE ROTH. Co-evolution in epistemic networks – reconstructing social complex systems. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*, 1(3):–art2, 2006. URL <http://repositories.cdlib.org/imbs/socdyn/sdeas/vol1/iss3/art2>.
- CAMILLE ROTH. Co-évolution des auteurs et des concepts dans les réseaux épistémiques: le cas de la communauté zebrafish. *Revue Française de Sociologie*, 48(2):333–367, 2008.
- CAMILLE ROTH, JEAN-PHILIPPE COINTET, SERGEI OBIEDKOV ET NIKITA ROMASHK. Analyse textuelle des motions du congrès de reims du ps. 2:1–3, 2009.

- GEORG SIMMEL. *Sociologie : Études sur les formes de la socialisation*. PUF, Paris, France, 1ère édition. quadrigé : grands textes édition, 2010 (1908).
- CHARLES TILLY. New school for social research. *Contemporary Sociology*, 22(3): 307–309, 1993.
- KARL M. VAN METER ET MATHILDE DE SAINT LEGER. German and french contemporary sociology compared: Text analysis of congress abstracts. *Bulletin de Méthodologie Sociologique*, 104(1):5–31, 2009.
- STANLEY WASSERMAN ET KATHERINE FAUST. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
- MAX WEBER. *L'Éthique protestante et l'esprit du capitalisme*. Gallimard, 2003.
- HARRISON C. WHITE. *Identity and Control: How Social Formations Emerge*. Princeton University Press, second edition édition, 2008.
- J. S. WITTE, G. URSIN, J. SIEMIATYCKI, W. D. THOMPSON, A. PAGININI-HILL ET R. W. HAILE. Diet and premenopausal bilateral breast cancer: a case-control study. *Breast Cancer Res Treat*, 42(3):243–251, Feb 1997.
- WORKING GROUP ON INEQUALITIES IN HEALTH. Black report: Inequalities in health. Rapport technique, D.H.S.S., London, 1980. URL <http://www.sochealth.co.uk/history/black.htm>. [Reproduit sur le site de la Socialist Health Association].

Annexe I

Liste des chercheurs

Tableau I.I: Liste des chercheurs et des disciplines en date du
1^{er} septembre 2009

ID	Chercheur	Statut		RTM					Code	Formation	
		R	A	1	2	3	4	5			
A1	ACHILLE, Marie		x						x	5	Ph.D Psychologie clinique et santé, philosophie
A2	ADRIEN, Alix		x		x					3	MD et MSc Epidémiologie
A3	ALBERT, Lucien		x					x		3	MD
A4	ARCHAM- BAULT, Isabelle	x		x						5	Ph.D Psychologie
A5	ARCHAM- BBAULT, Marie		x		x					6	Ph.D Microbiologie et immunologie
A6	AUBÉ, Denise		x						x	3	MD, MSc Santé communautaire
A7	AUGER, Nathalie		x		x					3	MD, MSc Épidémiologie et biostatistiques
A8	AVARD, Denise	x							x	2	Ph.D Sociologie

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A9	AYOTTE, Pierre		x		x				6	Ph.D Pharmacologie
A10	BARIL, Marc		x		x				6	Ph.D Pharmacologie et toxicologie
A11	BARNETT, Tracie		x	x					4	Ph.D Épidémiologie et biostatistiques
A12	BASTIEN, Robert	x				x			2	Ph.D Sciences de l'éducation
A13	BATTISTA, Renaldo	x				x			3	MD, Gestion et politiques de santé
A14	BEAULIEU, Marie-D.		x					x	3	MD, MSc Épidémiologie
A15	BEDOS, Christophe		x	x					1	Ph.D. Santé publique
A16	BÉKAL, Sad- jia		x		x				3	Ph.D. Sciences de l'alimentation
A17	BÉLAND, François	x			x				2	Ph.D Sociologie
A18	BÉLANGER, Denise		x		x				4	Ph.D Épidémiologie et biostatistiques
A19	BÉLANGER, Jean	x		x					5	Ph.D Psychologie de l'éducation
A20	BÉRARD, Annick	x		x					4	Ph.D Épidémiologie et biostatistiques
A21	BERGERON, Pierre		x			x			2	Ph.D Sciences politiques
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A22	BERGMAN, Howard		x					x	3	MD, Gériatrie
A23	BERNARD, Paul	x		x					2	Ph.D. Sociologie
A24	BERNIER, Nicole F.	x				x			2	Ph.D Sciences poli- tiques
A25	BERTRAND, Lise		x	x					1	M.Sc Santé publique
A26	BIGNAMI, Simona		x				x		2	Ph.D Démographie
A27	BIGRAS- POULIN, Michel		x	x					4	Ph.D Épidémiologie
A28	BILODEAU, Angèle	x				x			2	Ph.D Sciences hu- maines appliquées
A29	BLAIS, Lucie	x			x				4	Ph.D Épidémiologie et biostatistiques
A30	BLAIS, Régis	x						x	5	Ph.D Psychologie
A31	BOUCHARD, Michèle	x			x				1	Ph.D Santé publique
A32	BOULIANNE, Martine		x		x				7	Ph.D Médecine vétéri- naire
A33	BOURGAULT, Anne-Marie		x		x				3	MD Microbiologie médicale
A34	BOURQUE, Mélanie		x			x			2	Ph.D Sciences poli- tiques

