



HAL
open science

Dynamiques sociales et spatiales des collaborations scientifiques : une étude de cas de chercheurs en chimie

Marine Bernard

► **To cite this version:**

Marine Bernard. Dynamiques sociales et spatiales des collaborations scientifiques : une étude de cas de chercheurs en chimie. Economies et finances. Université de Poitiers, 2020. Français. NNT : 2020POIT4005 . tel-03772147

HAL Id: tel-03772147

<https://theses.hal.science/tel-03772147>

Submitted on 8 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Faculté
de Sciences
économiques



Université de Poitiers

**CRIEF EA2249 - Ecole Doctorale Sciences de la Société, Territoires, Sciences
Économiques et de Gestion**

**Dynamiques sociales et spatiales des collaborations scientifiques :
une étude de cas de chercheurs en chimie**

Thèse pour l'obtention du grade de docteur en sciences économiques

Présentée et soutenue publiquement par **Marine BERNARD**

Le 16 Décembre 2020

Membres du Jury

Alain Rallet, Professeur des Universités

Marie FERRU, Professeure des Universités, *directrice de thèse*

Bastien BERNELA, Maître de Conférences, *co-directeur de thèse*

Virginie JACQUIER-ROUX, Maître de Conférences

Virginie LETHIAIS, Maître de Conférences, *rapporteuse*

Rachel LEVY, Maître de Conférences, *rapporteuse*

L'Université n'entend ni approuver ni désapprouver les opinions particulières du candidat.

Remerciements

La finalisation de cette thèse est pour moi l'occasion d'en faire le bilan. La collaboration étant au cœur de mon sujet de recherche, il me semble légitime de remercier l'ensemble des personnes qui m'ont accompagnée et ont contribué directement et indirectement à la réalisation de ce projet qui me tenait à cœur depuis de nombreuses années. En effet, j'ai toujours gardé dans un coin de mon esprit ce désir fou de réaliser une thèse et de devenir docteur. Alors que je touche mon rêve du bout des doigts, je garde à l'esprit la dimension collective de ce projet scientifique et humain et remercie toutes les personnes qui m'ont aidé à atteindre la ligne d'arrivée.

Tout d'abord, je remercie très chaleureusement mes directeurs de thèse qui m'ont accompagnée avec beaucoup d'attention. Marie Ferru a été un soutien sans failles durant ces trois années de doctorat et durant mes années d'études universitaires. Ses cours en Master ont nourri mon intérêt scientifique pour la dimension territoriale de l'économie et son accompagnement dans la réalisation de mon mémoire a su me donner la confiance nécessaire pour prendre cette voie difficile qui est le doctorat. Son expertise et ses conseils ont guidé ma réflexion durant ces trois années et ont sans aucun doute contribué à la qualité de cette thèse. Je la remercie également pour la disponibilité, la patience et l'humanité qu'elle a montré durant ces mois de finalisation qui ont été riches en rebondissement. Je remercie également Bastien Bernela pour son soutien dans ce travail. Il a su être très présent au cours d'étapes délicates et fondamentales dans la réalisation de la thèse, telles que la création de la base de données : sa persévérance et sa disponibilité m'ont été très précieuses durant ces longs mois. Ce travail n'aurait pas pu être possible sans son investissement et son perfectionnisme, qui m'ont poussée chaque jour. J'ai eu la chance de passer de très bons moments avec mes deux directeurs de thèse, notamment au cours de colloques et d'une école d'été. Je garde en mémoire de très bons souvenirs de ces trois années et c'est en partie grâce à eux. Je leur souhaite beaucoup de satisfaction personnelle et professionnelle dans les divers projets qu'ils mèneront durant les prochaines années.

Je remercie ensuite les membres du jury, Alain Rallet, Virginie Jacquier-Roux, Virginie Lethiais et Rachel Levy, qui m'ont fait l'honneur de contribuer à ce travail en l'évaluant et en apportant leur expertise. Je remercie également Virginie Jacquier-Roux pour les précieux conseils qu'elle m'a fournis au début de ma thèse et qui ont guidé la poursuite de mon travail.

Je remercie également très chaleureusement les chercheurs qui m'ont accompagnée dans ce projet pour leur investissement et leur niveau d'expertise. Ils m'ont aidée à lever différents obstacles qui m'ont freiné pendant plusieurs mois. Je commencerai par remercier Liliane Bonnal dont l'aide a été centrale tout au long de la thèse. Sa disponibilité, en tant que responsable du laboratoire CRIEF et en tant que chercheuse, m'ont été un réel soutien. Son regard sur notre base de données et son niveau d'expertise

en économétrie nous ont aidé à envisager toutes les opportunités scientifiques que nous étions en mesure de saisir durant ces trois années. Je remercie également Béatrice Milard pour son aide tout au long de la thèse. Grâce à elle, j'ai eu la chance d'être initiée à l'analyse de réseaux : j'ai eu beaucoup de plaisirs à apprendre cette nouvelle méthodologie que j'espère remobiliser dans mes prochaines recherches. Son investissement et son expertise m'ont beaucoup aidé dans la réalisation de mes travaux empiriques.

Par la suite, je remercie les membres de la faculté de sciences économiques qui m'ont accompagnée dans la réalisation de travaux dirigés. Je remercie Pascal Chauchefoin pour la confiance qu'il m'a accordée durant ces deux années d'enseignement. J'ai beaucoup aimé cet aspect du métier d'enseignant-chercheur et je remercie tous les collègues qui m'ont accompagnée dans cet exercice exigeant qui est l'enseignement (Philippe Rogeon, Alexandra Popescu, Eliane Jahan, Oussama Houari et Nacera Yeddou).

Ce travail n'aurait pas pu être possible sans la coopération des chimistes qui ont constitué mon échantillon. Leur disponibilité m'a permis d'avoir accès à des informations qui n'étaient pas disponibles et de très bonne qualité. Je remercie également Jade et Yvan qui ont participé à la création de la base de données et dont l'aide a été très précieuse.

Je remercie le projet de recherche GARI (Géographie et Accès aux Ressources pour l'Innovation) et la région Nouvelle-Aquitaine pour le financement de mes travaux de recherche. Outre cet aspect financier, la participation à ce projet pluridisciplinaire et régional a été une réelle opportunité scientifique du fait de la richesse et de la diversité des travaux sur la géographie de l'innovation auxquels j'ai pu avoir accès. Je remercie également les membres du laboratoire CRIEF pour leur accueil et pour leurs conseils.

D'autres personnes ont aussi contribué à la réalisation de cette thèse. Je remercie tout particulièrement François Baty-Sorel pour sa disponibilité. Ses conseils et son accompagnement durant les formations doctorales m'ont été très précieux dans la construction de mon projet professionnel. Sa bienveillance m'a également été d'un très grand réconfort durant les périodes de doute. Je remercie également les collègues doctorants qui m'ont accompagnée au quotidien, Greivis Buitrago Gamez, Fatima Merzoug, Mihaela Nicolau, Oussama Houari, Eric Fina Kamani, Nacera Yeddou, Hamdi Jbir pour leur soutien et leurs conseils à différents égards. Je leur souhaite le meilleur pour l'avenir et d'atteindre les objectifs qu'ils se sont fixés. J'ai une pensée particulière pour Greivis Buitrago Gamez et Marion Mathé qui ont commencé cette aventure en même temps que moi et qui rendront bientôt, j'en suis sûre, une thèse de grande qualité. Je remercie également Annie pour sa bonne humeur quotidienne et pour toutes ses petites attentions qui ont fait que nous nous sommes sentis chez nous en B208.

Enfin, je finirai en exprimant un énorme merci à mon entourage. Merci à mes amies, Léa, Amandine et Ophélie pour leur soutien et pour leur bonne humeur qui m'accompagnent depuis plus de dix ans. Ces instants avec vous sont très précieux pour moi et ont été une ressource centrale dans ce travail. Je remercie également ma famille qui m'accompagne dans mes études depuis de nombreuses années. Je

remercie particulièrement mes parents et mon frère, Thomas, qui se sont toujours investis pour que je puisse réaliser mes rêves dans les meilleures conditions. Depuis l'école primaire, vous m'avez toujours poussée et soutenue malgré les difficultés que j'ai pu rencontrer. La route a été longue mais vous avez toujours été un soutien indéfectible : aujourd'hui, ma fierté se dirige également vers vous. Enfin, je remercie très affectueusement Clément pour son écoute, sa bienveillance et son affection durant ces dernières années. Malgré les kilomètres qui nous séparent, il a été un acteur de premier plan dans ce travail, au travers de longues soirées téléphoniques et de week-ends de décompression. Je le remercie particulièrement pour son encouragement à m'épanouir dans ma vie professionnelle et à repousser mes limites. A l'aube de nouveaux projets personnels et professionnels, j'ai hâte d'écrire cette nouvelle page à tes côtés.

Sommaire

Remerciements	3
Sommaire	7
Introduction générale	9
Chapitre 1 : Revue de la littérature la carrière scientifique et les comportements collaboratifs	24
I- Littérature et carrière scientifique.....	25
II- Politiques de mobilité géographique et comportement collaboratif du chercheur : interaction entre les enjeux politiques et individuels.....	41
III- Les dynamiques relationnelles au sein des réseaux de co-authorship.....	65
IV- Synthèse de la revue de littérature.....	80
Chapitre 2 : Présentation de la base de données	82
I- Création de la base de données et choix méthodologiques	82
II- Statistiques descriptives de la base de données	107
III- Réseaux personnels et temporalité de carrière	116
IV- Conclusion.....	132
Chapitre 3 : Carrières des chercheurs et dynamique des réseaux scientifiques : une mise en intrigue à partir des cas de cinq chimistes émérites	135
I- Contribution méthodologique : articulation des échelles d'analyse et des matériaux empiriques.....	135
II - Le cas de cinq chimistes émérites de l'Université de Poitiers	139
III- Résultats : contextualisation des relations collaboratives.....	145
IV-Conclusion.....	150
Chapitre 4 : Est-ce que la mobilité géographique façonne le réseau de co-authorship ? Une approche de panel des carrières scientifiques	153
I- Intégration de la mobilité géographique dans la trajectoire du chercheur	153
II- Présentation du modèle économétrique	159
III- Résultats	163
IV- Conclusion et implications politiques.....	166
Chapitre 5 : « Existe-t-il plusieurs façons de collaborer au fil de la carrière ? Une analyse de réseaux sociaux de co-authorship des chimistes »	168
I- Structures des réseaux et types de collaboration.....	169
II- Caractérisation des quatre structures de réseau	180
III- Styles de collaboration et carrière scientifique.....	195
IV-Conclusion.....	229
Conclusion générale	233
I- Caractériser la dynamique des collaborations scientifiques grâce à des données individuelles et longitudinales	233
II- Compréhension du comportement collaboratif des chercheurs au fil de la carrière	236
III- Limites et perspectives de recherche possible.....	239
Annexes	242
Bibliographie	271

Table des matières.....	292
Table des tableaux.....	297
Table des figures.....	300

Introduction générale

La géographie de la science apparaît comme un sujet économique crucial qui est au cœur des politiques publiques. En effet, l'Etat français a mis en place une stratégie visant à regrouper les universités en pôles scientifiques territoriaux au travers de diverses réformes successives. Cette stratégie nationale repose sur un investissement financier (les dépenses internes relatives à la recherche et développement dans le public étaient de 17,6 milliards d'euros en 2018) et humain (247 440 personnes travaillaient dans la recherche et développement publique en 2017). Elle vise à étendre la visibilité de la France dans les collaborations internationales scientifiques. **Nous proposons dans cette thèse de mieux comprendre la géographie de la science à partir de l'analyse de la dynamique sociale et spatiale des collaborations scientifiques, étudiée à partir du réseau de co-authorship des chercheurs.** Ainsi, nous appréhendons l'activité de publication des chercheurs **sous le prisme des relations collaboratives qu'elle génère.** Nous souhaitons nous focaliser sur les collaborations du fait des enjeux économiques et scientifiques que ces dernières soulèvent dans l'analyse de la géographie de la science. En effet, en permettant la circulation de connaissances, les collaborations favorisent la co-production scientifique et l'innovation (Mansfield, 1991 ; Salter *et al.*, 2000 ; Cohen *et al.*, 2002). La recherche est ainsi le fruit d'interactions sociales et les relations sociales apparaissent, dans ce cadre, déterminantes de l'activité scientifique du chercheur. Robert K. Merton, considéré comme l'un des fondateurs de la sociologie des sciences, a défini en 1942, ce qu'il a appelé « *la structure normative de la science* », montrant que la science est avant tout une activité sociale, normée par un *ethos*. Ce monde social est multiniveau puisqu'il est composé d'individus (les chercheurs), d'organisations (les équipes, laboratoires), d'institutions qui sont plus ou moins reliés les uns aux autres.

La littérature scientifique s'est intéressée à ce sujet et a participé à l'analyse de la géographie de la science en étudiant les liens collaboratifs. Le développement de bases de données bibliométriques et de travaux en scientométrie ont permis de mieux comprendre la construction et la dynamique des réseaux des chercheurs (Mullins, 1972 ; Bordons *et al.*, 1999 ; Wagner et Leydesdorff, 2005 ; Grossetti *et al.*, 2014). Dans ce cadre, des travaux ont étudié la géographie des collaborations scientifiques, essentiellement sous le prisme de la performance.

Certains de ces travaux ont cherché à évaluer la performance de ces liens collaboratifs, à une échelle macro-économique : est ce que les pays ou régions qui présentent un haut degré de collaborations scientifiques ont des performances supérieures ? Ces articles ont plus ou moins directement traité aux différentes réformes successives visant à faire émerger des pôles universitaires d'excellence en France. Plus précisément, deux types de réformes ont été mises en place : d'une part, des mesures de rationalisation des universités qui visent à accroître leur performance (par exemple la création de

l'AERES¹ en 2006, la mise en place d'un système d'évaluation² des universités basé sur des critères de « performances » ; d'autre part, des mesures de regroupement des universités qui cherchent à accroître leur visibilité (telles que la loi Savary en 1984, la création des « Initiative d'Excellence » en 2011, la loi Fioraso en 2013). Des chercheurs ont souligné les effets contreproductifs et limités de ces politiques cherchant à accroître la concentration spatiale des « élites » scientifiques et des ressources pour la recherche (Brocard, 1991 ; Langfeld *et al.*, 2015 ; Mongeon *et al.*, 2016 ; Zimmer *et al.*, 2020 ; Grossetti *et al.*, 2020 ; Métivier, 2020). Baudet-Michel *et al.* (2020) ont notamment souligné que ces réformes successives fragilisent le réseau de recherche français et surtout les petites structures qui sont elles-mêmes imbriquées dans des réseaux internationaux et locaux de coopération (Levy, 2015). Également, certains chercheurs interrogent la dimension spatiale du réseau de collaboration et ses enjeux en termes de diffusion des connaissances : ces travaux alimentent les politiques européennes, nationales incitant à la mobilité géographique des chercheurs (Morano-Foadi, 2005 ; Cañibano, 2006). De manière générale, ces travaux portant sur la performance des collaborations étudiées à l'échelle macroéconomique (i.e. pays) mobilisent la méthodologie des réseaux complets (Abbasi *et al.*, 2014 ; Arnaboldi *et al.*, 2016 ; Badar *et al.*, 2013 ; Bordons *et al.*, 2015) à partir de données sur courtes périodes.

D'autres travaux ont cherché à mesurer l'impact des collaborations scientifiques sur les performances au niveau, non plus des pays, mais des chercheurs. Dans cette littérature, la collaboration est considérée comme un facteur de productivité de l'activité scientifique du chercheur. Une partie de ces travaux teste l'effet de la mobilité géographique, appréhendée comme l'élargissement géographique du réseau de co-authorship du chercheur, sur la productivité individuelle. Il existe encore de nombreux débats sur les effets des collaborations sur la performance individuelle (Van Heeringen et Dijkwel, 1987), la quête de reconnaissance et de prestige (Mahroum, 2000), la progression de carrière (Lawson et Shibayama, 2015 ; Akers, 2008) et la géographie des collaborations (Melin, 2004 ; Cañibano, Otamendi et Andujar, 2008 ; Scellato *et al.*, 2015 ; Chinchilla-Rodriguez, 2017).

Nos travaux s'inscrivent dans cette littérature mais nous choisissons un angle d'analyse des collaborations décalé par rapport aux travaux existants sur la performance pour adopter une approche plus compréhensive. Ainsi, nos travaux apparaissent complémentaires à la littérature dominante sur le sujet en cherchant à mieux comprendre le comportement par lequel le chercheur choisit de collaborer au cours du temps. Nous choisissons d'adopter une **approche centrée sur l'individu, à savoir le chercheur**. En reconstruisant son réseau de co-authorship, nous avons une représentation précise des liens collaboratifs qu'il a réalisé avec l'ensemble de ses co-auteurs dans le cadre de co-publications communes. Quelques travaux ont d'ores contribué à la compréhension des collaborations

¹ Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur

² Cette évaluation repose sur des critères explicites qui prennent en compte la pluralité des missions de recherche, la diversité de la recherche et la dimension interdisciplinaire de l'établissement. Cette évaluation fait suite à plusieurs bilans tels que le rapport sur les universités de Claude Allègre en 1990 et la critique du modèle de gestion de la recherche publique en France faite par la Cour des Comptes en 2005.

scientifiques grâce à une analyse centrée sur l'activité de co-publication du chercheur et ont testé différentes caractéristiques, telles que l'homophilie de genre (Xie *et al.*, 2018) et le genre (Whittington, 2018). Néanmoins, à notre connaissance, très peu de travaux apportent une compréhension simultanée des **dimensions sociale et spatiale du comportement collaboratif** à l'échelle individuelle. Certains articles traitent de la localisation géographique des partenaires de collaboration (Bouba-Olga et Ferru, 2011 ; Bouba-Olga *et al.*, 2012 ; Maisonobe, 2015) ; d'autres portent sur la description des relations interpersonnelles collaboratives (Milard, 2014, Cabanac *et al.*, 2015 ; Bernela et Levy, 2017) mais peu de travaux à notre connaissance – à l'exception de Grossetti et Bès, 2001 ; Ferru, 2009; Bernela, 2016 - ne traite simultanément des deux aspects, pourtant liés : la géographie des collaborations et plus précisément le caractère local de ces dernières peut s'expliquer, au moins en partie, par l'existence de relations interpersonnelles locales entre les chercheurs. **Elargir l'analyse de la dynamique collaborative en appréhendant simultanément ces deux dimensions** intrinsèquement liées nous semble nécessaire. Il convient de caractériser les collaborations en tenant compte de la localisation et des caractéristiques sociales des co-auteurs. Nous pourrions ainsi vérifier si l'élargissement social du réseau (i.e. collaborations avec de nouveaux co-auteurs) se traduit simultanément par un élargissement géographique du réseau (i.e. collaborations avec de nouvelles villes). Ce type d'hypothèse sera centrales dans l'élaboration de notre problématique.

Par ailleurs, nous pensons que l'étude des collaborations scientifiques requiert l'introduction d'une **perspective dynamique**. En d'autres termes, il nous semble nécessaire de **contextualiser les liens collaboratifs dans la carrière du chercheur étudié** : cette précaution méthodologique est généralement négligée dans la littérature alors qu'elle nous semble déterminante à trois niveaux :

1) L'environnement scientifique dans lequel se déroule l'activité des chercheurs a fortement évolué au cours des dernières décennies face aux réformes importantes de l'ESR³ ;

2) La dynamique relationnelle exercée dans l'activité scientifique co-évolue avec la trajectoire individuelle (professionnelle, géographique, thématique) du chercheur ;

3) La collaboration apparaît comme une composante de la progression de carrière.

1) *L'environnement scientifique dans lequel se déroule l'activité des chercheurs a fortement évolué au cours des dernières décennies* : La quête de performance a pris de plus en plus de place dans l'activité scientifique du chercheur, incitant ce dernier à suivre une course aux publications et aux collaborations internationales. On observe tout d'abord une **concurrence accrue entre les chercheurs sur le marché du travail académique**. Déjà présentes dans les années 1970 (rapports de de Baecque (1974) et de Quermonne (1981)), les tensions sur le marché du travail scientifique se traduisent par une précarisation

³ L'Enseignement Supérieur et la Recherche

croissante des jeunes chercheurs soit parce qu'ils ne trouvent pas d'autres postes qu'assistant ou maître-assistant, soit parce qu'ils sont sur liste d'attente pour un emploi académique titulaire (Soulié, 1996 ; Foegle, 2013). Cette mise en concurrence s'est accélérée dès la candidature de titularisation. Le décret de 1992 promulgue la décomposition du processus de recrutement en tant que maître de conférences entre l'obtention de la certification préalable - la qualification⁴ - et le recrutement. Cette certification préalable, prérogative issue du CNU⁵, accentue la concurrence sur le marché du travail scientifique puisqu'elle n'est pas distribuée en fonction du nombre de postes vacants⁶. **Dans ce contexte de concurrence accrue, la progression au cours de la carrière est de plus en plus centrée sur des critères de performance bibliométrique.** En effet, dès 1979, les concours nationaux, basés sur les titres et les travaux des candidats, remplacent les listes d'aptitudes. Le processus de recrutement se décentralise à partir de 1985 : il ne repose plus sur une instance nationale mais sur un jury suivi par une audition avec la commission de spécialistes de l'établissement. En plus du rythme de publication, **la diversification des relations collaboratives devient un critère de performance**, via l'internationalisation des collaborations⁷ et le rapprochement entre la recherche publique et l'industrie⁸. **A partir des années 2000, l'évaluation de l'activité scientifique s'accroît tout au long de la carrière.** La décentralisation du recrutement scientifique franchit une nouvelle étape avec la loi sur les libertés et les responsabilités des universités (2007)⁹ : les chercheurs ne sont plus recrutés via une commission de spécialistes mais via un comité de sélection, instance ponctuelle nommée par le conseil d'administration de l'université. Le dossier de candidature est appréhendé à partir des évaluations de carrière réalisées par le CNU à partir du décret de 2009, restreintes sur les travaux de recherche du candidat et ne portant pas sur les autres missions du service public¹⁰. Cette évaluation a été longtemps critiquée par les membres académiques : Schwartz (2008) souligne à cet égard la nécessité d'élargir les critères d'évaluation de l'activité des chercheurs, et fait d'autres recommandations pour améliorer la gestion des carrières universitaires. A partir de 2014, cette évaluation prendra la forme d'un « suivi de carrière ». **Compte tenu de ces évolutions structurantes pour l'activité scientifique du chercheur,**

⁴ Les critères sont les suivants : l'inscription dans la discipline, la qualité de la thèse, les publications (nombre et qualité) et l'expérience d'enseignement.

⁵ Conseil National des Universités (créé en 1987) à la suite du Conseil Supérieur des Universités

⁶ Par exemple, selon le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, le nombre de qualifications accordées en 2005 sur l'ensemble des candidatures en lettres et en sciences (respectivement 52,1% et 61,9%) était bien supérieure au nombre de postes de maîtres de conférences ouverts (respectivement 29% et 15,1%)

⁷ Dans le cadre de l'Européanisation de l'Enseignement Supérieur et de la Recherche (Rapport de Jacques Attali (1998), Déclaration de Boulogne (1999)).

⁸ Loi sur l'innovation et la recherche (1999)

⁹ Cette loi propose de nouveaux systèmes de gouvernance, modes de recrutement et modalités d'évaluation des enseignants-chercheurs ; elle délègue par ailleurs de nouvelles compétences aux universités (compétences budgétaires, gestion, ressources humaines, recherche de financement privé).

¹⁰ Telles que l'enseignement, la performance pédagogique, l'engagement dans la vie collective et institutionnelle de l'établissement

il nous semble primordial d'intégrer une perspective dynamique pour appréhender les collaborations scientifiques.

2) *La dynamique relationnelle exercée dans l'activité scientifique co-évolue avec la trajectoire individuelle du chercheur* : Nous pressentons cette coévolution sous différentes dimensions de la trajectoire individuelle. Tout d'abord, **nous pressentons que les dynamiques de développement des réseaux scientifiques des chercheurs diffèrent en fonction des tâches occupées durant la carrière.** Certains chercheurs mettent en évidence l'existence d'un cycle de productivité scientifique des chercheurs au fil de leur carrière lié aux promotions (accès au premier poste titulaire puis au professorat). En effet, il existe un certain nombre de tensions dans l'activité professionnelle des enseignants-chercheurs (Musselin, 2008 ; Dahan et Mangematin, 2010 ; Kossi *et al.*, 2015) : l'engagement dans l'enseignement et dans des tâches administratives de plus en plus nombreuses influence le temps disponible pour la recherche au fil de la carrière. L'effort de production scientifique serait ainsi plus soutenu en début de carrière, jusqu'à stabilisation, puis redescendrait pour atteindre un pic avant le passage au professorat (Diamond, 1986 ; Levin et Stephan, 1991). Également, **nous pressentons que le pouvoir stratégique du chercheur dans la constitution de son réseau dépend de son stade de carrière.** En effet, la place des chercheurs dans l'activité scientifique collective est très dépendante du stade de carrière : notamment dans les disciplines expérimentales, le travail du chercheur dépend de sa place dans la hiérarchie de son équipe. Nous présumons que le jeune chercheur peut être dépendant du réseau de son équipe, jusqu'au jour où il pourra s'émanciper et construire son propre réseau en tant que chercheur confirmé. Enfin, **nous pressentons que la dynamique de construction du réseau de co-authorship dépend de la trajectoire géographique du chercheur.** De nombreux travaux ont mis l'accent sur le rôle de la mobilité des chercheurs dans la diffusion des connaissances à travers les réseaux scientifiques (Almeida et Kogut, 1999 ; Agrawal *et al.*, 2006 ; Breschi et Lissoni, 2009). Cañibano (Cañibano et Woolley, 2015 ; Cañibano, 2017) souligne notamment que la mobilité géographique reconfigure le réseau de co-authorship du chercheur dans une logique de circulation des connaissances (mobilité au moment des recrutements, mobilités temporaires). Elle peut modifier l'entourage relationnel du chercheur directement (par exemple, l'arrivée dans un nouveau laboratoire permet de créer de nouvelles relations et peut aussi fragiliser les anciennes relations avec qui le chercheur n'a plus de prise géographique) et indirectement (dans le cas où c'est le co-auteur qui est mobile et permet au chercheur d'enrichir son réseau).

3) *La collaboration apparaît comme un composant de la trajectoire individuelle en tant que tremplin dans la progression de carrière* : La carrière scientifique étant par nature régie de liens sociaux, plusieurs théories soulignent que la collaboration apparaît comme un **composant propre à la trajectoire de carrière, du fait du poids de l'institution de la communauté de pairs.** En effet, la

carrière scientifique diffère des autres carrières professionnelles du fait de l'importance du cadre de l'institution à la fois dans création, la diffusion des connaissances et dans la progression de la carrière scientifique. Dans une approche sociologique néo-institutionnelle, Laudel et Gläser (Gläser, 2001 ; Gläser et Laudel, 2007 et 2008 ; Laudel, 2001) définissent la carrière scientifique comme un processus multi-dimensionnel de création de connaissances dans le cadre de l'institution et comme une séquence d'emplois¹¹. Gläser (2001) explique notamment que **l'activité de recherche (telle que la collaboration) est structurée par les règles formelles et informelles de l'institution**. L'originalité de leurs travaux tient notamment dans la décomposition de la carrière scientifique sous trois dimensions analytiques parallèles¹². **La communauté des pairs est très structurante dans la progression de carrière**, notamment au début de carrière. **L'un des critères observés par la communauté de pairs est la collaboration scientifique**. En effet, Bozeman (Bozeman *et al.*, 2001) souligne que la carrière scientifique est une quête d'enrichissement de capital humain scientifique et technique pour acquérir de la reconnaissance auprès de la communauté de pairs¹³. **La collaboration serait donc un levier d'accès de ce type de compétences via l'accumulation de capital humain**¹⁴. Wolley *et al.* (2016) expliquent notamment que l'activité collaborative du chercheur accroît son volume de capital humain scientifique et technique et le rend plus attractif. C'est pour cette raison que nous ne considérons pas le comportement collaboratif comme fixe et unique pour tous les chercheurs : **la collaboration renvoie à une stratégie du chercheur pour faire carrière en fonction de ses ambitions** (Bozeman et Corley, 2004 ; Bozeman et Gaughan, 2011).

¹¹ "Series of interconnected work situations", Gläser et Laudel, 2015

¹² (i) La première dimension est la « cognitive career » : elle est définie comme une série de projets de recherche qui composent la trajectoire de l'activité scientifique et aboutit à la création de connaissances. Cette dimension est centrale dans la carrière puisqu'elle contribue aux deux autres.

(ii) La seconde dimension est la « career in research communities » : il s'agit d'une succession de statuts professionnels. L'évolution de la carrière scientifique se traduit donc par la transition d'un statut à un autre.

(iii) Enfin, la troisième dimension est la « organizational career » : elle représente la séquence d'emplois occupés par le chercheur au sein d'universités.

¹³ Il définit notamment trois types de capital humain scientifique et technique : i) les compétences cognitives (compétences propres à l'individu), ii) les connaissances scientifiques substantives et techniques (acquises au fil de l'expérience professionnelle et de la formation), iii) les compétences textuelles (connaissances tacites).

¹⁴ Le capital social est défini comme la somme des interactions personnelles et professionnelles dans lesquelles le chercheur est encadré et qui lui permettent d'avoir accès à des ressources innovatrices. Il résulte de la position du chercheur dans le réseau et son rôle.

L'objectif de la thèse sera donc de **comprendre la dynamique sociale et spatiale des collaborations scientifiques à partir de l'analyse du réseau de co-authorship du chercheur au fil de sa carrière.**

Plus précisément, nous tenterons de répondre aux questions suivantes :

1- Comment évoluent les collaborations des chercheurs au fil du temps ? Nous expliquons le comportement collaboratif en décrivant les tendances collaboratives du chercheur via la décomposition des équipes de publication au fil de la carrière. Cette décomposition se fera à l'échelle sociale (distinction entre les nouvelles relations et les relations renouvelées) et à l'échelle géographique (en fonction de la distance géographique entre le chercheur et les co-auteurs). Comment est-ce que la diversification sociale du réseau évolue au fil de la carrière ? Dans quelle mesure la création de relations collaboratives et le renouvellement des relations structurent l'activité de publication au fil de la carrière ? Comment est-ce que la diversification géographique du réseau évolue au fil de la carrière ? Dans quelle mesure le chercheur ouvre son équipe de publication à des co-auteurs externes de son laboratoire ou/et internationaux ? Le poids des collaborations avec des collègues du laboratoire est-il stable, croissant ou décroissant au fil de la carrière ? Dans quelle mesure la diversification sociale et la diversification géographique s'articulent au fil de la carrière ?

2- Est-ce qu'un évènement professionnel, telle que la mobilité du chercheur, a un impact sur la dynamique de son co-authorship ? Nous expliquons le comportement collaboratif en testant le rôle de la mobilité dans la construction et le développement du réseau scientifique à l'échelle individuelle : Quel est l'impact de la mobilité sur les dynamiques sociales et géographiques de collaboration ? Dans quelle mesure la géographie du co-auteur est-elle sensible aux trajectoires spatiales des chercheurs ? La mobilité favorise-t-elle la diversification sociale et géographique du réseau ? La mobilité a-t-elle un effet instantané ou de long terme sur l'extension du réseau de co-authorship ?

3- Existe-t-il plusieurs façons de collaborer ? Si, oui ces styles de collaboration sont-ils assimilés à des co-auteurs particuliers ou à des trajectoires professionnelles particulières ? Nous expliquons le comportement collaboratif en vérifiant s'il existe plusieurs manières de collaborer dans la carrière scientifique. Dans quelle mesure les chercheurs mobilisent différentes stratégies dans la gestion de leurs relations collaboratives ? Quels sont ces différents styles de collaboration ? A quelle structure de réseau sont-ils associés ? Une fois après avoir identifié cette typologie de styles collaboratifs, nous souhaiterions l'expliquer à partir de données relatives à la description (géographique et sociale) des co-auteurs et à la carrière du chercheur. Est-ce que chaque style collaboratif est associé à des co-auteurs particuliers ? Dans quelle mesure le passage d'un stade de carrière à l'autre se traduit par un bouleversement de la structure du réseau, et donc du style collaboratif ?

Le plan de la thèse s'articule en cinq chapitres. Nous proposons, dans un premier temps, une analyse conceptuelle de la littérature relative la carrière et aux collaborations scientifiques au travers d'une revue de littérature. Ensuite, nous détaillerons les choix méthodologiques suivis dans cette thèse en précisant les modalités de la création d'une base de données originale bibliométrique originale et permettant la reconstruction des réseaux de co-authorship dynamiques. Enfin, au travers des trois derniers chapitres empiriques à partir de cette base de données, nous contribueront à la compréhension de la dynamique collaborative géographique et sociale des chercheurs au fil de leur carrière.

Le chapitre 1 a pour but de faire un l'état des lieux de la littérature existante sur la carrière et les collaborations scientifiques et cherche à identifier les éléments potentiellement déterminants encore peu abordés dans ces travaux.

Avant d'aborder la collaboration scientifique, nous revenons sur le processus dans lequel elle s'inscrit : **la carrière scientifique**. Souvent abordée en termes de performance, nous soulignerons **l'importance de l'appréhender à l'échelle individuelle dans une approche dynamique**. En effet, bien que le chercheur soit considéré comme un agent autonome maximisateur de revenu, la carrière scientifique se déroule au fil des objectifs et des attentes individuelles : chaque chercheur dispose d'une carrière qui lui est propre. Aussi, du fait de l'évolution du cadre institutionnel de la recherche en France, la définition du métier d'enseignant-chercheur a évolué au fil des dernières décennies. Etre chercheur en 1970 et en 2010 ne correspond pas aux mêmes activités : l'environnement institutionnel de la recherche a modifié les pratiques de recherche.

Ensuite, nous nous attardons sur la **mobilité géographique, élément potentiellement déterminant de la carrière dans la formation des collaborations des chercheurs**. Nous montrons que le changement d'affiliation géographique résulte à la fois de motivations individuelles et aussi de l'environnement du chercheur (universitaire et national) qui incite à être mobile au cours de sa carrière. Nous présenterons ces différentes mesures politiques ainsi que le débat portant sur leurs effets à l'échelle territoriale (approches néoclassique et évolutionniste). Enfin, nous soulignerons les débats de la littérature relatifs à l'impact de la mobilité géographique sur la performance individuelle et la géographie du co-authorship. **Nous finirons par souligner que très peu d'articles évaluent l'impact de la mobilité sur la formation et la dynamique du réseau de co-authorship à l'échelle du chercheur.**

Enfin, nous étudierons les travaux développés sur le sujet des **réseaux sociaux qui rendent compte de la structure du réseau, du positionnement des acteurs et qui sont mobilisés pour comprendre collaborations de recherche**. Nous montrerons que beaucoup de ces travaux portent sur la performance des liens collaboratifs grâce aux caractéristiques structurelles des réseaux. Aussi, certaines études choisissent de décrire l'évolution des caractéristiques structurelles des réseaux de co-authorship à l'échelle individuelle sous une dimension précise telle que les dynamiques relationnelles

(Cabanac *et al.*, 2015), l'effet de genre et d'homophilie (Whittington, 2018), les styles de collaboration (Milard, 2014), la géographie des collaborations (Bernela et Milard, 2016). **Nous soulignerons que très peu d'articles décrivent la dynamique du réseau de co-authorship sous plusieurs dimensions simultanément.**

Le **chapitre 2** présente la base de données bibliométrique mobilisée dans la partie empirique de cette thèse. Répondre à notre problématique en prenant en compte les précautions soulevées précédemment suppose de mobiliser des **données individuelles relatives aux dynamiques collaboratives des chercheurs sur l'ensemble de leur carrière**. Cette base originale servira de support dans l'ensemble de nos travaux empiriques.

Nous commençons par décrire les **choix méthodologiques** qui ont permis de construire la base. Nous expliquons notamment notre choix de nous focaliser sur une seule discipline (la chimie) afin d'assurer l'homogénéité des pratiques scientifiques. Nous expliquons également comment nous avons constitué notre échantillon de recherche de 80 chercheurs issus de deux laboratoires de l'ouest de la France (Poitiers et Toulouse). Afin de nous assurer d'une certaine longévité de carrière, nous avons choisi de nous centrer sur des chercheurs qui occupaient des postes de directeur de recherche ou de professeur en 2018. A partir de cette liste de chercheurs, nous avons extrait sur le Web of Science les 9233 co-publications¹⁵ qu'ils ont réalisées sur l'ensemble de leur carrière. A partir de ces données bibliométriques, nous avons reconstitué les paires de co-auteurs sur l'ensemble de leurs publications afin de construire leurs réseaux personnels de co-authorship sur l'ensemble de leur carrière. **L'originalité de cette base de données tient dans l'intégration de données relatives à la localisation des co-auteurs, à la géographie des collaborations, aux caractéristiques individuelles des 80 chercheurs étudiés** (tel que le nom du directeur de thèse et la liste des doctorants sous leur supervision) **et de leurs co-auteurs** (genre, année de soutenance de thèse). Nous avons choisi de regrouper nos 80 chercheurs en cohortes générationnelles : les chercheurs ayant commencé à publier dans des décennies différentes, cette variable nous permet de prendre en compte le contexte historique de leur activité de collaboration.

Nous présentons ensuite **quelques statistiques descriptives relatives aux activités de publication et de collaboration de nos 80 chercheurs**. Nous montrerons qu'environ 60% des relations collaboratives sont ponctuelles (one-shot) et que la nature collaborative des articles augmente au fil des générations. A partir des données complémentaires ajoutées à celles du WOS, nous montrerons que le mentorship est très structurant dans les relations collaboratives : les équipes de publication des hommes sont structurées par la présence de doctorants qui sont/étaient sous la supervision du chercheur et les

¹⁵ Réalisées entre 1967 et 2018

publications des femmes sont structurées par la présence du directeur de thèse. L'accentuation de la dimension collaborative observée au fil des générations se traduit par **la collaboration croissante de chercheurs disposant d'un doctorat et le poids décroissant du directeur de thèse**. Enfin, à partir des données complémentaires relatives à la géographie du co-authorship, nous réaliserons une cartographie des principaux pôles français et internationaux avec qui notre échantillon collabore.

Enfin, à partir de ces données bibliométriques, nous **reconstituons le réseau de co-authorship des chercheurs que nous analyserons sous deux dimensions temporelles différentes : par année de publication (création de 2109 réseaux annuels) et par période (création de 259 réseaux ego-centrés)**. Après avoir présenté la méthodologie de chaque type de réseau, **nous décrivons l'activité de publication et de collaboration** pour chacun d'entre eux. Dans l'approche longitudinale, nous montrerons que le rythme annuel de publication dans la carrière est le même pour toutes les générations et augmente au fil des années de carrière. Nous montrerons aussi que la nature collaborative des articles augmente aussi bien au fil des générations que des années de carrière. Cette croissance collaborative se traduit par le maintien des relations collaboratives et l'internationalisation du co-authorship (la création de relations et le noyau de collègues du laboratoire étant stables au fil de la carrière). Dans l'approche périodique de la carrière (périodes de 10 ans en moyenne), nous montrerons que la taille du réseau dépend du nombre de publications faites, de la nature collaborative des articles et non du maintien des relations collaboratives. De ce fait, les réseaux sont de taille hétérogène.

Le chapitre 3 présente notre premier travail empirique réalisé à partir de la base de données. **Nous souhaitons comprendre le comportement collaboratif des chercheurs au travers des éléments de temporalité à l'échelle de chaque carrière individuelle**. Il s'agit d'une analyse exploratoire et compréhensive via l'étude de cas de cinq chercheurs émérites issus du laboratoire de Poitiers et appartenant à la même génération. Nous souhaitons identifier dans quelle mesure le contexte historique et le stade de carrière a structuré leurs dynamiques relationnelles entre 1967 et 2018.

Nous commencerons par décrire les contributions méthodologiques de cette étude, relatives à l'articulation de différentes échelles analyse (action et temps) et de différents matériaux empiriques. Nous soulignons **l'importance de la dynamique de carrière pour comprendre le processus social qu'est la collaboration**. Ce type d'analyse requiert une méthodologie appropriée, permettant d'identifier les dynamiques collaboratives au travers de données individuelles très fines (de nature géographique et sociale). **Nous proposons une méthodologie mixte via l'articulation de données bibliométriques avec des données qualitatives** (issues de curriculum vitae et d'entretiens semi-directifs). Les paires de co-authorship issues de la base de données permettent de retracer la dynamique collaborative du chercheur et d'apporter une première contextualisation à partir des données sociales et géographiques complémentaires. Les Curriculum Vitae apportent des informations supplémentaires

relatives à la trajectoire professionnelle du chercheur, à la liste de ses publications et aux données relationnelles pour caractériser les relations collaboratives. Enfin, les entretiens semi-directifs permettent d'expliquer les déterminants de la dynamique des réseaux scientifiques au fil de la carrière grâce au matériau narratif. **A partir de cette méthodologie, nous sommes en mesure de donner une épaisseur sociale aux résultats issus de la base de données bibliométrique.**

L'homogénéité générationnelle permet de conserver le même moment d'émergence des réseaux scientifique permettant de dégager des déterminants récurrents des dynamiques relationnelles, les chercheurs ayant progressé dans leur carrière dans un même contexte historique. **Nous montrons l'homogénéité des dynamiques collaboratives du fait d'un contexte de recherche et de trajectoires professionnelles communes.** Cette homogénéité se traduit par un co-authorship local qui s'internationalise tardivement et une autonomisation scientifique suivie d'une prise de responsabilité assurant une forme de centralité dans le collectif. Nous soulignons néanmoins quelques particularités relatives à l'appartenance au CNRS ou à l'Université, à l'orientation thématique et aux liens du chercheur avec l'industrie. Il s'agit ici d'une « mise en intrigue » des collaborations à partir d'un petit groupe de chercheurs du contexte historique et professionnel de leurs collaborations : les résultats n'ont pas pour but d'être généralisés.

Le chapitre 4, la seconde étude empirique à partir de la base de données bibliométrique, cherchera à **comprendre le comportement collaboratif des chercheurs en se centrant sur l'impact de la mobilité géographique des chercheurs sur la construction de leur réseau de co-authorship.** En appréhendant le réseau collaboratif selon une approche longitudinale (par année de publication), nous souhaitons comprendre dans quelle mesure l'évènement de mobilité géographique modifie les dynamiques sociales et géographiques de co-authorship : **est-ce que le changement d'affiliation du chercheur favorise ou limite la dynamique (sociale et spatiale) de son réseau de co-authorship ?**

Nous commencerons par **intégrer la mobilité géographique dans la trajectoire professionnelle du chercheur.** Pour ce faire, nous avons mobilisé les données géographiques relatives à la localisation des 80 chercheurs à chaque collaboration pour retracer leur trajectoire géographique. Nous abordons donc la mobilité comme le changement d'affiliation géographique d'une année à l'autre. **L'originalité de ce travail tient également dans le fait que nous ne concevons pas la mobilité de manière binaire** comme cela est réalisé dans beaucoup de travaux : nous tenons compte à la fois de sa nature (nationale ou internationale) et du moment d'apparition dans la carrière. A partir de premiers résultats statistiques, nous montrons **l'effet restreint de la mobilité sur la collaboration avec les villes extérieures** : le laboratoire du chercheur est la principale source de création de relations et de collaborations du chercheur. Aussi, nous soulignons que **la mobilité n'est pas un prérequis à**

L'ouverture du réseau puisque le chercheur n'a pas été affilié avec la plupart des villes avec qui il collabore et maintient très peu ses relations avec ses anciennes villes fréquentées.

A partir d'un modèle de panel, nous souhaitons vérifier si **la mobilité géographique se traduit par la création ou la rupture des relations collaboratives au cours du temps et dans quelle mesure**. Nous réalisons différents modèles à effet fixe pour nous focaliser sur l'étude des dynamiques collaboratives au sein des carrières individuelles et non sur la comparaison des carrières individuelles entre elles. Nous identifions la mobilité sous deux variables : en fonction de sa nature et du moment où la dernière mobilité a eu lieu. Nous **étudions les collaborations en termes d'entrées et de sorties de co-auteurs** dans le réseau. L'objectif est d'identifier **l'effet direct de la mobilité** (relatif aux co-auteurs affiliés aux villes fréquentées par le chercheur) et **l'effet indirect** (relatif aux co-auteurs affiliés aux villes non fréquentées par le chercheur).

Enfin, nous avons montré que **l'expérience de mobilité augmente l'intégration de nouveaux co-auteurs par article. Cet effet est limité** pour trois raisons. Tout d'abord, la mobilité a un effet instantané sur la diversification sociale du réseau. Aussi, cet effet est restreint aux co-auteurs qui sont intégrés dans la trajectoire géographique du chercheur : il y a peu d'impact sur la diversification géographique du réseau en dehors de la trajectoire du chercheur. Enfin, la mobilité fragilise les anciennes relations du chercheur issus des anciennes villes d'affiliation du chercheur : cet effet de rupture est plus long dans la carrière que l'effet de création de relations.

Le chapitre 5, la dernière étude empirique faite à partir de la base de données bibliométrique, cherchera à **mieux comprendre les réseaux de co-authorship en interrogeant l'existence de « styles collaboratifs » à partir de données de carrière**. A partir des 259 réseaux *ego*-centrés créés dans le chapitre 2, nous souhaitons contribuer à la compréhension des dynamiques collaboratives en vérifiant s'il existe plusieurs manières de collaborer.

Nous commencerons par identifier les styles collaboratifs via une analyse des structures de réseaux de co-authorship. Tout d'abord, nous décrirons treize indicateurs structurels issus d'études dans l'analyse de réseaux sociaux. Ces indicateurs sont souvent utilisés dans le cadre de réseaux complets : ainsi, un travail de réinterprétation dans le cadre de réseaux personnels a été nécessaire. Ces indicateurs portent d'une part sur les relations *ego-alter* (via la position des alter dans le réseau) et l'organisation des collaborations (via la structure du réseau). Au travers de ces treize indicateurs structurels, nous avons identifié, grâce à une analyse de composantes principales, deux dimensions déterminantes pour comprendre l'hétérogénéité des relations et la segmentation des collaborations. Ces deux dimensions seront intégrées aux treize variables dans la classification hiérarchique ascendante. Cet algorithme regroupe les 259 réseaux en groupes homogènes relatifs à leur structure. **A l'issue de ce processus, nous identifions donc quatre styles collaboratifs différents**.

Ensuite, nous nous attacherons à **décrire chacun de ces styles collaboratifs à partir de des caractéristiques structurelles** du réseau de co-authorship. L'objectif est d'interpréter chaque style collaboratif, soit la manière par laquelle chaque chercheur gère ses relations collaboratives au sein de son réseau de co-authorship. A partir des variables les plus contributrices, nous avons commencé par faire une caractérisation statique de chaque classe. Nous avons constaté que ces styles s'organisent en matrice en fonction des deux axes soulevés dans l'analyse de composantes principales (hétérogénéité des relations et segmentation des collaborations). Après avoir décrit chacun de ces styles de manière indépendante les uns des autres, nous avons choisis de reconstituer le processus de classification afin de comprendre chaque étape de création de classe. L'objectif n'est plus de comprendre l'identité de chaque classe (variance intra-classe) mais plutôt pourquoi chaque classe est structurellement différente des autres (variance inter-classe). Nous avons notamment montré que la classification s'est déroulée en fonction des critères de densité du réseau, à l'échelle globale et locale.

Enfin, **nous avons estimé chacun de ces styles collaboratifs en fonction de données de carrière**. A partir de modèles *logit*, nous expliquons la probabilité d'appartenir à chaque classe de réseau via deux types de modèles. Nous commencerons par expliquer le style collaboratif en fonction du **stade de carrière, de la génération et des caractéristiques du co-authorship** (type de co-auteur, composition des équipes de publication, durée des relations, etc.). Ainsi, nous serons en mesure de «contextualiser» davantage chaque style collaboratif. Dans un second temps, nous avons estimé chaque style collaboratif en fonction de données relatives à **la dynamique structurelle du réseau et à la dynamique de publication du chercheur**. L'objectif dans cette seconde approche n'est plus de considérer les réseaux comme indépendants les uns des autres mais de reconstituer les trajectoires collaboratives individuelles pour évaluer dans quelle mesure le chercheur change son style collaboratif au fil de sa carrière. Nous apportons une attention particulière à la caractérisation de la dynamique structurelle du réseau, via i) le changement de classe, ii) la mutation des coordonnées dans le plan des deux dimensions collaboratives de l'ACP, iii) l'ampleur de ces mutations dans le plan.

Tableau 1 : Organisation de la thèse

Objectif	Données empiriques	Méthodologie	Principaux résultats
Chapitre 1 Positionnement la thèse au sein des travaux existants	Corpus de 315 résumés	- Statistiques lexicales - Revue de littérature	Le chercheur scientifique est un agent maximisateur particulier (spécificités sociologique et psychologique, évolution du métier durant les dernières décennies). Une grande partie des travaux étudient la collaboration sous le prisme de la performance scientifique (effet de la mobilité géographique, caractéristiques structurelles du réseau de co-authorship). Très peu d'articles portent sur les dynamiques collaboratives au fil de la carrière scientifique.
Chapitre 2 Présentation de la base de données bibliométrique mobilisée	Données bibliométriques relatives à 80 chimistes : 9233 co-publications, 14603 collaborations, 14603 co-auteurs	- Création d'une base de données bibliométrique - Statistiques descriptives - Création de réseaux ego-centrés	Création de la base de données bibliométrique : création de paires de co-authorship et ajout de données individuelles complémentaires Statistiques descriptives Construction de réseaux de co-authorship : approche longitudinale ou par période d'activité
Chapitre 3 Dynamique relationnelle des chimistes au fil de la carrière et du contexte scientifique	Données relatives à 5 chimistes émérites : 893 co-publications, 1 073 co-auteurs et 3 343 cosignatures	- Statistiques descriptives - Analyse de données bibliométriques, de Curriculum Vitae, et d'entretiens semi-directifs	Homogénéité des dynamiques collaboratives du fait d'un contexte de recherche et de trajectoires professionnelles communes. Cette homogénéité se traduit par un co-authorship local qui s'internationalise tardivement et une autonomisation scientifique suivie d'une prise de responsabilité assurant une forme de centralité dans le collectif.
Chapitre 4 Impact de la mobilité spatiale des chercheurs sur la dynamique du co-authorship	Données bibliométriques relatives à 80 chimistes, 9233 co-publications, 14603 collaborations, 14603 co-auteurs	- Statistiques descriptives - Modèles économétriques de panel (modèles à effet fixe)	La mobilité favorise la création de relations collaboratives mais cet effet est limité : effet spontané, restreint aux co-auteurs intégrés dans la trajectoire géographique du chercheur, ruptures des relations passées
Chapitre 5 Typologie des collaborations scientifiques	Données bibliométriques relatives à 80 chimistes, 259 réseaux ego-centrés	- Analyse de réseaux sociaux - Analyse de composantes principales - Classification hiérarchique - Modèles <i>logit</i>	Création d'une typologie de quatre styles collaboratifs différents : - Style collaboratif 1 : collaborations non compartimentés et relations internes homogènes - Style collaboratif 2 : collaborations compartimentés en sous-espaces collaboratifs et relations internes homogènes - Styles collaboratif 3 : collaborations non compartimentés et relations internes diverses - Style collaboratif 4 : collaborations compartimentés en sous-espaces collaboratifs et dualité des relations internes

Au fil de ces cinq chapitres, trois résultats méthodologiques et empiriques ressortent dans la compréhension du comportement collaboratif des chercheurs au fil de leur carrière.

i) **Contextualisation des collaborations dans la carrière scientifique** : Nous montrerons dans la revue de littérature que la **collaboration est pratiquée dans un contexte particulier, évoluant au fil de contraintes et d'objectifs individuels** (spécificité de ses mécanismes de progression de carrière et de la définition dynamique de son métier). Il nous semble intéressant d'étudier la carrière scientifique en adoptant un point de vue individuel dans le but de **contextualiser le plus finement possible l'activité de collaboration** du chercheur en fonction de ces différentes dimensions. Nous percevons la **carrière scientifique comme une source d'informations inexploitées** que nous voudrions mobiliser pour décrire l'activité scientifique. Pour réaliser ce type d'analyse, une **approche longitudinale** de la carrière scientifique nous semble être une bonne opportunité pour appréhender l'évolution des contraintes individuelles (ex : besoins et des attentes du chercheur), des contraintes institutionnelles (ex : autonomie croissante des universités) et issues de la communauté de pairs (ex : procédures de gestion de carrière, course à la publication).

ii) **Explication des dynamiques collaboratifs à partir de données individuelles** : Nous avons également montré que la compréhension du contexte de création des relations sociales et géographiques dans la trajectoire individuelle est un sujet intéressant et très peu abordé dans la littérature. Ce type d'analyse requiert **des données individuelles riches et variées**, obtenues à partir d'une méthodologie qualitative (entretiens, CVs) auprès des chercheurs concernés par la collaborations et à partir d'une base de données bibliométriques (géographiques et sociales) pour pouvoir interpréter chaque relation collaborative.

iii) **Description de la pluralité du comportement collaboratif des chercheurs** : Nous avons souligné que beaucoup de travaux de l'analyse des réseaux sociaux étudient les caractéristiques structurelles du réseaux dans le prisme de la performance individuelle ou macro. Quelques travaux contextualisent les collaborations du réseau de co-authorship au travers de données longitudinales et traitent de leur évolution structurelle à l'échelle individuelle (dynamiques relationnelles, styles de collaboration, effets de genre et d'homophilie). Dans la continuité de ces travaux, **il nous semble intéressant d'expliquer le comportement collaboratif en mettant en lumière la multiplicité des façons de collaborer possible au travers de l'analyse de la structure du réseau de co-authorship.**

Chapitre 1 : Revue de la littérature la carrière scientifique et les comportements collaboratifs

L'objectif de cette revue de littérature est d'expliquer le positionnement de la thèse, au travers de sujets peu abordés dans les travaux existants. Pour rappel, la thèse porte sur le comportement collaboratif des chercheurs au fil de leurs carrières en termes de dimension géographique et sociale.

Premièrement, nous commencerons par mettre en contexte la carrière scientifique pour comprendre l'environnement dans lequel le scientifique développe son activité et construit son réseau de co-authorship. Nous commencerons par une analyse lexicale exploratoire qui illustrera les principaux enjeux soulevés par la littérature dans l'étude de la carrière académique. Ensuite, nous présenterons plusieurs théories relatives à la définition de la carrière qui soulignent la nécessité d'appréhender la carrière scientifique à l'échelle individuelle pour mieux comprendre l'activité scientifique. Après avoir soulevé ces différentes dimensions de carrière, nous montrerons que la définition du métier d'enseignant-chercheur a évolué au fil des dernières années : la prise en compte de la dimension temporelle est donc nécessaire pour comprendre le comportement collaboratif du chercheur.

Ensuite, nous affinerons notre analyse de la carrière scientifique sur un évènement professionnel qui semble structurer le réseau de co-authorship : la mobilité géographique. Nous insisterons sur le fait que la mobilité n'est pas que le résultat d'une décision individuelle mais aussi une composante de l'environnement de recherche. En effet, nous commencerons par montrer que la mobilité scientifique résulte d'incitations institutionnelles (que ce soit à l'échelle nationale ou au niveau du milieu universitaire). Ainsi, nous dresserons l'ensemble des mesures qui incitent le chercheur à changer d'affiliation géographique au cours de sa carrière. Ensuite, nous présenterons les effets attendus et les effets réels de ces politiques, à l'échelle du territoire. Cette littérature repose sur un débat sur les effets positifs et négatifs induits dans deux champs de vision de la mobilité géographique : la théorie néoclassique et la théorie évolutionniste. Enfin, nous détaillerons les contributions scientifiques relatives aux effets de la mobilité à l'échelle individuelle de la carrière scientifique, notamment sur l'activité scientifique et sur la géographie des collaborations.

Enfin, nous étudierons comment les liens de collaboration des chercheurs ont été étudiés dans la littérature de l'analyse de réseaux sociaux (SNA). Cette littérature s'est développée dans les années 1970 et permet d'aborder la structure et le positionnement dans le réseau de co-authorship pour mieux capter les mécanismes de production et de circulation de connaissances. Nous commencerons par décrire

les principales théories relatives à l'optimisation de la performance des collaborations à partir de la structure du réseau. Ensuite, nous présenterons les principales contributions empiriques et méthodologiques dans l'étude de la corrélation entre la structure du réseau et l'activité scientifique via des données de court terme ou longitudinales. Nous dresserons ainsi l'état des lieux des principaux travaux qui portent sur la compréhension de la collaboration scientifique au fil de la carrière.

I- Littérature et carrière scientifique

Cette partie concerne l'**appréhension de la carrière scientifique** au travers de différentes dimensions. Avant de traiter plus explicitement des collaborations scientifiques, nous souhaitons décrire le cadre dans lequel elles se déroulent. Nous commencerons par une analyse lexicale afin d'appréhender les principaux enjeux relatifs à la carrière académique soulevés dans la littérature empirique. Ensuite, nous décrirons plusieurs cadres théoriques (économique, sociologique et psychologique) de la carrière scientifique pour comprendre les différentes dimensions dans laquelle le comportement collaboratif s'établit. Enfin, nous contextualiserons la carrière scientifique des enseignants-chercheurs en montrant que les attributs de ce métier sont très contextuels et ont fortement évolué au cours des dernières années.

A) Carrière scientifique : état des lieux de la littérature

Nous souhaitons faire un premier état des lieux des principaux sujets de recherche au sein des articles portant sur **la carrière des chercheurs**. Cette analyse n'est pas exhaustive : elle vise à **identifier les principaux enjeux soulevés dans la littérature** via l'analyse d'un corpus de résumés d'articles scientifiques.

1- Méthodologie de l'analyse lexicométrique

L'analyse lexicométrique permet de compresser une grande variété d'informations lexicales issues d'un corpus de texte. Les statistiques textuelles et la classification de segments de textes sont des outils mobilisés dans les sciences sociales (Lebart *et al.*, 1998 ; Marchand et Ratinaud, 2012 ; Jean-Jacques, 2013 ; Loubère, 2016) : cette méthode nous permet de mettre en lumière les principales tendances et similitudes lexicales de la littérature relative à la carrière scientifique. Nous avons utilisé le logiciel IRaMuTeQ (Ratinaud et Marchand, 2012) afin de faire les différents traitements.

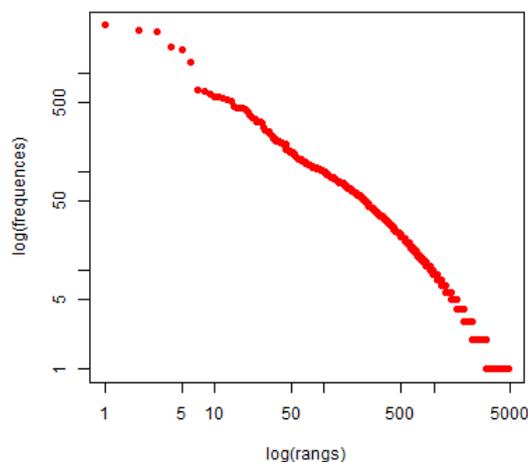
Dans le moteur de recherche « Service commun de la documentation » de l'Université de Poitiers¹⁶, nous avons recherché l'ensemble des articles qui contenaient les expressions « *scientific career* » ou

¹⁶ Nous avons choisi ce moteur de recherche puisqu'il nous offre un accès gratuit et illimité aux résumés des articles scientifiques via une recherche avancée. En effet, la plupart des autres bases de données gratuites (telles que Google Scholar, Pascal Francis, Cairn) ne permettent pas l'export massif de résumés scientifiques. Or, le

« *academic career* » dans leur titre ou dans la description de leur sujet. Nous avons restreint notre recherche avancée aux articles anglophones. En effet, l'analyse lexicale requiert l'uniformité linguistique au sein du corpus de texte, le logiciel utilisant un dictionnaire pour identifier les formes lexicales : l'anglais est le meilleur choix pour capter le maximum de travaux puisqu'il s'agit de la langue universelle du monde scientifique.

Cette recherche ne cible que les travaux pour lesquels **l'étude de la carrière scientifique est centrale**. Nous avons retenu 315 articles, publiés entre 1966 à 2020. Nous avons extrait les résumés de ces articles afin de les compiler dans un même corpus de texte¹⁷ et fragmenter en 1925 segments de texte¹⁸. L'importation du texte a suivi une procédure de lemmatisation (permettant d'obtenir les racines lexicales de chaque mot) et un encodage à partir d'un dictionnaire anglais. Seuls les adjectifs, les verbes, les noms communs et les noms non identifiés par le dictionnaire sont retenus comme actifs dans l'analyse. 68 613 mots sont considérés comme des co-occurrences (répétés plusieurs fois) et 1809 hapax¹⁹ (représentant 2.64% des occurrences).

Figure 1 : Distribution rang-fréquences des formes lexicales



Le graphique suivant présente en abscisse le logarithme des rangs et en ordonnée les logarithmes des fréquences de formes (Loubère et Ratinaud, 2014). La courbe est décroissante : plus nous nous déplaçons sur l'axe des abscisses, moins les formes textuelles sont récurrentes (Lebart et Salem, 1988). Dans ce corpus traitant de la carrière scientifique et académique, nous notons 6 formes actives qui sont plus récurrentes que les autres : il s'agit des termes « *student* » (utilisé 679 fois), « *career* » (utilisé 624

résumé nous semble plus adéquat à notre objectif d'état des lieux de la littérature parce qu'il est plus riche lexicalement que le titre ou les mots-clefs.

¹⁷ Un travail de nettoyage et de regroupement textuel a été réalisé afin de conserver les expressions relatives à la carrière scientifique : cela concerne les postes académiques (par exemple « *associate_professor* », « *assistant_professor* », « *full_professor* », « *tenured_position* »), certaines expressions sociologiques (par exemple « *social_capital* ») et certaines abréviations (par exemple « *h-index* », « *g-index* », « *PhD* », « *CV* »).

¹⁸ Dans IRaMuTeQ, la segmentation du texte se fait à partir d'un critère de taille (40 mots) et en fonction de la ponctuation, afin d'obtenir des segments de taille homogène.

¹⁹ Ce sont les mots qui n'apparaissent qu'une seule fois dans les segments de texte.

fois), « *research* » (utilisé 567 fois), « *academic* » (utilisé 564 fois), « *study* » (utilisé 405 fois) et « *program* » (utilisé 375 fois).

2- Analyse des similitudes : détection des communautés lexicales relatives à la carrière scientifique

L'analyse de similitude permet de montrer les **liens entre les différentes formes textuelles**. Elle mobilise les concepts de la théorie des graphes utilisés en Analyse des Données Relationnelles. Cette étape est préliminaire à la catégorisation des segments de texte puisqu'elle permet de **mettre en lumière les différents termes associés au concept de carrière scientifique**. Nous utilisons l'indice de la co-occurrence pour construire cette analyse : il s'agit du nombre de segments de texte dans lesquels une forme textuelle est associée à une autre²⁰.

Nous réalisons l'analyse sur les 50 formes lexicales les plus récurrentes dans le corpus. Graphiquement, les formes lexicales (mots) sont les sommets du graphe : plus leur taille est importante, plus ils sont fréquents dans le corpus. Aussi, les arrêtes représentent la co-occurrence entre chaque couple de nœuds : plus elle est épaisse, plus les deux mots sont co-occurents. Nous avons également associé la détection de communautés afin de regrouper les termes qui sont co-occurents entre eux et très peu avec les autres.

L'arbre est composé d'une branche principale dotée de cinq formes lexicales : « *academic* », « *career* », « *research* », « *student* », « *medical* », « *study* ». Chacun de ces termes est associé à un ensemble de mots récurrents différent.

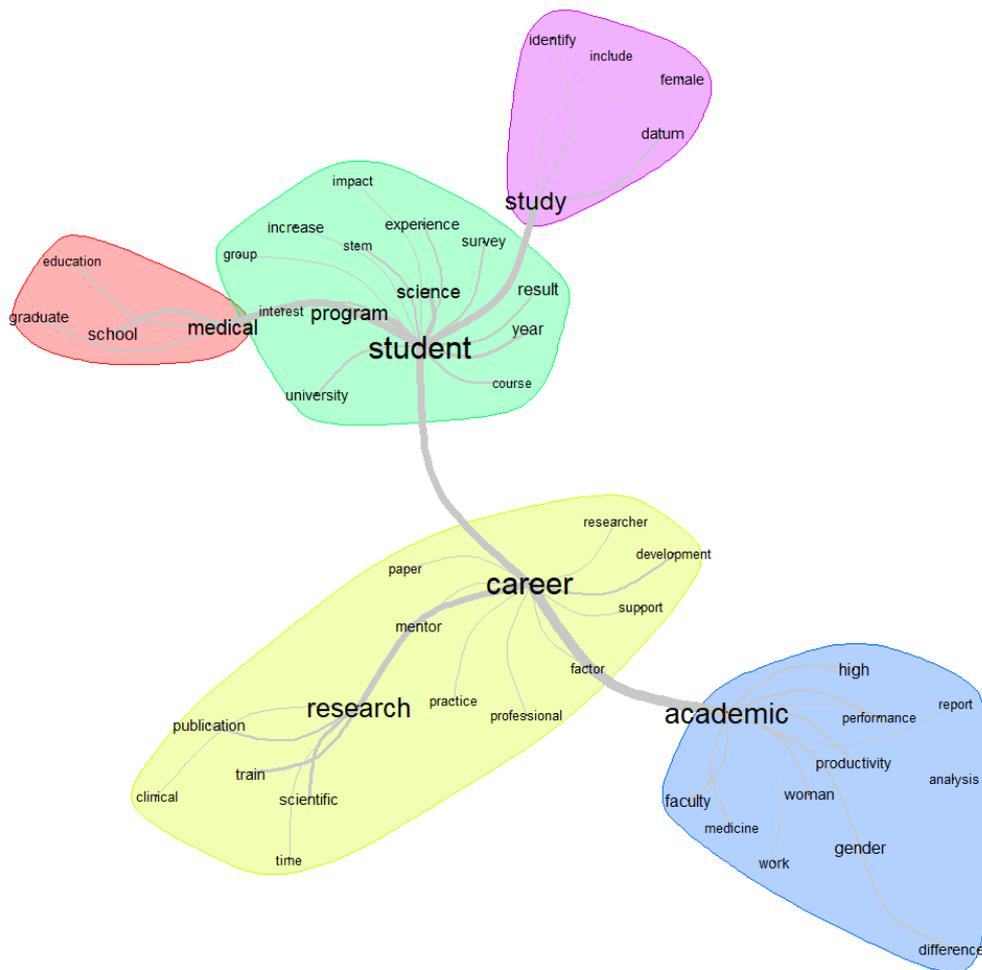
i) Premièrement, nous constatons que l'analyse de la **carrière académique** est récurrente avec l'analyse de la **performance scientifique** (« *performance* », « *productivity* », « *work* ») du chercheur, notamment en fonction de son **genre** (« *gender* », « *difference* », « *woman* »). Ces travaux portent souvent sur une discipline particulière (« *faculty* », « *medecine* »).

ii) Ensuite, nous notons une communauté lexicale relative à la carrière et à l'activité de recherche. Une première co-occurrence concerne la **progression de carrière** (« *professional* », « *support* », « *development* », « *mentor* », « *researcher* », « *practice* ») et une seconde concerne **l'activité de recherche** (« *train* », « *scientific* », « *publication* », « *clinical* »).

iii) Enfin, nous notons une communauté lexicale relative à la carrière de **l'étudiant scientifique** (« *student* », « *course* », « *experience* », « *university* », « *school* », « *graduate* », « *education* »), notamment dans des **domaines de science exactes** (« *medical* », « *STEM* »). Ce champ sémantique est associé avec des analyses quantitatives (« *datum* », « *survey* ») qui ont pour but de mettre en lumière les principaux mécanismes (« *identify* »).

²⁰ Par exemple, si les formes textuelles « *career* » et « *physical* » apparaissent 20 fois ensemble dans des segments de texte du corpus, leur indice de co-occurrence vaudra 20.

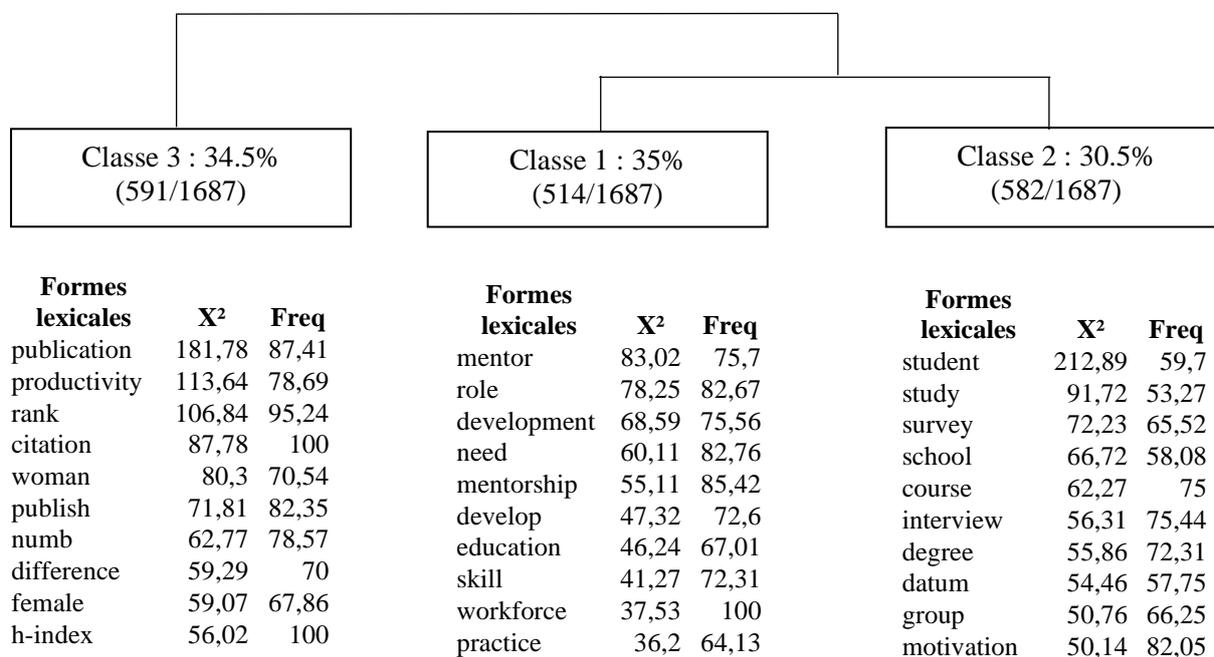
Figure 2 : Arbre maximum de l'analyse de similitude (indice de co-occurrence) des 50 formes les plus récurrentes dans le corpus



3- Classification des enjeux de la carrière scientifique soulevés dans la littérature

Nous opérons la classification de Reinert (Reinert, 1993) pour avoir une large vision **des principaux enjeux de recherche relatifs à la carrière scientifique**. Il s'agit d'une classification hiérarchique ascendante des segments de texte en fonction de la présence et de l'absence de mots (les pronoms, les conjonctions et les adjectifs sont exclus). Avec la méthode ALCESTE, nous avons regroupé les formes lexicales en fonction de leur indépendance, mesurée avec le test au Chi². A partir des 315 résumés scientifiques (composés de 1925 segments de texte), l'algorithme a classifié 87.64% des segments de textes dans trois classes différentes.

Figure 3 : Dendrogramme issu de la classification hiérarchique descendante de Reinter



Le dendrogramme suivant présente les trois classes proposées par l’algorithme ainsi que les 10 formes lexicales les plus significatives de chaque classe. Pour chaque forme lexicale, deux variables expliquent le lien du mot avec la classe : la valeur de χ^2 atteste le résultat du test de dépendance entre le mot et la classe (pour un risque d’erreur de 5%) et la fréquence exprime la part des résumés de la classe dans lesquels le mot en question est écrit au moins une fois.

L’algorithme a donc permis d’identifier **trois sujets récurrents** dans la littérature relative à la carrière scientifique académique :

i) **La relation de mentorship** : La première classe concerne les résumés issus d’une littérature **d’Amérique du Nord** spécialisée sur le mentorat dans la communauté des **métiers de la santé**. Elle regroupe 514 des 1687 segments de texte classés. Les principales revues dont sont extraits ces résumés sont spécialisées dans la formation du personnel médical²¹. Dans les travaux de cette classe, la carrière scientifique est appréhendée sous les prismes du mentorat : “*mentor*”, “*play [...] role*”, “*mentorship*”, “*mentee*”, “*support*”, “*need*” et de la formation “*education*”, “*skill*”, “*practice*”, “*teach*”, “*culture*”. Sambunjak et al (2006) définissent le mentorat comme un **partenariat au cœur de la carrière universitaire** en Amérique du Nord. Sambunjak et Marušić (2009) décrivent l’importance du **rôle du mentor dans la carrière du clinicien** universitaire en Amérique du Nord. Ils

²¹ Medical Teacher, Innovation in Aging, Nurse Education Today, Journal of Continuing in Nursing

expliquent qu'il s'agit d'une relation basée sur des intérêts mutuels : cette relation nécessite l'engagement et les compétences interpersonnelles aussi bien auprès du mentor que du mentoré. Dans cette première thématique de recherche, **la collaboration avec le mentor est considérée comme le moteur du développement personnel et professionnel du chercheur** (Kalén *et al.*,2010) **et de la progression de carrière** (Morrison *et al.*,2014).

ii) **Les choix d'orientation professionnelle** : La seconde classe regroupe 582 des 1687 segments de textes classés. Les principales revues concernent la science de l'éducation appliquée avec comme terrain d'études les sciences technologiques et exactes²². Les résumés de cette classe sont caractérisés par deux champs thématiques. Tout d'abord, la forme lexicale « *student* » dispose d'un Chi² beaucoup plus élevé que les autres : **l'orientation professionnelle des étudiants** est donc centrale dans ce type de travaux. En effet, nous notons plusieurs formes lexicales relatives aux choix professionnels : « *choose* », « *motivation* », « *response* », « *participant* », « *enrol* », « *degree* ». Enfin, les formes lexicales traduisent une récurrence **méthodologique relative aux méthodes qualitatives ou mixtes** : « *survey* », « *interview* », « *datum* », « *qualitative* », « *test* ». En d'autres termes, les articles portant sur les choix professionnels au sein des carrières scientifiques et académiques se distinguent par leur méthodologie de recherche, basée principalement sur des entretiens et des questionnaires et dirigée vers des démarches exploratoires et inductives. Cette littérature se décline sous divers enjeux :

- Tout d'abord, des travaux portent sur les **inégalités d'accès aux carrières scientifiques entre hommes et femmes** dans les disciplines de science, de technologie, ingénierie et mathématiques et ont pour but de comprendre la réticence des femmes à suivre ce type de carrière. Powell *et al.* (2012) pointent une certaine contradiction dans les récits féminins portant sur leurs choix de carrière, du fait de la persistance de stéréotypes. La littérature tente alors de souligner les particularités de la perception féminine de la carrière scientifique auprès d'enfants (Koul *et al.*, 2011 ; Kang *et al.*, 2019) et d'adultes (Dabney et Tai, 2014 ; Fouad *et al.*, 2016).