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A35	BOWEN, François	x		x					5	Ph.D Psychologie
A36	BRASSARD, Paul		x			x			4	MD, MSc Épidémiologie
A37	BRAULT, Is- abelle		x					x	1	Ph.D Santé publique
A38	BRODEUR, Jean-Marc	x				x			1	Ph.D Santé communautaire
A39	BROUSSELLE, Astrid	x				x			1	Ph.D Santé publique
A40	BRUNEAU, Julie		x			x			3	MD, M.Sc Épidémiologie et biostatistiques
A41	CANTIN, Stéphane	x		x					5	Ph.D Psychologie
A42	CARON, Jean		x			x			5	Ph.D Psychologie
A43	CARTIER, Sylvie	x				x			5	Ph.D Psychopédagogie
A44	CAUX, Chan- tal		x	x					1	Ph.D Santé publique
A45	CAZALE, Linda		x			x			1	Ph.D Santé publique
A46	CHAILLET, Nils	x						x	2	Ph.D. Anthropologie
A47	CHAMPAGNE, François	x				x			1	Ph.D Santé communautaire
A48	CHOINIÈRE, Manon		x					x	5	Ph.D Psychologie

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A49	CHOINIÈRE, Robert		x		x				2	MSc Démographie
A50	CHOUINARD, Roch	x				x			5	Ph.D Psychologie
A51	CHUNG, Ryoa		x			x			2	Ph.D Philosophie
A52	CLAPPERTON Irma		x					x	1	M.Sc Santé commu- nautaire
A53	COLIN, Christine		x	x					3	MD, MSc Santé com- munautaire
A54	COLLIN, Jo- hanne	x						x	2	Ph.D Histoire
A55	CONTANDRIO- POULOS, AP	x						x	2	Ph.D Économie
A56	CONTANDRIO POULOS, D	x				x			1	Ph.D Santé publique
A57	CÔTÉ, José		x					x	3	Ph.D Sciences infir- mières
A58	DAGENAIS, Christian	x				x			5	Ph.D Psychologie com- munautaire
A59	D'AMOUR, Danielle	x						x	1	Ph.D Santé publique
A60	DANEAULT, Serge		x					x	8	Ph.D Sciences environ- nementales
A61	DANIEL, Mark		x		x				4	Ph.D Épidémiologie et soins de santé
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A62	DASCAL, André		x		x				3	MD
A63	DASSA, Clément	x		x					5	Ph.D Psychométrie
A64	DEDOBBE- LEER, Nicole	x						x	1	DSc Santé publique
A65	DEL CASTILLO, Jérôme		x		x				6	Ph.D Pharmacologie
A66	DELISLE, Hélène	x						x	3	MSc Nutrition, Ph.D Sc. cliniques
A67	DEMERS, Andrée	x		x					2	Ph.D Économie des ressources humaines
A68	DENIS, Jean- Louis	x						x	1	Ph.D Santé communau- taire
A69	DESBIENS, Nadia	x		x					5	Ph.D Psychopédagogie
A70	DESCHENES, Marthe	x		x					1	Ph.D Santé publique / Promotion de la santé
A71	DION, Chan- tal		x		x				6	Ph.D Chimie inorgani- que
A72	DROUIN, Louis		x	x					3	MD, MSc Santé pu- blique
A73	DUBOIS, Carl-Ardy	x						x	1	Ph.D Santé publique
A74	DUBREUIL, Daniel	x			x				6	Ph.D Microbiologie et Immunologie

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A75	DUFOUR, Sarah	x		x					5	Ph.D Psychologie communautaire
A76	DUFRESNE, André		x	x					1	Ph.D Santé travail
A77	DUMONT, Alexandre	x					x		3	Ph.D Études médicales
A78	DURAND, Pierre	x		x					4	Ph.D Épidémiologie et biostatistiques santé et sécurité au travail
A79	FAIRBROTHER, John Morris	x			x				6	Ph.D Microbiologie
A80	FALLU, Jean-Sébastien	x			x				5	Ph.D Psychologie
A81	FARAND, Lambert	x						x	5	MD, Ph.D Sciences cognitives
A82	FELDMAN, Debbie	x						x	4	MD, Ph.D Epidémiologie
A83	FIRBANK, Oscar		x					x	2	Ph.D Sociologie
A84	FLEURY, Marie-Josée		x					x	1	Ph.D. Santé publique
A85	FOREST, Pierre-Gerlier		x			x			2	Ph.D Histoire et sociopolitiques des sciences
A86	FOURNIER, Louise	x				x			1	Ph.D. Santé publique

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A87	FOURNIER, Michel		x		x				2	MSc Éducation, mesure et évaluation
A88	FOURNIER, Pierre		x				x		3	MD Santé tropicale, MSc Santé communau- taire
A89	FRIGAULT, Louis-Robert		x		x				2	Ph.D Anthropologie
A90	FROHLICH, Katherine	x							1	Ph.D Santé publique
A91	GAUVIN, Lise		x		x				5	Ph.D Science du com- portement
A92	GENDRON, Sylvie	x				x			1	Ph.D Santé publique
A93	GÉRIN, Michel	x			x				6	Ph.D Chimie
A94	GIRARD, Vincent	x			x				6	Ph.D Biochimie
A95	GODARD, Béatrice	x				x			2	Ph.D Sociologie
A96	GOULET, Lise	x		x					4	MD, Ph.D Épidémiolo- gie
A97	HADDAD, Slim	x					x		3	Ph.D Économie, MD santé communautaire
A98	HAGGERTY, Jeannie	x						x	2	Ph.D. Philosophie
A99	HAINES, Vic- tor		x	x					2	Ph.D Relations indus- trielles
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A100	HAMEL, Pierre J.		x	x					2	Ph.D Économie et sociologie du travail
A101	HAREL, Josée	x			x				6	Ph.D Microbiologie et Immunologie
A102	HARTZ, Zulmira		x			x			1	Ph.D Santé communautaire
A103	HARVEY, Steve	x		x					2	Ph.D Philosophy
A104	HATEM, Marie		x					x	1	Ph.D Santé publique
A105	HÉBERT-CROTEAU, Nicole		x					x	1	Ph.D Santé communautaire
A106	JANOSZ, Michel	x		x					2	Ph.D Sciences humaines appliquées
A107	JENSON, Jane		x			x			2	Ph.D Sciences sociales
A108	JOHRI, Mira	x					x		2	Ph.D Philosophie, MPH Épidémiologie et santé publique
A109	JOUBERT, Pierre		x			x			5	Ph.D Psychologie sociale
A110	KAIROUZ, Sylvia		x	x					5	Ph.D Psychologie sociale
A111	KARP, Igor		x		x				4	Ph.D Épidémiologie
A112	KEMPENEERS Marianne		x	x					2	Ph.D Démographie