- Ensuite, de nombreux articles traitent des cas où les **chercheurs renoncent à une carrière universitaire**. Ces travaux ont pour but d'identifier les causes et de proposer des alternatives pour attirer davantage de chercheurs vers ce type de carrière. Harrison *et al.* (2011) expliquent que les étudiants ont changé **leur mécanisme de choix de carrière** : notamment, le stage professionnel se traduit davantage par un changement de trajectoire professionnelle alors qu'il était une étape de consolidation de trajectoire durant les dernières décennies. Gibbs et Griffin (2013) indiquent que très peu de chercheurs choisissent de poursuivre leur carrière jusqu'au grade de professeur dans les sciences biomédicales et détaillent la complexité du choix de carrière. Enfin, Hayter et Parker (2019) ont réalisé une étude inductive pour identifier les facteurs qui influencent la phase de transition entre le post-doctorat et la

²² Journal of Science Education, CBE Life Sciences Education, Physical Review Special Topic, Journal of Engineering Education

réorientation vers une carrière non universitaire. Plusieurs travaux proposent **des pistes pour attirer les étudiants vers la carrière scientifique**, tels que les programmes parascolaires (Markowitz, 2004) et communautaires (VanMeter-Adams *et al.*, 2014), l'accompagnement dans la lecture d'articles scientifiques (Hubbard, 2017) ou le développement de compétences en communication scientifique (Cameron *et al.*, 2020).

- Enfin, certains articles portent sur les **motivations des chercheurs à diversifier leur carrière scientifique** avec une activité atypique. Par exemple, Duberley *et al.* (2007) détaillent les **initiatives entrepreneuriales** au sein des carrières scientifiques de biologie à partir de récits de carrière. Aussi, Van Rijnssoever et Hessels (2011) énumèrent les facteurs qui incitent les chercheurs à faire des **collaborations inter-discipline**, à savoir le genre féminin, l'expérience en entreprise et dans des organisations gouvernementales.

iii) **La productivité scientifique** : La troisième classe regroupe 591 des 1687 segments de texte classés. Les principales revues dont sont extraits ces résumés sont *Scientometrics*, *Journal of Surgical Education* et *Plos One*. Dans ces travaux, la carrière scientifique est appréhendée sous le prisme de la **performance du chercheur** : le terme « *publication* » dispose du plus gros coefficient de Chi², ainsi que d'autres termes relatifs à la **bibliométrie du chercheur** comme « *rank* », « *citation* », « *publish* », « *numb* », « *h-index* », « *journal* », « *author* », « *collaboration* ».

Ces travaux portant sur la productivité scientifique soulèvent différents enjeux :

- Premièrement, de nombreux travaux ont pour but **d'identifier les facteurs de la productivité** du chercheur afin de **prédire la portée scientifique du travail du chercheur**. Parker *et al.* (2012) étudient les scientifiques en science environnementale et en écologie qui ont été les plus cités au niveau mondial pour déterminer les facteurs de leur réussite, à savoir un fort rythme de publication et la publication dans des revues scientifiques très spécialisées. Ces articles mobilisent donc la bibliométrie des chercheurs afin d'identifier les caractéristiques des plus performants (Gyórfy *et al.*, 2018 ; Kwiek, 2018 ; Susarla *et al.*, 2016 ; Pfeiffer *et al.*, 2016 ; Cope, 2017 ; Nowak *et al.*, 2018). La productivité du chercheur est aussi bien étudiée à l'échelle individuelle (Petersen *et al.*, 2011) qu'à l'échelle de disciplines précises (Swihart *et al.*, 2016 ; Diem *et al.*, 2012 ; Boudreaux *et al.*, 2019 ; Doja *et al.*, 2014 ; Pagel et Hudetz, 2011 ; Biswal, 2013).

- Certains articles **approfondissent** l'analyse bibliométrique en cherchant les **liens entre la progression de carrière, la réputation et la productivité scientifique du chercheur**. Par exemple, Jensen *et al.* (2009) ont testé la relation d'indépendance d'indicateurs bibliométriques sur les promotions de 600 chercheurs du CNRS de plusieurs disciplines. **L'analyse dynamique de la carrière permet de tester l'impact du début de carrière sur la progression du chercheur**, et donc de vérifier si la productivité du chercheur est déterminée par ses premières années d'activité. Par exemple, Chan *et al.* (2018) montrent que l'obtention de la médaille John Bates Clark (à la suite des performances de début

de carrière) augmentait la probabilité de recevoir le prix Nobel au cours de la carrière. Aussi, Reyes Gonzalez *et al.* (2018) ont montré que dans le contexte de pays émergents, tels que le Mexique, commencer sa carrière de chercheur avec le soutien d'un éminent chercheur augmente la probabilité de productivité future. Certains articles expliquent plutôt la **productivité du chercheur à partir de données de carrière**. Par exemple, Peterson (2015) montre via l'étude longitudinale du réseau de co-authorship d'un chercheur que de fortes relations collaboratives favorise la productivité et les citations du chercheurs grâce aux effets de confiance, de conviction et d'engagement qui en découlent. Aussi, Peterson *et al.* (2014) expliquent que la réputation du chercheur acquise durant sa carrière accroît significativement le nombre de citation de son article au début de son cycle de vie et décroît au fil des années. Abramo *et al.* (2011) ont évalué les liens entre la performance individuelle du chercheur et son rang académique auprès de chercheurs italiens en sciences exactes. Enfin, certains chercheurs ont montré que **d'autres aspects de la carrière ne sont pas corrélés aux performances scientifiques du chercheur**. Par exemple, Sandnes (2018) montre que les performances de publication des chercheurs en Norvège n'impactent pas la fixation du salaire du personnel ayant peu de responsabilité.

- Enfin, **d'autres facteurs** sont identifiés pour **caractériser la productivité scientifique du chercheur**. De nombreux articles soulignent **l'impact du genre** du chercheur dans sa position académique (Santos *et al.*, 2019 ; Puuska, 2010 ; Odel *et al.*, 2019 ; Pashkova *et al.*, 2013 ; Valsangkar *et al.* 2015) et dans la diffusion scientifique des travaux (Van den Besselaar, 2017 ; Van den Besselaar *et al.*, 2015 ; Mayer *et al.*, 2018 ; Nguyen *et al.*, 2018). La comparaison de la productivité des hommes et des femmes caractérise fortement les résumés de cette classe, au travers des formes lexicales « woman », « female », « difference », « male ». Ensuite, l'effet de la **mobilité géographique** sur la productivité du chercheur est étudié par certains articles (Horta, 2012 ; Deville *et al.*, 2014 ; Albarrán *et al.*, 2017 ; Long *et al.*, 2009). Enfin, d'autres articles pointent le rôle des **facteurs institutionnels** sur la productivité bibliométrique du chercheur. Par exemple, Lissoni *et al.* (2011) montrent que les politiques de recrutement et de promotion des systèmes universitaires italiens et français influençaient la productivité de la recherche sur le long terme.

Cette étude préliminaire de la carrière scientifique souligne l'intérêt de ce sujet d'étude qui lève différents enjeux dans la littérature. En effet, l'activité et les choix du scientifique sont abordés sous les thématiques du mentorship, des choix professionnels des chercheurs en début de carrière et de la productivité d'un point de vue bibliométrique. A partir de là, il convient de décrire les principaux cadres théoriques permettent d'appréhender la carrière scientifique et l'activité de collaboration.

B) Différentes approches de la carrière scientifique

Dans cette seconde sous-partie, nous nous intéresserons à **la définition de la carrière scientifique au travers de différents prismes** : la théorie **économique** (Becker, 1964 ; Biddle et Roberts, 1994 ; Sent, 1999 ; Stephan *et al.*, 2007), celle de la **sociologie des sciences** (Merton, 1968) et celle de **sciences psychologiques** (Dalton *et al.*, 1977 ; Lent *et al.*, 1994 ; Hall, 1996 ; Arthur et Rousseau, 1996 ; Sullivan et Baruch, 2009). L'objectif est d'adopter une appréhension de la carrière scientifique la plus riche possible à l'échelle individuelle pour mieux comprendre l'activité de collaboration qui se déroule dans son cadre.

1- Approche économique : progression de carrière d'un agent autonome

Nous présenterons deux grandes théories économiques de la carrière scientifique : celle des Néo-classiques (Arrow, 1962 ; Becker, 1964) et de Stephan (Stephan *et al.*, 2007). Ces théories économiques insistent sur la perception du chercheur comme un **agent autonome dans la gestion de sa carrière**. La théorie néo-classique décrit le chercheur comme un agent maximisateur de revenu et Stephan souligne le désintérêt de certains post-doctorats pour la carrière universitaire.

Appréhension de la carrière scientifique : d'une approche néoclassique à celle de l'économie de la science

La théorie des sciences économiques relative aux carrières scientifiques a beaucoup évolué au cours de XXème siècle. L'appréhension de la **carrière scientifique** est passée de restreinte au **cadre du marché du travail** pour être ensuite considérée comme imprégnée dans le **monde socio-économique**.

L'analyse néoclassique du marché du travail s'est développée dans le modèle théorique de la concurrence pure et parfaite. Dans cette approche, le travail est considéré comme un bien échangé sur le marché entre l'offreur (le salarié) et la demande (l'entreprise). **Le travail scientifique** est alors conçu comme la **rencontre sur un marché** entre l'offre (le chercheur) et la demande (l'institution de recherche) : la connaissance scientifique est le résultat de cette interaction. Selon la branche néoclassique de l'économie du bien-être (Samuelson, 1954), le savoir scientifique est considéré comme un bien commun (Arrow 1962) qui est non excluable et rival.

Également, la théorie néoclassique du travail et du capital humain (Becker, 1964) décrit la **carrière** comme une **succession d'emplois issus du marché du travail**. La carrière se dessine au fil des interactions entre la demande et l'offre, à savoir que le scientifique accumule du capital humain scientifique au fil de ses formations et de ses expériences. Ainsi, **la progression de carrière** dépend en partie des investissements faits par le scientifique dans **l'enrichissement de sa dotation de capital humain scientifique**. Comme tout autre offreur de travail, le scientifique est donc un **agent autonome maximisateur de revenu souhaitant accroître son capital humain**. Par exemple, Biddle et Roberts

(1994) décrivent la dualité de carrière entre les tâches techniques et managériales. Ils expliquent dans un modèle d'équilibre sur le marché du travail que certains scientifiques et ingénieurs choisissent de s'orienter vers des carrières en dehors de la sphère académique afin d'avoir accès à une plus grande autonomie, à des postes managériaux et à des revenus plus en adéquation avec leur productivité.

De nombreux travaux ont par la suite pointé les limites de l'approche néoclassique. Sent (1999) a notamment souligné la nécessité de ne pas seulement appréhender la carrière scientifique dans le cadre du marché du travail. Selon elle, l'activité scientifique est imbriquée dans le monde socio-économique et la carrière doit être contextualisée dans les réalités économiques et sociales dans lesquelles elle se situe.

L'économie de la science selon Paula Stephan

Dans la poursuite des travaux néo-classiques, Paula Stephan (Stephan et Levin, 1997, 2001 ; Stephan *et al.*, 2007 ; Stephan, 2008, 2013 ; Stephan *et al.*, 2014) est une économiste qui a beaucoup contribué dans l'économie de la science.

Sa première question de recherche porte sur **la productivité scientifique du chercheur**. Elle mobilise la théorie néo-classique du capital humain afin d'introduire la productivité scientifique comme **un levier de progression de carrière, via le gain de réputation et de prestige**. Son originalité réside dans la **mise en perspective dynamique de la productivité**, au travers des conditions de marché évolutives et dans l'introduction de facteurs de marché (le stade de carrière et le genre) pour déterminer la productivité du chercheur. Ainsi, elle étudie la **productivité du chercheur en termes de cycle de vie** (Levin et Stephan 1991 ; Stephan 1996 ; Stephan et Levin 1997) et non comme une variable continue. Par exemple, elle démontre que la productivité du chercheur s'accroît au fil des premières années avant de suivre une tendance décroissante à partir d'un certain seuil.

Sa seconde question de recherche porte sur les **motivations qui incitent les post-doctorants à ne pas poursuivre une carrière académique**. En effet, Stephan appréhende la carrière académique comme une succession de trois étapes : le doctorat, le post-doctorat et le poste permanent. Elle travaille sur la phase transitoire entre le post-doctorat et le poste permanent. Elle souligne **deux facteurs qui handicapent les post-doctorants** dans l'accès à ce type de poste. Le premier concerne la **dépendance vis-à-vis de fonds externes** dans le financement de leur activité de recherche (« soft-money ») : le projet de recherche du post-doctorant est très vulnérable vis-à-vis d'une éventuelle baisse de cette ressource. Le second obstacle provient de la tendance de certains professeurs à **prolonger la période de post-doctorat du chercheur** au lieu de lui proposer un poste permanent (Stephan et Levin 1997).

C'est notamment pour ces raisons que les doctorants projettent négativement la carrière académique (Fox et Stephan, 2001). Ces perceptions négatives diffèrent en fonction de la **discipline du chercheur et en fonction de son genre** : les femmes seraient plus enclines à développer une carrière dans les tâches

d'enseignement à l'université que dans la recherche. Stephan (2008) souligne les conséquences de ce type de changement de trajectoire professionnelle dans les sciences techniques, l'ingénierie et les mathématiques, à savoir la chute de la recherche fondamentale (au lieu de sujets appliqués avec l'industrie) et l'attraction de doctorants vers l'industrie en Europe. Sauermann et Stephan (2013) décrivent les similitudes et les différences dans la recherche dans l'industrie et le milieu universitaire, et montrent que les chercheurs qui recherchent l'autonomie et une forte rémunération choisiront davantage l'industrie que la sphère académique.

Applications empiriques

Le tableau suivant présente quelques applications empiriques de ces théories économiques. Les données recueillies sont principalement issues de la bibliométrie et d'enquêtes. Les principales thématiques abordées sont les **facteurs de progression de carrière** ainsi que l'identification des **aspirations professionnelles des chercheurs en début de carrière**.

Tableau 2 : Applications empiriques des théories économiques relatifs à la carrière scientifique

Etude	Enquête	Discipline	Pays	Résultats
Pezzoni et al., 2012	Panel de carrière de 813 d'enseignants-chercheurs à partir de données bibliométriques	Physique	France et Italie	<p><i>Rôle du capital social (capital humain scientifique et technique et capital politique) dans la progression de carrière :</i></p> <ul style="list-style-type: none"> - Le capital social est un mécanisme de diffusion des connaissances mais aussi d'échanges de pouvoir et politiques - En Italie, la collaboration avec des seniors de la discipline et de l'université (capital politique) favorise la progression de carrière - En France, les relations avec des instituts prestigieux (capital humain scientifique et technique) favorisent la progression de carrière - Le sexe freine la progression alors que l'ancienneté et la productivité l'accroissent
Sauermann et Roach, 2012	Enquêtes auprès de 4109 doctorants	Sciences de la vie, chimie, physiques	Etats-Unis	<p><i>Aspirations de carrière des doctorants et inadéquation avec la réalité des opportunités :</i></p> <ul style="list-style-type: none"> - Nombre excédentaire de docteurs qui souhaitent occuper un poste de professeur permanent - Cela est dû à un manque d'informations sur les trajectoires professionnelles alternatives

Fitzenberger et Schulze, 2013	Enquêtes auprès de 1 169 post-doctorants	Sciences économiques, économie des affaires, sociologie, sciences sociales	Allemagne	<i>Différentes perceptions de carrière des post-doctorants :</i> <ul style="list-style-type: none"> - Les chercheurs optimistes et motivés (vis-à-vis de leur carrière académique ou non, jeunes chercheurs) - Les académiques confiants (optimistes vis-à-vis de la carrière académique, chercheurs âgés) - Les pessimistes vis-à-vis de la carrière en général - La perception positive de la carrière académique est liée aux conditions de travail et aux perspectives de carrière - La perception négative de la carrière académique est liée à la précarité de l'emploi et au fait d'avoir un enfant - Suggestions : mécanismes de soutien pour les professeurs assistants et amélioration des conditions de travail des post doctorants
Peterson et al., 2014	412 publications issues de 23 revues : 258 626 scientifiques dont les trajectoires ont été reconstituées au sein des revues	Sciences sociales et naturelles	International	<i>Lien entre la publication dans des revues à haut rang et la progression de carrière :</i> <ul style="list-style-type: none"> - L'intervalle de temps moyen entre les publications successives dans les meilleures revues diminue au fil du temps - Les publications répétées dans les meilleures revues ont comme effet de baisser les citations des publications ultérieures - Il y a donc un biais dans l'analyse de l'impact de la productivité sur la progression de carrière

2- Approches psychologiques et sociologiques de la carrière

Alors que le chercheur est considéré par les économistes comme un agent autonome maximisateur, certaines théories sociales ont une tout autre perception de la carrière scientifique. Ces approches **complètent notre appréhension de la trajectoire professionnelle** ainsi que de l'activité collaborative du chercheur. Ces travaux soulignent **le rôle central des objectifs et perceptions individuels** dans le déroulement de l'activité scientifique, ce qui explique pourquoi **chaque carrière professionnelle est unique**.

Tout d'abord, les attentes et les besoins du chercheur évoluent en fonction du contexte de carrière. Dalton *et al.* (1977) ont été les précurseurs dans la **caractérisation dynamique et psychologique de la carrière professionnelle** en proposant une typologie des stades de carrière et des situations psychologiques correspondants. Ce premier modèle montre que le **rôle interpersonnel et des compétences attendues du chercheur évoluent au fil de son activité**.

- i) Le premier stade est celui d'« *apprentice* » : l'individu est dans une situation de dépendance et ses activités consistent à se former et suivre les directions.
- ii) Ensuite, le second stade est celui de « *colleague* » : l'individu se déleste de sa dépendance professionnelle dans la structure pour devenir un contributeur indépendant.
- iii) La troisième phase est celle de « *mentor* » : l'individu assume des responsabilités et forme son équipe.
- iv) Enfin, la dernière phase est celle de « *sponsor* » : l'individu dispose d'un fort pouvoir dans la direction de l'institution.

L'évolution des objectifs et des perceptions du chercheur va **structurer sa trajectoire individuelle** ainsi que son activité scientifique. Bandura (1977) propose le concept de *self-efficacy* comme la propension individuelle à organiser et exécuter ses actions pour atteindre des objectifs précis. A la suite de cette contribution, la théorie social-cognitive de la carrière (Lent *et al.*, 1994 ; Lent *et al.*, 2002 ; Lent et Brown, 2008) a été appliquée pour **traduire les croyances des chercheurs pour expliquer le déroulement de leur carrière**. Par exemple, Inda *et al.* (2013) ont appliqué cette théorie sur un échantillon d'étudiants en ingénierie au Mexique. Ils ont montré que les femmes avaient moins de confiance et d'intérêt que les hommes dans le *self-efficacy*, qu'elles sont plus susceptibles de recevoir un soutien personnel que les hommes et qu'elles identifient moins les contraintes familiales que les hommes.

Ainsi, à partir de ces besoins et attentes individuels, **les travaux ont adopté plus de flexibilité dans la définition de la carrière scientifique**. Par exemple, certains ont intégré la mobilité organisationnelle et sectorielle pour décrire la trajectoire professionnelle **en fonction des caractéristiques psychologiques du chercheur concerné**. En 1996, Hall propose le concept de « *protean career* » qu'il définit comme « *une série d'expériences individuelles variées, en matière d'éducation, de formation et de contenu d'emploi* » (Hall et Moss, 1998). Ce type de carrière correspond à un apprentissage continu des connaissances et du savoir-faire dans l'objectif de s'adapter aux exigences évolutives sur le lieu de travail : il concerne davantage les chercheurs autonomes et ambitieux (Hall, 2004). Aussi, en 1996, Arthur et Rousseau proposent le concept de « *boundaryless career* », défini comme « *une série d'opportunités d'emplois qui dépasse les frontières d'un seul lieu de travail* » (DeFillippi et Arthur, 1994). Ce type de carrière est caractérisé par la mobilité organisationnelle : il concerne les chercheurs extravertis et capables de travailler dans divers contextes organisationnels (Briscoe *et al.*, 2006). Enfin, Sullivan et Baruch (2009) proposent le concept de « *hybrid career* », soit la carrière caractérisée simultanément par des aspects de la carrière traditionnelle et des « *protean careers* » ou « *boundaryless careers* ». Ce type de carrière est structuré par des motivations de carrière traditionnelle (par exemple, gravir les échelons) et non traditionnelle (par exemple, avoir accès à une formation au sein et en dehors de la structure) (Granrose et Baccili, 2006).

Au vue de la diversité des trajectoires professionnelles scientifiques, **la collaboration apparaît comme un comportement individuel réalisé en fonction des objectifs et des attentes du chercheur**. En effet, le chercheur a tout intérêt à privilégier ses collaborations avec des chercheurs à forte notoriété, dans le but de progresser plus rapidement dans leur carrière. **L'effet Matthieu** (aussi appelé « avantage cumulatif »), développé par Merton en 1968, traite de **l'allocation de la reconnaissance scientifique**, et notamment dans quelle mesure le succès favorise le succès dans la communauté scientifique. Merton montre à quel point la réputation acquise par les scientifiques et les établissements détermine la visibilité de leurs travaux et le crédit qui leur est porté. Merton souligne, par exemple, que si un jeune chercheur **collabore avec un chercheur à forte notoriété**, sa publication gagnera en visibilité et le jeune chercheur se verra doté d'une meilleure réputation par procuration. Cette théorie a soulevé de nombreuses critiques. Par exemple, Petersen et Penner (2014) ont montré que cet effet n'était pas illimité. Stephan (2008) ayant prouvé que la productivité du chercheur chutait au fil des années (à partir d'un certain seuil), Petersen et Penner (2014) soulignent que le cumul de réputation dans le temps n'est pas corrélé avec la qualité des articles produits et crée des effets négatifs sur la science. Hicks *et al.* (2015) soulignent les stratégies perverses et les biais d'évaluation de ce type de mesure. De plus, de nouvelles métriques d'évaluation bibliométrique ont été proposés par certains chercheurs, le décompte des citations (h-index) étant influencé par l'effet Matthew et donc non représentatif de la qualité de la production scientifique selon certains. Meijer *et al.* (2016) plaident notamment pour un usage de « responsible metrics ».

La carrière scientifique apparaît donc comme un phénomène complexe et multidimensionnel. Dans la littérature économique, le chercheur est considéré comme un agent autonome sur le marché du travail qui souhaite maximiser son revenu via l'accumulation de capital humain et de prestige. Néanmoins, les contributions sociologiques et psychologiques soulignent **le rôle moteur des besoins et des objectifs de l'individu dans le déroulement de sa carrière**. Ainsi, la trajectoire professionnelle scientifique est définie dans par des définitions de plus en plus flexibles, soulignant l'intérêt de **caractériser la particularité de ce cadre dans lequel l'activité de collaboration se déploie**. La carrière scientifique n'est plus appréhendée comme une trajectoire normée mais comme une **multiplicité de trajectoires possibles**. En effet, il est difficile de classifier les carrières académiques du fait de leur grande diversité (Musselin, 2009 ; Dowd et Kaplan, 2005 ; Gatignol, 2014). Enders et Kaulish (2006) soulignent même que ces carrières sont de plus en plus liées au contexte professionnel et sont donc uniques. Nous complétons cette appréhension de la carrière scientifique avec **la dimension temporelle du métier d'enseignant-chercheur**.

C) Impact de l'évolution de l'environnement universitaire sur le métier d'enseignant-chercheur

Le monde académique a été caractérisé par une **succession de réformes** relatives à la **gestion de carrière** des enseignants-chercheurs et au **statut de l'université** durant les dernières décennies. En effet, les **méthodes de management** du secteur public mutent pour se rapprocher de celles du privé, où l'évaluation de la **performance est la clef de voûte** (Commeiras et Georgescu, 2012). Musselin (2008) souligne que les établissements d'enseignement supérieur sont soumis à de nouveaux impératifs, à savoir la transformation du rôle de l'Etat, le regroupement des universités en organisations, le rôle accru des parties prenantes, la logique de privatisation et la perspective mondiale de la science. Au sein de ce nouvel environnement, **l'autonomie individuelle de l'enseignant-chercheur va être remise en cause** (Enders et Karulish, 2006) **et la définition de son métier va évoluer.**

Le métier d'enseignant-chercheur introduit de plus en plus de **tâches dites administratives**. Le Rapport Ponthieux et Berthelot (1992) souligne « l'invasion » de ces nouvelles tâches dans le travail des enseignants-chercheurs. Faure et Soulié (2005) parlent de leur côté d'« inflation » et de « constante augmentation » des charges administratives. Ces tâches concernent tout d'abord la **recherche de financement privé**, notamment auprès du monde industriel. Gingras (2003) souligne la pression exercée à ce sujet par la réduction des budgets publics d'enseignement et de recherche. Ainsi, l'enseignant-chercheur doit réaliser des tâches administratives relatives à la **valorisation de leur recherche**, à ses **responsabilités administratives** et à la professionnalisation de son **rôle d'enseignant.**

Plusieurs chercheurs ont décrit l'impact de cette nouvelle activité sur le métier et la progression de carrière. Dahan et Mangematin (2010) ont souligné que la profession a évolué : elle intègre désormais **une dimension de gestion et de pilotage de projet**. Selon eux, le métier d'enseignant-chercheur serait donc composé de deux types d'activité : les **activités centrales et jugées légitimes** (expérimentation, lecture d'articles, encadrement de thèse) et les **activités périphériques et jugées illégitimes** (direction de laboratoire, recherche de financement, activité organisationnelle et transversale). Ces nouvelles tâches sont intégrées dans la **stratégie de progression de carrière du chercheur** : elles seraient jugées comme un gage de pouvoir, d'autonomie et l'engagement des enseignants-chercheurs dans ces tâches est indispensable au bon fonctionnement de l'établissement. Ainsi, face à cet « élargissement du cœur de métier », certains enseignants-chercheurs assumeraient ces nouvelles tâches afin de progresser dans leur carrière. Également, Courpasson et Guedri (2007) parlent d'une transformation de l'*ethos* de l'enseignant-chercheur et donc de ses valeurs internes. Selon eux, **la performance et les traits caractérisés des enseignants-chercheurs** ont évolué à la suite de la transformation institutionnelle du paysage universitaire et aux paradoxes qu'elle soulève. En d'autres termes, le **centrage sur soi-même et l'individualisation des enseignants-chercheurs** seraient le résultat de la hausse de la compétitivité entre les établissements, des activités multidimensionnelles (pédagogiques, administratives et

bureaucratiques), le management économique de la production de la science, l'évaluation individuelle de l'activité scientifique.

La carrière du scientifique se déroule dans un **nouvel environnement** régi par de multiples **tensions**. En effet, Becquet et Musselin (2004) parlent d'un environnement source de contraintes et Gingras (2008) souligne sa pression accrue. La carrière scientifique est en proie à une **certaine incertitude** (Louvel et Valette, 2014) du fait de la concurrence entre les chercheurs et de la fragilité de la sécurité de l'emploi (Baruch, 2013). La progression de carrière n'est plus régie par l'Etat mais l'établissement universitaire, qui est autonome quant à sa gestion des ressources humaines. Ainsi, le recrutement du chercheur dépend à la fois des **stratégies d'établissement propres** (Musselin, 2010) et des **pressions isomorphiques universitaires**. En effet, DiMaggrio et Powell (1983) décrivent l'isomorphisme institutionnel comme la convergence des organisations d'un même champ, soit comme une homogénéisation interne des établissements. Cette création de norme commune prend trois dimensions : i) l'isomorphisme normatif (structuration d'un champ organisationnel, standardisation des cursus, réseaux professionnels...), ii) l'isomorphisme mimétique (imitation des organisations entre elles à cause de l'incertitude de l'environnement et de leur rationalité limitée), iii) l'isomorphisme coercitif (homogénéisation des organisations par la pression de l'Etat via par exemple le financement publication).

Ayant suscité l'intérêt scientifique chez de nombreux chercheurs de sciences sociales, la carrière scientifique est appréhendée dans la littérature sous différents enjeux. La relation de mentorship, la productivité scientifique ainsi que les mécanismes de choix professionnels sont notamment au cœur de la littérature en tant que déterminants de la progression dans la carrière. Néanmoins, la trajectoire professionnelle du chercheur est un processus complexe et multidimensionnel. Bien que beaucoup d'économistes l'appréhendent comme un processus unique et propre au marché du travail scientifique, certaines études de sciences sociales nous invitent à prendre en compte la multiplicité des carrières dans lesquelles le comportement collaboratif opère. **La prise en compte de la dimension individuelle de la carrière scientifique semble primordiale : la carrière et l'activité de collaboration se déroulent en fonction i) des objectifs et attentes individuels, ii) du contexte historique de la carrière.**

II- Politiques de mobilité géographique et comportement collaboratif du chercheur : interaction entre les enjeux politiques et individuels

Cette seconde partie porte sur un événement professionnel structurant dans la carrière scientifique : la mobilité géographique. Elle correspond à un déplacement du chercheur entre plusieurs universités (mobilité inter-universitaire) ou plusieurs pays (mobilité internationale) avec ou sans retour. Nous montrerons que la mobilité géographique ne peut être seulement associée à un choix individuel puisque de mesures dans l'environnement scientifique incitent le chercheur à être mobile. L'objectif de cette partie est triple : i) **contextualiser le choix de mobilité scientifique** au sein des différentes incitations politiques, ii) comprendre dans quelle mesure la mobilité géographique soulève des **enjeux politiques territoriaux**, notamment en termes d'allocation et de création de connaissances, iii) comprendre dans quelle mesure la mobilité géographique soulève **des enjeux individuels**, notamment en termes de productivité scientifique, de progression de carrière et de géographique des collaborations.

A) Les différentes mesures politiques à la mobilité scientifique

La mobilité géographique (nationale et internationale) des enseignants-chercheurs est un enjeu politique de grande envergure et occupe une place très importante dans la stratégie de développement des établissements d'enseignement supérieur et de la recherche, à l'échelle nationale et internationale. Les programmes de mobilité concernent tous les personnels de l'enseignement supérieur (doctorants, post-doctorants, enseignants-chercheurs). Nous reviendrons sur les **incitations à la mobilité géographique dans leur carrière**, dans le cadre de mesures internationales, nationales et universitaires.

1- La création d'une communauté de chercheurs européens

À la suite du processus de Lisbonne dans les années 2000, la carrière du chercheur s'est vue intégrée dans deux nouveaux espaces collaboratifs européens, ayant pour but de favoriser la mobilité, la création et la diffusion du savoir, et l'attractivité scientifique du territoire européen.

Le premier espace est **l'Espace Européen de l'Enseignement Supérieur**. Officialisée en 2010 à Vienne, la construction de l'Espace Européen de l'Enseignement Supérieur (E.E.E.S) est née d'un accord intergouvernemental européen. Au travers de réunions régulières entre les ministres européens chargés de l'enseignement supérieur, cet espace européen poursuit deux objectifs : i) la création d'un espace « sans frontières » où la mobilité des étudiants et des enseignants chercheurs est encouragée, et ii) la dotation des ressources nécessaires à cet espace afin qu'il gagne en lisibilité et en attractivité mondiale.

Ce nouvel espace est issu d'une succession chronologique d'accords à la suite desquels les Etats européens ont adhéré au projet. En 1998, la France, le Royaume-Uni, l'Allemagne et l'Italie concluent une « harmonisation de l'architecture de l'enseignement supérieur en Europe » dans la Déclaration de la Sorbonne. En 1999, 29 pays européens s'accordent autour du « processus de Boulogne » dans la coordination et la convergence de leurs systèmes d'enseignement supérieur. Plusieurs déclarations ministérielles ont élargi progressivement ce nouvel espace qui compte 48 pays européens depuis 2015. Chacun de ses membres suit les mêmes procédures d'harmonisation des cursus : i) l'établissement d'un système de diplômes comparable et lisible, ii) la refondation du cursus en trois cycles (Licence, Master, Doctorat), iii) la reconnaissance des diplômes avec le système de crédits ECTS, iv) l'orientation du 3^{ème} cycle (doctorat) vers l'espace européen de la recherche. **L'accroissement de la mobilité et le développement de l'ouverture internationale des établissements** font partie des lignes directrices des perspectives d'avenir de l'EEES.

En parallèle, **l'Espace européen de la recherche** est un projet qui a été introduit dans les années 1970 avant d'être concrétisé à la fin des années 1990. Il est officialisé en 2000 lorsque la Commission européenne annonce la création de ce nouvel espace, visant à favoriser la coopération entre les scientifiques européens et lutter contre le morcellement des efforts de recherche. Cette politique européenne est menée dans un souci de cohérence et de concertation, dans le but de nourrir la compétitivité, l'innovation, la coopération en Europe. Dans le cadre de la stratégie de Lisbonne (2000), ce nouvel espace bouleverse la sphère de la recherche européenne avec la **création d'un « marché intérieur » européen de la recherche**, la coordination des politiques de recherche des 27 Etats membres et le financement des activités par l'Union. Ce projet repose sur cinq principes²³ fondateurs issus de l'European Research Area Committee (2007), afin de faire de la connaissance une compétence partagée.

En 2012, la Commission européenne propose des mesures de mise en œuvre de cette politique européenne. Ces mesures visent à encourager les chercheurs à faire carrière en Europe (accroître la transparence et l'ouverture du recrutement sur le marché du travail, réduire les inégalités de genre), optimiser la création et la diffusion des connaissances à l'échelle européenne (améliorer l'efficacité des systèmes nationaux de recherche, accroître la coopération et la concurrence internationale, améliorer la circulation et le transfert du savoir scientifique) pour contribuer à la croissance et à la création d'emplois durables.

La construction de ces deux espaces est fondée sur plusieurs mesures destinées à faciliter les flux migratoires entre les chercheurs européens, telles que l'harmonisation des pratiques de ressources humaines, les services EURAXESS et les actions Marie Skłodowska-Curie.

²³ i) Création de programmes conjoints de recherche entre les états membres, ii) Développement des carrières et de la mobilité des chercheurs via le partenariat entre les chercheurs, iii) Accompagnement dans le transfert de savoir et sur la gestion de propriété intellectuelle, iv) Stratégie de coopération internationale scientifique et technique, v) Cadre juridique européenne pour les infrastructures de recherche des états membres.

L'harmonisation des pratiques de ressources humaines

En 2005, la Commission Européenne a publié la « Charte européenne du chercheur » et le « Code de conduite pour le recrutement des chercheurs » dans le but de favoriser l'attraction des meilleurs chercheurs internationaux et d'encourager les jeunes chercheurs à poursuivre une carrière scientifique européenne. La charte a pour objectif de décrire les rôles, les responsabilités et les droits des chercheurs et des employeurs. De son côté, le code de conduite propose des solutions pour accroître l'équité et la transparence des processus de recrutement des chercheurs. L'adhésion à ces deux documents vise à renforcer l'attractivité des universités et des organismes de recherche vis-à-vis des chercheurs européens dans la mise en place de leur projet de recherche. La commission européenne délivre également le logo « HH excellence in research » pour récompenser les établissements qui stimulent un environnement de travail favorable pour le personnel de recherche. Selon le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, en 2013, 1200 établissements (dans 35 pays) ont signé la charte et le code de conduite et 102 établissements ont reçu le label « RH excellence pour la recherche ». Ces dispositifs s'inscrivent dans la Stratégie de gestion des ressources humaines pour les chercheurs mise en place par l'Union Européenne.

Les services EURAXESS : Faciliter la mobilité internationale entrante et sortante

Le réseau EURAXESS est une démarche européenne qui vise à faciliter la mise en œuvre de partenariats européens de recherche. Il se décline en plusieurs branches²⁴ et constitue un réseau de 42 centres localisés en France qui offrent des services aux chercheurs étrangers qui travaillent en France concernant la préparation du séjour (tâches administratives, recherche de logement) et l'intégration durant le séjour en France (intégration culturelle, cours de français, vie quotidienne...).

Les actions Marie Skłodowska-Curie : Financement des projets de mobilité

Créées par l'Union Européenne et la Commission européenne, les actions Marie Skłodowska-Curie sont un ensemble de programmes de bourses visant à financer plusieurs projets scientifiques dans lesquels la mobilité est incluse. Entre 2014 et 2020, ces bourses sont financées dans le cadre du programme Horizon 2020. Ces actions représentent cinq outils de financement de mobilité.

i) Innovative Training Network : Soutien des réseaux de formation innovante (formations conjointes de recherche, formations doctorales) et des partenariats entre les établissements de recherche (université, instituts de recherche) et les entreprises

²⁴ Euraxess Jobs (offres d'emplois), Euraxess Services (accompagnement des chercheurs dans la préparation d'un séjour à l'étranger), Euraxess Rights (informations relatives à la gestion de carrière), Euraxess Links (services pour les chercheurs européens qui travaillent à l'étranger).

- ii) Research and Innovation Staff Exchange : Soutien de programmes d'échanges de personnel de recherche, de collaborations internationales et intersectionnelles
- iii) Individuel fellowships : Soutien de la formation, de la mobilité internationale et intersectorielle de chercheurs souhaitant enrichir leur potentiel créatif et innovant
- iv) Cofund : Soutien de programmes régionaux, nationaux, internationaux (par exemple : co-financement de programmes doctoraux, post-doctorats...)
- v) European Researchers'Night : Rencontre de chercheurs avec le grand public pour encourager la sensibilisation aux activités scientifiques et encourager la vocation scientifique chez les jeunes.

2- La France et ses programmes internationaux non européens

En parallèle des mesures européennes, la France suit également une stratégie de **développement de partenariats internationaux ayant pour but de favoriser les flux de chercheurs.**

Cette stratégie s'inscrit dans un environnement concurrentiel où chaque pays met en œuvre des **mesures d'attraction de talents afin de valoriser l'innovation interne** : *“Mobility—and international mobility—of skilled human resources plays an important role in innovation. It contributes to the creation and diffusion of knowledge, particularly tacit knowledge, which is more effectively shared within a common social and geographical context.”* (OECD, 2010). Par exemple, l'Etat Français a mis en place un **titre de séjour flexible et spécifique aux chercheurs étrangers** entrants sur son sol. Cette mesure est une initiative menée à la suite de la directive de la Commission Européenne (2016) afin de faciliter les conditions d'entrée et de séjours de chercheurs non ressortissants de l'Union Européenne (harmonisation de la procédure d'accueil, délais supplémentaires de 9 mois sur le territoire pour recherche d'emploi ou création d'entreprise, remplacement de la demande de visa). Ainsi, en 2016, l'Etat octroie à ces chercheurs étrangers la carte « passeport talent », dotée d'une durée maximale de 4 ans. A partir de 2019, de nouvelles mesures plus souples sont inscrites dans la loi. La carte de séjour « passeport talent » est créée avec la mention « chercheur-programme de mobilité ». Selon le ministère de l'enseignement supérieur, de la recherche et de l'innovation français, 7150 visas scientifiques ont été délivrés en France : 34% d'entre eux correspondant à des séjours de court terme (inférieurs à 3 mois) et 66% à des séjours de long terme.

Également, la France est impliquée dans de **nombreux programmes d'échange scientifique et de mobilité**. Tout d'abord, elle participe à des **programmes de mobilité internationale créés à l'initiative d'autres pays**.²⁵ D'autre part, nous trouvons les **programmes déployés à l'initiative de l'Etat**

²⁵ C'est par exemple le cas du programme Fulbright créé par les Etats-Unis en 1946. Il s'agit d'un système de bourses d'études pour les ressortissants des pays partenaires (souhaitant étudier aux Etats-Unis) ou pour les ressortissants américains (souhaitant étudier dans l'un des pays partenaires).

français. Ils sont financés sous la tutelle du ministère des Affaires étrangères et européennes, avec le soutien du ministère de l'enseignement supérieure et de la recherche et sont gérés par l'Agence Campus France (créée en 2010). Ces programmes ont comme objectif d'enrichir et de diversifier le réseau scientifique français via la création et le renforcement de relations collaboratives durables. Le tableau 3 décrit les principaux programmes multilatéraux et bilatéraux de mobilité de chercheurs développés par la France.

Tableau 3 : Programmes nationaux de mobilité de personnel d'enseignement supérieur

	Personnel, pays et discipline concernés	Objectif	Actions du programme
Programmes multilatéraux			
Les partenariats Hubert Curien (PHC) ²⁶	Tout le personnel et toutes les disciplines sont concernés	Développement de projets de recherche entre les laboratoires français et des laboratoires internationaux (privés ou publics)	Financement du « surcoût international » (frais de mobilité encourue par le projet)
Le programme Hermès	Postdoctorants originaires d'Asie, d'Amérique Latine et d'Afrique du Sud Sciences humaines et sociales	Favoriser l'attraction de jeunes chercheurs pour des séjours de recherche de court terme (entre 3 et 9 mois) en France	Financement
Les Bourses Diderot (1989)	Postdoctorants étrangers Sciences humaines et sociales	Soutenir les chercheurs postdoctoraux de étrangers souhaitant réaliser un séjour de long terme dans un laboratoire de recherche français	Financement
Programme PAUSE (2017)	Scientifiques étrangers en situation d'urgence	Inciter les établissements d'enseignement supérieur et les organismes de recherche publique accueillir des scientifiques en situation d'urgence	Financement et un accompagnement dans les démarches administratives.
Programmes bilatéraux			
Accord CAPES-COFECUB (1979)	Doctorants français et brésiliens	Encourager l'échange de doctorants entre le Brésil et la France	Financement des échanges, de missions de court terme et de missions de professeurs universitaires
Accord USP-COFECUB (1994)	Enseignants-chercheurs français et brésiliens	Création d'un système d'échanges académiques entre la France et le Brésil (recherches conjointes)	Financement de projets conjoints et de missions de court terme

²⁶Les principaux critères sont l'excellence scientifique, la nécessité du partenariat international, la capacité de l'établissement français à accueillir ce partenariat (compétences, qualité de l'infrastructure, complémentarité des équipes, participation de doctorants et jeunes chercheurs) ainsi que les résultats (création d'une nouvelle relation collaborative, transfert mutuel de compétences, retombées scientifiques et industrielles). Selon le ministère de l'Europe et des affaires étrangères français, 60 pays avaient mis en place des PHC bilatéraux avec la France en Aout 2019 : les plus anciens programmes datent de plus de 30 ans.

<https://www.diplomatie.gouv.fr/fr/politique-etrangere-de-la-france/diplomatie-scientifique-et-universitaire/le-rayonnement-scientifique-de-la-france-a-l-etranger/article/les-outils-d-encouragement-a-la-mobilite-des-chercheurs>

Programme Zhang Heng	Jeunes chercheurs français spécialisés en maladies infectieuses ou émergentes	Initier les chercheurs français au savoir-faire chinois au cours d'un séjour d'une semaine	Financement du séjour
Programme Bernardo Houssay	Postdoctoraux français et argentins	Favoriser les projets de recherche communs via la mobilité postdoctorale entre la France et l'Argentine	Financement des mobilités
Partner University Fund (2007)	Chercheurs de France et des Etats-Unis	Enrichir la durée des partenariats académiques via des appels à projet de grande ampleur	Financement sous plusieurs formes (diplômes conjoints, mobilité étudiante, visites professorales, publications communes, codirection de thèse...).

Source : https://eacea.ec.europa.eu/national-policies/eurydice/france/mobility-higher-education_fr

3- Autonomie du programme de mobilité universitaire

À la suite des lois Faure (1968) et LRU (2007), les établissements d'enseignement supérieur et de recherche sont chargés de **développer en toute autonomie des partenariats internationaux pour favoriser la mobilité étudiante et des chercheurs**. Ainsi, chaque établissement dresse sa propre stratégie de politique internationale en fonction de ses objectifs de développement. La politique internationale est considérée par ces établissements comme un **levier de compétitivité** face à la compétition internationale et de gain en visibilité. Elle constitue l'un des critères mobilisés dans l'évaluation du Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (HCERES). En effet, dans son référentiel d'évaluation externe des établissements d'enseignement supérieur et de recherche (Campagne d'évaluation 2019-2020), le HCERES note explicitement l'organisation et le **pilotage de la politique d'internationalisation** de l'établissement dans ses critères d'évaluation: « *En matière de recherche, l'établissement soutient la coopération multilatérale, ainsi que l'activité menée sur projets et contrats, notamment dans le cadre de l'espace européen de la recherche et de l'innovation (...) L'établissement encourage et soutient la mobilité internationale, entrante et sortante, des étudiants et de son personnel. Il développe des dispositifs de préparation à la mobilité sortante et d'accueil dans le cadre de la mobilité entrante. Il intègre une dimension sociale à sa politique de mobilité étudiante* ».

La sensibilisation à la mobilité internationale est mise en œuvre dès lors que le chercheur est au stade d'étudiant. L'établissement d'enseignement supérieur peut choisir de participer à **des programmes et des réseaux de partenariat internationaux** qui facilitent la mobilité de ses étudiants. En effet, il existe des programmes multilatéraux²⁷ ou bilatéraux entre la France et un autre pays²⁸.

²⁷ Par exemple le programme Erasmus, Assistant Comenius

²⁸ Par exemple le programme Fulbright, Entente Cordiale

Les universités développent également de nouvelles mesures afin d'enrichir la **dimension internationale des compétences** de leurs étudiants :

i) Les premières mesures concernent **l'harmonisation des cursus européennes**. Depuis 2002, le système universitaire européen est structuré autour de trois diplômes : la licence (bac+3), le master (bac+5) et le doctorat (bac+8). Cette harmonisation européenne des cursus d'enseignement supérieur a pour but de favoriser la mobilité des étudiants européens et d'accroître la visibilité des formations à l'échelle européenne. Le système d'accréditation des ECTS a été développé dans la même optique d'harmonisation des cursus européens au travers de crédits qui sont transférables et capitalisables.

ii) Ensuite, certaines mesures concernent le développement de **la visibilité internationale et l'attractivité de la formation doctorale**. Depuis le décret de 2005, la cotutelle de thèse et le diplôme conjoint avec un autre pays que la France est possible dans le cadre de partenariats internationaux. Cette nouvelle réforme est une opportunité pour les universités pour accroître ses collaborations internationales et l'accueil de doctorants, de maîtres de conférences et de professeurs étrangers. Plusieurs mesures sont donc suivies par les établissements pour atteindre cet objectif, telles que le recrutement d'enseignants étrangers, la mise en place d'un dispositif d'accueil et d'accompagnement des étudiants et personnels étrangers, le renforcement de la lisibilité de l'offre doctorale (cahier des charges, regroupement des pôles de recherche et d'enseignement supérieur).

iii) De plus, les universités ont la possibilité de **développer des certifications linguistiques** afin d'accroître la visibilité des compétences linguistiques en tant **compétences transversales**, aussi bien pour les étudiants que pour le personnel de recherche. Créée en 1995, le diplôme de compétences en langues (DCL) est une certification nationale professionnelle pour laquelle le candidat est évalué via une mise en contexte dans le monde du travail. Ce diplôme concerne 13 langues et est aligné sur les niveaux du Cadre Européen Commun de Référence en Langues (CECRL). Aussi, le certificat de compétences en langues de l'enseignement supérieur (CLES) a été créé en 2000 qui est aussi dans le cadre du CECRL. Il a pour but d'évaluer les compétences opérationnelles des étudiants : le diplôme concerne 9 langues. Enfin, d'autres diplômes sont proposés aux étudiants, tels que le TOEIC ou le TOEFL en anglais.

Concernant la mobilité des chercheurs, l'université développe également **un réseau de partenariats scientifiques**. Elle gère de manière autonome **les projets de mobilité** (via un organisme spécialisé au sein de l'université) et **le financement de ces programmes** (via une « aide financière à la mobilité » issue du budget de l'université). Le financement des programmes de mobilité des chercheurs est également octroyé par l'Agence Nationale de la Recherche et des bourses d'aides émises par les collectivités territoriales (aux échelles municipale, départementale et régionale).

Ces différentes mesures incitatives sont au cœur de la politique de développement scientifique et d'innovation de la France et des universités. Elles **structurent les carrières de chercheurs** en favorisant **la mobilité scientifique** dans l'espoir de développer des externalités positives. Les effets de ces politiques ont fait débat dans la littérature, notamment sur les effets de réallocation et de transformation des connaissances à l'échelle territoriale. Ces effets induits structurent l'environnement dans lequel les chercheurs développent leur carrière et leurs relations de co-authorship.

B) Incitations politiques de la mobilité : effets attendus à l'échelle universitaire et territoriale

La mobilité scientifique est donc au **cœur de plusieurs dispositifs politiques**. Par exemple, Morano-Foadi (2005) souligne l'« expectation of mobility » commune à de nombreux programmes européens centrés sur l'accentuation de la mobilité scientifique au nom d'enjeux individuels et territoriaux. Néanmoins, de nombreux chercheurs soulignent les limites de ces politiques, du fait d'une inéquation entre les **mesures employées** (les résultats attendus) et les **effets territoriaux induits** (résultats réels). Tout d'abord, certains chercheurs dénoncent le **manque de coordination des politiques européennes de migration scientifique**. En effet, Morano-Foadi (2005) pointe ce manque de coordination au niveau géopolitique (entre les pays membres) et institutionnel (réintégration difficile dans le système universitaire des chercheurs mobiles à la suite de leur retour). La mobilité européenne apparaît donc comme une expérience non valorisée dans la sphère universitaire, ce qui incite les chercheurs à s'orienter vers le privé. Cañibano (2006) souligne également l'importance de la coordination et de la gestion des liens entre les organisations impliquées dans le flux de mobilité : selon elle, il s'agit d'une condition nécessaire pour que l'expérience de mobilité scientifique permanente soit efficiente. Arrieta *et al.* (2017) soulignent les limites de l'expansion de l'Espace Européen de la Recherche entre 2004 et 2007 : ils ont montré un impact négatif sur la collaboration européenne, notamment due à un déséquilibre migratoire et à un manque de politiques de coordination. D'autres chercheurs dénoncent **la rigidité de ces politiques quant à la perception de la mobilité**. En effet, Ackers (2008) traduit la perception de la mobilité en tant que phénomène stable et unique comme une erreur politique et une pression contreproductive : elle invite à la valorisation qualitative des différentes expériences de mobilité. Aussi, Cañibano (2006) préconise une **meilleure reconnaissance des mobilités temporaires**.

Bien que la mise en place de ces politiques fasse débat, la mobilité scientifique est considérée par les Etats et les territoires comme un **vecteur de transmission de connaissances**, via la proximité physique des chercheurs (Collins, 1974). En effet, bon nombre de chercheurs ont souligné la dimension géographique de la transmission de savoir tacite qui est propre à un contexte géographique (Polanyi, 1958, 1962, Feldman et Lichtenberg, 1998, Grossetti et Bès, 2002). Également, la proximité géographique permet la circulation de connaissances interdisciplinaires et la résolution de problèmes de

coordination au sein de l'équipe de recherche (Rallet et Torre, 2001). Finalement, Cañibano (2006) décrit la mobilité scientifique comme un processus de propagation de flux de connaissances de deux dimensions : i) les connaissances tacites (propres à l'individu) qui évoluent entre les individus au fil des mobilités sous le « modèle de segmentation » (Sanz-Menendez et Cruz-Castro, 2004), ii) les connaissances « sociales » (propres à la communauté) qui évoluent entre les entités au fil des mobilités sous le « modèle de circulation ».

Outre les répercussions à l'échelle individuelle, la mobilité scientifique est considérée comme une **clef de développement économique** via la circulation de connaissances à l'échelle du territoire. En effet, Cañibano (2006) écrit : « *Cette culture serait essentiellement dédiée à faire de la mobilité professionnelle un mécanisme d'enrichissement personnel des chercheurs « mobiles », mais aussi à faire profiter la société autant que possible de ces « mouvements »*. Nous présenterons deux perceptions politiques de la mobilité scientifique, chacune associée à un mécanisme de diffusion de connaissances différent. La première conception est l'approche néoclassique. Issue des années 1960, elle associe la mobilité à une **réallocation mondiale des connaissances** au travers du concept de « *brain drain* ». La seconde conception, inspirée de l'approche évolutionniste et développée dans les années 1990, associe la mobilité à une **transformation territoriale des connaissances**.

1- Approche néoclassique : la mobilité en termes de réallocation des connaissances

La théorie néoclassique est fondée sur le principe d'un **mécanisme autorégulateur** du modèle d'équilibre général, dans un cadre de concurrence pure et parfaite (Knight, 1921). Dans le cas où les individus sont considérés comme des agents homogènes, le chercheur est perçu comme un être autonome, maximisateur de revenu et de capital humain. Il répond aux signaux du marché et fait ses propres choix professionnels : la mobilité scientifique est une expérience professionnelle qu'il choisit (ou non). Selon Beckers (1993), « *knowledge is not subject to diminishing returns in the same way as is physical capital because greater knowledge raises the productivity of further investment in knowledge* ». Dans cette optique, les **chercheurs** ne sont plus homogènes mais **hétérogènes** en fonction de leur **niveau de capital humain** (qui n'est plus considéré comme un bien physique). Les économies d'échelles de la production scientifique varient selon leur stock de capital humain et de connaissances. En effet, la compétence du chercheur augmente sa propre productivité individuelle (bénéfices privés) ainsi que celle de la société dans laquelle il se trouve (bénéfices sociaux). Le marché peut devenir inefficace dans le cas où le chercheur n'est pas rémunéré en fonction de l'ensemble des bénéfices qu'il génère. Dans les pays à faible revenu, cela se traduit par une perte de bien-être et le départ des chercheurs vers des économies où leur revenu sera davantage corrélé aux bénéfices privés et sociaux qu'il génère. Ainsi, dans cette approche, **la mobilité scientifique est le résultat d'une quête de maximisation du revenu**

suivie par les chercheurs qui sont sous-rémunérés compte tenu des externalités qu'ils génèrent dans leur société d'origine (Cañibano et Wolley, 2015).

La connaissance est perçue comme un bien attaché à l'individu : la mobilité scientifique traduit donc une relocalisation des connaissances. En effet, Cañibano écrit "*In this approach, embodied knowledge travels automatically in the same direction than its human career and is automatically and efficiently used in its new destination*". Désireux d'accroître sa productivité et son revenu, le chercheur choisit de réaliser plusieurs expériences professionnelles (telles que la mobilité) pour enrichir son capital humain, créant ainsi des flux de connaissances entre son établissement d'origine et sa nouvelle destination. Gibson et McKenzie (2012) décrivent ce processus comme étant unilatéral: «*diffusion of knowledge occurs in part because movers bring knowledge to the destination and in part because mobile scientists establish collaborations with and across the local communities they bridge* ».

Vertus territoriales attendues de la mobilité

Les territoires sont donc en concurrence dans **l'attraction des scientifiques** afin d'accroître leur stock de capital humain et les externalités positives internes qui en découlent. Grubel et Scott (1966) font partie des premiers chercheurs à parler de « *brain drain* » (traduit comme la « fuite des cerveaux »). Ils décrivent **l'autorégulation de ce système de migration** de personnel qualifié : les pertes de capital humain du pays d'origine (pays en développement) seraient temporaires dans le sens où elles seraient vite remplacées par de nouveaux talents internes (via l'éducation) ou immigrés. Ce système de flux encouragerait donc l'éducation au sein des pays en développement : les individus y voient une opportunité d'accroître leur revenu via l'accumulation de capital humain individuel (Stark *et al.*, 1998). Cattaneo *et al.* (2019) ont d'ailleurs montré que la mobilité ne se traduisait pas seulement par le départ des chercheurs les plus qualifiés (vers des instituts d'excellence) mais aussi par celui des moins compétents (vers des établissements non axés sur la recherche). A partir d'une analyse de trajectoires de carrières de docteurs italiens en économie, finance et sciences de gestion, ils ont montré que la mobilité n'amputait pas forcément le niveau de capital social du pays d'origine : dans le cas de l'Italie, le niveau des chercheurs immigrants était proche des chercheurs italiens restant dans le pays. Ainsi, ces mécanismes d'autorégulation permettent aux économies de départ et d'arrivée de développer une croissance durable de leur productivité (Stark, 2004, Saxenian, 2002).

Les territoires et les établissements de recherche élaborent donc des **politiques d'attractivité de talents étrangers dans un climat concurrentiel**. Mahroum (2000) décrit dans quelle mesure l'intégration de nouveaux talents accroît le prestige des instituts. Tout d'abord, les instituts accumulent la réputation et la notoriété associées aux chercheurs entrants. Ensuite, les instituts ont accès à de nouvelles ressources d'innovation (réseaux de co-authorship des chercheurs immigrants, ressources financières, ressources physiques et humaines). Il décrit cette quête de prestige des instituts comme un cercle vertueux à partir du mécanisme de l'« effet Matthieu » (Merton, 1973) : la croissance de la réputation de l'institut et

l'accès à de nouvelles ressources nourrissent de nouvelles opportunités quant à l'intégration de nouvelles élites. Le « brain drain » se traduit donc par la **formation de « global spaces »** (Mahroum, 2000) : il s'agit de réseaux de pairs territoriaux, construits par la mobilité, disposant d'un fort pouvoir magnétique quant à l'attraction d'élites (forte réputation, forte dotation de chercheurs à forte notoriété).

Le « brain drain » serait donc un mécanisme qui **profiterait à la fois au pays d'origine** (via l'autorégulation du marché) **et au pays d'arrivée**. Le modèle internationaliste porté par Jonshon (1965, 1968) dresse le bilan global des effets de ce processus en termes de gains et de pertes pour les deux pays. L'objectif de ces travaux est d'apporter une vision globale du phénomène, non centrée sur les effets directs négatifs sur le pays d'origine. Tout d'abord, il réfute l'expression « brain waste » en soulignant que le « brain drain » ne traduit pas une perte de talents pour le pays d'origine. En effet, le chercheur migrant s'insère dans un nouvel environnement plus en adéquation avec son capital humain alors qu'il n'avait pas de place dans le pays d'origine (Jonshon, 1968). De plus, le flux migratoire ne traduit pas une perte de connaissance globale mais plutôt un enrichissement via l'accès à de nouvelles ressources. Les bienfaits de ce flux découlent sur les deux pays via de nouvelles découvertes scientifiques (Jonshon, 1965). Enfin, le chercheur migrant peut toujours influencer la politique de son pays d'origine (Johnson, 1965) et garder des liens avec lui (Jonshon, 1968) pour compenser les pertes temporaires occasionnées. Des travaux empiriques s'inscrivent dans ce modèle internationaliste en prouvant que les gains mondiaux du « brain drain » sont supérieurs aux pertes. Par exemple, Beine *et al.* (2008) ont fait une analyse empirique sur 127 pays en développement pour évaluer les effets du « brain drain » : ils ont notamment montré que les gains pour la Chine, l'Inde et le Brésil sont supérieurs aux pertes supportées par les autres pays.

Dommages territoriaux de la mobilité

De nombreux travaux critiquent les politiques de « brain drain » mises en place : ils portent surtout sur **les effets négatifs induits dans les pays en développement**. Tout d'abord, certains chercheurs pointent le risque de « brain waste » (traduit comme « gaspillage des cerveaux »). Bhagwati et Dellafar (1973) notent notamment le risque d'inflation des salaires dans les pays en développement. Dans ces territoires, le niveau des salaires des chercheurs hautement qualifiés est aligné sur celui des pays développés afin d'accroître l'attractivité de talents. Cette hausse salariale va avoir trois conséquences directes. Le salaire étant au-dessus du prix d'équilibre dans le marché du travail, le chômage des chercheurs nationaux va s'accroître. Aussi, la hausse salariale va inciter les étudiants à s'orienter vers les carrières scientifiques, ce qui va nourrir le chômage de cette profession. Enfin, la hausse salariale académique risque de contaminer le reste de l'économie. Le « brain drain » se traduirait par i) un gaspillage des ressources scientifiques dans les pays en développement, où la diffusion de la science se restreindrait (Bhagwati, 1979) et ii) le développement du chômage et de l'inflation sur l'ensemble de l'économie. D'autres chercheurs **soulignent les enjeux d'endogénéité de la croissance**. Wong et Yip (1999) élaborent un

modèle endogène de croissance dans lequel la migration a un effet négatif, du fait de la chute du capital humain. Cette fuite des ressources innovantes peut également entraîner un changement de la composition du capital humain dans les pays en développement. En effet, afin d'être attractif sur le marché international, les chercheurs des pays en développement ont tendance à développer un capital humain qui n'est pas adéquat avec les enjeux de leur pays d'origine mais avec ceux des pays développés (Di Maria et Stryzowski, 2009).

En parallèle du modèle internationaliste de Johnson, Boulding (1968) développe un modèle nationaliste pour décrire les effets du « brain drain » en termes de **perte de capital humain dans les pays en développement**. Ce modèle est basé sur des hypothèses préliminaires à l'approche évolutionniste : i) les individus sont considérés comme hétérogènes, indépendants, interconnectés et incertains, ii) l'encastrement des individus, de la productivité et de la valeur du capital humain dans un contexte temporel, social et économique, iii) les facteurs de production sont complémentaires. Boulding (1968) ne considère pas le développement d'un pays comme dépendant de son stock de connaissances mais plutôt de sa capacité d'apprentissage. Ainsi, le système de création et de diffusion de connaissances du pays est dynamique et interactif : il repose sur l'intégration des chercheurs dans la structure de capital humain. En d'autres termes, la fuite de certains scientifiques qui étaient au préalable centraux dans ce système peut paralyser le développement scientifique du pays. Les conséquences sont de plusieurs ordres : la fuite de l'investissement public (par lequel le chercheur émigrant a été formé), le bouleversement de la structure du capital humain du pays (connexion et complémentarité des chercheurs), la perte d'apprentissage possible auprès des étudiants. Les flux migratoires de chercheurs bouleversent le système scientifique du pays, soit via le départ de chercheurs centraux qui structurent le système (Kidd, 1965 ; Patinkin, 1968), soit via l'introduction de nouvelles recherches non adaptées à la structure et qui ne sont pas transférables au pays d'origine (Kidd, 1965).

La mise en évidence des limites de l'approche néoclassique quant à la conception de l'activité scientifique a encouragé le **développement des approches évolutionnistes**. Bozeman *et al.* (2001) souligne notamment que le capital humain a une dimension connectée (issu de réseaux sociaux et contextuels). Aussi, Meyer (2001) décrit l'activité scientifique comme l'interconnexion d'individus dont les connaissances proviennent de leurs interactions (par exemple dans le cadre de diaspora de connaissances).

2- Approche évolutionniste : la mobilité est source de création de connaissances

Contrairement à l'approche néoclassique, l'approche évolutionniste contextualise **l'activité scientifique dans le cadre de réseaux** où les composants interagissent. L'approche de Latour (1987, 1992, 1996, 1999) illustre cette nouvelle conception de la pratique scientifique : il la décrit comme un processus de construction de réseaux, dotés de composants humains et non humains. Dans la théorie de l'acteur-

réseau, le réseau scientifique est considéré comme une mobilisation systématique de ressources humaines non homogènes au sein de « centers of calculation » : l'accumulation des ressources dans ces espaces se traduira par la création de nouvelles connaissances. Les ressources mobilisées dans la création de connaissances (dites « actants ») sont très diverses : il peut s'agir de composants humains, matériels et immatériels (incorporés dans le processus via la mobilisation, la traduction, la transformation : on parlera donc de ressources socio-matérielles). **La mobilité accroît donc le potentiel accès à des ressources nécessaires au processus de création de connaissance.** En effet, le désir individuel de mobilité du chercheur est considéré en soi-même comme une ressource immatérielle d'innovation, tout comme les infrastructures, les ressources matérielles et socio-matérielles issues de la mobilité et intégrées dans le processus de production de connaissance.

Ainsi, c'est dans ce nouveau climat scientifique que l'approche évolutionniste a été développée, en adoptant une approche méso-économique. Dans cette perspective, Cañibano (2017) définit **la mobilité comme une transformation des connaissances.** Le chercheur est considéré comme le nœud d'un réseau scientifique et interagit avec d'autres composants du réseau avec qui il va échanger des connaissances. La connaissance n'est plus considérée comme un bien propre à chaque individu mais comme une structure : le réseau d'interaction de différents contextes. Callon (1991) définit le capital humain comme « *un réseau de composants complémentaires au sein de réseaux techno-économiques* ».

Vertus territoriales attendues de la mobilité

Plusieurs travaux insistent sur les bienfaits de la mobilité scientifique dans la **création de nouvelles connaissances.** Tout d'abord, certains chercheurs considèrent la sédentarité géographique comme néfaste pour l'activité scientifique en termes de reproduction de connaissances acquises ou d'encastrement dans une loyauté institutionnelle (Aschenbruck *et al.*, 2008). Ensuite, certains chercheurs soulignent la dimension géographique des flux de connaissances et voient dans la mobilité un élargissement territorial du réseau de création de connaissances. Agrawal *et al.* (2006) soulignent que les villes fréquentées sont beaucoup plus structurantes pour les flux de connaissances que les villes non fréquentées à cause des relations personnelles développées dans un contexte de co-localisation. Dans le cadre d'une analyse de 469 questionnaires et CV de chercheurs portugais, Fontes *et al.* (2013) ont montré que la co-localisation crée une proximité sociale, cognitive et organisationnelle qui favorise la transmission de connaissances.

La mobilité est un **canal de transmission de connaissances aussi bien pour le territoire d'origine que pour le territoire d'accueil.** Gibson et McKenzie (2013) ont notamment évalué l'impact de l'émigration du personnel hautement qualifié du Pacifique (Nouvelle Zélande, Papouasie Nouvelle Guinée, Tonga) sur les réseaux de connaissances nationaux. A partir de 1808 réponses à un questionnaire, ils ont étudié le cas de mobilité dite « nécessaire » étant donné les faibles perspectives de poursuite de carrière sur le territoire, et ont constaté que le retour de ces chercheurs était un vecteur de

connaissance crucial pour ces territoires : les chercheurs migrants réintègrent leur pays d'origine avec des réseaux de co-authorship et de connaissances enrichis. Également, Jöns (2009) souligne le rôle clef qu'ont eu les scientifiques et les ingénieurs américains immigrants en Allemagne entre 1954 et 2000. Dans ce contexte, la mobilité scientifique est apparue comme un processus vertueux qui a permis la reconstruction scientifique du pays et sa réintroduction dans la communauté scientifique internationale.

La mobilité **redéfinit les frontières et l'identité du capital humain via l'instauration de réseaux transfrontaliers**. Kim (2010) présente le concept de « transnational identity capital » comme le capital propre aux académiques qui sont mobiles à l'échelle internationale. Bien que la mobilité soit une expérience commune et partagée par beaucoup de scientifiques, elle crée de nouvelles connaissances en fonction de « l'identité transnationale » du réseau. Kim (2017) note la naissance de « tribus et territoires transnationaux » dans le processus de création de connaissances.

Limites de la mobilité à l'échelle territoriale

Le retour du chercheur mobile dans son institution d'origine n'est pas toujours perçu par l'institution comme une opportunité de développement de réseau de connaissances. Melin (2005) étudie les difficultés liées au retour de séjour de 284 post-doctorants suédois dans leur institution via des entretiens et des questionnaires. Il exprime la déception ressentie par ces chercheurs mobiles, notamment sur la difficulté liée au transfert de connaissances avec leur institution d'origine suite au retour. Gilbson et McKenzie (2014) soulignent que les gains de mobilité ne sont pas conservés une fois que le chercheur retourne à son institution d'origine (en termes de financement de la recherche, manque de flexibilité du poste pour maintenir les relations avec d'autres pays). Ainsi, certains travaux montrent que le retour est très peu pratiqué dans certains pays, le transfert de connaissance suivant donc un autre canal que la mobilité. Par exemple, Fontes (2007) a montré que parmi les 55 inventeurs portugais étudiés, très peu considèrent le retour dans leur pays d'origine comme une option viable. Ils maintiennent des relations collaboratives avec le Portugal via d'autres moyens que la mobilité permanente (contrats occasionnels, diaspora, échanges temporaires de chercheurs...). Gedeshi et King (2019) étudient le cas de doctorants albanais (l'Albanie connaissant un fort taux d'émigration) : les échanges de diaspora sont privilégiés au retour dans leur pays d'origine. Certains chercheurs font des préconisations politiques pour favoriser le retour de chercheurs sur le territoire d'origine. Thorn et Holm-Nielsen (2008) préconisent deux types de politique pour favoriser le retour des scientifiques : l'une est relative aux individus (politique de court terme qui structurent les décisions des chercheurs) et l'autre est relative à l'environnement de recherche (politique de long terme qui apporte des solutions aux problèmes institutionnels et contextuels). Barrufaldi et Landoni (2012) insistent sur le maintien du lien entre le chercheur mobile et son pays d'accueil (afin que les relations persistent une fois que le chercheur est rentré dans son pays d'origine) via la diaspora ou la mobilité temporaire (Tani, 2008).

Enfin, d'autres travaux mettent en lumière **d'autres formes de la proximité dans le processus de création de connaissance**, soulignant que la mobilité géographique n'est pas toujours utile. Fontes et al (2013) identifient trois autres formes de proximité qui permettent la transmission de connaissances : i) la proximité sociale (attributs sociaux, liens sociaux), ii) la proximité cognitive (compréhension et absorption de connaissances), iii) la proximité organisationnelle (codes de comportement, hiérarchie, règles). Ainsi, la **création de « communautés épistémiques » et la mobilité temporaire** sont d'autres vecteurs de création de connaissances que la mobilité permanente. Hussler et Ronde (2005) ont étudié les déterminants de la diffusion spatiale des connaissances à l'université. Ils apportent une dimension polymorphique au concept de « proximité ». En effet, ils soulignent que les chercheurs universitaires disposent d'une forte proximité organisationnelle et donc que la proximité géographique n'est pas nécessaire à la diffusion et création technologique. Jöns (2007) explique que chaque matériel est plus ou moins contextualisé géographiquement. Ainsi, la mobilité n'est pas nécessaire pour avoir accès aux ressources qui sont immatérielles, aux processus standardisés, aux processus de productions simples de connaissances. Enfin, Horta et Yudvich (2016) soulignent les vertus que peut avoir la **sédentarité géographique** dans le développement de la structure de recherche (promotions internes, identité institutionnelle, faibles attentes de salaires, sensibilisation aux spécialités du laboratoire). Ainsi, à partir d'entretiens avec des experts portugais et russes, ils montrent que la mobilité forcée peut s'avérer contre-productive et préjudiciable au développement des laboratoires.

3- Synthèse : deux conceptions politiques de la mobilité

Le tableau ci-dessous synthétise les principaux concepts de la mobilité utilisés dans l'élaboration des politiques universitaires, nationales et internationales de mobilité scientifique. L'objectif est de résumer les différents mécanismes opérés, ainsi que les attentes et les limites de ces stratégies politiques à l'échelle territoriale.

Tableau 4 : Synthèse des conceptions néoclassique et évolutionniste de la mobilité géographique

	Approche néoclassique :	Approche évolutionniste :
	La mobilité est une réallocation des connaissances	La mobilité est une transformation des connaissances
<i>Apparition du courant</i>	Années 1960 Chicago School of Economics	Années 1990
<i>Angle d'approche</i>	Méthodologie individualiste	Méso-économique
<i>Perception du scientifique</i>	Agent économique individuel qui réalise des choix rationnels	Le nœud d'un réseau

<i>Perception de la connaissance</i>	Résultat (« output ») des choix de formation, d'activités professionnelles, de mobilité réalisées par l'agent économique Il s'agit d'une agrégation de compétences	Résultat de différents contextes , elle est dépendante du réseau des hommes et de ses supports Il s'agit d'une structure
<i>Définition du capital humain</i>	Un bien appropriable, un « input » de la fonction de production scientifique	« Un réseau de composants complémentaires au sein de réseaux techno-économiques » (Callon)
<i>Mécanismes de la mobilité scientifique</i>	En amont, le capital humain suit automatiquement les flux de carrière des individus sur le marché du travail . En aval, ces mouvements de marchés assurent la productivité du marché scientifique	Il s'agit d'une réadaptation et recombinaison des composants complémentaires du réseau via le changement de la structure du réseau
<i>Enjeux des politiques publiques</i>	Face au risque du « jeu à somme nulle », le territoire accroît le stock de son capital humain (formation, attraction et maintien de scientifiques stars). Faire face à la concurrence et à la course internationale aux talents scientifiques	Prendre en compte l' hétérogénéité des acteurs et réseaux. Approche « connectionniste » : création de diasporas et mise en valeur de la mobilité temporaire
<i>Gains territoriaux attendus de la mobilité</i>	<ul style="list-style-type: none"> - Les pertes de capital humain sont temporaires : autorégulation dans les pays d'origine - Elaboration de pôles d'attractivité magnétiques dans les pays d'accueil - Réallocation optimales des connaissances globales (gains mondiaux) 	<ul style="list-style-type: none"> - Elargissement du réseau de connaissance national pour les territoires d'origine et d'accueil - Vertus de la proximité géographique dans le transfert de connaissances - Création de capital et de communautés transnationaux
<i>Domages et limites territoriaux de la mobilité</i>	<ul style="list-style-type: none"> - « Brain waste » (inflation, chômage, gaspillage des ressources innovantes) - Endogénéité de la croissance (perte et transformation du capital humain) - Bouleversement de la structure du capital humain dans le pays d'origine - Non prise en compte de la dimension connectée des connaissances 	<ul style="list-style-type: none"> - Difficulté du retour - Alternatives à la proximité géographique - Effets pervers de la mobilité forcée dans la structure du réseau

Réalisé par l'auteur

Sources : Cañibano, C., & Woolley, R. (2015). Towards a socio-economics of the brain drain and distributed human capital. *International Migration*, 53(1), 115-130.

Cañibano, C. (2017). Scientific mobility and economic assumptions: From the allocation of scientists to the socioeconomics of network transformation. *Science as Culture*, 26(4), 505-519.

Les politiques de mobilité ont des effets territoriaux positifs et négatifs en termes de réallocation et de transformation des connaissances. Outre ces enjeux à l'échelle méso et macroéconomique, les politiques de mobilité structurent également **les choix individuels des chercheurs**. En effet, la mobilité scientifique est une partie intégrante de la sphère académique, présentée aux chercheurs comme un

vecteur de productivité, d'accès aux promotions et d'internationalisation des collaborations. Ces facteurs individuels de mobilité font débat dans la littérature, notamment entre les effets escomptés et les effets réels sur la performance scientifique et la géographie des collaborations.

C) La mobilité scientifique : quels effets sur l'activité scientifique et la géographie des collaborations ?