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A113	KESTENS, Yan		x		x				2	Ph.D Aménagement
A114	KISHCHUK, Natalie		x			x			5	Ph.D Psychologie sociale
A115	KNOPPERS, Bartha Maria	x				x			2	Ph.D Droit
A116	KRISHNAN, Kannan	x			x				1	Ph.D Santé communautaire
A117	KUATE- DEFO, Barthelemy	x		x					2	Ph.D Démographie et épidémiologie
A118	LABERGE, Suzanne	x				x			2	Ph.D Anthropologie
A119	LABRÈCHE, France		x		x				4	Ph.D Épidémiologie et biostatistiques
A120	LACHAINE, Jean		x					x	2	Ph.D Pharmacoeconomie
A121	LAFONTAINE, Céline	x						x	2	Ph.D Sociologie
A122	LAFORET, Sophie	x		x					4	Ph.D épidémiologie et biostatistiques
A123	LALONDE, Lyne	x						x	4	Ph.D épidémiologie et biostatistiques et santé au travail
A124	LAMARCHE, Paul	x						x	1	Ph.D Organisation des soins de santé
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A125	LAMBERT, Gilles	x			x				3	MD
A126	LAMBERT, Marie		x	x					3	MD
A127	LAMOTHE, Lise	x						x	2	Ph.D Théorie des organisations
A128	LANGLEY, Ann		x					x	9	Ph.D Administration
A129	LARDOUX, Solène		x	x					2	Ph.D Démographie
A130	LAROCQUE, Richard		x	x					3	MOA Audiologie
A131	LAURENDEAU, Marie-Claire		x			x			5	Ph.D Psychologie communautaire
A132	LAURIER, Claudine		x					x	1	Ph.D Santé communautaire
A133	LAURIN, Is- abelle		x	x					5	Ph.D Psychologie
A134	LAVOUÉ, Jérôme		x		x				6	DEA Chimie analytique
A135	LECLERC, Pascale		x		x				4	MSc Épidémiologie
A136	LEDUC, Nicole		x					x	1	Ph.D Santé communautaire
A137	LEFFRONDRÉ, Karen		x		x				4	Ph.D Biostatistiques

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A138	LE GRAND, Thomas		x				x		2	Ph.D Démographie
A139	LEHOUX, Pascale	x				x			1	Ph.D Santé publique
A140	LEIBING, Annette	x					x		2	Ph.D Anthropologie
A141	LE LORIER, Jacques		x		x				6	Ph.D Pharmacologie et biostatistiques
A142	LEMAY, Anne	x							1	Ph.D Santé communau- taire
A143	LEMIÈRE, Catherine		x		x				3	MD, MSc Sciences bio- médicales
A144	LEMIRE, Marc	x				x			2	Ph.D Sciences poli- tiques
A145	LETELLIER, Ann	x			x				6	Ph.D Sciences micro- biologie et immuno- logie
A146	LEVALLOIS, Patrick		x		x				3	MD Santé communau- taire
A147	LÉVESQUE, Jean-Frédéric	x						x	1	Ph.D Santé publique
A148	MARCHAND, Alain	x		x					2	Ph.D Sociologie
A149	MASSÉ, Richard		x			x			3	MD
A150	MAYRAND, Marie-Hélène		x					x	3	MD

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A151	MIAUX, Sylvie		x			x			2	Ph.D Géographie
A152	MICHEL, Pascal		x		x				4	Ph.D Épidémiologie
A153	MONNAIS,- Laurence	x						x	2	Ph.D Histoire
A154	MONTPETIT, Éric	x				x			2	Ph.D Science po- litiques
A155	MORENCY, Patrick		x		x				1	Candidat Ph.D Santé publique
A156	MORIZOT, Julien	x		x					5	Ph.D Psychologie
A157	NGUYEN, Vinh-Kim	x					x		2	Ph.D Anthropologie
A158	NOËL, Alain	x				x			2	Ph.D Études interna- tionales
A159	OGDEN, Nicholas	x			x				7	Ph.D Zoologie
A160	O'LOUGHLIN, Jennifer		x		x				4	Ph.D épidémiologie et biostatistiques
A161	OTERO, Marcelo	x						x	2	Ph.D Sociologie
A162	PAGANI, Linda	x				x			5	Ph.D Psychologie de l'éducation
A163	PAPINEAU, - Élizabeth		x	x					2	Ph.D Anthropologie
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation	
		R	A	1	2	3	4	5			
A164	PAQUETTE, Marie-Claude		x	x						3	Ph.D Community nutrition
A165	PAQUIN, Sophie		x	x						2	Ph.D Études urbaines
A166	PARADIS, Gilles		x		x					3	MD, MSc Épidémiologie et biostatistiques
A167	PARENT, Marie-Élise		x		x					3	M.Sc Nutrition
A168	PARENT, Sophie	x		x						5	Ph.D Psychologie du développement humain
A169	PERREAULT, Sylvie	x				x				6	Ph.D Pharmacologie
A170	PERRON, Stéphane		x	x						1	MSc Santé communautaire
A171	PINEAULT, - Raynald	x								3	MD, Ph.D Organisation des soins de santé
A172	POIRIER, Léo-Roch	x						x		9	M.Sc.Administration des services de santé
A173	POISSANT, Julie		x	x						5	Ph.D Psychologie communautaire
A174	POMEY, Marie-Pascale	x						x		1	Ph.D Santé publique
A175	POTVIN, Louise	x				x				1	Ph.D Santé communautaire
A176	POULIN, Carole		x	x						1	MSc Santé communautaire

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A177	PROULX, Renée		x			x			1	Ph.D Psychologie communautaire
A178	QUESNEL- VALLÉE, Amélie		x			x			2	Ph.D Sociologie
A179	QUESSY, Sylvain		x		x				6	Ph.D Microbiologie
A180	RAVEL, André		x		x				3	Ph.D Sciences biomédicales
A181	RAYNAULT, Marie-France		x			x			4	MD, M.Sc Épidémiologie
A182	RECEVEUR, Olivier		x		x				3	Ph.D Nutrition
A183	RENAUD, Lise		x			x			1	Ph.D Éducation la santé
A184	RICHARD, Lucie	x		x					1	Ph.D Santé communautaire
A185	RIDDE, Valéry		x			x			1	Ph.D Santé communautaire
A186	RIVARD, Michèle	x				x			4	ScD Biostatistique
A187	RIVEST, Paul		x		x				3	MD, MSc Santé communautaire
A188	ROBERGE, Danièle		x					x	1	Ph.D Santé communautaire
A189	ROBITAILLE, Yvonne	x			x				4	Ph.D Épidémiologie
... Suite la prochaine page										

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A190	ROCHON, Jean		x			x			3	MD, Ph.D Santé publique
A191	ROCK, Melanie		x	x					2	Ph.D Anthropologie
A192	RODRIGUEZ, Charo		x					x	3	MD, Ph.D Santé publique
A193	ROSSIGNOL, Michel		x		x				3	MD, MSc Épidémiologie
A194	ROUSSEAU, Louise		x			x			1	Ph.D Santé publique
A195	ROUSSEAU, Marie-Claude		x		x				4	Ph.D. Epidémiologie, biostatistiques
A196	ROUSSEAU, Vincent		x	x					5	Ph.D Psychologie industrielle et organisationnelle
A197	ROY, Denis A.		x					x	3	MD, MSc Épidémiologie, MSc santé pub.
A198	ROY, Élise		x			x			3	MD, MSc Santé communautaire
A199	ROY, Jean-Pierre		x		x				3	MD
A200	SAINT-MARTIN, Denis		x			x			2	Ph.D Science politiques
A201	SÉGUIN, Louise	x		x					3	MD

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation	
		R	A	1	2	3	4	5			
A202	SÉNÉCAL, Gilles		x	x						2	Ph.D Géographie
A203	SICOTTE, Claude	x							x	1	Ph.D Santé communautaire
A204	SIEMIATYCKI Jack	x			x					4	Ph.D. Épidémiologie et biostatistiques
A205	SMARGIASSI, Audrey	x			x					8	Ph.D Sciences de l'environnement
A206	SOTO, Julio		x	x						1	Ph.D Santé communautaire
A207	SOUBHI, Hassan	x							x	1	Ph.D Santé publique
A208	STOCK, Suzanne		x	x						4	MSc Épidémiologie clinique
A209	STRYCHAR, Irene		x		x					2	Ph.D Éducation
A210	TARDIF, Robert	x			x					1	Ph.D Santé communautaire
A211	TOUATI, Nassera	x							x	9	Ph.D Gestion
A212	TOURIGNY, André		x						x	9	MBA Administration
A213	TOUSIGNANT, Pierre	x							x	3	MD, MSc Épidémiologie et santé
A214	TREMBLAY, Michèle		x			x				3	MD

... Suite la prochaine page

ID	Chercheur	Statut		RTM					Code	Formation
		R	A	1	2	3	4	5		
A215	TURGEON, Lyse	x			x				5	Ph.D Psychologie clinique
A216	UGLAND, Trygve		x			x			2	Ph.D Politiques
A217	VAILLAN- COURT, Jean-Pierre		x		x				7	Ph.D Médecine vétérinaire
A218	VALLIÈRES, André		x		x				7	Ph.D Médecine vétérinaire
A219	VIAU, Claude	x			x				3	M.Sc Chimie
A220	VISSANDJEE, Bilkis		x				x		1	Ph.D Santé internationale
A221	VYSKOCIL, Adolf	x			x				6	Ph.D Biochimie
A222	WEINSTOCK, Daniel		x			x			2	Ph.D Philosophie
A223	WHITE, Deena	x				x			2	Ph.D Sociologie
A224	WILLIAMS- JONES, Bryn	x				x			1	Ph.D Études interdisciplinaires
A225	ZAYED, Joseph	x			x				1	Ph.D Santé communautaire
A226	ZUNZUNE- GUI, Maria	x					x		4	Ph.D Épidémiologie