Les enjeux politiques de la mobilité se dessinent aussi bien à l'échelle territoriale qu'individuelle. En effet, la mobilité géographique est une composante de l'activité scientifique et est encouragée pour la progression de la carrière scientifique. Ackers (2008, p413) souligne que les politiques européennes visent à **associer l'excellence du chercheur à sa mobilité géographique** "*European policy at the present time tends to conflate different forms of mobility and promotes the use of the concept as a proxy for internationalization, excellence and competitiveness. ... This implies both that 'excellence' is mobile (that 'excellent' scientists are equally footloose*". Également, **plusieurs mécanismes de récompense** nationaux sont mis en place pour valoriser l'expérience de mobilité dans la carrière. Par exemple, dans les articles 39 et 55 du décret de 1984 (fixant les dispositions statutaires des enseignants chercheurs ainsi que le statut particulier des maîtres de conférences et professeurs), une mobilité internationale d'au moins deux ans ou d'au moins un an dans la communauté européenne ou dans l'espace économique européen est valorisée par une bonification d'ancienneté d'un an. Cette valorisation concerne également tout le personnel des établissements publics scientifiques et technologiques (article 11 du décret de 1983). La mobilité serait donc associée comme une caractéristique de la productivité du chercheur (Van Heeringen et Dijkwel, 1987) qui dégagerait de multiples externalités locales via le développement de relations personnelles (Powell et Brantley, 1992), Almeida et Kogut (1997) et Zucker *et al.* (1994, 1998).

Outre la quête de prestige et de reconnaissance (Mahroum, 2000), **plusieurs motivations** incitent le chercheur à choisir la mobilité géographique (Bauffaldi et Landoni, 2016 ; Saint-Blancat, 2017). Le choix individuel de mobilité s'articulerait autour d'une **dimension structurelle, contextuelle et individuelle** (Delicado, 2010).

Tout d'abord, cette pratique est profondément ancrée dans la **sphère universitaire**. Polanyi (1958, 1962) décrit la mobilité comme un composant de la culture de la recherche : elle développe la créativité du chercheur via l'accès à la diversité et à l'hétérogénéité des connaissances. Aussi, Bourdieu (1986, 1988) décrit la mobilité comme un « habitus » de la recherche qui permet l'accumulation de « propriétés » (capital culturel et social) et accroît la compétitivité du chercheur sur le marché. La nécessité de mobilité s'exprime ainsi via des facteurs propres (Mahroum, 2000) relatifs à la socialisation professionnelle et à la quête de références professionnelles, et des facteurs motivationnels (Rouse, 1987) relatifs au système méritocratique universitaire.

Ensuite, **la stratégie de mobilité diffère en fonction du pays d'origine du chercheur**. Par exemple, Ackers (2005) souligne que tout chercheur mobile est à la recherche d'une optimisation de revenu et des conditions de travail. En revanche, selon elle, les chercheurs issus d'un pays peu développé scientifiquement sont à la recherche de meilleures conditions de travail alors que ceux issus d'un pays développé souhaitent gagner en crédibilité (un haut niveau de mobilité étant demandé dans la progression de carrière).

Enfin, **les facteurs individuels influencent également la stratégie de mobilité**. Ackers (2005) exprime par exemple que la temporalité des flux de mobilité sont influencés par le parcours de vie, la vie familiale et le genre du chercheur. Cañibano *et al.* (2016) montrent notamment que les comportements de mobilité diffèrent entre les genres : les femmes sont plus mobiles que les hommes mais à plus faible fréquence, plus tôt dans la carrière, dans des séjours plus courts et des destinations plus proches.

Plusieurs méthodologies quantitatives, qualitatives et mixtes sont utilisées pour identifier et interpréter la trajectoire géographique des chercheurs. Le curriculum vitae (CV) est un outil mobilisé dans plusieurs travaux (Wolley et Turpin, 2009 ; Cañibano et Bozeman, 2009 ; Cañibano *et al.*, 2011) qui permet de détecter aussi bien les flux migratoires de court terme que de long terme. Également, de nouvelles approches méthodologiques misent sur la création d'indicateurs bibliométriques de mobilité via la reconstruction de la trajectoire géographique à partir des affiliations des chercheurs dans les publications (Robinson-Garcia *et al.*, 2016 ; Aman, 2018 ; Robinson-Garcia *et al.*, 2019). Enfin, le recours aux enquêtes et aux entretiens permet l'interprétation des mobilités et ses effets induits (Gibson et McKenzie, 2014 ; Veugelers et Van Bouwel, 2015 ; Saint-Blancat, 2017).

1- Débats autour du recours de la mobilité géographique comme indicateur de performance

Une grande partie des travaux vise à tester la corrélation entre la mobilité et la productivité du chercheur. Pour décrire cette littérature, nous utiliserons les indicateurs relatifs à la performance scientifique issus des travaux de Costas et Bordons (2005) :

i) Le premier concerne le **niveau d'activité** : la productivité scientifique sera donc évaluée en fonction du nombre de publications nationales, internationales du chercheur ainsi que la part de publications internationales dans l'ensemble des travaux.

ii) Le second type d'indicateurs concerne la **visibilité des publications**. Ils dissocient la visibilité « espérée » (avec les indicateurs d'« impact factor » et de part des articles dans les revues à fort impact) et la visibilité « réelle » (associée au nombre de citations par auteur et par article).

iii) Enfin, la troisième série d'indicateurs concerne les **habitudes de publication** (associées à l'ordre d'apparition dans la liste des co-auteurs, à la durée de la vie professionnelle).

A partir de ces indicateurs de performance, le tableau 5 synthétise des articles empiriques portant sur la corrélation entre la mobilité et la productivité du chercheur. Nous constatons que ces articles mobilisent des méthodologies diverses (qualitatives, quantitatives, mixtes) sur des **échantillons de grande taille** qui sont souvent **multidisciplinaires**. Les mobilités inter-universitaire et internationale sont toutes les deux étudiées. En revanche, nous remarquons que la productivité scientifique est à chaque fois abordée en fonction du nombre de publications générées (le **niveau d'activité**).

Le recours à la mobilité comme **indicateur de performance** est utilisé dans la carrière pour évaluer la progression de carrière du chercheur. Le recours à cet indicateur fait débat dans la littérature. Certains travaux prouvent une corrélation positive entre l'exercice de mobilité internationale et **l'accès à une promotion plus rapide**. Lawson et Shibayama (2015) ont réalisé une étude sur 370 chercheurs japonais en biosciences à partir de questionnaires et de CV. Ils ont montré que les visites de recherche internationales ont pour effet d'accélérer la promotion et cela sur l'ensemble de la carrière. Cet effet est d'autant plus fort lorsqu'il s'agit de la première mobilité dans la carrière. Néanmoins, cet effet stratégique ne concernerait que la promotion vers des instituts de rang inférieur à celle d'origine et les séjours post-doctoraux n'auraient aucun effet. La promotion sur la fin de carrière (par exemple au poste de professeur titulaire) ou l'accès à une université de haut rang en milieu ou fin de carrière est davantage favorisé par le mérite que par l'exercice de la mobilité. Également, Cañibano *et al.* (2020) ont montré, à partir d'une enquête à grande échelle, que la mobilité internationale à mi-carrière augmentait la probabilité d'une promotion dans le pays d'origine, dans la mesure où le retour n'est pas trop tardif.

En revanche, d'autres travaux ont montré que la mobilité en tant que telle n'était **pas un bon indicateur pour estimer le mérite et la progression de carrière**. Ackers (2008) souligne que la mobilité permanente n'est pas un indicateur méritocratique car elle peut être nécessaire (manque de ressources, d'opportunités internes, pression externe du contexte) ou entravée par des obstacles non méritocratiques (genre, situation familiale, santé, problèmes financiers...). Selon elle, la mobilité temporaire serait un meilleur indicateur quant aux performances du chercheur. Horta et Yonezawa (2013) soulignent que la corrélation entre la mobilité et la productivité devient négative à compter d'un certain stade : ils préconisent donc qu'à partir d'un seuil de carrière, la mobilité ne soit plus utilisée comme indicateur de productivité. Aussi, Cruz-Castro *et al.* (2010) soulignent que la croyance selon laquelle la mobilité accélère la carrière est propre au système universitaire américain et non adéquat au système espagnol, dans lequel une université composée de chercheurs non mobiles n'est pas dépourvue de scientifiques qualifiés. Ils insistent sur le fait que la mobilité ne lance pas la carrière. Veugelers et Van Bouwel (2015) ont montré que le système universitaire était beaucoup plus méritocratique que le système européen du fait d'une forte sélection. Ainsi, les effets positifs issus de la mobilité (en termes de productivité scientifique, de réputation et de développement de carrière) sont majorés chez les chercheurs qui ont été mobiles aux États-Unis par rapport à ceux qui ont été en Europe. L'originalité de ce travail est de considérer la mobilité comme un accélérateur de carrière dans le cas où la productivité du chercheur est

déjà garantie. Enfin, certains chercheurs pointent la rigidité du critère de mobilité comme source d'effets négatifs contreproductifs. En effet, Bernela (2016) souligne les effets pervers de la règle tacite de mobilité dans les mathématiques sur la vie privée des chercheurs et Laudel et Bielick (2019) montrent que la standardisation du critère de mobilité dans la poursuite de la carrière nuit aux carrières scientifiques dans certaines disciplines.

Tableau 5 : Articles traitant de la corrélation entre la mobilité et la performance scientifique individuelle

Corrélation entre la mobilité inter-universitaire et ...					
	Echantillon	Données mobilisées	<i>Le niveau d'activité</i>	<i>La visibilité espérée</i>	<i>La visibilité réelle</i>
Fernandez-Zubieta et al. (2013)	171 carrières (entre 1982 et 2005) de chercheurs universitaires de Grande-Bretagne	CV, bibliométrie	La mobilité vers une université de rang supérieur a un effet faiblement positif sur le nombre annuel de publications Cependant, une mobilité vers une université de rang inférieur réduit sur le nombre annuel de publications		Une mobilité vers une université de rang inférieur réduit sur le nombre moyen de citations par articles durant les 5 années qui suivent la parution
Horta et Yonezawa (2013)	3925 chercheurs Japonais et étrangers affiliés à 34 universités publiques ou privées au Japon	Questionnaire (taux de réponse : 16,2%)	La mobilité intra-sectorielle accroît le nombre de publications dans les revues internationales jusqu'à un certain seuil à partir duquel la stabilité professionnelle est préférée		
Halevi et al. (2016)	100 chercheurs productifs	Bibliométrie	Une mobilité entre deux affiliations dans un même pays a un impact positif sur le nombre de publications dans les revues internationales à comité de lecture	Une mobilité entre deux affiliations dans un même pays a un impact positif sur le h-index et le Field Weighted Citations Impact	
Di Lorenzo et Tartari (2017)	80 chercheurs issus d'universités du Royaume-Uni Focus sur les sciences de la vie	CV, liste des publications	La mobilité interuniversitaire augmente le nombre de publications annuelles		
Ejermo et al. (2020)	35 361 chercheurs universitaire en Suède	Scopus publications	La mobilité interuniversitaire augmente le nombre de publications du chercheur en		La mobilité interuniversitaire augmente les citations du chercheur en médecine, sciences

médecine, sciences naturelles,
ingénierie et technologie

naturelles, ingénierie et
technologie

Corrélation entre la mobilité internationale et ...

	Echantillon	Données mobilisées	<i>Le niveau d'activité</i>	<i>La visibilité espérée</i>	<i>La visibilité réelle</i>
Aknes et al. (2013)	11 000 chercheurs norvégiens Focus sur les sciences de la terre, les mathématiques, la sociologie, la géographie	Bibliométrie CV, données personnelles (genre, âge, position, éducation, institution)	La mobilité accroît le taux de publication (résultats peu significatifs)		La mobilité accroît le taux de citation (résultats peu significatifs)
Jonkers et Cruz-Castro (2013)	124 chercheurs en Argentine issue des centres CONICET de Buenos Aires Focus sur les sciences de la vie	Entretiens, Bibliométrie	L'expérience internationale en Europe ou en Amérique Latine a un effet structurant sur les co-publications internationales (collaborations avec les chercheurs du pays « hôte »	L'expérience internationale accroît la capacité à publier avec des journaux à forte renommée	
Reiner et Meyer (2015)	211 chercheurs autrichiens Focus sur les sciences humaines, les sciences nationales, les sciences sociales et économiques, le droit	CV, Bibliométrie	La mobilité a un effet positif sur le nombre de publications en langue étrangère		

2- Débats sur l'articulation entre la mobilité géographique et la géographie des collaborations

En complément des travaux sur la productivité des chercheurs mobiles, **certains articles portent sur l'impact de la mobilité géographique sur les habitudes de collaboration des chercheurs**. Ces travaux sont beaucoup plus récents et traduisent un intérêt nouveau pour cette corrélation : « *there have been few attempts to link mobility to research networks and to map the extent to which mobility networks and research networks overlap* » (Scellato *et al.*, 2015). Une partie de ces travaux montrent notamment que la mobilité provoque **un effet d'ouverture internationale du réseau de co-authorship**. Cañibano *et al.* (2008) montrent par exemple que la mobilité internationale accroît la probabilité de participer à un projet de recherche financé à l'échelle internationale. Plusieurs travaux considèrent que la mobilité est plus structurante en début de carrière. En effet, Horta et Yonezawa (2013) soulignent que les chercheurs japonais ayant réalisé leur doctorat à l'étranger communiquent davantage avec leurs pairs internationaux (ils sont considérés comme des « gardiens internationaux ») alors que ceux qui l'ont fait au Japon communiquent davantage avec leurs pairs nationaux (ils sont considérés comme des « gardiens nationaux »). Le **séjour post-doctoral** est également décrit comme très structurant dans la constitution du réseau de co-authorship, bien que ces effets varient selon le pays d'origine (Wolley *et al.*, 2008), le genre ou la discipline du chercheur (Marmolejo-Leyva et Perez-Angon, 2015). L'enrichissement des réseaux scientifiques avec des organisations étrangères issu de la mobilité (Fontes *et al.*, 2013) s'explique notamment via le maintien de relations une fois la mobilité faite et aux systèmes de diaspora (Barruffaldi et Landoni, 2012 ; Scellato *et al.*, 2015).

Bien que certains travaux mettent en évidence une corrélation positive entre la mobilité géographique et l'internationalisation des collaborations, d'autres travaux tendent à **relativiser cet effet et le rôle de la mobilité dans l'enrichissement du réseau collaboratif**. Melin (2004) montre que les séjours postdoctoraux se font principalement dans des destinations scientifiques dites traditionnelles au détriment de l'exploration de « nouveaux environnements scientifiques ». Ainsi, le jeune chercheur va hériter du réseau de ses collègues ayant déjà des contacts avec cette nouvelle destination. L'université va donc centraliser le réseau de ses chercheurs, qui n'auront pas besoin de mobilité pour y avoir accès. Bernela (2016) montre également qu'en mathématiques, les comportements de collaboration sont très hétérogènes et indépendants de la trajectoire géographique permanente du chercheur. Également, il souligne d'autres logiques sociales de mise en relation hors mobilité permanente, telle que la modalité de « système professionnel » (issue de la mobilité temporaire) et la modalité d'« intermédiation » (issue d'invitations croisées et de relations croisées) (Bernela et Milard, 2016). Dans une étude bibliométrique à grande échelle, Chinchilla-Rodriguez *et al.* (2017) montrent également que la dispersion géographique du réseau de co-authorship est trois fois plus élevée que celle de la trajectoire géographique du chercheur. Aussi, il souligne que le réseau issu de la mobilité est beaucoup plus étroit et pauvre en termes de liens que le réseau de co-authorship. Enfin, Wang *et al.* (2019) ont montré à partir d'une étude bibliométrique de 378 chercheurs universitaires très mobiles de Singapour que le changement

d'affiliation dans un nouveau pays accroît la collaboration locale aux dépens des collaborations internationales. Ainsi, la mobilité internationale se traduit par un renforcement du localisme à la suite de l'intégration d'un nouvel environnement.

A la vue de ces principales contributions dans la littérature, **très peu de travaux évaluent l'impact de la mobilité géographique sur le réseau de co-authorship au travers de données longitudinales** (Scellato *et al.*, 2015). Fontes (2013) souligne : « *We still know very little about the role played by mobility on the creation of relationships* ». Cañibano *et al.* (2008) montrent la nécessité d'une analyse dynamique des carrières des chercheurs, notamment dans la compréhension de la construction du réseau scientifique. En effet, cette dimension dynamique semble pourtant souligner plusieurs intérêts scientifiques, tels que l'impact du changement d'affiliation sur la création et la rupture de relations, l'évaluation du poids du pays d'origine (Baruffaldi & Landon, 2012) ou des pays fréquentés dans le réseau de co-authorship (Chinchilla-Rodriguez *et al.*, 2017 ; Wang *et al.*, 2019).

Les carrières scientifiques se déroulent dans un cadre politique structurant et dynamique. En effet, elles sont gérées par l'Etat (CNU) et les universités : ces institutions développent des politiques qui encadrent l'activité du chercheur afin d'en tirer le plus d'externalités positives possible. C'est dans cette dynamique que les politiques incitatives à la mobilité scientifique se sont développées à partir des années 1990. En effet, ces mesures sont clefs dans le développement stratégique scientifique et innovant de la France et des universités : elles visent à mobiliser les leviers de réallocation et de transformation des connaissances. En revanche, l'effet concret de la mobilité scientifique sur la carrière scientifique fait débat dans la littérature et très peu de travaux traitent de son impact sur la diversification du réseau de co-authorship. **Ainsi, il existe une réelle opportunité scientifique quant à la compréhension de la manière à laquelle le chercheur ajuste ses collaborations à la suite de la mobilité.** La dernière partie de la revue de littérature relative à la carrière scientifique et au comportement collaboratif se focalise sur la théorie de l'analyse de réseaux sociaux. Cette nouvelle littérature a émergé dans les années 1970 et permet d'identifier la structure des collaborations, et leur dynamique au fil de la carrière scientifique.

III- Les dynamiques relationnelles au sein des réseaux de co-authorship

L'objectif de cette dernière partie est de faire l'état de la littérature relative aux collaborations dans **l'analyse de réseaux sociaux**. L'outil du réseau a longtemps été mobilisé dans la littérature managériale dans une optique de gestion de la coordination des flux. La littérature de SNA (Social Network Analysis) se distingue de ces travaux : elle est issue de la sociologie et vise à étudier les interactions sociales grâce à la théorie des réseaux. En effet, à partir des années 1970, la « nouvelle sociologie économique » se développe et invite les économistes à intégrer une approche sociologique dans la compréhension des comportements individuels économiques. Nous commencerons par décrire ce courant littéraire ainsi que ses **théories fondatrices fournissant** des conditions structurelles de qualité de réseau. Dans un second temps, nous détaillerons la littérature empirique existante relative à la structure du réseau et à ses effets sur la **performance scientifique et la dynamique collaborative**.

A) Réseau scientifique : de l'outil managérial à l'analyse du réseau de co-authorship

L'utilisation du réseau pour appréhender les interactions sociales s'est développée au cours des dernières années. D'abord utilisé comme un outil de management, le réseau a été ensuite mobilisé dans le cadre de l'analyse de réseau dans de nombreuses études sociologiques. Les travaux fondateurs de ce courant fournissent des **conditions structurelles à la qualité des interactions dans le cadre de réseau communautaire** (réseau complet d'un laboratoire, d'un champ scientifique, etc.). Contrairement aux réseaux personnels, ces réseaux capturent les interactions sociales dans le cadre d'une structure. Ces théories fondatrices, qui **nourriront par la suite une grosse partie des travaux empiriques de l'analyse de réseaux**, portent sur la taille du réseau, sur la force des liens, le positionnement dans le réseau et la cohésion entre les collaborateurs.

1- Le réseau est un outil managérial en sciences sociales

Le mot « réseau » est issu du terme latin *retis*, qui signifie « filet », soit un « *ouvrage formé d'un entrelacement de fils* ». Il a été repris dans la langue française au XII^{ème} siècle sous le terme de *resel*. Le réseau représente un **ensemble de nœuds reliés entre eux par des canaux** qui matérialisent les flux et les échanges. Pendant de nombreux siècles, ce terme n'a pas été théorisé mais appliqué dans de nombreux domaines : cette représentation métaphorique est mobilisée par les entreprises et les institutions publiques pour mieux appréhender la complexité des flux et des interactions économiques, techniques et sociologiques. Le terme sera théorisé dans le domaine de la sociologie par Barnes en 1954. Le réseau social permet alors de retracer le **modèle d'organisation et d'action des individus à l'échelle micro, meso, macro et globale** (Bagla-Gökalp, 2000). Il se définit comme un ensemble d'entités sociales (représentées par les nœuds) qui agissent ensemble : les arrêtes du graphe permettent de

modéliser les différentes interactions sociales qui lient chaque nœud ensemble. Cet outil est notamment mobilisé dans la sociologie économique (étude du marché du travail), dans la sociologie des organisations et de l'entreprise (relations intra et inter organisationnelles) et la sociologie de l'innovation.

Bagla-Gökalp (2000) décrit le réseau comme le « *nouveau paradigme des représentations du social* » à la suite des travaux de Kuhn (1983). En effet, les travaux d'analyse de réseaux relatifs à l'organisation sociale et aux relations humaines prolifèrent en anthropologie (Levi-Strauss, 1967 ; Bollig, 1998), en histoire économique (Landes, 1993) et en sociologie (Durkheim, 1897 ; Elias, 1991). En dépit de ces diverses applications, les **principales approches sociologiques de réseau** n'ont été développées qu'à partir des années 1970. A cette époque, le champ sociologique recentre « **l'acteur** » **au cœur de ses études : les structures sociales sont considérées comme des contraintes matérielles et des déterminants des choix individuels des agents**. De plus, la conjoncture économique des années 1970 oriente la littérature managériale vers l'élaboration de nouveaux scénarios de survie des entreprises et des salariés (Thorelli, 1986). Le **réseau est un outil** plébiscité par les entreprises (Biggart et Hamilton, 1992) puisqu'il permet de représenter des liens informels et de consolider sa position stratégique dans une période d'instabilité et de compétitivité croissante. Le modèle d'interaction du réseau est basé sur une stratégie d'intermédiaire et de coopération afin de réduire les coûts de transaction. Par opposition au système pyramidal (Biggart et Hamilton 1992), l'organisation en réseau permet une communication et une coordination plus fluides entre les différentes parties prenantes de l'entreprise (Badaracco *et al.*, 1991). La stratégie organisationnelle des entreprises est évaluée en fonction de la qualité du réseau dans lequel elle est insérée : les bienfaits attendus sont la flexibilité, l'accès aux ressources, le transfert d'information et d'expérience, un meilleur contrôle de l'opportunisme et de l'incertitude. De nombreuses études ont portées sur les bienfaits de la proximité géographique et organisationnelle au sein des réseaux d'entreprises, tels que les districts industriels italiens (Piore et Sabel, 1986 ; Powell et Smith-Doerr, 1994). La Silicon Valley est également un exemple en termes de proximité géographique et de coopération entre les universités et les laboratoires de recherche (Angel, 1991 ; Saxenian, 1994). Certaines études portent également sur les clusters d'Asie, comme les Keiretsu au Japon (JR Lincoln *et al.*, 1996) et les Chaebol en Corée du Sud (EM Kim, 1996).

2- Taille du réseau de co-authorship et capital social du chercheur

Le **capital social** est un concept mobilisé dans les sciences sociales associé à la **performance individuelle**. Il est défini comme l'ensemble des **ressources associées au sein du réseau de relations**. Étudié pour la première fois dans les travaux de Hanifan (1916), le capital social a fait l'objet de nombreuses études à partir des années 1980, notamment à partir des travaux de Bourdieu (1980, 1986, 1992).

Dans son ouvrage *Les formes du capital*, Bourdieu (1986) énumère plusieurs formes de capital :

Tableau 6 : Formes de capital selon Bourdieu

Type de capital	Définition	Forme de naturalisation
Capital économique	Patrimoine d'un individu, revenus	Convertible (ex : outil monétaire) Institutionnalisé (ex : droits de propriété)
Capital culturel	Ensemble des ressources culturelles d'un individu	Incarné (par les individus) Objectivité (ex: capital artistique) Institutionnalisé (ex : diplômes universitaires)
Capital social	Ensemble des relations personnelles d'un individu	Ancré au sein des échanges dans les réseaux de personnels durables
Capital symbolique	Manifestation des autres formes de capital	

Plus précisément, le capital social est décrit comme « *la somme des ressources, réelles ou virtuelles, qui reviennent à un individu ou à un groupe du fait de la possession d'un réseau durable de relations plus ou moins institutionnalisées de connaissance et de reconnaissance mutuelles* » (Bourdieu et Wacquant, 1992, p.119). Cette forme de capital est mesurée par la **taille du réseau** (Bourdieu, 1986) et non par les échanges effectifs. En d'autres termes, le volume de ce type de capital est appréhendé **par le nombre de co-auteurs différents** dans le cadre d'un réseau de co-authorship : plus le réseau dispose de ressources humaines collaboratives différentes, plus la diffusion et la création de connaissances seront accrues. Le capital social est donc dynamique : il est issu **de l'accumulation de relations durables et utiles**. Bourdieu (1986) souligne également que le capital social a un rôle primordial dans le sens où il **démultiplie la dotation des autres formes de capital**. Dans le cadre du réseau de co-authorship, la dotation d'un réseau riche en co-auteurs différents (capital social) démultiplie l'accès d'autres réseaux de co-authorship variés (capital économique), les alliances collaboratives (capital culturel) et donc la valeur du capital symbolique.

Tout comme Bourdieu (1986), Lin (1995) décrit le capital social comme un ensemble de ressources sociales issues de l'investissement fait au sein des relations du réseau (« *Le capital social n'est possible que lorsqu'on examine les relations d'un individu et les ressources accessibles grâce à ces relations* »). Il considère que ces ressources peuvent être évaluées avec des « indicateurs empiriques » à l'échelle de chaque individu : « *Ces ressources peuvent être acquises (l'éducation, le prestige ou l'autorité) ou héritées (l'appartenance ethnique, le sexe, parfois la religion ou les ressources des parents). Elles peuvent être classées en deux catégories : ressources personnelles et ressources sociales* » (Lin, p. 687). Lin ne s'intéresse pas à la force des échanges au sein du réseau mais plutôt comment et avec qui les échanges sont fait. Selon lui, la position dans le réseau doit être étudiée pour puisqu'elle joue un rôle clef dans l'accès aux ressources circulantes au sein du réseau.

3- Force des collaborations et positionnement dans le réseau

Le courant de *l'analyse structurale*, né aux Etats-Unis dans les années 1970, étudie **l'impact de la nature des relations et la position au sein du réseaux sur les choix individuels**. L'analyse structurale a pour but de **donner un sens aux relations** et d'identifier les **régularités de comportement** à partir du réseau. Deux méthodes sont alors utilisées. La première méthode est la **théorie des graphes** (relative à la densité des relations et à la détection de cliques dans le réseau) : elle est appliquée dans la théorie de la « force des liens faibles ». La seconde méthode est **l'analyse des positions** (à partir d'outils algébriques et de matrices pour trouver les équivalences structurelles) : elle est appliquée dans la théorie des « trous structuraux ».

« *La force des liens faibles* »

La théorie de Granovetter (1973, 1985) évalue la **qualité du réseau** en fonction de la **force de ses liens**²⁹. Cette théorie est basée sur l'hypothèse selon laquelle l'interaction entre les individus est une source d'opportunités dans la création et le partage de savoir. Granovetter (1973) définit deux types de liens dans le réseau, qui sont illustrés dans la figure 4 :

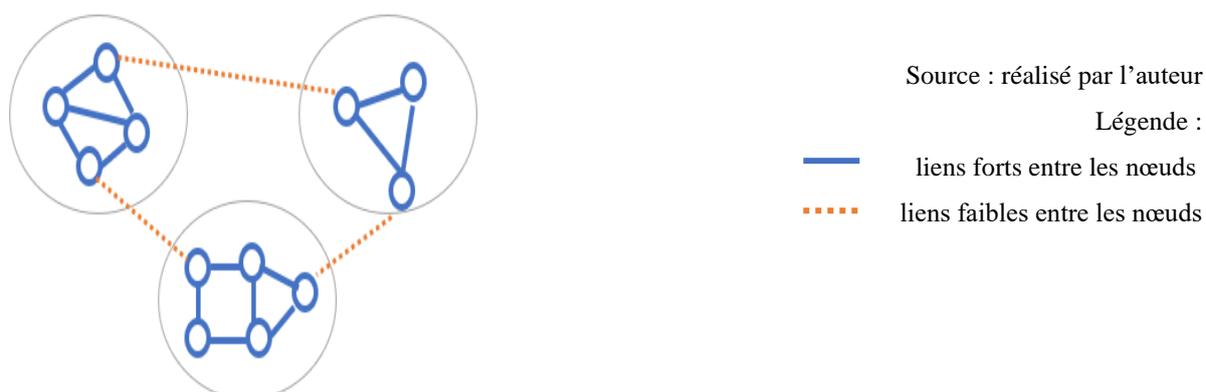
i) Les **liens forts** correspondent à des relations récurrentes et soutenues. Un réseau doté de liens forts est dense (structuré en cliques ou en îlots isolés) et **l'information véhiculée** par les arrêtes **a plus de chance d'être redondante** (du fait de la faible distance sociale et géographique entre les nœuds). Par exemple, dans la figure 4, nous remarquons que les liens forts structurent trois microcosmes dans le réseau dans lesquels chaque nœud partage la même information avec les autres. Ainsi, la circulation et l'interaction de connaissances hétérogènes est limitée.

ii) Les **liens faibles** correspondent à des relations ponctuelles et non soutenues³⁰. Granovetter (1973) explique que les liens faibles ont un double intérêt. Tout d'abord, ils favorisent la **circulation d'informations non redondantes et la cohésion sociale** en servant de pont entre plusieurs individus hétérogènes qui ne seraient pas liés sans eux. Dans la figure 4, nous remarquons que les liens faibles permettent la circulation d'information entre les trois entités indépendantes et donc l'interaction de connaissances hétérogènes. Ensuite, ils permettent **l'échange d'informations peu complexes** dans relations qui nécessitent peu d'entretien et moins soumises à la pression sociale (Hansen, 1999 ; Granovetter, 1995).

²⁹ Dans cette théorie, la force des liens interpersonnels dépend de la temporalité de la relation, du degré d'intensité émotionnelle et d'intimité ainsi que de la réciprocité de la relation.

³⁰ Granovetter (1973) détaille : « *Les liens faibles, souvent dénoncés comme source d'aliénation pour les individus (...), apparaissent ici au contraire comme des instruments indispensables aux individus pour saisir certaines opportunités qui s'offrent à eux, ainsi que pour leur intégration au sein de la communauté ; au contraire, les liens forts, qui engendrent la cohésion sociale, se traduisent par une fragmentation de l'ensemble social* » (Granovetter, 1973).

Figure 4 : Illustration des liens faibles et forts



Certains auteurs réfutent cette théorie en montrant les avantages des **liens forts** dans la structure du réseau. Krackhardt (1992) montre notamment qu'ils sont des **vecteurs de confiance** plus importants que les liens faibles. Aussi, certains soulignent que les liens forts ont des vertus en termes d'échanges de connaissance, notamment dans les domaines à **forte intensité de flux** (Levin et Cross, 2004) et où **l'information est complexe** (Hansen, 1999 ; Reagans et Zuckerman, 2001).

La théorie des trous structuraux

A partir de la théorie de Granovetter (1973), Burt (1992) montre dans quelle mesure **l'absence de relations** entre deux groupes constitue un **avantage compétitif** dans la structure du réseau. Burt n'évalue pas la taille du réseau en fonction du nombre d'acteurs mais en fonction du **nombre d'acteurs non redondants**.³¹ Ainsi, il définit un **réseau efficace** comme étant composé de plusieurs **groupes d'acteurs connectés qui émettent des informations non redondantes**.

Le **trou structurel**, correspondant au manque de relations entre deux acteurs du réseau, est une caractéristique structurelle qui favorise la **circulation d'informations non redondantes**. Cette position stratégique fournit deux types d'avantages :

i) Gestion de l'information non redondante : Le trou structurel permet de compartimenter les sources d'informations non redondantes, en composant d'acteurs qui sont structurellement équivalents et cohésifs. Le connecteur sert de liaisons entre ces différents composants : le retrait de cet acteur du réseau

³¹ En 1995, il définit : « Deux contacts sont redondants lorsqu'ils donnent accès aux mêmes personnes, et procurent ainsi les mêmes bénéfices en informations » (p.602).

entraînerait une rupture de connexion à l'échelle du réseau. Ainsi, le connecteur a **accès à un gros volume d'informations non redondantes** (accès indirect à une grande partie des acteurs, diversité des contacts). Aussi, cette position de carrefour permet au connecteur de **connaître les spécificités et les activités** de chaque groupe et donc de mieux gérer la diffusion des nouvelles idées.

ii) Contrôle des liens collaboratifs : le connecteur étant central dans le réseau doté de trous structuraux, il dispose d'une **bonne connaissance des autres acteurs du réseaux** (objectifs, compétences, identité). Ainsi, il peut négocier les relations et les collaborations. Ce « *tertius gaudens* » (Simmel 1955, Merton 1957) détient un **pouvoir d'autorité corporative et a la capacité d'optimiser la gestion des flux au sein du réseau**. Aussi, cela lui permet d'individualiser l'approche avec chacun de ses nœuds, selon les besoins et les compétences de chaque sous-entité, plutôt que d'opter pour une approche standardisée.

Figure 5 : Illustration de trous structuraux

Figure 5a. Présence de trous structuraux

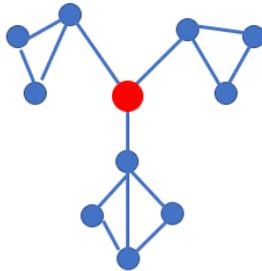
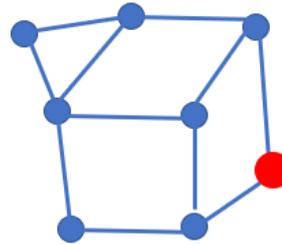


Figure 5b. Absence de trous structuraux



Source : réalisé par l'auteur

La figure 5 illustre un réseau composé de trous structuraux et un autre qui n'en a pas. Ainsi, dans le graphique de gauche, nous notons que le connecteur (en rouge) occupe une position de centralité d'intermédiation (*betweenness centrality*, Freeman, 1979) et que la part des arrêtes redondantes est très faible. En revanche, dans le graphique de droite, l'absence de trous structuraux souligne une faible centralité des nœuds et des arrêtes très redondantes.

4- Réseau de co-authorship et cohésion inter-individuelle

Coleman (1988) s'intéresse à la **densité du réseau**, qui est selon lui génératrice de **capital social**. Selon lui, le capital social réside dans les **relations entre les individus** contrairement aux autres ressources qui sont des attributs individuels. Le capital social est intégré dans les relations individuelles (Shen,

2010) et génère le capital humain³². Selon Coleman (1988), la source du capital humain est la **cohésion entre les individus du réseau social**. Cette cohésion est favorisée dans les réseaux qui sont densément interconnectés puisqu'ils facilitent la **confiance, la coopération, le partage d'information**. Cette forte densité interne permet notamment la circulation de normes et sanctions communes, de relations réciproques et une forte proximité entre les acteurs du réseau. Le contrôle social instauré permet d'identifier et de sanctionner les comportements opportunistes (Grandori, 1997 ; Rowley *et al.*, 2000). Grâce à ces comportements, les acteurs du réseau profitent de deux avantages (Baret *et al.*, 2006) : **une meilleure circulation de l'information et une structure bienveillante** (solidaire et coopérative).

A partir de cette conception du réseau, la théorie de l'acteur-réseau s'est développée dans les années 1980 en France (Callon, 1991 ; Callon *et al.*, 1995 ; Callon et Latour, 1985). Elle ne s'intéresse pas à la structure du réseau mais plutôt au **processus collectif et complexe d'innovation** par des acteurs qui créent simultanément la technologie et la société (processus de co-production)³³. Le **réseau mobilisé** dans les projets innovants résulte du processus de **recrutement des laboratoires**. Ils enrichissent le réseau disponible au sein de leur structure via l'intégration de nouveaux chercheurs et de leur réseau associé (Callon, 1988) et fournissent **un espace de cohésion et de circulation de l'information**. La particularité de cette théorie tient également dans le **souci de coordination du réseau**. Ainsi, le réseau n'est plus seulement un maillage d'acteurs mais aussi **un maillage d'intérêts, de problèmes divers** (relatifs à chacun des acteurs). Le *traducteur* aura donc comme rôle de **stabiliser le projet collaboratif en coordonnant ces différents acteurs hétérogènes**.

³² « *Ce n'est pas une entité unique, mais une variété d'entités différentes qui ont deux caractéristiques en commun : elles constituent toutes un aspect de la structure sociale, et elles facilitent certaines actions d'individus qui sont au sein de la structure. Comme d'autres formes de capital, le capital social est productif, rendant possible la réalisation de certains buts qui ne pourraient être réalisés en son absence* » (Coleman, 1990, p. 302-303).

³³ Callon (1986) décrit les chercheurs comme des acteurs qui « font à la fois le contexte social et les choses qui sont dans ce contexte ; les ingénieurs font à la fois le marché et les produits qui tiennent ce marché ». Ainsi, tout objet est vu comme une combinaison d'éléments issus de l'identité des acteurs et de la société (métiers, contraintes, représentations). Par exemple, le texte scientifique est « *un réseau à soi tout seul* » puisqu'il naît de l'interaction entre des acteurs hétérogènes : ses caractéristiques (titre, revue, liste des auteurs, affiliations, bibliographie) informent les sources d'informations collaboratives et bibliographiques dont il résulte (Callon, 1991).