Légendes:

Statut:

- R Chercheur régulier
- A Chercheur associé

RTM = « Regroupements thématiques majeurs » de l'IRSPUM:

- 1 Déterminants sociaux de la santé
- 2 Environnement et santé
- 3 Connaissances à l'action
- 4 Santé mondiale
- 5 Système de santé

Code des disciplines:

No. Disciplines

- 1 Santé publique et communautaire
- 2 Sciences sociales
- 3 Médecine et sciences biomédicales
- 4 Épidémiologie et biostatistiques
- 5 Psychologie
- 6 Biologie et chimie
- 7 Médecine vétérinaire et zoologie
- 8 Sciences environnementales
- 9 Administration et gestion

Annexe II

Liste des *topics*

Voici la liste des 150 *topics*:

ID **Étiquettes des *topics***

- T0 Fertilité et troubles de la reproduction (3)
- T1 Ostéoporose (16)
- T2 Infections par les tiques (18)
- T3 leucocytes (18)
- T4 Soins en santé mentale (61)
- T5 Promotion de l'exercice physique et d'une alimentation saine (39)
- T6 VIH/sida (54)
- T7 Expression génétique (26)
- T8 Système de régulation du phosphate chez Escherichia Coli (5)
- T9 Pratiques en matière de vaccination (4)
- T10 NI (0)
- T11 Effets de l'Aspirine (14)
- T12 Zoonoses liées à l'élevage (16)
- T13 Peur conditionnée (3)
- T14 Expression des cellules (36)
- T15 Troubles de l'alimentation (24)
- T16 Mitochondropathies (21)
- T17 NI (0)
- T18 Effets des antioxydants, acides gras et vitamines (10)
- T19 Fonctions thyroïdiennes (11)
- T20 Exposition à des contaminants dans l'environnement(94)
- T21 Flux d'azote dans les porcheries (10)
- T22 Ixodes scapularis (23)
- T23 Hormones parathyroïdiennes (8)

- T24 Cliniques d'anticoagulant (2)
- T25 Aide humanitaire (5)
- T26 Lignes directrices cliniques*** epidemio (110)
- T27 Escherichia Coli entérotoxinogènes du porc (11)
- T28 Suicide (19)
- T29 Pédiatrie et maladies infantiles (74)
- T30 Activité physique (65)
- T31 NI (0)
- T32 NI (0)
- T33 Évaluation des technologies de la santé(37)
- T34 Composées hépatotoxiques (6)
- T35 Pleuropneumonies (animaux) (17)
- T36 Maternité mortalité et pathologies infantiles (99)
- T37 Démence (56)
- T38 Coronavirus du dindon (14)
- T39 Réadaptation psychologique et physique (20)
- T40 Usagers de drogues illicites et « joueurs » (20)
- T41 Prise en charge des patients (223)
- T42 Cancers occupationnels (travail, maison, école) (138)
- T43 Excipient pharmaceutique (2)
- T44 Aedes aegypti (2)
- T45 Séquençage génétique (77)
- T46 Services de santé pour les personnes âgées fragiles (87)
- T47 Facteurs résidentiels (19)
- T48 Résistance aux bactéroïdes (34)
- T49 NI***** (0)
- T50 Santé et soins de santé dans les pays en développement (28)
- T51 Problèmes de comportement (6)
- T52 Techniques d'imagerie médicale (3)
- T53 Intoxication des espèces aquatiques (4)

- T54 Virus du papillome humain (35)
- T55 Maturité dentaire (20)
- T56 Nématodes parasites des plantes (10)
- T57 Conflits d'intérêts en sciences biologiques (1)
- T58 NI (0)
- T59 Politiques publiques (135)
- T60 Marquage des protéines (5)
- T61 Études sur les jeunes (121)
- T62 Dépression (16)
- T63 Risques liés aux thérapies médicales (8)
- T64 Régimes et habitudes alimentaires (114)
- T65 Développement des jeunes enfants (2)
- T66 Estimation des coûts et de l'efficacité des thérapies (59)
- T67 Déterminants sociaux de la santé (314)
- T68 Alzheimer (13)
- T69 ****(0)
- T70 Éthique en recherche génétique (77)
- T71 Silicose (20)
- T72 Campagnes médiatiques en santé publique*** (0)
- T73 Accouchements et mortalité maternelle (50)
- T74 Contaminants dans le lait(4)
- T75 Exposition au pyrène (43)
- T76 Pratique médicale (252)
- T77 Donneurs de soins aux personnes âgées (76)
- T78 Expositions professionnelles (79)
- T79 NI (0)
- T80 Souches d'Escherichia Coli de type attachant et effaçant (34)
- T81 Transport intestinal (10)
- T82 Programmes et promotion de la cessation tabagique (11)
- T83 Pollution industrielle (14)

- T84 Contamination alimentaire par les bactéries (18)
- T85 Promotion et prévention de la santé (103)
- T86 Comportements sexuels à risque (60)
- T87 Adhérance aux traitements et/ou aux prescriptions médicales (231)
- T88 Facteurs sociogéographiques (19)
- T89 Consommation d'alcool (42)
- T90 Troubles coronariens (5)
- T91 Mortalité liée à la pollution atmosphérique (36)
- T92 Hygiène dentaire (14)
- T93 Pliage des protéines (8)
- T94 Santé des immigrants (19)
- T95 Antigénicité des leptospires (4)
- T96 Cancers (71)
- T97 Streptococcus suis (21)
- T98 Insuffisance cardiaque (7)
- T99 Résistance aux anti-microbiens (88)
- T100 Gestion des systèmes de soins de santé (315)
- T101 Tuberculose (34)
- T102 Souches mutantes d'Escherichia Coli (42)
- T103 Fonctions hépatiques (7)
- T104 Intentions et comportements (12)
- T105 Exposition aux composés organochlorés (28)
- T106 Promotion de la santé au travail (2)
- T107 Hépatites virales (16)
- T108 Poids corporel (27)
- T109 Risques cardiovasculaires (89)
- T110 Usagers de drogues par voies intraveineuses (61)
- T111 Maladies nosocomiales (7)
- T112 Critique des nouvelles technologies (32)
- T113 NI (0)

- T114 Qualité de vie des patients (53)
- T115 Exposition aux hydrocarbures aromatiques polycycliques (13)
- T116 NI (0)
- T117 Pathologies animales (49)
- T118 Nausées et vomissements (6)
- T119 Exposition au manganèse (42)
- T120 Inégalités sociales de santé (239)
- T121 Tabagisme (51)
- T122 Asthme (58)
- T123 Contamination de l'eau potable (53)
- T124 Intoxication au manganèse (31)
- T125 Diabète (surtout chez les amérindiens) (24)
- T126 Détresse psychologique liée au travail (66)
- T127 Séquences des structures de protéines (30)
- T128 Anti-inflammatoire non stéroïdien (16)
- T129 Gestion de la douleur (33)
- T130 Asthme professionnel (46)
- T131 Régimes de protection sociale et citoyenneté (12)
- T132 Cancers du sein et/ou hormono-dépendants chez les femmes (8)
- T133 NI (0)
- T134 NI (0)
- T135 Médecine préventive (123)
- T136 Vaccination (33)
- T137 Salmonellose (23)
- T138 Toxicité du lait (10)
- T139 Intoxications professionnelles (90)
- T140 Amiantose (45)
- T141 Métaplasie *** (0)
- T142 Statut socioéconomique (87)
- T143 Cyanoses des volailles (2)

- T144 Risques liés à la grossesse et au développement foetal (69)
- T145 Transplantation et don d'organes (17)
- T146 NI (0)
- T147 Médecine de première ligne**** (0)
- T148 Facteurs familiaux (35)
- T149 Pharmacocinétique et toxicocinétique (79)

Légende:

NI = Non identifiable

(*n*) = Nombre d'articles liés au *topic* correspondant

Annexe III

Légendes des graphes

Formations

-  Médecine et sciences médicales
-  Épidémiologie et biostatistiques
-  Biologie et chimie
-  Sciences vétérinaires et zoologie
-  Sciences environnementales
-  Psychologie
-  Sciences sociales et humaines
-  Santé publique
-  Administration et gestion

Thématiques

-  (Th1) Santé dentaire
-  (Th2) Médecine préventive
-  (Th3) Soins de santé
-  (Th4) Développement des jeunes et adolescents
-  (Th5) Santé des femmes et des enfants
-  (Th6) Santé internationale
-  (Th7) Soins aux personnes âgées
-  (Th8) Autres troubles médicaux
-  (Th9) Maladies chroniques
-  (Th10) Maladies infectieuses
-  (Th11) Lignes directrices en santé
-  (Th12) Médicaments
-  (Th13) Immunologie
-  (Th14) Recherche en génétique et protéomique
-  (Th15) Zoonoses
-  (Th16) Maladies professionnelles
-  (Th17) Santé environnementale
-  (Th18) Santé mentale
-  (Th19) Déterminants sociaux de la santé
-  (Th20) Comportements sociaux
-  (Th21) Politiques publiques
-  (Th22) Éthique en recherche génétique
-  (Th23) Promotion de la santé
-  (Th24) Évaluation des technologies de la santé
-  (Th25) Gestion des services de soins de santé

Formes des sommets *

-  (Carré) Présent dans la base de données à partir de T-3
-  (Triangle) Présent dans la base de données à partir de T-2
-  (Diamant) Présent dans la base de données à partir de T-1
-  (Trapèze) Présent dans la base de données qu'au temps T0

* Ne s'applique qu'aux sommets représentant des chercheurs.

Bordures*

-  Statut de chercheur régulier à l'IRSPUM
-  Statut de chercheur associé à l'IRSPUM

* La bordure ou son absence peuvent s'appliquer à l'une des quatre formes présentées auparavant. La couleur peut aussi varier.

Intensité des arrêtes

-  Copublication au temps donné
-  Copublication pendant deux temps consécutifs
-  Copublication pendant trois temps consécutifs

Les pointillés rouge servent à différencier les liens vers les chercheurs associés