5- Synthèse : Conception du capital social et structure du réseau de co-authorship

Tableau 7 : Conception du capital social et structure du réseau de co-authorship

		Structure optimale du réseau	Indicateurs d'analyse de réseaux référents	Avantages structurels au sein du réseau de co-authorship
	Bourdieu (1986)	La diversité des co-auteurs	- Nombre de nœuds (Réseau personnel) - Centralité de degré (Réseau complet)	- Diversité des ressources de collaboration - Effet démultiplicateur sur les autres formes de capital
Capital social à l'échelle personnelle	Granovetter (1973)	- Forte dotation de liens faibles - Présence de trous structuraux	- Taille moyenne des arrêtes - Contrainte	- Accès à information non redondante - Echange d'informations non complexes à faible coût d'entretien
	Burt (1992)	- Nature des ressources accessibles (capital social mobilité et latent)	- Centralité d'intermédiarité (Freeman, 1979)	- Management et contrôle des flux d'information
	Lin (1995)		- Taille effective du réseau	
Capital social à l'échelle communautaire	Coleman (1988)	Forte interconnexion entre les nœuds	- Densité	- Circulation diffuse de l'information - Structure coopérative et bienveillante - Comportements normés, sanctions communes - Bénéfices mutuels

B) Analyses empiriques de la structure du réseau de co-authorship

Une large partie de la littérature mobilise ces théories structurelles pour **évaluer la performance scientifique** dans le cadre de réseaux communautaires, dits complets. Ces études reposent sur une **méthodologie commune**, basée sur le recours aux réseaux complets.

1- Réseaux complets et performance scientifique

Une grande partie de la littérature évalue l'impact d'indicateurs structurels sur la performance scientifique au travers de **réseaux complets**. Dans ces articles, la performance scientifique est souvent appréhendée en termes de visibilité des travaux scientifiques (*g-index*, *h-index*, facteur d'impact de revue, citations par article). Le tableau 8 présente les résultats de dix de ces études.

Les résultats récurrents sont :

- i) La relation positive entre la **taille du réseau** du chercheur (Bourdieu, 1986) et la **visibilité des travaux** ;

ii) La relation positive entre la **force des relations collaboratives** (Granovetter, 1973 ; Krackhardt, 1992) et la **visibilité des travaux** ;

iii) La relation positive entre la **position de d'intermédiaire entre plusieurs co-auteurs** (Freeman, 1979 ; Burt, 1992) et la **visibilité des travaux**.

Tableau 8 : Articles portant sur la performance scientifique et la structure de réseaux de co-authorship complets

Indicateurs personnels structuraux	Résultats
Indicateur de performance : le g-index	
<p>Abbasi <i>et al.</i> (2011)</p> <p>Centralité au sein du réseau (Freeman, 1979 ; Scott, 1991) :</p> <ul style="list-style-type: none"> - Centralité de degré normalisée - Centralité de proximité normalisée - Centralité normalisée d'intermédiarité - Centralité normalisée des vecteurs propres <p>Théorie des « liens faibles » (Granovetter, 1973) :</p> <ul style="list-style-type: none"> - Force moyenne des arrêtes <p>Théorie des trous structuraux (Burt, 1995, 2009) :</p> <ul style="list-style-type: none"> - Efficacité individuelle 	<p>Facteurs qui favorisent la hausse du g-index :</p> <ul style="list-style-type: none"> - Collaborer avec de nombreux co-auteurs distincts (centralité de degré) - Maintenir des relations collaboratives récurrentes (force moyenne des arrêtes) - Maintenir une relation forte avec une seule personne d'un groupe lié plutôt qu'avec l'ensemble du groupe (Efficacité individuelle) <p>Facteurs qui favorisent la chute du g-index :</p> <ul style="list-style-type: none"> - Collaborer avec des chercheurs non centraux (avec des étudiants que des chercheurs performants) (centralité de vecteurs propres)
<p>Abbasi <i>et al.</i> (2012)</p> <p>Interconnexion interne :</p> <ul style="list-style-type: none"> - Densité <p>Centralité au sein du réseau :</p> <ul style="list-style-type: none"> - Centralité de degré - Centralité d'intermédiarité <p>Théorie des trous structuraux (Burt, 1995, 2009) :</p> <ul style="list-style-type: none"> - Taille effective - Efficacité individuelle - Contrainte 	<p>Facteurs qui favorisent la hausse du g-index :</p> <ul style="list-style-type: none"> - Collaborer avec de nombreux co-auteurs distincts (centralité de degré) - Disposer d'une position d'intermédiarité au sein du réseau (centralité d'intermédiarité) - Maintenir une relation forte avec une seule personne d'un groupe lié plutôt qu'avec l'ensemble du groupe (Efficacité individuelle)
<p>Badar <i>et al.</i> (2015)</p> <p>Centralité au sein du réseau :</p> <ul style="list-style-type: none"> - Centralité de degré - Centralité d'intermédiarité - Centralité de proximité 	<p>Relation curviligne (en forme de U inversé) entre la centralité de degré et l'indicateur de performance <i>g-index</i>.</p>
<p>Bordons <i>et al.</i> (2015)</p> <p>Centralité au sein du réseau :</p> <ul style="list-style-type: none"> - Centralité de degré - Centralité d'intermédiarité - Centralité de vecteurs propres - Centralisation 	<p>Facteurs qui favorisent la hausse du g-index :</p> <ul style="list-style-type: none"> - Nombre de co-auteurs distincts (centralité de degrés) - Maintenir des relations collaboratives récurrentes (force moyenne des arrêtes)

Mesures de cohésion :

- Force des liens
- Contrainte de réseau
- Coefficient de clustering

Facteurs qui favorisent la baisse du g-index :

- La cohésion locale dans les domaines expérimentaux

Autres indicateurs de performance relatifs à la visibilité scientifique

Abbasi *et al.*
(2014)

Centralité au sein du réseau :

- Centralité de degré
- Degré de centralité pondéré
- Centralité d'intermédiarité

La performance scientifique (h-index) est significativement et positivement corrélée aux mesures de capital social.

Théorie des « liens faibles » (Granovetter, 1973) :

- Force moyenne des arrêtes

Théorie des trous structuraux (Burt, 1992) :

- Efficacité individuelle

Diversité et pouvoir du capital social (Arnaboldi et al., 2016) :

- Individual power–diversity index (PDI)
- Individual power–tie–diversity index (PTDI)

Arnaboldi *et al.*
(2016)

Interconnexion interne :

- Densité

La taille du réseau (centralité de degré) accroît la performance scientifique (le h-index).

Centralité au sein du réseau :

- Centralité de degré

Théorie des « liens faibles » (Granovetter, 1973) :

- Force moyenne des arrêtes

Mesures de cohésion :

- Ego network circles

Badar *et al.*
(2013, 2014)

Centralité au sein du réseau :

- Centralité de degré
- Centralité d'intermédiarité
- Centralité de proximité

Performance scientifique évaluée avec un indicateur de facteur d'impact de revue (5 ans) :

- Effet modéré du genre du chercheur sur la corrélation positive entre la centralité (degrés et proximité) sur la performance

Badar <i>et al.</i> (2016)	<ul style="list-style-type: none"> - Degré de centralité - Indicateurs d'homophilie 	<ul style="list-style-type: none"> - Effet modérateur de l'âge universitaire et du secteur institutionnel sur la corrélation positive entre la centralité des diplômés et la performance scientifique - La performance du chercheur (facteur d'impact) favorise la centralité du degré (attraction de co-auteurs et croissance de la taille du réseau) - Effet modérateur positif de l'homophilie d'âge et de province - Effet modérateur négatif de l'homophilie de ville
Li <i>et al.</i> (2013)	<p>Centralité au sein du réseau :</p> <ul style="list-style-type: none"> - Centralité de degré - Centralité d'intermédiarité - Centralité de proximité <p>Capital cognitif :</p> <ul style="list-style-type: none"> - Exploitation d'équipe - Mandat de publication <p>Capital relationnel :</p> <ul style="list-style-type: none"> - Nombre de co-auteurs prolifiques 	<p>Effets directs sur la performance scientifique (citations par article)</p> <ul style="list-style-type: none"> - La centralité est le meilleur moyen d'accéder aux ressources non redondantes et d'accroître les citations des publications. <p>Effets indirects sur la performance scientifique (citations par article)</p> <ul style="list-style-type: none"> - La collaboration avec des co-auteurs prolifiques aide à développer des centralités - Les chercheurs dotés d'un mandat d'édition aide à développer des centralités - La collaboration conjointe avec plusieurs savants accroît la centralité de proximité et d'intermédiarité
Uddin <i>et al.</i> (2013)	<p>Centralité au sein du réseau :</p> <ul style="list-style-type: none"> - Centralité de degré - Centralité d'intermédiarité - Centralité de proximité 	<p>Le nombre de citations par article est positivement corrélé aux degrés de centralité et d'intermédiarité des auteurs.</p>

2- Méthodologie commune

Les articles portant sur la corrélation entre la performance scientifique et les indicateurs structuraux de réseaux au travers de réseaux complets disposent de plusieurs **points communs méthodologiques**. En effet, le tableau 9 détaille la méthodologie de dix de ces articles. Les particularités de ces études sont les suivantes : i) La construction de l'échantillon via **l'extraction de publications au sein d'une organisation plus ou moins définie** (université, revue, champ disciplinaire) ; ii) La constitution d'un **échantillon de grande taille**, aussi bien en termes de publications que d'auteurs impliqués ; iii) Publications extraites sur des **périodes relativement courtes**.

Tableau 9 : Méthodologie de 10 articles qui mobilisent des réseaux complets dans l'évaluation de la performance scientifique

	Sources de données	Choix de l'échantillon	Dates des publications	Taille de l'échantillon	Disciplines
Abbasi <i>et al.</i> (2011)	DBLP, Google scholar, portail ACM	Publications issues de de 5 universités américaines	2001-2005	1806 auteurs 2139 publications 5310 co-auteurs	Gestion et des systèmes d'information
Abbasi <i>et al.</i> (2012)	Scopus	Publications issues de 9 revues à fort impact	2000-2009	8069 auteurs 4837 publications 12783 collaborations 93 660 citations	Information science & library science
Abbasi <i>et al.</i> (2014)	Scopus	Extraction de publications avec sélection de discipline	2001-2010	10255 auteurs 4579 publications dans 1392 revues	Science de information
Arnaboldi <i>et al.</i> (2016)	Google Scholar	Extraction d'auteurs sous condition de discipline	1900-2013	313 207 auteurs 19410220 publications	Machine learning, physics, neuroscience, bioinformatics, ecology, molecular biology Chimie
Badar <i>et al.</i> (2013)	Web of Science	Chercheurs affiliés aux universités et instituts de recherche au Pakistan	2002-2009	2027 auteurs 1699 publications dans 151 revues	
Bordons <i>et al.</i> (2015)	Web of Science, Publish or Perish, CV	Extraction de publications sous condition de discipline pour chercheurs affiliés en Espagne	2006-2008	4888 auteurs 15176 publications	Nanosciences, pharmacologie, statistiques

Li <i>et al.</i> (2013)	Social Science Citation Index (SSCI)	Extraction de publications dans 5 revues	1999-2003 Période de citations après la publication : 2004-2001	137 auteurs 330 co-auteurs 308 articles	Systèmes d'information
Sarigöl <i>et al.</i> (2014)	Microsoft Academic Search	Extraction de publications dans une discipline	1995-2008	160891 auteurs 108758 publications	Sciences informatiques
Uddin <i>et al.</i> (2013)	Scopus	Extraction de publication dans deux universités (Australie et Singapour)	2005-2009	1825 auteurs 888 publications	Steel structure
Yan <i>et al.</i> (2010)	Chinese Social Science Citation Index	Extraction de publications dans 18 revues chinoises	2002-2007	11 067 auteurs 14471 collaborations	Library and information science

Ainsi, ces études n'ont pas pour but de décrire les comportements de collaboration des chercheurs sur l'ensemble de leur carrière mais plutôt de faire **une photographie des liens collaboratifs dans le cadre d'une structure et sur une courte période donnée**. En parallèle à cette littérature, quelques travaux s'attachent à aborder le réseau collaboratif au travers de données longitudinales. La reconstitution de réseaux collaboratifs sur une **longue période** permet d'identifier **l'évolution des flux et du comportement collaboratif scientifique**.

C- Dynamiques de collaboration au fil de la carrière

Plusieurs études ont intégré **les données longitudinales** dans l'analyse de **réseau complet** afin d'enrichir la compréhension de collaboration au sein d'une communauté scientifique. En effet, Grossetti (2008) explique que le réseau complet est étudié dans le cadre d'une organisation (« *existence de ressources communes entre certains acteurs, et donc de liens probables entre eux* ») alors que le réseau personnel est étudié à l'échelle individuelle (« *on étudie des voisinages relationnels autour d'acteurs choisis sans faire d'hypothèse sur les relations qu'ils pourraient avoir entre eux* »). Ainsi, dans un réseau complet, le réseau de co-authorship d'un chercheur est étudié dans une **structure commune aux réseaux du reste de l'échantillon**.

Tout d'abord, **certaines travaux conservent le prisme de la performance**. Hou *et al.* (2008) reconstituent un large réseau dynamique de 1927 documents de scientométrie issus de 1620 auteurs entre 1978 et 2004 pour évaluer la structurelle individuelle des réseaux via diverses analyses³⁴. Yan et Ding (2009) ont aussi montré que la centralité des auteurs (centralité de degré, d'intermédiarité, de

³⁴ Analyse de réseaux sociaux, analyse de cooccurrence, analyse de clusters, analyse de fréquence de mots

proximité) est significativement corrélée avec le nombre de citations au travers de données bibliométriques issues de 16 revues scientifiques entre 1988 et 2007. Également, Mohammadamin *et al.* (2017) ont construit un réseau de co-authorship évolutif via 3125 articles issues de la revue *Scientometrics* entre 1980 et 2012. Ce réseau complet dynamique (doté de 3024 auteurs différents répartis dans 1207 institutions, 68 pays et 22 disciplines différentes) est évalué par quelques indicateurs de centralité pour montrer que les liens collaboratifs sont structurés autour de quelques chercheurs centraux à forte notoriété.

Certains travaux mobilisent les données longitudinales afin d'étudier le **comportement collaboratif scientifique dans une dimension géographique au travers de réseaux communautaires**. Li (2015) étudie notamment l'évolution du réseau national scientifique en Chine en sciences humaines et sociales au fil de différents contextes scientifiques entre 1978 et 2001. Aussi, Maisonobe *et al.* (2016) montrent que les liens collaboratifs mondiaux s'articulent autour des contextes nationaux et des aires culturelles et linguistiques au travers de publications extraites dans le Web of Science entre 1999 et 2007. Ils soulignent également la diversité des échelles géographique des collaborations et réfutent l'hypothèse d'un renforcement des échanges autour de quelques agglomérations internationales.

L'utilisation de données longitudinales est également pratiquée dans **l'analyse de réseaux personnels**. Cette niche d'articles porte sur le **comportement collaboratif des chercheurs à l'échelle micro**, dans la perspective où la collaboration individuelle est soumise au pragmatisme du chercheur en fonction d'objectifs et espérés escomptés (Melin, 2000) et sous des motivations différentes en fonction des champs institutionnels ou nationaux (Glänzel et Schubert, 2004). La mobilisation de données longitudinales à l'échelle individuelle permet de mieux **décrire le comportement collaboratif du chercheur, de souligner la pluralité des comportements collaboratifs et de décrire quelques mécanismes de formation de relations collaboratives** (Mali *et al.*, 2012 ; Milard, 2014 ; Hayat et Lyons, 2017).

Les dynamiques relationnelles sont étudiées en fonction du renouvellement social du réseau de co-authorship à l'échelle individuelle. Hossain et Fazio (2009) montrent notamment que la visibilité scientifique (au travers du nombre de citations reçues) n'influence pas la propension future à collaborer via une analyse d'articles entre 1980 et 2004. Également, Cabanac *et al.* (2015) ont étudié les carrières de 3860 chercheurs de sciences informatiques (ayant commencé à publier entre 1980 et 1985) pour reconstituer l'évolution du co-authorship au fil de leur carrière. Ils ont notamment montré le caractère dynamique du renouvellement social : au fil des stades de carrière, le nombre de nouveaux co-auteurs par article chute alors que le nombre moyen de relations renouvelées par article s'accroît. En d'autres termes, les dynamiques collaboratives sont marquées i) par la croissance des maintiens de

relations avec des chercheurs qui sont à mi-carrière ou confirmés au fil de la carrière, et ii) par la chute des créations de relation avec les chercheurs débutants au fil de la carrière.

D'autres chercheurs soulignent la particularité des dynamiques relationnelles en fonction de **caractéristiques personnelles du chercheur**, telles que son genre (Whittington, 2018) ou sa proximité avec ses co-auteurs via des indicateurs d'homophilie (Xie *et al.*, 2018).

L'analyse de réseaux sociaux a émergé dans les années 1970 et permet une meilleure compréhension de la structure des flux entre différents acteurs. Plusieurs théories relatives à l'optimisation du capital social issu des relations ont été développées (Bourdieu, 1986 ; Granovetter, 1973 ; Coleman, 1988 ; Burt, 1992), pointant l'importance de certaines caractéristiques structurelles du réseau (taille, force des liens, position, densité, trous structuraux). Ces théories sont mobilisées dans de nombreux travaux empiriques, notamment dans l'analyse de réseaux complets de co-authorship. Ces travaux empiriques testent souvent la corrélation entre la structure du réseau et la performance scientifique et mobilisent principalement de larges réseaux sur de courtes périodes. Quelques travaux mobilisent néanmoins des données longitudinales dans l'étude de l'évolution structurelle de réseaux macros nationaux ou micros (dynamiques relationnelles, styles de collaboration, effets de genre et d'homophilie), ce qui nous semble le plus intéressant.

IV- Synthèse de la revue de littérature

Ce premier chapitre nous a permis de faire un premier état des lieux de la littérature existante sur la carrière scientifique et les . Nous avons commencé par décrire le cadre dans lequel le comportement collaboratif se déploie, à savoir la carrière scientifique. La contextualisation des nous semble d'autant plus importante du fait de la dimension individuelle et dynamique de la trajectoire professionnelle scientifique. Ensuite, nous avons affiné notre perception de la carrière scientifique en nous focalisant sur un évènement précis qui apparaît comme prépondérant dans les dynamiques collaboratives : la mobilité géographique. Nous retenons la dimension politique de cet évènement (à l'échelle universitaire et nationale) et donc la nécessité le contextualiser dans la trajectoire scientifique pour comprendre ses effets individuels. Aussi, nous avons souligné que très peu d'articles traitent de l'effet de la mobilité sur les dynamiques collaboratives à l'échelle individuelle. Enfin, nous avons une nouvelle fois affiner notre perception du comportement collaboration en étudiant la littérature de l'analyse des réseaux sociaux. Nous avons montré que la structure du réseau et la position des acteurs ont été largement mobilisés dans le cadre de réseaux complets et d'études relatives à la performance. Très peu d'articles mobilisent ces outils pour apporter une description du comportement collaboratif sous différentes dimensions (notamment sociale et géographique). La figure 6 synthétise les principaux apports de ce chapitre.

Figure 6 : Synthèse du chapitre 1 : revue de la littérature sur la carrière scientifique et le comportement collaboratif



<p>Carrière scientifique</p>	<ul style="list-style-type: none"> • Etudiée dans la littérature relativement aux enjeux de hiérarchie (mentorat), de choix professionnels et de productivité bibliométrique • Dans les études économiques, le chercheur est considéré comme un agent autonome et maximisateur de revenus. Les sciences sociales nous invitent à prendre en compte la dimension individuelle de la carrière du fait du rôle central des objectifs et attentes individuelles. Aussi, la carrière scientifique est intégrée dans un environnement où le métier d'enseignant-chercheur a évolué dans le temps (concurrence, tâches administratives, performance scientifique, mobilité)
<p>Exemple d'évènement professionnel : la mobilité géographique</p>	<ul style="list-style-type: none"> • Dimension politique de la mobilité géographique scientifique : différentes mesures nationales et universitaires • Enjeux territoriaux de la mobilité scientifique • La littérature évalue les effets de la mobilité sur la performance scientifique et la géographie des collaborations. Peu de contributions sur l'effet de la mobilité sur les dynamiques collaboratives
<p>Comportements collaboratifs au sein des réseaux sociaux</p>	<ul style="list-style-type: none"> • Plusieurs théories détaillent les conditions structurelles d'optimisation des liens collaboratifs à l'échelle individuelle et communautaire • Une grande partie des travaux porte sur l'évaluation de la performance scientifique en fonction de la structure de réseaux complets (réseaux personnels intégrés dans une structure) • Quelques travaux mobilisent des données longitudinales pour étudier l'évolution structurelle de réseaux macro et micro (dynamiques relationnelles, styles de collaboration, effets de genre et d'homophilie)

Chapitre 2 : Présentation de la base de données

Afin d'étudier le comportement collaboratif des chercheurs au fil de la carrière, nous avons créé une base de données bibliométrique originale. Le chapitre 2 détaille la méthodologie suivie dans la création de cette base ainsi que ses principales statistiques descriptives. Dans un premier temps, nous décrirons les choix méthodologiques sur lesquels repose la base de données. Ces choix concernent la constitution de l'échantillon (champ disciplinaire, sélection des laboratoires et des chercheurs) et des données bibliométriques. Dans un second temps, nous présenterons les statistiques descriptives relatives à l'activité de publication des chercheurs et aux tendances collaboratives des chercheurs (dans une dimension générale, sociale et géographique). Enfin, nous présenterons deux approches complémentaires pour l'analyse des dynamiques de réseaux personnels : l'économétrie de panel et l'Analyse de Réseaux Sociaux.

I- Création de la base de données et choix méthodologiques

Cette première partie a pour but de décrire les choix méthodologiques qui ont conduit à la création de notre base de données bibliométrique. Cette base apparaît originale tant du fait de la diversité que de la finesse des données mobilisées. Nous commencerons par décrire les choix méthodologiques relatifs à la constitution de l'échantillon de recherche. Tout d'abord, nous décrirons les raisons pour lesquelles la chimie est un champ disciplinaire approprié pour étudier le comportement collaboratif des chercheurs. Aussi, nous choisissons de focaliser notre échantillon sur deux laboratoires de chimie dans l'Ouest de la France, à la fois comparables et disposant de spécificités. Enfin, nous avons sélectionné au sein de ces structures 80 chercheurs à partir de critères précis. Dans un second temps, nous détaillerons la méthodologie relative à la construction de la base à partir de données bibliométriques. Nous commencerons par décrire l'extraction des publications et la création des paires de co-authorship. Ensuite, nous détaillons les informations complémentaires intégrées dans la base de données et les opérations de codages suivies. Enfin, nous expliquerons pour quelles raisons et dans quelle mesure nous avons regroupé les chercheurs de notre échantillon en cohortes générationnelles.

A) *Choix de l'échantillon*

La création de la base de données repose sur plusieurs choix méthodologiques, à savoir le choix de la discipline étudiée, le choix des laboratoires et des chercheurs sélectionnés.

1- Etude de cas de la recherche en chimie

Dans cette thèse, nous avons fait le choix de nous focaliser sur une unique discipline de recherche, afin d'assurer la comparabilité des activités et des pratiques scientifiques : **la chimie**. Cette discipline est définie par le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation français comme « *une science de la création et de la transformation de la matière [...] source d'innovations et un moteur de développement économique, capable d'apporter des solutions aux défis liés à l'énergie, la santé, l'eau, l'alimentation, la préservation des ressources et l'environnement* ». Cette discipline est traditionnellement décomposée en six sous-disciplines selon l'Institut de chimie (Annexe 1).

Nous choisissons cette discipline pour trois raisons :

i) La chimie française montre de fortes performances scientifiques mondiales, notamment concernant la valorisation économique sur le marché européen ;

ii) Cette performance scientifique provient d'une stratégie nationale de spécialisation thématique. Ainsi, les chimistes français sont réputés pour la qualité de leurs travaux en chimie minérale et nucléaire ainsi qu'en chimie des matériaux ;

iii) La bibliométrie des travaux en chimie offre de très belles opportunités dans l'étude du comportement collaboratif des chercheurs (taille des équipes de publication, visibilité des travaux, dimension géographique de l'activité et des collaborations).

La France et la recherche en chimie : Enjeux scientifiques mondiaux

La France occupe une place importante dans la discipline. Elle fait partie des 10 plus importants pays producteurs de publications en chimie. Les chercheurs affiliés à la France ont réalisé 2,7% des publications mondiales de chimie entre 2015 et 2017³⁵. Le tableau 10 classe les 15 pays mondiaux qui contribuent le plus dans la recherche en chimie selon trois critères de contribution :

i) Contribution dans le co-authorship mondial : Le premier critère concerne la part de chaque pays dans le volume mondial de publications de chimie. Cet indicateur est calculé en compte fractionnaire (à partir de la localisation de chaque membre d'équipe de publication) : il s'agit d'une approximation du poids de chaque pays dans le co-authorship mondial.

³⁵ Sources : Clarivate Analytics, Web of Science, traitements OST du HCERES

ii) Valorisation scientifique : Le second critère évalue la valorisation scientifique des articles selon la nationalité des co-auteurs. L'indice d'impact national en chimie est le ratio entre le nombre moyen de citation par publication de chimie d'un pays sur le nombre moyen de citation par publication de chimie pour le monde. Ainsi, il évalue si les publications de chimie d'un pays sont en moyenne plus citées (si supérieur à 1), également citées (indice égal à 1) ou moins citées que la moyenne des publications de chimie mondiales (indice compris entre 0 et 1). Cet indicateur n'a pas pour but de quantifier les travaux nationaux mais plutôt de qualifier la contribution qu'ils ont dans la recherche chimique mondiale.

iii) Valorisation économique : Le troisième critère porte sur la valorisation économique des travaux en chimie selon les pays. Nous évaluons donc le poids du dépôt de brevet en chimie de chaque pays dans l'ensemble des demandes de brevets en chimie de l'OEB. L'objectif d'un brevet est de promouvoir le développement technique et/ou industriel de travaux, soit en vendant les droits des inventeurs aux producteurs, soit en incitant les producteurs à innover eux-mêmes, sous le statut protecteur du monopole durant une période. Sachant que le dépôt de brevet n'est validé que sur un territoire donné, nous avons choisi de ne retenir que les dépôts de brevets européens. Cet indicateur évalue donc les pays qui contribuent à l'innovation en chimie sur le territoire européen.

En termes de **co-authorship mondial**, la **France fait partie des pays européens qui contribuent à moins de 3% du volume de publications internationales** entre 2015 et 2017, tels que le Royaume-Uni (2,60%), l'Espagne (2,20%), l'Italie (2,00%) et la Pologne (1,70%). L'Allemagne est une exception européenne puisqu'elle est le 5^{ème} pays mondial en contribuant à 4,30% du stock d'articles de chimie mondial. Nous notons de grandes disparités géographiques entre les contributions des différents pays. En effet, l'Asie est le continent qui a le plus produit d'articles en chimie, soit 52,10% du volume mondial entre 2015 et 2017. En effet, parmi les 10 pays qui ont le plus contribué en la matière, 6 sont asiatiques. La Chine est le pays qui a produit le plus d'articles en chimie entre 2015 et 2017 avec près d'un tiers du volume mondial. A ses côtés, d'autres pays asiatiques ont une contribution non négligeable dans le volume mondial de publications : l'Inde (6,30%), le Japon (4,80%), la Corée du Sud (4,20%), la Fédération de Russie (3,40%), l'Iran (2,70%). Ensuite, les Etats-Unis est le second pays qui contribue le plus au volume mondial d'articles en chimie (11%).

Comprendre le positionnement scientifique d'un pays ne peut se limiter à une approche quantitative mais doit être complété par une dimension qualitative des différents travaux. En effet, la contribution scientifique en termes de citations suit une toute autre répartition que celle du volume de publications. Bien que certains pays européens n'aient qu'une faible contribution individuelle dans le co-authorship mondial, ils montrent de meilleures performances relativement à l'impact de leurs publications. Par exemple, **la France est au 7^{ème} rang mondial relativement à la valorisation scientifique et dispose d'un impact normalisé supérieur à la moyenne mondiale**. C'est également le

cas du Royaume-Uni, l'Allemagne, l'Espagne et l'Italie. Bien que l'Asie soit le principal producteur mondial d'articles de chimie, il s'avère que ses travaux sont très peu cités. En effet, de nombreux pays asiatiques, dotés d'un rang élevé en termes de volume de publication, ont un taux moyen de citation inférieur à la moyenne mondiale. Par exemple, entre 2015 et 2017, bien que l'Inde soit au rang de 3ème mondial concernant le volume de publication, son indice normalisé est de 0,79 (il détient donc la 12ème place concernant le nombre moyen de citation par article). Le Japon, l'Iran et la Fédération d Russie sont dans la même situation. La Chine et la Corée du Sud ont un indice d'impact légèrement supérieur à 1 : leurs articles en chimie sont en moyenne légèrement plus cités que la moyenne mondiale (ce qui leur vaudra les rangs respectifs de 9ème et 8ème). Sur le continent nord-américain, les Etats-Unis est le pays dont les articles sont les plus cités (indice d'impact de 1,45). Le pays cumule l'avantage d'être le second plus grand producteur d'article de chimie et le plus cité.

La chimie française est caractérisée par une forte valorisation économique de ses travaux. En effet, entre 2015 et 2017, elle se situe au **4^{ème} rang mondial grâce au dépôt de 6,7% des brevets de chimie à l'OEB**. En Europe, elle se situe comme intermédiaire entre l'Allemagne (2^{ème} rang mondial grâce au dépôts de 18% des brevets) et le Royaume-Uni (5^{ème} rang grâce au dépôt de 3,9% des brevets). **Ainsi, entre 2015 et 2017, la France est la seconde puissance européenne en termes de valorisation économique de sa recherche en chimie**. En dehors du continent européen, les Etats-Unis restent le pays qui a le plus contribué entre 2015 et 2017 dans le dépôt de brevets européens (25,7% des dépôts de brevets en chimie à l'OEB). Ce pays apparaît comme un acteur de premier plan dans la recherche en chimie mondiale, que ce soit en termes de source de co-authorship ou de valorisation scientifique et économique. Enfin, très peu de pays asiatiques contribuent au dépôt de brevets en chimie à l'OEB. Le Japon est le principal pays asiatique qui valorise ses travaux sur le marché européen : entre 2015 et 2017, il a déposé 17,6% des brevets. Bien que ce pays ait un impact normalisé inférieur à la moyenne mondiale, il constitue tout de même une forte source de production mondiale de publications et de brevets européens.

Tableau 10 : Positionnement de la France dans la recherche en chimie

		Co-authorship mondial (2015-2017)		Valorisation scientifique (2013-2015)		Valorisation économique en Europe (2015-2017)	
		Rang mondial	Part mondiale de publications en chimie	Rang mondial	Impact normalisé à 3 ans en chimie	Rang mondial	Part mondiale des demandes de brevets à l'OEB en chimie
Europe	Allemagne	5	4,3%	3	1,17	2	18,0%
	Autriche					12	1,3%
	Belgique					11	1,6%
	Espagne	11	2,2%	5	1,11	15	1,0%
	<u>France</u>	9	2,7%	7	1,04	4	6,7%
	Italie	12	2,0%	6	1,06	9	2,6%
	Pays-Bas					10	2,4%
	Pologne	13	1,7%	14	0,64		
	Royaume-Uni	10	2,6%	2	1,32	5	3,9%
	Suède					14	1,0%
	Suisse					8	2,6%
Amérique	Canada	14	1,6%	4	1,15	13	1,2%
	États-Unis	2	11,0%	1	1,45	1	25,7%
Asie	Chine	1	29,3%	9	1,02	7	2,8%
	Corée du Sud	6	4,2%	8	1,03	6	3,2%
	Fédération de Russie	7	3,4%	15	0,37		
	Inde	3	6,3%	12	0,79		
	Iran	8	2,7%	13	0,78		
	Japon	4	4,8%	11	0,83	3	17,6%
	Taïwan	15	1,4%	10	0,83		

Source : Clarivate Analytics, Web of Science, traitements OST du HCERES - Données extraites de la publication État de l'Enseignement supérieur, de la Recherche et de l'Innovation en France n°12

Les champs thématiques de la chimie française

La France se positionne également vis-à-vis des autres pays en fonction **des thématiques de recherches** dans lesquelles elle se spécialise. Nous comparerons la spécialisation française dans les domaines de la chimie avec les deux pays les plus contributeurs en termes de publications : la Chine et les Etats-Unis. Pour identifier ces domaines de la chimie, nous utilisons deux indices :

i) *L'indice de spécialisation du pays* correspond au rapport entre la part d'une discipline dans les publications nationales sur la part de la même discipline dans les publications internationales. Cet indice est calculé en compte fractionnaire sur la période 2015-2016-2017. Ainsi, un indice de spécialisation supérieur à 1 indique une forte spécialisation thématique du pays relativement aux publications mondiales alors qu'un indice inférieur à 1 sous-entend une faible spécialisation.

ii) *L'indice d'activité nationale d'un domaine de recherche dans le 1^{er} décile de citations (10% des articles les plus cités)* correspond au ratio entre la part des articles du domaine nationaux dans cette classe sur la part des articles du domaine internationaux dans cette classe. Cet indice est calculé en compte fractionnaire sur la période 2013-2014-2015³⁶ (citations à 3 ans). Ainsi, un indice d'activité supérieur à 1 signifie que les travaux nationaux dans ce domaine sont surreprésentés dans le décile d'articles les plus cités par rapport à la moyenne internationale et un indice d'activité inférieur à 1 signifie que les travaux nationaux sont sous-représentés dans cette classe d'articles.

Dans la figure 7, nous constatons que la **France a une spécialisation thématique** différente de celles des Etats-Unis et de la Chine.

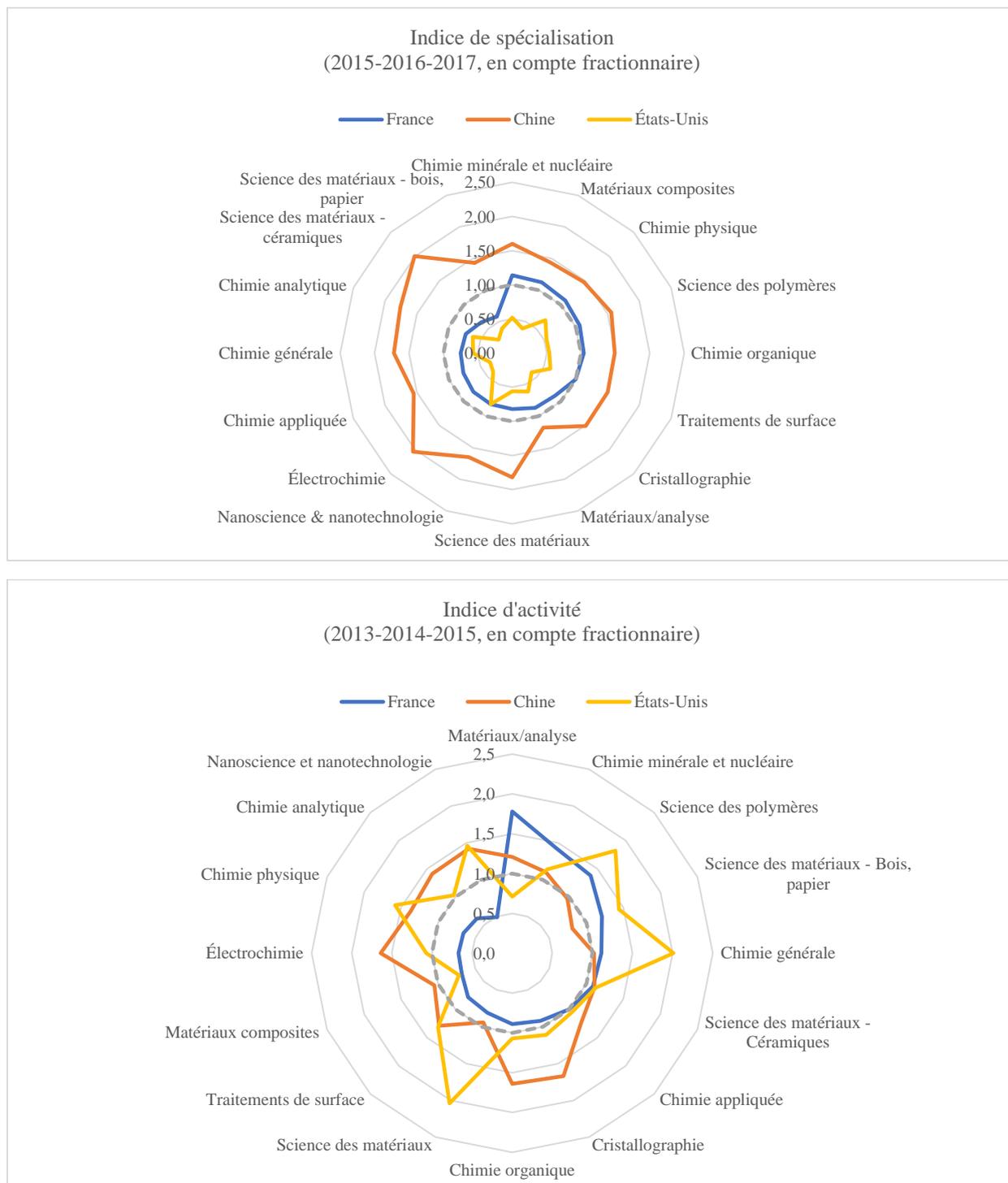
i) Tout d'abord, nous constatons que la France dispose d'un **indice de spécialisation** supérieur à la moyenne mondiale dans **cinq des sous domaines de la chimie**. Les indices les plus élevés concernent la « **chimie minérale et nucléaire** » (1,14) et les « **matériaux composites** » (1,12). Ainsi, la part des publications de ses sous-domaines dans l'ensemble des publications nationales est supérieure à la part mondiale. En revanche, le sous-domaine de « science des matériaux, bois, papier » est associé à l'indice de spécialisation le plus bas du pays. En comparaison, les Etats-Unis sont très peu spécialisés dans la chimie : en effet, tous leurs indices de spécialisation sont inférieurs à 1. La Chine a de son côté choisit une stratégie inverse puisqu'elle dispose d'indices de spécialisation supérieurs à la moyenne mondiale pour chaque sous-domaine (particulièrement pour l'« électrochimie » et la « science des matériaux céramiques »).

ii) Ensuite, la France présente un **indice d'activité** supérieur à la moyenne mondiale et à ceux de la Chine et des Etats-Unis pour les sous-domaines de « **matériaux/analyse** » (1,8) et « **chimie minérale et nucléaire** » (1,4). Ainsi, les travaux français dans ces deux thématiques chimiques sont

³⁶ Bien que ces deux indices ne correspondent pas aux mêmes périodes, nous les comparons sous hypothèse de stabilité des spécialisations thématiques de la France

surreprésentés dans les travaux **à forte qualité relativement aux publications mondiales**. Les Etats-Unis sont davantage spécialisés d'un point de vue qualitatif dans les thématiques « sciences des matériaux » et « chimie générale » alors que la Chine se distingue via la « cristallographie » et la « chimie organique ».