Surface occupée par les sommets

La surface occupée par les sommets indique leur degré de centralité

Annexe IV

Liste complète des thèmes fusionnés

Tableau IV.I: Liste des fusions de *topics*

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
Th1	Santé dentaire	T55 T92	Maturité dentaire Hygiène dentaire
Th2	Médecine préventive	T135 T52	Médecine préventive Techniques d'imagerie médicale
Th3	Soins de santé	T4 T39 T114 T129 T24 T111 T63	Soins en santé mentale Réadaptation psychologique et physique Qualité de vie des patients Gestion de la douleur Cliniques d'anticoagulants Maladies nosocomiales Risques liés aux thérapies médicales
Th4	Développement des enfants et adolescents	T61 T51 T65	Études sur les jeunes Problèmes de comportement Développement des jeunes enfants
Th5	Santé des femmes et des enfants	T29 T36	Pédiatrie et maladies infantiles Maternité mortalité et pathologies infantiles
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
		T73	Accouchements et mortalité maternelle
		T144	Risques liés à la grossesse et au développement foetal
Th6	Santé internationale	T25	Aide humanitaire
		T50	Santé et soins de santé dans les pays en développement
Th7	Soins des personnes âgées	T46	Services de santé pour les personnes âgées fragiles
		T68	Azzheimer
		T77	Donneurs de soins aux personnes âgées
		T37	Démence
Th8	Autres troubles médicaux	T0	Fertilité et troubles de la reproduction
		T16	Mitochondropathies
		T118	Nausées et vomissements
		T19	Fonctions thyroïdiennes
		T81	Transport intestinal
		T94	Santé des immigrants
		T103	Fonctions hépatiques
Th9	Maladies chroniques	T1	Ostéoporose
		T141	Métaplasie
		T132	Cancers du sein et/ou hormono-dépendants chez les femmes
		T96	Cancers
		T109	Risques cardiovasculaires
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
		T90 T98 T122 T125	Troubles coronariens Insuffisance cardiaque Asthme Diabète (surtout chez les amérindiens)
Th10	Maladies infectieuses	T6 T8 T27 T80 T102 T101 T107 T54	VIH/sida Système de régulation du phosphate chez Escherichia Coli Escherichia Coli entérotoxigènes du porc Souches d'Escherichia Coli de type attachant et effaçant Souches mutantes d'Escherichia Coli Tuberculose Hépatites virales Virus du papillome humain
Th11	Lignes directrices cliniques	T26	Lignes directrices cliniques
Th12	Médicaments	T43 T11 T18 T128 T87	Excipient pharmaceutique Effets de l'Aspirine Effets des antioxydants, acides gras et vitamines Anti-inflammatoire non stéroïdien Adhérence aux traitements et/ou aux prescriptions médicales
Th13	Immunologie	T3 T48	Leucocytes Résistance aux bactéroïdes
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
		T99 T95 T9 T136	Résistance aux anti-microbiens Antigénicité des leptospires Pratiques en matière de vaccinations Vaccination
Th14	Génétique et protéomique	T7 T23 T14 T93 T45 T127 T60	Expression génétique Hormones parathyroïdiennes Expression des cellules Pliage des protéines Séquençage génétique Séquences des structures de protéines Marquage des protéines
Th15	Zoonoses et parasites	T12 T21 T117 T143 T35 T38 T53 T97 T56 T2 T22 T56 T44	Zoonoses liées à l'élevage Flux d'azote dans les porcheries Pathologies animales Cyanoses des volailles Pleuropneumonies (animaux) Coronavirus du dindon Intoxication des espèces aquatiques Streptococcus suis Nématodes parasites des plantes Infections par les tiques Ixodes scapularis Nématodes parasites des plantes Aedes aegypti
Th16	Maladies professionnelles	T139	Intoxications professionnelles
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
		T78	Expositions professionnelles
		T42	Cancers occupationnels (travail, maison, école)
		T71	Silicose
		T140	Amiantose
		T130	Asthme professionnel
		T49	Pneumoconioses
Th17	Santé environnementale	T20	Exposition à des contaminants dans l'environnement
		T75	Exposition au pyrène
		T105	Exposition aux composés organochlorés
		T115	Exposition aux hydrocarbures aromatiques polycycliques
		T83	Pollution industrielle
		T119	Exposition au manganèse
		T124	Intoxication au manganèse
		T91	Mortalité liée à la pollution atmosphérique
		T34	Composées hépatotoxiques
		T69	Pollution environnementale
		T149	Pharmacocinétique et toxicocinétique
		T84	Contamination alimentaire par les bactéries
		T123	Contamination de l'eau potable
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
		T138	Concentrations de toxines dans le lait
		T74	Toxicité du lait de vache
		T137	Salmonellose
Th18	Santé mentale	T126	Détresse psychologique
		T13	Peur conditionnée
		T15	Troubles de l'alimentation
		T28	Suicide
		T62	Dépression
Th19	Déterminants sociaux de la santé	T67	Déterminants sociaux de la santé
		T120	Inégalités sociales de santé
		T142	Statut socioéconomique
		T148	Facteurs familiaux
		T88	Facteurs sociogéographiques
		T47	Facteurs résidentiels
Th20	Comportements sociaux	T30	Activité physique
		T40	Usagers de drogues illicites et <i>gamblers</i>
		T110	Usagers de drogues par voies intraveineuses
		T86	Comportements sexuels à risque
		T89	Consommation d'alcool
		T104	Intentions et comportements
		T64	Régimes et habitudes alimentaires
		T108	Poids corporel
		T121	Tabagisme
... Suite à la prochaine page			

ID	Thèmes fusionnés	id	Anciens <i>topics</i>
Th21	Politiques publiques	T131	Régimes de protection sociale et citoyenneté
		T59	Politiques publiques
Th22	Éthique en recherche et en sciences	T57	Conflits d'intérêts en sciences biologiques
		T70	Éthique en recherche génétique
		T112	Critique des nouvelles technologies
Th23	Promotion de la santé	T5	Promotion de l'exercice physique et d'une alimentation saine
		T85	Promotion et prévention de la santé
		T72	Campagnes médiatiques en santé publique
		T82	Programmes et promotion de la cessation tabagique
		T106	Promotion de la santé au travail
Th24	Évaluation des technologies de la santé	T33	Évaluation des technologies de la santé
Th25	Gestion des services de soins de santé	T100	Gestion des systèmes de soins de santé
		T41	Prise en charge des patients
		T26	Validation des mesures
		T147	Soins de première ligne
		T76	Pratique médicale
		T66	Estimations des coûts et de l'efficacité des thérapies

Exclusions (non identifiables): T10, T17, T31, T32, T58, T79, T113, T116, T133, T134, T146, T145

Annexe V

MouliNet

MouliNet est un programme de traitement automatisé des données. Nous avons conçu ce programme afin de manipuler et transformer plus aisément des bases de données bibliographiques et les résultats des analyses. Écrit en TCL, langage de programmation adapté aux données textuelles, il permet l'utilisation de bibliothèques variées ouvrant de multiples possibilités. De plus, il est multiplateforme.

MouliNet est modulaire ; on peut lui ajouter des fonctionnalités au besoin: nouveaux formats de fichiers, autre type d'entrées ou de sorties ; génération d'images graphiques et plus encore. Les applications possibles sont nombreuses: conversion de fichier ; *data mining*, statistiques, analyse de réseau...

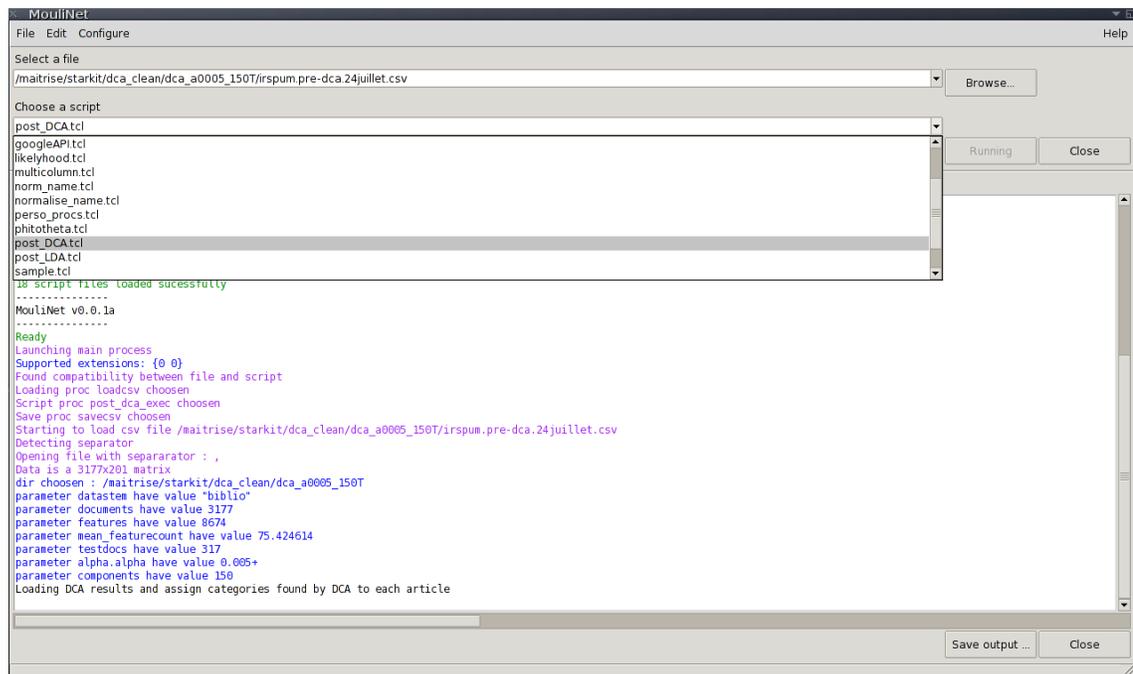


Figure V.1 – MouliNet en action...

Description des scripts et pluggins de *MouliNet* utilisés dans la recherche:

pluggin Bibtex : Ce fichier permet à *MouliNet* d'importer un fichier bibtex sous forme de tableau. Il ne renvoie que les informations pertinentes : liste des auteurs, année, titre, résumé.

script Multicolumn : Ce script sépare la liste des auteurs contenue sur la première colonne du tableau en un auteur par cellule. Les trois dernières colonnes sont toujours : année, titre, résumé.

script g_translate : Ce script demande au service en ligne « *Google Translate* » la langue de chaque titre et résumé. si la langue est différente de l'anglais, il demande la traduction complète.

script DCA : Ce script prépare le répertoire et les fichiers pour le traitement DCA (*discrete component analysis*)

script post-DCA : Ce script récupère les résultats du DCA et l'ajoute au tableau d'où étaient extraites les informations. Pour chaque article sont ajoutés les trois topics les plus probables, avec leur score (six colonnes supplémentaires)

script norm_name : Ce script essaye de rassembler sous un même nom les personnes présente dans la liste des auteurs dont le nom est écrit différemment : exemple fictif : Damien RILLY pourrait aussi apparaître sous le nom D. L. RILLY, RILLY Damien L. ou RILLY Damien Laurent.

La liste des auteurs à suivre est un tableau contenant : un numero identifiant, le nom, le prenom, le domaine de recherche et la discipline de formation, pour chaque chercheur.

Ce script génère les matrices auteurs/auteurs et auteurs/topics

script random_sample : Ce script renvoie 50 articles pris au hasard dans la liste des articles, nous permettant de vérifier la pertinence du classement DCA.

scripts statsn , *statistics* et *stats_aut* : Ces scripts renvoient différentes statistiques sur le corpus, les auteurs...

MouliNet est libre de droit et disponible sur demande.