Figure 7 : Indices de spécialisation et d'activité de la chimie



Source : Clarivate Analytics, Web of Science, traitements OST du HCERES - Données extraites de la publication État de l'Enseignement supérieur, de la Recherche et de l'Innovation en France n°12

Enfin, il nous semble que la chimie est une discipline au sein de laquelle le comportement collaboratif des chercheurs est particulièrement dynamique :

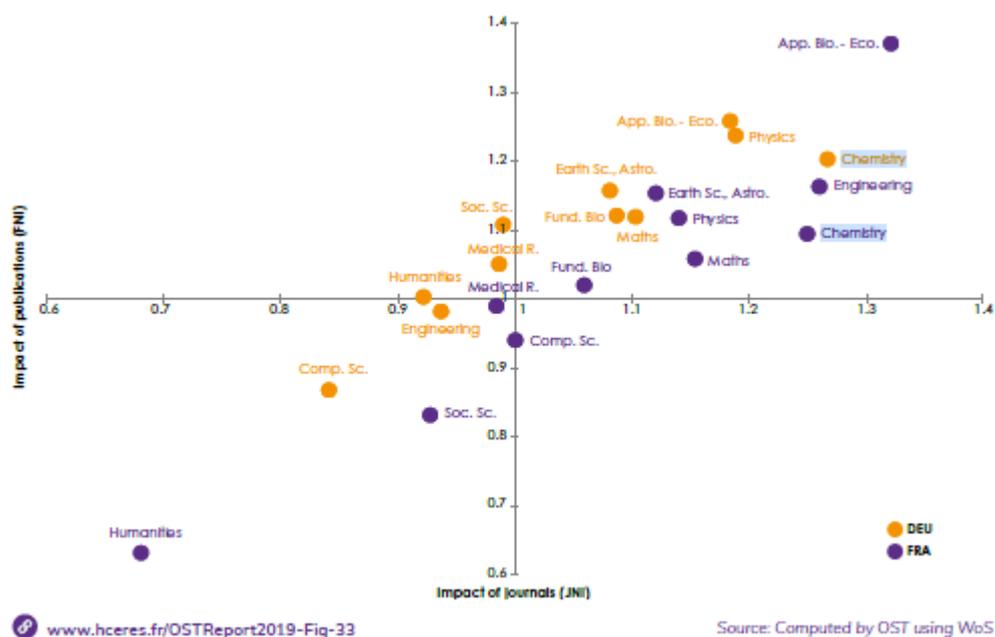
i) Les travaux de chimie sont associés à un **fort impact de citation** par article relativement aux autres disciplines

ii) Les équipes de publication sont de **taille importante mais modérée relativement à d'autres disciplines**

iii) La **géographie des collaborations** est proche de la moyenne nationale et les **collaborations internationales sont structurantes**

iv) L'élargissement de la dimension géographique de l'équipe traduit une faible hausse de l'impact de citation de l'article : **la collaboration internationale est soumise à d'autres motivations que la performance scientifique.**

Figure 8 : Impact de citation normalisé des publications et revues



Le graphique précédent permet de comparer les travaux scientifiques issus de chaque discipline en fonction de l'impact de citation normalisé de l'article et de la revue. Il présente les résultats de la France ainsi que ceux de l'Allemagne. Il nous semble intéressant d'analyser la position particulière des travaux de chimie français dans ce graphique. En effet, **les articles français de chimie disposent à la fois d'un très fort impact normalisé de publication et d'un impact de revue supérieur à la moyenne.**

Tableau 11 : Nombre moyen de co-auteurs par publication par discipline entre 2000 et 2004 et 2012 et 2016

	Monde		France	
	2000-04	2012-16	2000-04	2012-16
Physique	4.5	7.6	10.0	52.6
Sciences de la terre, astrophysique	3.5	5.8	5.0	23.7
Recherche médicale	4.6	5.8	5.7	8.2
Biologie fondamentale	4.4	5.7	5.6	7.9
Biologie appliquée, écologie	3.4	4.6	4.3	6.4
Chimie	3.8	4.7	4.5	5.8
Ingénierie	3.2	3.8	4.3	4.9
Sciences sociales	2.3	3.1	3.3	4.3
Sciences informatiques	2.7	3.4	2.9	3.9
Sciences humaines	2.0	2.5	2.2	3.1
Mathématiques	2.0	2.4	2.0	2.5

Source : www.hceres.fr/OSTReport2019-Tab-3

Également, la recherche en chimie est caractérisée par sa **nature particulièrement collaborative** au regard des publications des autres disciplines (Kannappanavar *et al.*, 2004, Viera et Gomes, 2010). Selon l’HCERES, les publications françaises de chimie étaient co-écrites en moyenne par 5,8 auteurs entre 2012 et 2016. Les équipes de publication de cette discipline nous semblent être de taille suffisamment importante pour évaluer les dynamiques collaboratives. Relativement à **des disciplines où les équipes sont plus petites** (mathématiques, sciences humaines, sciences informatiques, sciences sociales), les équipes de chimie sont suffisamment riches en liens collaboratifs pour reconstituer des dynamiques collaboratives au fil de la carrière. En revanche, nous pensons que certaines disciplines disposent d’équipes trop grosses pour notre étude (physique, science de la terre et astrophysique). En effet, il nous semble difficile d’appréhender la cohésion de l’équipe de publication, la contribution de chaque membre ainsi que la relation que chaque chercheur entretient avec chaque membre de l’équipe. **La chimie est donc un bon compromis** puisqu’elle est constituée par des équipes de recherches qui sont à la fois **riches et de taille raisonnable**.

La figure 9 décrit la distribution des publications françaises en 2016 de chaque discipline en fonction de la géographie du co-authorship (même adresse, même pays, équipe européenne, équipe hors Europe). Alors que certaines disciplines sont caractérisées par un co-authorship resserré géographiquement (par exemple, les sciences humaines), d’autres sont marquées par un co-authorship très internationalisé (par exemple, les sciences de la terre et l’astrophysique). **La distribution des équipes en chimie est très proche de la distribution moyenne de l’ensemble des disciplines**. Sur l’ensemble des publications en chimie françaises en 2016, 14% d’entre elles ont été réalisées par des chercheurs d’une même adresse et 25% d’entre elles sont issues de collaborations nationales. Ainsi, 61% des publications sont issues de collaborations internationales (21% sont des équipes européennes et 40% sont des équipes hors Europe). **La chimie est donc structurée par la collaboration internationale**.

Figure 9 : Distribution des publications de France selon la géographie de l'équipe et la discipline, 2016

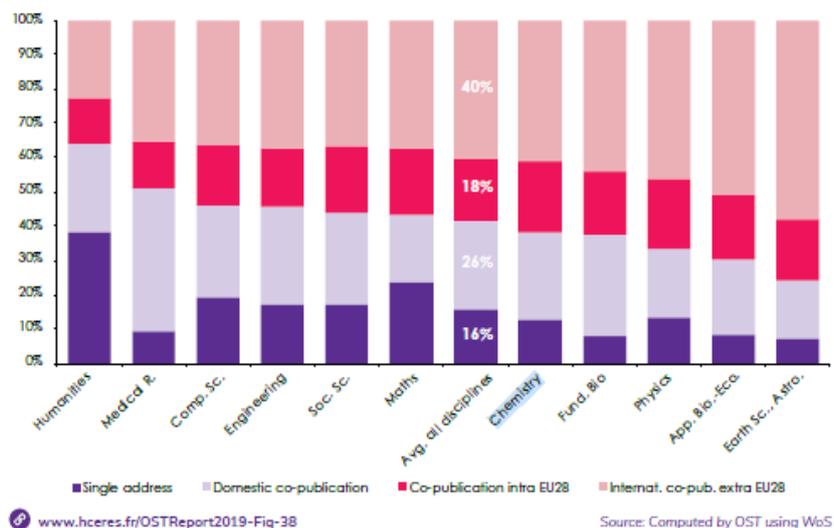
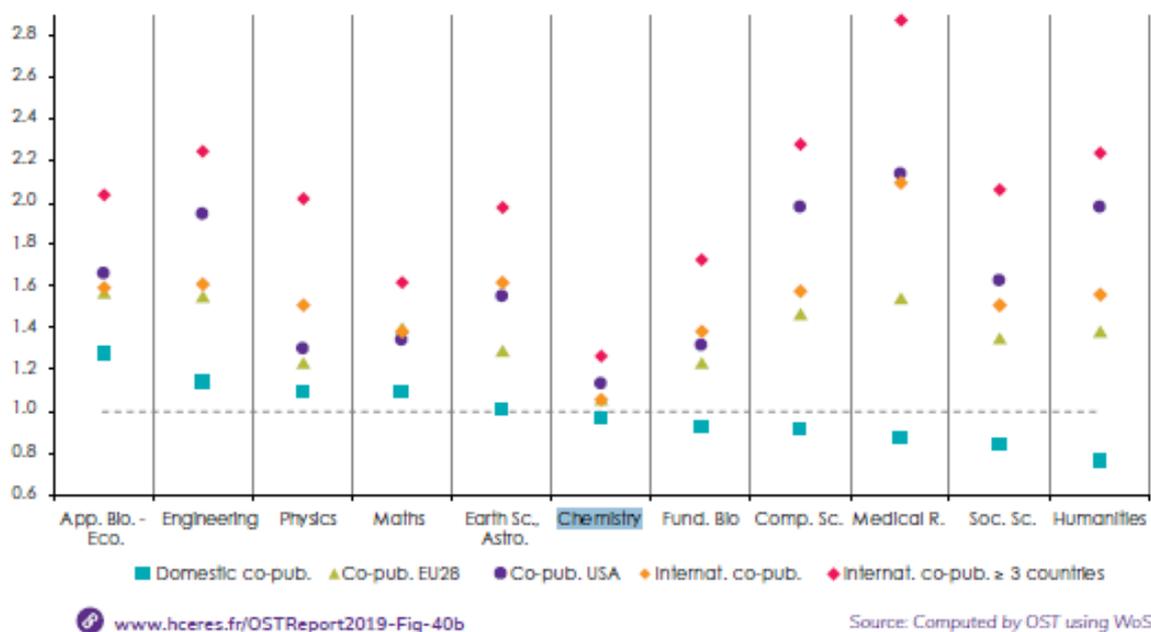


Figure 10 : Impact des co-publications françaises par type de collaboration et discipline, 2014

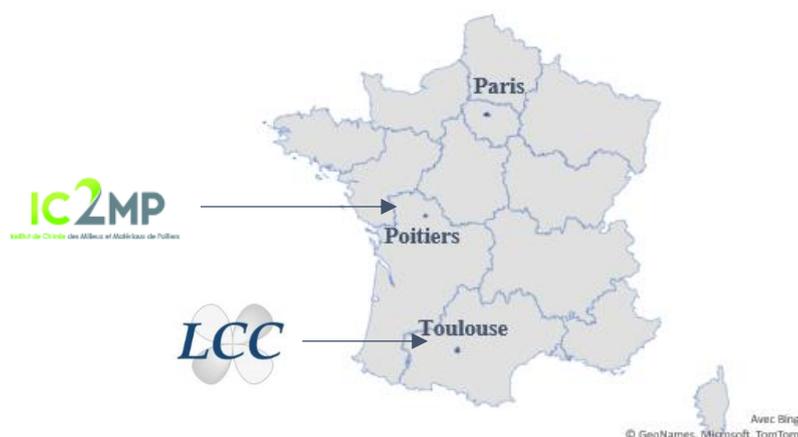


Contrairement aux autres disciplines, l'impact de citations des publications de chimie n'augmente pas significativement avec la dimension géographique de l'équipe. En effet, toutes les disciplines montrent des publications issues d'équipe domestiques de bien moindre impact que celles d'équipes internationales. Or, en chimie, les différentes dimensions géographiques ont très un faible écart d'impact

de citation. Ainsi, **la chimie est un champ disciplinaire adapté pour étudier le comportement collaboratif du chercheur puisqu'il implique d'autres motivations que la visibilité scientifique des travaux.**

2- Etude de cas de deux laboratoires

Figure 11 : Carte de la France et laboratoires IC2MP et LCC



Nous choisissons de cibler notre analyse bibliométrique sur **deux laboratoires de chimie français** : IC2MP de Poitiers (Institut de Chimie des Milieux et Matériaux, UMR CNRS 7285) et le LCC de Toulouse (Laboratoire de Chimie de Coordination, UPR 8241). Ces deux structures, localisées sur la façade ouest de la France, disposent toutes les deux d'une forte réputation (nationale et internationale) et comportent chacune certaines spécificités en termes de spécialisation thématique et de structuration interne.

Comparabilité des laboratoires

i) Les deux structures sont spécialisées sur des thématiques de recherche précises :

Né en 2012 à la suite de la fusion de quatre laboratoires, l'**IC2MP** est une structure interdisciplinaire en chimie et géoscience, dont les travaux sont articulés autour de la thématique de la **chimie durable**. En effet, dans son rapport de la campagne 2016-2017, l'HCERES décrit ce laboratoire comme « *une unité présentant désormais une offre intégrée de recherche autour de la chimie durable, ce qui est un gage de nouvelles opportunités* ». Ainsi, ce laboratoire a choisi de se positionner sur les thématiques environnementales (préservation de la quantité et la qualité des ressources, intégration de la chimie dans

les filières d'avenir) et ses travaux concernent une grande partie de la chaîne de valeur des procédés. L'institut est segmenté en cinq équipes de recherche :

Tableau 12 : Equipes de l'IC2MP

Equipes	Thématique de recherche	Recherches et chimie durable
E-BICOM : Biomarqueurs, Contaminants Organiques, Milieux	Eaux, Chimie de l'Eau Géochimie Organique	Impact de l'activité humaine sur les sols et les eaux Utilisation des milieux comme ressource
HydrASA : Hydrogéologie, Argiles, Sols et altérations	Minéraux argileux et leurs agrégats Profils des sols argileux Grands systèmes géologiques	Compréhension et modélisation des transferts dans les géosystèmes réactifs
SAMCat : du Site Actif au Matériau Catalytique	Contrôle nano morphologique des matériaux catalytiques et développement d'outils de caractérisation spécifique Traitement des polluants et matériaux pour l'absorption Valorisation de la biomasse et activation des gaz à effet de serre Electro-catalyse pour l'énergie Propulsion spatiale	Développement de matériaux catalytiques et électro catalytiques Applications dans les domaines de l'énergie et de l'environnement
MediaCat : Catalyse et Milieux non- Conventionnels	Mécanismes réactionnels Milieux non-conventionnels Méthodes d'activation Matériaux dédiés	Mise en œuvre de réactions catalytiques en milieux non- conventionnels pour améliorer le rendement ou la sélectivité
SynthOrg : Synthèse Organique	Synthèse Glycochimie Synthèse en milieu superacide Systèmes moléculaires programmés	Chimie verte Vectorisation Reconnaissance moléculaire Domaines thérapeutiques

Source : <http://ic2mp.labo.univ-poitiers.fr/index.php/institut/>

De son côté, le LCC est un laboratoire beaucoup plus ancien puisqu'il est né en 1974. Historiquement, il s'articule autour de **trois thématiques de recherche** : i) la chimie fine, la chimie de coordination et la catalyse (application dans le développement durable, l'environnement et l'énergie), ii) les matériaux moléculaires (application dans les technologies quantiques : la physique, les nanosciences et les nanotechnologies), iii) la chimie bio-organique et le rôle des métaux en biologie (application dans la santé et les sciences de la vie). Le LCC a structuré ces thématiques autour de la conception de molécules et d'édifices polyatomiques originaux. Le tableau suivant décrit les 16 équipes de recherche de l'institut en fonction de leur thématique de recherche et des activités de recherche mobilisées :

Tableau 13 : Equipes du LCC

	Activités de recherche				
	<i>Synthèse moléculaire</i>	<i>Catalyse</i>	<i>Conception de nano-objects</i>	<i>Etude des propriétés physiques et chimiques des objets et matériaux</i>	<i>Conception de molécules à visées thérapeutiques</i>
Equipes de l'interface Environnement et Energie (chimie fine, chimie de coordination et catalyse)					
Ingénierie moléculaire des pré-catalyseurs	x	x			
Catalyse et chimie fine	x	x	x		
Molécules instaurées pour la physique, la biologie et la chimie	x	x	x	x	x
Ligands, architectures complexes et catalyse	x	x	x		
Dendrimères et hétérochimie	x	x	x	x	x
Activation de petites molécules	x	x			
Equipes de l'interface Physique, nanosciences et nanotechnologies (Matériaux moléculaires)					
Molécules et matériaux			x	x	
Ingénierie des nanoparticules métalliques	x	x	x	x	
Matériaux moléculaires commutables			x	x	
Molécules et composites pour l'optique	x			x	x
Matériaux moléculaires et supramoléculaires	x		x	x	
Nanochimie, organisation et capteurs	x		x	x	
Equipes de l'interface Santé et Sciences de la vie (chimie bio-organique et rôle des métaux en biologie)					
Alzheimer et amyloïdes : bases moléculaires, thérapie par chélation et détection	x	x			x
Agents anti-infecteurs moléculaires et macromoléculaires	x				x
Chimie médicinales pour l'oncologie					
Nouvelles molécules antipaludiques et approches pharmacologiques	x	x			x

Source : <https://www.lcc-toulouse.fr/article53.html>

ii) Les deux structures ont mis en place des plateformes de collaboration interne :

Également, les stratégies de développement des deux laboratoires visent à favoriser la **collaboration interne entre les équipes au travers d'actions et de plateformes transversales**. En effet, dans son rapport de campagne 2016/2017, l'HCERES écrit concernant l'IC2MP que « *les interactions entre équipes sont encouragées au travers d'actions transverses, qui commencent à porter leurs fruits. Cependant, le bilan de ces actions transverses n'est pas complètement abouti* » et

concernant le LCC que « *Par ailleurs, l'organisation interne de l'unité s'appuie sur des plateaux techniques mutualisés de haut niveau, très complémentaires, et concentrant une part importante des moyens (financiers et humains) de l'unité, avec des personnels fortement motivés pour le service aux projets de recherche* » et « *La structuration proposée en trois axes devra permettre de développer la synergie entre les équipes au sein de chaque axe* ». Plus concrètement, l'IC2MP a mis en place cinq plateformes transverses (« Evabio », « Ingénierie, développement et support technique », « Plateforme Eaux », « PreMiMAT », « Site Expérimental Hyfrogéologique ») afin de fournir les ressources nécessaires aux recherches des différentes thématiques du laboratoire. Ces plateformes reposent sur la transversalité des thématiques de recherche. Aussi, le LCC fait partie de l'Institut de Chimie de Toulouse (avec les laboratoires CEMES, IMRCP, LCA, LHFA, LPCNO, Pharma-Dev, SPCMIB). Cet institut fédère huit laboratoires de Toulouse dans tous les domaines de la chimie moléculaire (de l'étude fondamentale aux applications). Il met à disposition une plateforme scientifique (ouverte aux académique et aux industriels) afin de favoriser l'accès aux ressources, la collaboration et la valorisation des travaux scientifiques.

iii) Les deux structures disposent d'une forte visibilité nationale et internationale :

Les deux laboratoires disposent d'une forte visibilité à l'échelle nationale et internationale, ce qui se traduit par divers partenariats.

Tout d'abord, l'IC2MP détient une forte réputation pour ses travaux en chimie verte. En effet, le laboratoire fait partie de la Fédération INCREASE (*International Consortium on Renewable Resources*), du Global Network of Green Chemistry et de la Fédération FREDD (*Fédération de Recherche en Environnement pour le Développement Durable*). Cette notoriété concerne également le monde industriel puisque le laboratoire a également participé au congrès ISGC 2019 (plus grande concentration of de chercheurs en chimie verte et d'entreprises de chimie industrielle) qui a débouché sur la signature avec le groupe SOLVAY d'une convention de partenariat entre Shangai et Poitiers (en partenariat avec CNRS et de l'Université de Poitiers). Dans cet accord international, l'IC2MP devient le site miroir de l'Eco-Efficient Products & Processes Laboratory de Shanghai afin de favoriser et renforcer les collaborations scientifiques entre les deux structures. L'IC2MP détient également une forte notoriété pour ses travaux relatives à l'argile (partenariats avec des sociétés savantes nationales et internationales, pilotage d'un programme international d'étude des paleo-environnements à la suite de découvertes de fossiles au Gabon (2010)), dans les piles à combustible (membre du groupement de recherche Hydrogène, systèmes et Piles à combustible" et du « Natural Sciences and Engineering Research Council » du Canada), et dans les milieux fluorés et fluorants (membre du réseau GIS Fluor).

De son côté, le LCC est considéré comme moteur dans la création de plateformes communes d'équipement dans la région Midi-Pyrénées (plan MRN Midi-Pyrénées, plateforme « NANMAT », plateforme « TEMSCAN », projet « FEDER NANOMAT 2019 »). Également, le LCC est aussi associé

à la création de laboratoires internationaux associés (LIA). En Octobre 2013, le Laboratoire de Chimie Moléculaire avec applications dans les Matériaux et la Catalyse (LCMMC) est créé à la suite d'un accord entre l'Institut de Chimie de Toulouse et un consortium d'universités mexicaines. Aussi, en janvier 2018, le laboratoire de Chimie de Coordination pour la polymérisation radicalaire contrôlée est créé à la suite d'un accord entre le LCC, le laboratoire IMRCP de Toulouse et les départements de chimie des universités Carnegie Mellon et Duquesne de Pittsburg (Etats-Unis). Enfin, cette notoriété concerne également le monde industriel puisque le LCC valorise ses travaux via le renforcement de ses relations industrielles (Minakem, Cyclopharma, Etienne Lacroix, Sanofi, Palumed, Bayer, STMicroelectronics SAS, Renault, Sasol, Alpha MOS, Blue Star Silicones, Astrium, Gelfor...) et via la création de start-up ou sociétés anonymes ("Nanomeps", "Dendris", "Biodendrimer International").

iv) Les deux laboratoires sont reconnus pour la qualité de leurs travaux scientifiques :

Enfin, ces deux laboratoires sont également reconnus pour **la qualité de leur activité scientifique** : cela se traduit dans le rapport d'évaluation du HCERES et dans leurs données relatives à la valorisation scientifique et économique des travaux.

Dans sa campagne 2016-2017, le HCERES qualifie le laboratoire IC2MP de la manière suivante : « *L'unité présente un excellent bilan scientifique, avec un taux de publications très satisfaisant. L'ensemble des équipes présente des activités de recherche de très bonne qualité, dont certaines relèvent de l'excellence dans leur discipline. Il est à souligner que l'unité fait preuve d'un engagement sans faille dans la vulgarisation scientifique et dans la dissémination de leurs activités au grand public* ». En effet, entre 2015 et 2018, 740 publications ont été réalisées par les chercheurs du laboratoire et 19 brevets ont été déposés. Aussi, depuis sa création en 2012, les travaux scientifiques ont été valorisés économiquement via la création de 7 start-ups par les chercheurs du laboratoire.

Concernant le LCC, le HCERES a qualifié dans sa campagne d'évaluation de 2014-2015 ce laboratoire de la façon suivante : « *le LCC est à l'origine, en chimie de coordination et hétérochimie, d'avancées conceptuelles majeures dans ses trois domaines d'activité principaux, à savoir Chimie fine et catalyse, Matériaux moléculaires et Chimie bio-inorganique. Sa production scientifique est très importante en nombre et qualité, et même exceptionnelle pour certaines équipes. Elle se combine à une activité de valorisation en forte croissance (ressources propres, contrats industriels, brevets, start up, ...) et une politique volontariste de diffusion de la culture scientifique, localement et au travers de structures nationales*. ». Dans ce rapport, le HCERES souligne que le LCC a réalisé plus de 7210 articles à comité de lecture entre 1974 et 2013, soit environ 183 articles par an.

i) Des histoires différentes

Les deux laboratoires bénéficient d'une **ancienneté différente**. En effet, le LCC est un laboratoire qui a été créé en 1974, en tant qu'unité de recherche spécialisée dans la chimie de coordination en chimie inorganique. Au fil des années, le laboratoire a intégré de nouvelles thématiques de recherche via l'intégration de nouveaux chercheurs dans la structure et la constitution de nouvelles équipes de recherche. En 1979, le laboratoire rapproche sa ligne directrice de chimie de coordination vers des applications biologiques (réactivité des complexes, chimie organométallique). Dans les années 1980, de nouvelles équipes se développent dans l'étude des matériaux moléculaires et une équipe en catalyse homogène émerge en 1986. A la suite d'un incendie dans un bâtiment de chimie de l'Université Toulouse III-Paul Sabatier, le LCC accueille de nouveaux chercheurs chimistes issus de ces équipes sinistrées. L'arrivée de ces nouvelles équipes va élargir les thématiques du laboratoire à la suite de l'intégration de la chimie des hétéro-éléments (dendrimères, chimie des carbènes) et va favoriser la collaboration entre les différents laboratoires de chimie de Toulouse. Cet événement va impulser l'émergence d'une communauté de chimistes interconnectée dans la ville. Au début des années 1990, le laboratoire réoriente ses thématiques clefs vers la micro-électronique (catalyse asymétrique, nanoparticules). Entre 1998 et 2006, le laboratoire a connu une extension et une réhabilitation partielle. Les thématiques de recherche se renforcent sur la biologie, la chimie organométallique et la catalyse. L'intégration de nouveaux chercheurs a permis de constituer de nouvelles équipes dans la thématique des Matériaux Moléculaires. A partir de 2007, de nouvelles équipes affiliées à l'institut National Polytechnique de Toulouse et à l'IUT de Castres intègrent le laboratoire.

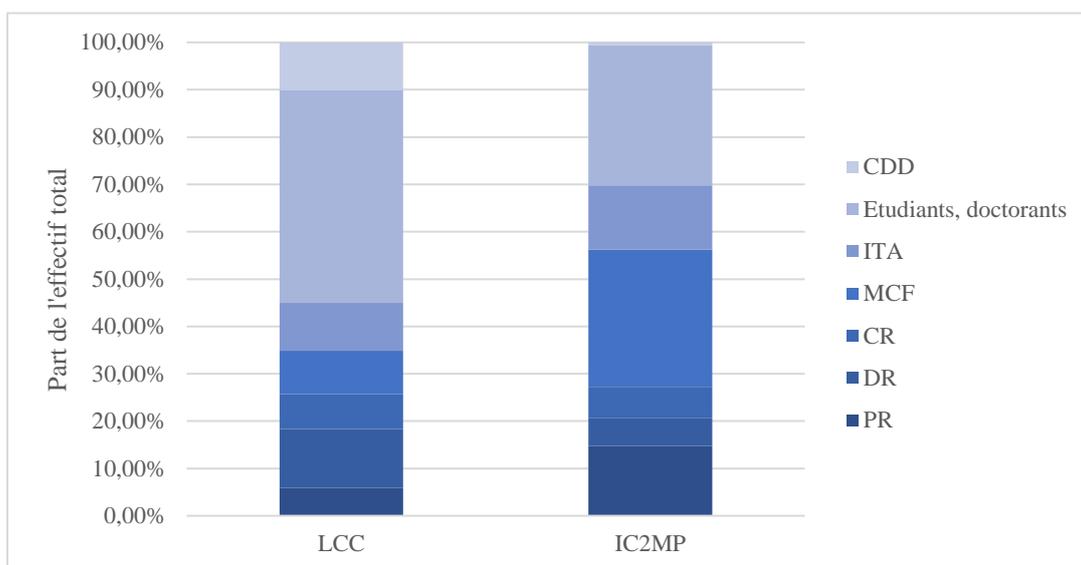
L'IC2MP est une structure plus récente, bien qu'elle dispose d'une forte expérience dans la chimie. En effet, le laboratoire est né en 2012 de la fusion de quatre équipes spécialisées et présentes à Poitiers depuis une quarantaine d'années : i) HydrASA (Hydrogéologie, Argiles, Sols, Altération), ii) LACCO (Laboratoire Catalyse en Chimie Organique), iii) LCME (Laboratoire Chimie et Microbiologique de l'Eau), et iv) SRSN (Synthèse et Réactivité des Substances Naturelles).

ii) Unités de recherche et composition du personnel

Bien que l'IC2MP et le LCC soient de même taille (effectif d'environ 250 personnes, dont 100 chercheurs permanents), **la composition du personnel de recherche diffère en lien avec le label d'unité de recherche des deux structures**. En effet, l'IC2MP est une Unité Mixte de Recherche (UMR) : il s'agit d'une unité multi-tutelles qui est rattachées à la fois au CNRS et l'Université de Poitiers. Cette co-gestion se traduit par une contractualisation commune (notamment dans la gestion des ressources humaines et financières). En revanche, le LCC est une Unité Propre de Recherche (UPR) : il s'agit d'une entité qui est entièrement sous la tutelle du CNRS.

A partir des annuaires des deux laboratoires, nous pouvons appréhender la composition de chaque laboratoire en fonction du grade de chaque chercheur. Nous constatons que l'IC2MP est davantage structuré par la présence de Professeurs (PR) et de Maîtres de Conférences (MCF), qui sont des personnels directement affiliés à l'Université, alors que l'effectif du LCC est davantage structuré par la présence de Directeurs de Recherche (DR), d'étudiants et de doctorants.

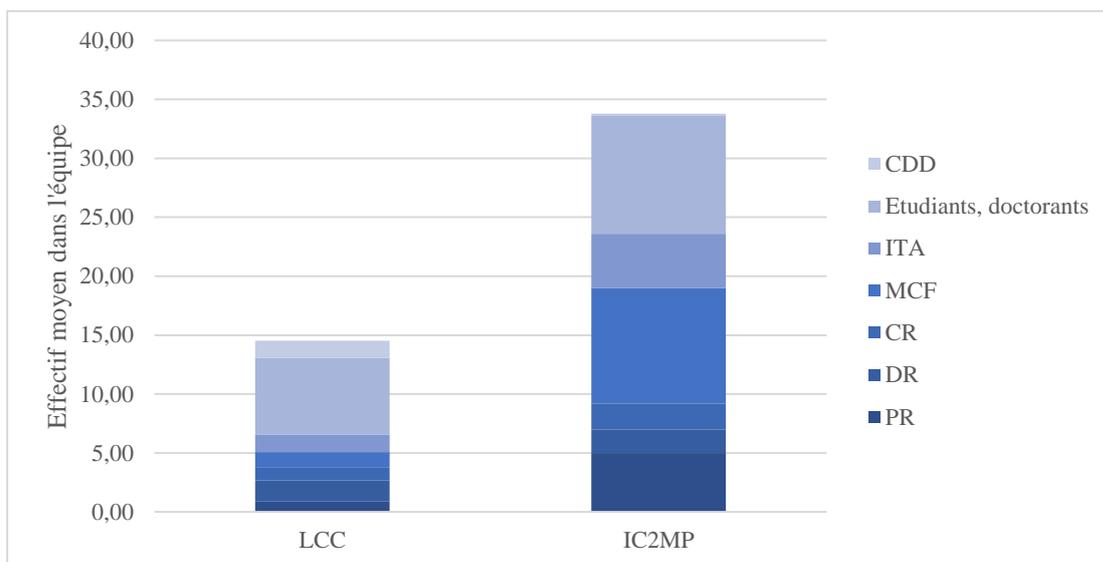
Figure 12 : Composition de l'effectif des deux laboratoires (Annuaire de Juin 2020)



iii) Structure interne et équipes de recherche

Enfin, la **structure interne des laboratoires est différente**. Le laboratoire IC2MP est composé de 5 équipes différentes, chacune composée de 33 personnels en moyenne (dont 19 chercheurs permanents). Le LCC est quant à lui structuré en 15 petites équipes, chacune composée en moyenne de 15 membres en moyenne (dont 5 chercheurs permanents en moyenne).

Figure 13 : Composition moyenne de chaque équipe de publication par laboratoire



3- Choix de l'échantillon

Nous avons sélectionné **les chercheurs** qui occupaient les positions de **Directeur de Recherche CNRS et Professeur** dans ces deux laboratoires au printemps 2018. Il s'agit de sélectionner des chercheurs qui sont jugés productifs scientifiquement du fait de leur progression de carrière et sont comparables en termes de **notoriété et de réputation professionnelle**. Ils disposent d'une **longévité de carrière suffisante** pour étudier leurs dynamiques de collaboration. En effet, en France en 2012, l'âge moyen de recrutement de maître de conférences en chimie était de 31 ans et celui de professeur en chimie était de 42 ans³⁷.

Tableau 14 : Description de l'échantillon

	Directeurs de recherche CNRS	Professeurs	Total
IC2MP – Poitiers	<i>Hommes: 6</i>	<i>Hommes: 26</i>	<i>Hommes: 32</i>
	<i>Femmes: 3</i>	<i>Femmes: 3</i>	<i>Femmes: 6</i>
	Total: 9	Total: 29	Total: 38
LCC - Toulouse	<i>Hommes: 14</i>	<i>Hommes: 12</i>	<i>Hommes: 26</i>
	<i>Femmes: 12</i>	<i>Femmes: 4</i>	<i>Femmes: 16</i>
	Total: 26	Total: 16	Total: 42
Total	<i>Hommes: 20</i>	<i>Hommes: 38</i>	<i>Hommes: 58</i>
	<i>Femmes: 15</i>	<i>Femmes: 7</i>	<i>Femmes: 22</i>
	Total: 35	Total: 45	Total: 80

³⁷ Selon le Ministère de l'Enseignement Supérieur et de la Recherche : https://cache.media.enseignementsup-recherche.gouv.fr/file/statistiques/05/3/orig2012_302053.pdf

Le tableau 14 présente la distribution de notre échantillon en 2018 en fonction du genre, du laboratoire de rattachement et du poste occupé. Nous avons sélectionné 38 chercheurs du laboratoire IC2MP de Poitiers et 42 du laboratoire LCC de Toulouse. Également, 45 chercheurs de l'échantillon occupaient le grade de Professeur en 2018 : 29 d'entre eux étaient affiliés à l'unité mixte de recherche IC2MP. Aussi, 35 des 80 chercheurs de l'échantillon occupaient le grade de Directeur de Recherche en 2018 : 26 d'entre eux étaient affiliés à l'unité propre de recherche LCC. Enfin, 72,5% des chercheurs de notre échantillon sont des hommes. Cela est cohérent avec la distribution nationale : en effet, 62% des chercheurs académiques français en chimie étaient masculins en 2018³⁸. Plus précisément, les hommes représentaient 56% des maîtres de conférences et 75% des professeurs : ceci illustre le concept de « plafond de verre » dans les carrières académiques (Sabatier, 2010).

B) Construction de la base de données

Le processus de construction de la base de données bibliométrique suit trois étapes : i) l'extraction des publications et la création des paires de co-authorship, ii) le codage d'informations supplémentaires, iii) le regroupement de chercheurs en cohortes.

1- Extraction des publications et la création des paires de co-authorship

A partir de la base de données Web of Science Core Collection (Thomson Reuters), nous avons extrait toutes les publications (articles, reviews, letters, notes and editorial materials³⁹) des 80 chercheurs étudiés. Cette base de données bibliométrique est une source reconnue dans les sciences sociales et fréquemment utilisée dans la collecte de données relatives aux co-publications. À la suite d'un nettoyage initial afin d'éviter les homonymes, nous avons retenu 9336 publications, dont 9233 co-publications. La plupart des variables relatives aux publications ont été gardées telles quelles le titre, l'année, la revue scientifique, la liste des co-auteurs et certaines ont fait l'objet d'un codage et de recherches plus approfondies.

³⁸ Selon le Ministère de l'Enseignement Supérieur et de la Recherche : <http://www.enseignementsup-recherche.gouv.fr/cid129560/fiches-demographiques-des-sections-de-sciences-2017.html>

³⁹ Les articles représentent 99,4% du nombre total des publications

Figure 15 : Réseau complet des publications réalisées par les 38 chercheurs qui étaient affiliés à l'IC2MP en 2018

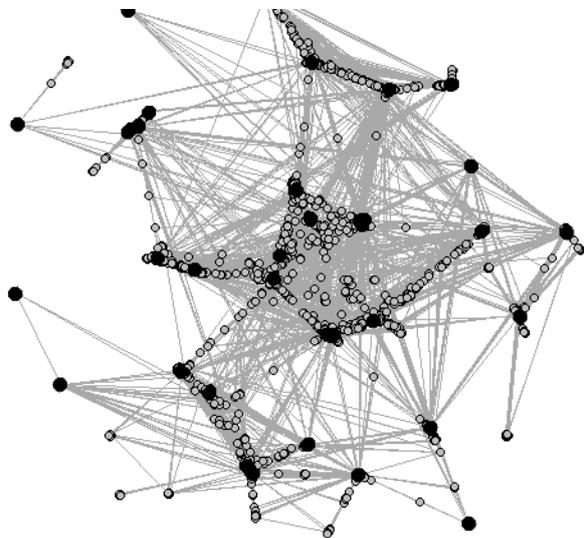
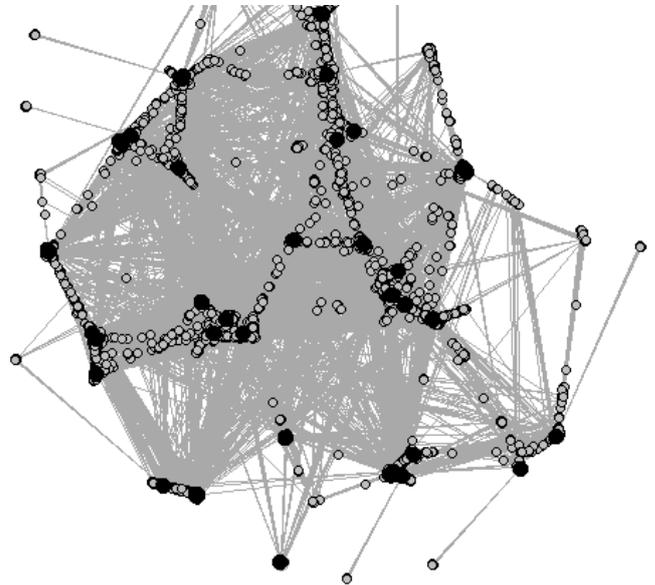


Figure 14 : Réseau complet des publications réalisées par les 42 chercheurs qui étaient affiliés au LCC en 2018



Les réseaux complets des chercheurs affiliés à Poitiers et Toulouse sont représentés par les figures 14 et 15. Ces graphiques nous permettent d'appréhender la structure des liens collaboratifs dans le cadre des deux laboratoires. Graphiquement, les nœuds noirs représentent chacun des 80 chercheurs de notre échantillon et les nœuds gris représentent chacun des co-auteurs avec qui ils ont co-publié. Nous notons la forte interconnexion au sein des réseaux, soulignant la dimension locale des collaborations entre les chercheurs étudiés. Néanmoins, cette approche des réseaux ne nous permet pas d'aborder le comportement collaboratif des chercheurs puisque nous sommes à une échelle structurelle. Ainsi, nous **choisissons d'opter une approche micro via la reconstitution des réseaux personnels des chercheurs**. Décrit dans le chapitre 1 (Section III), un réseau personnel de co-authorship représente l'ensemble des d'un chercheur principal (nommé *ego*) avec ses co-auteurs (nommés *alter*).

Dans le but de reconstituer les 80 réseaux personnels de nos chercheurs, nous avons créé des paires de co-authorship à partir de la liste des publications (Melin et Persson, 1996). Dans cette perspective de réseaux sociaux, nous ne nous intéressons qu'aux relations *ego-alter* et non aux relations *alter-alter*. Ainsi, pour toute publication dans laquelle *ego* a collaboré avec x co-auteurs, nous avons reconstitué x paires de co-authorship entre *ego* et chacun de ses *alter*. La reconstitution de ces paires repose sur deux échelles d'analyse : le co-auteur et la collaboration. Le co-auteur (*alter*) correspond à l'individu avec qui *ego* réalise de publications. La collaboration correspond au nombre de publications qu'*ego* réalise avec chacun de ses co-auteurs (*alter*) sur la période considérée. Chaque relation collaborative est évaluée via le fractionnement du nombre distinct d'*alter* qui contribuent dans chaque publication. Par exemple, une co-publication qui implique *ego* et trois de ses *alter* générera trois relations collaboratives *ego-alter*

qui pèsent chacune un tiers⁴⁰. Cette méthode réfère au « whole normalized counting » issu des travaux de Gauffriau *et al.* (2008). Au total, les 80 chercheurs ont collaboré 41 720 fois avec 14 603 co-auteurs différents.

2- Codage et informations supplémentaires

La seconde étape de la construction de notre base de données concerne la collecte de données supplémentaires pour mieux caractériser les paires de co-authorship. Notre objectif est de contextualiser chaque collaboration en mobilisant des informations relatives à la relation *ego-alter* ou des caractéristiques propres à *ego* et *alter*. Le tableau 15 synthétise l'ensemble des variables supplémentaires en fonction de la source d'informations mobilisée.

Tableau 15 : Données supplémentaires à l'échelle individuelle

	Données supplémentaires	Source d'information
Dimension géographique des relations ego-alter	Localisation géographique d' <i>ego</i> et <i>alter</i> à chaque collaboration	<u>Données Web of Science avant 2008</u> : Recherche web des notices des articles et codage manuel de chaque couple "auteur-adresse"
		<u>Données Web of Science après 2008</u> : Association automatique des couples "auteur-adresse"
Description d'<i>alter</i>	Genre	Recherche web des notices des articles
	Titulaire de doctorat, année du diplôme	Recherche des CV et au sein de moteurs de recherche (Linkedin, sudoc.abes.fr ...)
Description d'<i>ego</i>	Nom du (des) directeur(s) de thèse d' <i>ego</i>	Questionnaires Recherche des CV et au sein de moteurs de recherche (Linkedin, sudoc.abes.fr ...)
	Nom des doctorants d' <i>ego</i>	Questionnaires Recherche des CV et au sein de moteurs de recherche (Linkedin, sudoc.abes.fr ...)

Dimension géographique des relations ego-alter :

La plus grande opération de codage concerne la **localisation du chercheur et du co-auteur à chaque collaboration**. En effet, notre objectif est d'enrichir la compréhension des collaborations en intégrant de l'information géographique à l'échelle micro pour pouvoir comparer la localisation d'*ego* et *alter* au fil de la carrière. La figure 16 décrit ce processus de recherche et de codage d'information.

i) Concernant les publications réalisées avant 2008, les données géographiques issues des publications du Web of Sciences sont présentées sous forme de liste d'adresses sans aucun lien

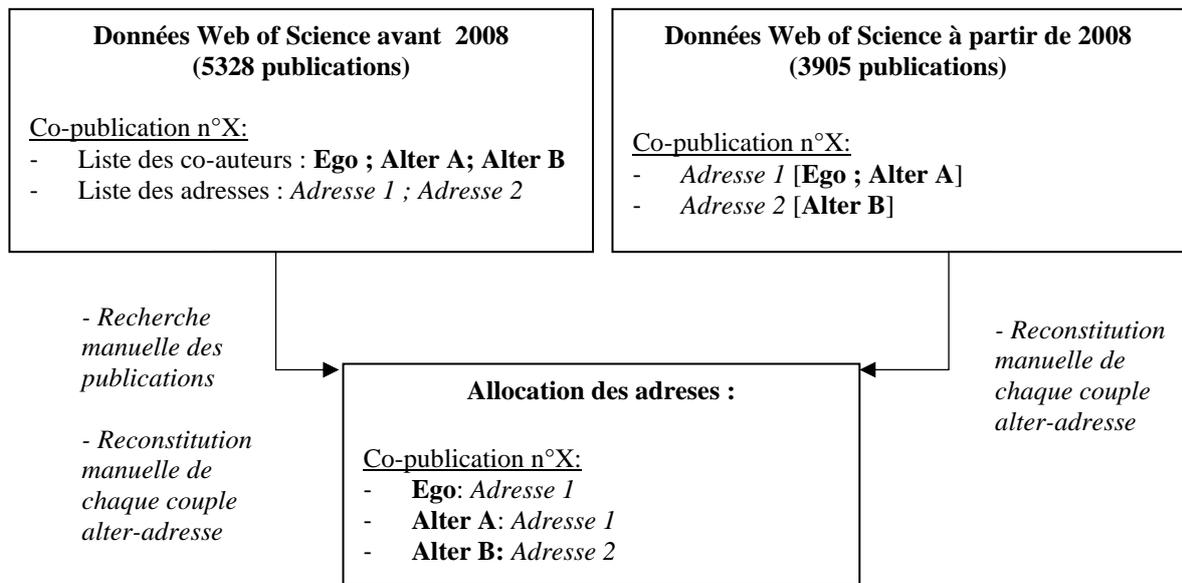
⁴⁰ Par construction, la somme de toutes ces relations fractionnées correspond au nombre total de publication de la base de données

d'affiliation avec la liste des co-auteurs. Par conséquent, le nombre d'adresses géographiques ne correspond pas toujours au nombre de co-auteurs (affiliation commune de plusieurs co-auteurs à une même adresse, multi-affiliation d'un co-auteur...). D'un point de vue méthodologique, l'absence de lien entre le co-auteur et son adresse explique pourquoi beaucoup d'études en sciences sociales se sont focalisées sur la dimension géographique de la publication – combien de villes ou de pays ont collaboré entre eux dans cette publication ? – pour contribuer à la compréhension de la géographie de la science (Grossetti *et al.*, 2014, Maisonobe, 2015). En revanche, cela apparaît comme un obstacle dans l'étude des dynamiques géographiques des trajectoires scientifiques individuelles. Pour résoudre ce problème, nous avons recherché les notices de ces articles pour collecter l'ensemble des localisations associées à chaque co-auteur (à l'échelle du laboratoire, de la ville et du pays). A notre connaissance, c'est la seule méthode qui permet d'obtenir une information géographique de qualité des collaborations à l'échelle micro pour les données avant 2008 : par ailleurs, elle n'a jamais été appliquée pour un si grand volume de publications (Bernela et Milard, 2016). Concrètement, la recherche manuelle de la notice de l'article a été nécessaire pour 5328 co-publications avant 2008. L'affiliation des données géographiques est automatique dans le Web of Science pour les publications réalisées à partir de 2008 (cela concerne 3905 co-publications).

ii) Une fois les données collectées, un long travail de nettoyage des homonymes, de standardisation des données et de codage de l'information a été nécessaire pour obtenir une information géographique claire et comparable. Pour chaque affiliation géographique, nous retenons trois échelles d'information géographique : le laboratoire, la ville et le pays. Le nettoyage a notamment été nécessaire pour les données relatives au laboratoire. Par exemple, dans le cas du laboratoire IC2MP de Poitiers, il a aussi été enregistré sous le nom « UMR 7285 » ou par les différents noms des anciens laboratoires avant la fusion de 2012.

iii) Enfin, nous avons choisi de conserver les multi-affiliation géographiques. Cela concerne les publications au sein desquelles *ego* et/ou *alter* est affilié à plusieurs laboratoires. Pour chaque couple de co-authorship présentant cette situation, nous avons donc choisis de le répéter pour chaque affiliation géographique (afin que la somme de ces couples multi-affiliés pondérés soit égale à 1).

Figure 16 : Allocation des localisations géographiques à l'échelle micro



Une fois que les **localisations géographiques** ont été associées à *ego* et *alter* pour chaque collaboration, nous avons utilisé ces données pour deux types de traitement. Le premier concerne la reconstruction de la **trajectoire géographique d'ego** à partir de ses affiliations. Le second traitement concerne la **distance géographique au sein de chaque couple**. Cette variable est décrite par quatre niveaux de distance : local (*ego* et *alter* partagent le même laboratoire), ville (*ego* et *alter* ne partagent pas le même laboratoire mais la même ville), pays (*ego* et *alter* ne partagent pas la même ville mais le même pays), international (*ego* et *alter* ne partagent pas le même pays). L'objectif est d'identifier dans quelle mesure *ego* et *alter* partagent le même environnement géographique. **Cette collecte et ce codage d'informations géographiques à l'échelle micro pour chaque couple de co-authorship fait de notre base de données un outil original dans l'étude de la géographie de la science.**

Description d'alter

Également, nous avons enrichi notre base de données à partir de **données descriptives d'alter**. Notre objectif est de comprendre les dynamiques collaboratives au fil de la carrière d'*ego* en qualifiant *alter*.

Dans l'optique où nous considérons chacun des 80 réseaux personnels indépendamment les uns des autres, notre base de données contient 14603 co-auteurs différents. Néanmoins, au vue des données géographiques et des graphiques de réseaux complets, nous constatons que ces réseaux partagent des *alter* communs. Ainsi, si nous soustrayons les *alter* répétés au sein des 80 réseaux, nous obtenons 9345 *alter* qui sont des individus différents. La recherche de données complémentaires servant à qualifier *alter* concerne donc ce nombre réduit d'individus : les informations seront donc répétées si besoin dans le cas où des *ego* partagent des collaborations avec les mêmes *alter*.

Tout d'abord, nous avons recherché le prénom de chaque co-auteur via les notices de chaque publication étudiée. A partir de ces données et de recherches complémentaires pour les prénoms mixtes, nous avons pu créer la **variable « genre » pour des alter**. Seuls 272 *alter* (1,86% des co-auteurs) n'ont pas pu être affilié à un genre, du fait de données manquantes. Également, nous avons recherché **la date d'obtention de doctorat** de chaque *alter*. L'objectif de cette seconde démarche est de déterminer les stades de carrière des *alter* (par exemple, un *alter* est jugé doctorant si la date de son doctorat est supérieure ou égale à celle de la publication). La date de doctorat a été recherchée via la recherche individuelle de chaque CV. Dans les cas où les CV n'ont pas été trouvés, nous avons recherché la date de soutenance de chaque co-auteur sur des sites professionnels (Linkedin, sudoc.abes.fr, theses.fr...). Seuls 1458 *alter* (9,98% des co-auteurs) n'ont pas pu être affiliés à une éventuelle date d'obtention de doctorat, du fait de données manquantes.

Description d'ego

Enfin, nous avons recherché des informations complémentaires relatives à *ego*. Ces données ont pour but de **contextualiser chaque relation collaboration ego-alter en fonction du rôle joué par alter dans la carrière d'ego**. Plus précisément, nous avons envoyé un questionnaire auprès de chacun des 80 *ego* afin de prendre connaissance du **nom du/des directeur(s) de thèse d'ego⁴¹ et de ses doctorants**. L'objectif est de tester l'impact du lien d'encadrement sur les dynamiques collaboratives durant le doctorat d'*ego/alter* ou durant les années suivantes. Parmi les 80 questionnaires envoyés, le taux de réponse a été de 76,25%. Dans les cas où *ego* n'a pas répondu au questionnaire, des recherches complémentaires ont été faites sur les réseaux sociaux professionnels afin de déterminer le directeur de thèse d'*ego* et les soutenances d'*alter* pour lesquelles *ego* est directeur de thèse. Finalement, chaque *ego* est associé à un/deux directeur(s) de thèse et à une liste de doctorants.

3- Regroupement en cohortes

Puisque nous souhaitons appréhender les dynamiques de publication et de collaboration au fil de la carrière, il convient de définir le critère à partir duquel la carrière commence d'un point de vue bibliométrique. Nous choisissons **l'année de la première publication** comme repère de début de carrière. Bien que les 80 chercheurs occupent tous la même position au printemps 2018, ils ont des durées de carrière différentes puisqu'ils ont commencé à publier durant des décennies différentes. En effet, la plus longue carrière de l'échantillon dure 51 ans (première publication en 1967) et la plus courte

⁴¹ Concernant les chercheurs ayant commencé leur thèse avant 1984 (loi Savary), nous retiendrons comme référence le directeur de thèse du doctorat de troisième cycle plutôt que celui du doctorat d'Etat. En effet, la thèse de troisième cycle est un diplôme relatif aux pratiques scientifiques alors que la thèse d'Etat évalue la capacité du chercheur à mettre en place une recherche scientifique originale de haut niveau. La thèse d'Etat est davantage aligné sur le diplôme d'Habilitation à Diriger des Recherches (HDR)

de 12 ans (première publication en 2006). Ainsi, il nous semble important de contextualiser chaque collaboration afin de mieux comprendre le comportement collaboratif du chercheur. **Nous constituons des cohortes en regroupant les chercheurs sur des périodes de 10 ans en fonction de l'année de leur première publication.** L'intérêt des cohortes est de prendre en compte l'évolution historique de la pratique de la science et de l'activité de publication depuis les années 1970. Également, les cohortes nous permettent d'identifier un éventuel impact de la période de début de carrière du chercheur sur le comportement collaboratif qu'il aura ensuite. La figure 17 illustre le procédé de regroupement en cohortes. En effet, chaque barre représente un chercheur étudié : il comporte donc 80 barres. Chaque carrière individuelle est représentée à partir de la première année de publication jusqu'en 2018. Nous confirmons que les carrières n'ont pas la même durée puisqu'elles n'ont pas commencé aux mêmes périodes.

Figure 17 : Regroupement en cohortes

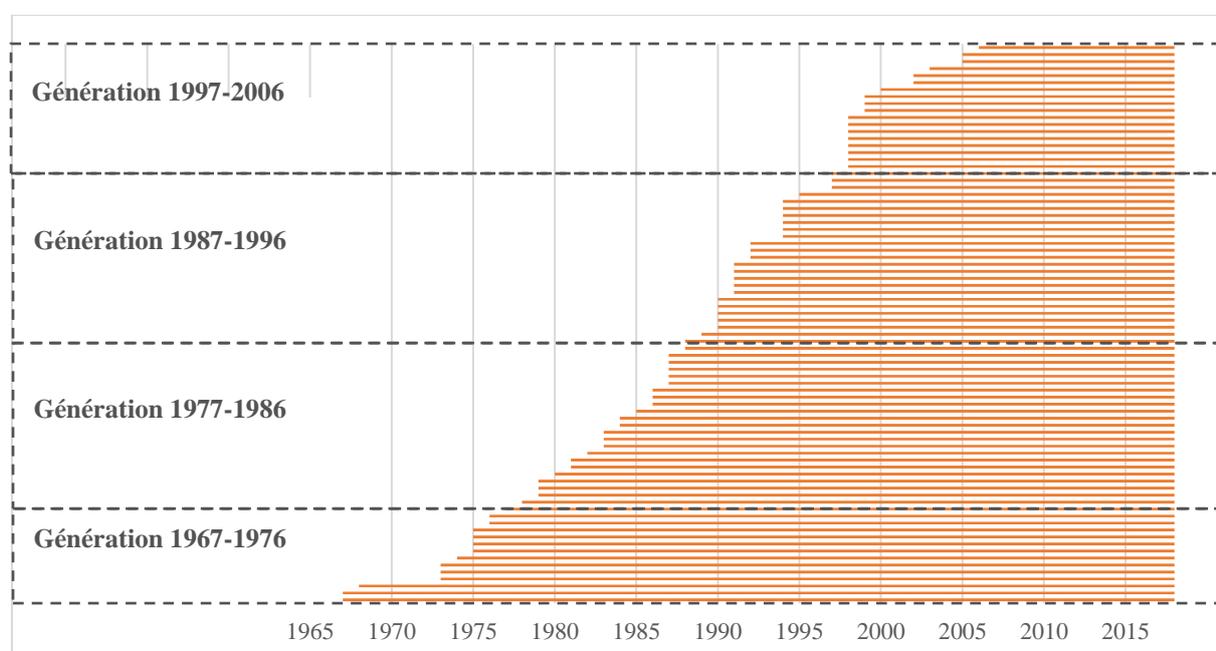


Tableau 16 : Composition des cohortes de génération des chercheurs

Année de la première publication		1967-1976	1977-1986	1987-1996	1997-2006	Total
Durée de carrière		42-51 ans	32-41 ans	22-31 ans	12-21 ans	
Laboratoires	Poitiers	6	7	13	12	38
	Toulouse	7	11	15	9	42
Genre	Hommes	13	13	14	18	58
	Femmes	0	5	14	3	22
Position en 2018	Professeurs	8	9	15	13	45
	Directeurs de recherche CNRS	5	9	13	8	35
Total		13	18	28	21	80

Enfin, le tableau 16 décompose chaque cohorte générationnelle en fonction de la provenance, du genre et du type de carrière des chercheurs. On constate que seulement 13 chercheurs ont commencé leur carrière entre 1967-1976 : 49 chercheurs ont commencé leur carrière après 1987 et ont moins de 22 années d'exercice de publication. Également, nous constatons que les femmes sont sous-représentées dans les générations d'avant 1987.

II- Statistiques descriptives de la base de données

Après avoir détaillé la méthodologie suivie dans la création de la base de données, nous présenterons plusieurs statistiques descriptives. Nous commencerons par décrire l'échantillon étudié en termes d'activité de publication et de tendances collaboratives en fonction du genre, du type de carrière et de la génération du chercheur. Ensuite, nous décrirons les tendances collaboratives en qualifiant les co-auteurs. Enfin, nous présenterons la géographie des réseaux.

A- Description de la population étudiée

Le tableau 17 détaille les statistiques descriptives relatives aux activités de publication et de co-authorship : les 80 chercheurs ont co-publié 9233 articles sur l'ensemble de leur carrière dans lesquels ils ont co-signés 41270 fois avec 14603 co-auteurs différents. A l'échelle individuelle, chaque chercheur a publié en moyenne 115,4 co-publications durant sa carrière, soit 3.86 publications par an. Au sein de chaque co-publication, un chercheur collabore en moyenne avec 4.52 co-auteurs, dont 1,58 nouveaux co-auteurs. Ces statistiques descriptives soulignent une activité de publication soutenue et sa nature fortement collaborative. En termes de volume, la majorité des co-auteurs (58%) sont des « one shot » : il s'agit des *alter* avec qui *ego* n'a collaboré qu'une fois. Bernela et Milard (2016) ont montré que les co-auteurs « one shot » correspondent souvent aux co-auteurs de second rang qui entretiennent très peu de liens avec *ego*. Du fait de la structuration de la recherche en équipe, il arrive que certains co-auteurs contribuent à la publication sans avoir d'interaction avec *ego* (les doctorants d'une autre équipe par exemple). Enfin, nous remarquons que les femmes ont un taux d'homophilie de genre plus faible que la moyenne : en moyenne, 66,50% des co-auteurs des femmes sont des hommes.

Tableau 17 : Activité de publication et tendances collaboratives selon le genre et la position

	Tous les chercheurs	Genre		Position		
		Hommes	Femmes	Professeurs	Directeurs de recherche	
	Nombre de chercheurs	80	58	22	45	35
	Durée Moyenne de carrière (années)	30	30,50	28,45	29,18	30,91
Activité de publication	Nombre de publications	9336	7316	2020	4208	5128
	Nombre de co-publications	9233	7219	2014	4137	5096
	Nombre moyen de co-publication par chercheur	115,41	124,47	91,55	91,93	145,60
	Nombre moyen de co-publications annuelles par chercheur	3,86	4,08	3,22	3,15	4,71
Tendances collaboratives	Nombre de collaborations	41720	31868	9852	18180	23540
	Nombre moyen de collaborations par chercheur	521,50	549,45	447,82	404,00	672,57
	Nombre moyen de collaborations par article	4,52	4,41	4,89	4,39	4,62
	Homophilie de genre (% dans les collaborations)	59,55%	68,95%	29,15%	66,47%	54,21%
	Nombre de co-auteurs	14603	10964	3639	6861	7782
	Nombre moyen de co-auteurs par chercheurs	182,54	189,03	165,41	152,47	222,34
	Nombre moyen de nouveaux co-auteurs par publications	1,58	1,52	1,81	1,66	1,53
	Homophilie de genre (% des co-auteurs)	58,75%	67,13%	33,50%	63,86%	54,27%
	Co-auteurs "one-shot"	57,30%	56,38%	60,07%	57,38%	56,93%

Le tableau 18 détaille les activités de publication et de co-authorship en fonction des générations. En dépit des volumes différents de co-publications (dû à l'hétérogénéité des durées de carrière), les **générations partagent le même rythme annuel moyen de publication** : la quantité annuelle de publications réalisées ne dépend pas de la génération du chercheur. En revanche, **la propension à collaborer augmente d'une génération à l'autre** : le nombre moyen de co-auteurs par co-publication est de 3,94 pour la première cohorte et de 5,30 pour la quatrième cohorte. Ainsi, la dimension collaborative des articles s'accroît au fil des générations. Enfin, nous constatons que le poids des one-shot est relativement stable au fil des générations.

Tableau 18 : Activité de publication et tendances collaboratives selon la génération

	Tous les chercheurs	Génération				
		1967-1976	1977-1986	1987-1996	1997-2006	
Nombre de chercheurs	80	13	18	28	21	
Durée Moyenne de carrière (années)	30	45	36	28	18	
Activité de publication	Nombre de publications	9336	2867	2432	2634	1406
	Nombre de co-publications	9233	2832	2369	2634	1398
	Nombre moyen de co-publication par chercheur	115,41	217,85	131,61	94,07	66,57
	Nombre moyen de co-publications annuelles par chercheur	3,86	4,84	3,64	3,40	3,63
Tendances collaboratives	Nombre de collaborations	41720	11157	9890	13263	7410
	Nombre moyen de collaborations par chercheur	521,50	858,23	549,44	473,68	352,86
	Nombre moyen de collaborations par co-publication	4,52	3,94	4,17	5,04	5,30
	Homophilie de genre (% dans les collaborations)	59,55%	64,41%	53,14%	56,43%	66,40%
	Nombre de co-auteurs	14603	3660	3314	4805	2824
	Nombre moyen de co-auteurs par chercheurs	182,54	281,54	184,11	171,61	134,48
	Nombre moyen de nouveaux co-auteurs par co-publication	1,58	1,29	1,40	1,82	2,02
	Homophilie de genre (% des co-auteurs)	58,75%	65,16%	55,43%	53,63%	63,03%
	Co-auteurs "one-shot"	57,30%	56,78%	56,70%	57,02%	59,14%

B) Description d'alter et tendances collaboratives

Le tableau 19 détaille les tendances collaboratives des chercheurs en fonction des données sociales complémentaires récoltées pour décrire les *alter*. Nous constatons que les collaborations avec des docteurs pèse davantage qu'avec des doctorants.

Nous notons des différences en fonction du genre d'*ego*. Notamment, nous remarquons que les *ego* hommes collaborent en moyenne avec 0,44 doctorants qui sont/étaient sous leur supervision, contre 0,19 pour les *ego* femmes. Ainsi, **les équipes de collaboration des hommes sont plus structurées par la relation de mentorship sur les co-auteurs que celles des femmes**. En parallèle, on constate que les publications des femmes sont beaucoup plus structurées par la présence du directeur de thèse que les hommes. En effet, en moyenne, le directeur de thèse collabore dans 28.15% des co-publications des femmes et seulement dans 14.38% des co-publications des hommes.

Également, nous rappelons que la **dimension collaborative des articles évolue au fil des générations**. Cette croissance se traduit par la hausse du poids des docteurs dans les co-auteurs. En moyenne, les chercheurs de la première génération collaborent avec 2,87 docteurs par article et ceux de la quatrième

génération collabore avec 4,08 chercheurs docteurs. Ainsi, **l'extension de l'équipe de publication se traduit par l'enrichissement de co-auteurs expérimentés**. Aussi, le poids du directeur de thèse dans les publications s'accroît au fil des générations. Dans la première génération, *ego* collabore en moyenne avec son directeur de thèse dans 4,41% des co-publications, contre 18,67% dans la quatrième génération.

C) Description de la géographie des réseaux

A partir des données géographiques collectées, nous pouvons appréhender finement la géographie des réseaux de co-authorship des chercheurs.

A partir des affiliations géographiques des publications, nous pouvons identifier les principales sources géographiques du co-authorship des chercheurs. Nous choisissons de raisonner en cosignatures fractionnées⁴² afin de mieux capter les contributions des différentes villes dans le travail de publication d'*ego*. Nous constatons que les chercheurs ont mobilisé la majorité de leur co-authorship en France. En effet, environ 7001 des cosignatures fractionnées sont issues de villes françaises : en somme, les relations avec les chercheurs français ont contribué à la réalisation de 75,82% des 9233 co-publications. Plusieurs pôles territoriaux ont été des sources de co-auteurs entre 1967 et 2018 pour les 80 chimistes. Sans grande surprise, les principales villes de co-authorship sont Poitiers et Toulouse. En effet, 1810 cosignatures fractionnées (19,59% des relations pondérées) ont été réalisées avec des *alter* localisés à Poitiers et 3511 (38,02% des relations pondérées) ont été réalisées avec des alter de Toulouse. Ces données illustrent le poids du **localisme des collaborations**. Néanmoins, il nous semble intéressant d'identifier les autres villes françaises avec qui les chercheurs ont collaboré durant ces 51 années de publication. A partir de notre base de données, nous pouvons contextualiser les collaborations avec ces **pôles territoriaux français en** fonction des thématiques de recherche des laboratoires et avec l'année de la première collaboration avec la ville. A partir de la carte de la figure 18, nous détaillons les 10 principaux pôles collaboratifs français sources de co-authorship :

i) Dans le cadre de notre étude bibliométrique, la première ville de co-authorship française (en dehors de Poitiers et Toulouse) est Paris. En effet, les chercheurs entretiennent une longue relation collaborative avec cette ville (depuis 1967) au travers de multiples thématiques de recherche. En somme, les relations de co-authorship entre *ego* et un co-auteur affilié à Paris ont contribué à 3,4% des co-publications faites entre 1967 et 2018.

⁴² Nous mesurons la cosignature fractionnée d'une collaboration *ego-alter* dans une co-publication en divisant cette contribution par le nombre d'*alter* inclus dans la co-publication. Par exemple, dans une co-publication où *ego* collabore avec 4 *alter* différents, chacun de ces co-auteurs entretient une cosignature fractionnée d'un quart avec *ego*. En somme, plus l'équipe de collaboration est riche, plus la contribution individuelle des alter est faible. Par définition, la somme des cosignatures fractionnées vaut 9233 (le nombre total de co-publications étudiées).

Tableau 19 : Tendances collaboratives en fonction du grade et du rôle d’alter dans la carrière d’ego

		Tous les chercheurs	Genre		Position		Génération			
			Femmes	Hommes	Directeurs de recherche	Professeurs	1967-1976	1977-1986	1987-1996	1997-2006
	Nombre de co-publications	9233	2014	7219	5096	4137	2832	2369	2634	1398
Grade d’alter	Nombre moyen de doctorants en cours de formation par co-publication	0,60	0,59	0,61	0,62	0,59	0,52	0,57	0,65	0,72
	Nombre moyen de docteurs par co-publication	3,36	3,68	3,27	3,42	3,28	2,87	3,03	3,80	4,08
	Nombre moyen de co-auteurs non-doctorants et non docteurs par co-publication	0,24	0,29	0,23	0,27	0,21	0,20	0,26	0,28	0,23
	Nombre moyen de co-auteurs non trouvés par co-publication	0,31	0,34	0,31	0,31	0,31	0,34	0,31	0,30	0,27
Directeur de thèse	Poids des co-publications dans lesquelles <i>ego</i> a collaboré avec son directeur de thèse ⁴³	17,38%	28,15%	14,38%	18,88%	15,54%	4,41%	29,84%	19,44%	18,67%
Doctorant d’ego	Nombre moyen de doctorants d’ <i>ego</i> par co-publication ⁴⁴	0,39	0,19	0,44	0,31	0,48	0,51	0,32	0,35	0,34

⁴³ Directeur de thèse actuel ou ancien directeur de thèse

⁴⁴ Doctorant actuel sous la supervision d’ego ou ancien doctorant sous la supervision d’ego

ii) Ensuite, le département de l'Essonne contribue à 1,60% des co-publications réalisées. Certaines villes ont contribué aux travaux dès les années 1970 via des collaborations avec des spécialistes (Gif-sur-Yvette : collaborations avec l'Institut de la chimie des substances naturelles) ou plutôt multithématiques (spectrochimie, chimie inorganique, chimie théorique, physique des solides...). A partir des années 1990, certains chercheurs ont également noué des relations collaboratives avec des industries, telle que l'entreprise Renault (Laudry) ou des établissements d'ingénieurs telle que l'Ecole Polytechnique (Palaiseau).

iii) Le département du Rhône a contribué à 1,48% des travaux réalisés. Certains des chercheurs ont commencé à collaborer avec la ville de Villeurbanne à partir de 1978, principalement sur des thématiques de catalyse. Ensuite, à partir des années 1980, ils ont collaboré avec des groupes chimiques de St-Fons (Solvay, Rhodia, Rhône-Poulenc) et de Solaize (Centre de recherche des pétroles Elf Aquitaine, IFP Énergies nouvelles). Enfin, ils ont noué des relations collaboratives avec des co-auteurs issus de Lyon (notamment avec le laboratoire de chimie de l'Ecole Normale Supérieure).

iv) Le quatrième territoire français source de co-authorship est la Gironde (contribution à 1,24% des co-publications). Certains chercheurs entretiennent une forte relation collaborative avec la ville de Pessac depuis 1981 dans les sciences moléculaires et avec l'Institut Chimique de la Matière condensée. Les collaborations avec Bordeaux ont commencé en 1992 dans diverses thématiques (pharmacologie, cristallographie, chimie des sciences des polymères...).

v) Les chercheurs spécialisés dans la chimie physique ont commencé à collaborer avec la ville de Rennes en Ille-et-Vilaine en 1984. Dans les années 1980 et 1990, les collaborations avec cette ville portaient majoritairement sur la physique cristalline, la chimie du solide et de l'inorganique. A compter des années 2000, les chercheurs collaborent également avec l'Institut de physique de la ville.

vi) Les collaborations avec l'Hérault ont contribué à 0,98% des publications. Montpellier est une ancienne source de co-authorship pour les chercheurs puisqu'ils ont commencé à collaborer avec elle en 1973. Jusqu'aux années 1980, les chercheurs collaboraient surtout avec les co-auteurs de cette ville sur les thématiques de chimie organométallique et de chimie bio-organique. A compter des années 1980, cette ville s'est développée en véritable pôle multithématique (hydrogéologie, géologie structurale, chimie moléculaire, chimie bio-organique...).

vii) A partir de 1985, certains chercheurs ont noué des relations collaboratives avec les chercheurs de Dijon dans le département de Côte d'Or. Ces collaborations portent massivement sur les synthèses et l'électrosynthèse organométallique et avec l'Institut de chimie moléculaire (depuis les années 2000).

viii) Les chercheurs entretiennent également une longue relation collaborative avec la ville de Strasbourg dans le Bas-Rhin depuis 1974. Jusque dans les années 1980, les thématiques de recherche

avec cette ville concernent la géologie, la spectroscopie moléculaire et la cristallographie. Durant les années 1980 et 1990, les chercheurs collaborent toujours avec cette ville en géologie et élargissent leurs thématiques collaboratives avec la cristallographie, les synthèses métallo-induites, la chimie de coordination, la catalyse de chimie des surfaces. La ville de Strasbourg se développe comme une source nationale multithématique et structurante dans les collaborations des chercheurs.

ix) Dans les années 1980, les chercheurs ont noué des relations avec certaines villes des Hauts-de-Seine (contribution à 0,73% des co-publications). Les collaborations sont spécifiques en fonction des villes en question. Par exemple, les collaborations avec Meudon concernent le thème des terres rares, les collaborations avec Rueil-Malmaison se font avec l'Institut français du Pétrole et les collaborations avec Fontenay-aux-Roses se font avec le Commissariat à l'Energie Atomique et l'Institut de radioprotection et de sûreté nucléaire.

x) Enfin, les chercheurs collaborent avec des co-auteurs de l'Isère. Depuis 1975, Grenoble structure les liens collaboratifs dans différents domaines. Jusqu'aux années 1980, les collaborations concernent surtout la chimie physique (Centre d'Etudes Nucléaires et Institut Laue-Langevin). A partir des années 1980, les thématiques des collaborations se sont élargies à la catalyse (Laboratoire de Physiologie Cellulaire Végétale, European Molecular Biology Laboratory), à la science des matériaux (Département de Recherche Fondamentale sur la Matière Condensée du CEA) et à la santé (Laboratoire Louis Néel). Aussi, depuis 1981, les collaborations avec la ville de St Martin d'Hères sont spécialisées dans l'électrochimie, le génie physique et la science des matériaux.

Les cosignatures avec des co-auteurs hors France ont contribué à 24,18% des 9233 co-publications. La distribution internationale des cosignatures fractionnées est illustrée dans la carte de la figure 19. Ne disposant pas d'informations suffisamment explicites sur la thématique de recherche des laboratoires internationaux, nous ne pouvons que contextualiser les relations collaboratives avec les villes en fonction de la thématique des laboratoires. Néanmoins, nous constatons que la majorité des relations collaboratives internationales ont été tissées à compter des années 1980. En dehors de la France, **très peu de pays ont une contribution structurante dans les travaux : beaucoup ont une contribution de co-authorship très ponctuelle.** Les deux plus grosses sources de co-authorship sont l'Espagne et les Etats-Unis : ces deux pays contribuent chacun dans 2,63% des co-publications. Dans ces deux pays, la plupart des relations ont été tissées dans les années 1980 (Barcelone, Zaragone, College Station, College Park, Newark...), durant les années 1990 (Grenade, Valence) et les années 2000 (Belaterra, Pittsburg, Riverside). L'Allemagne (contribution à la hauteur de 1,59% des co-publications) est la troisième source de co-authorship (Berlin, Göttingen, Mainz, Würzburg) suivie par la Russie (Moscou, Chernogolovka) avec 1,32% des co-publications. Dans les années 1990, les chercheurs nouent de nouvelles relations avec des pays en développement hors Europe tels que le Maroc (Casablanca, Fes, Marrakech, Rabat), la Chine (Whuan, Shangai, Guangzhou), le Brésil (Sao Carlos, Belo Horizonte, Ribeirao Preto), le

Mexique (Mexico, Leon, Coyoacan), l'Algérie (Alger, Tilimsen, Blida, Bejaia). Durant cette décennie, les relations européennes se sont également développées, notamment avec la Pologne (Cracovie, Sienkiewicza), l'Italie (Bologne, Milan, Monserranto, Sesto Fiorento) et l'Angleterre (York, Londres, Bristol).

Figure 18 : Distribution des collaborations fractionnées parmi les villes de la France métropolitaine (Hormis Poitiers et Toulouse)

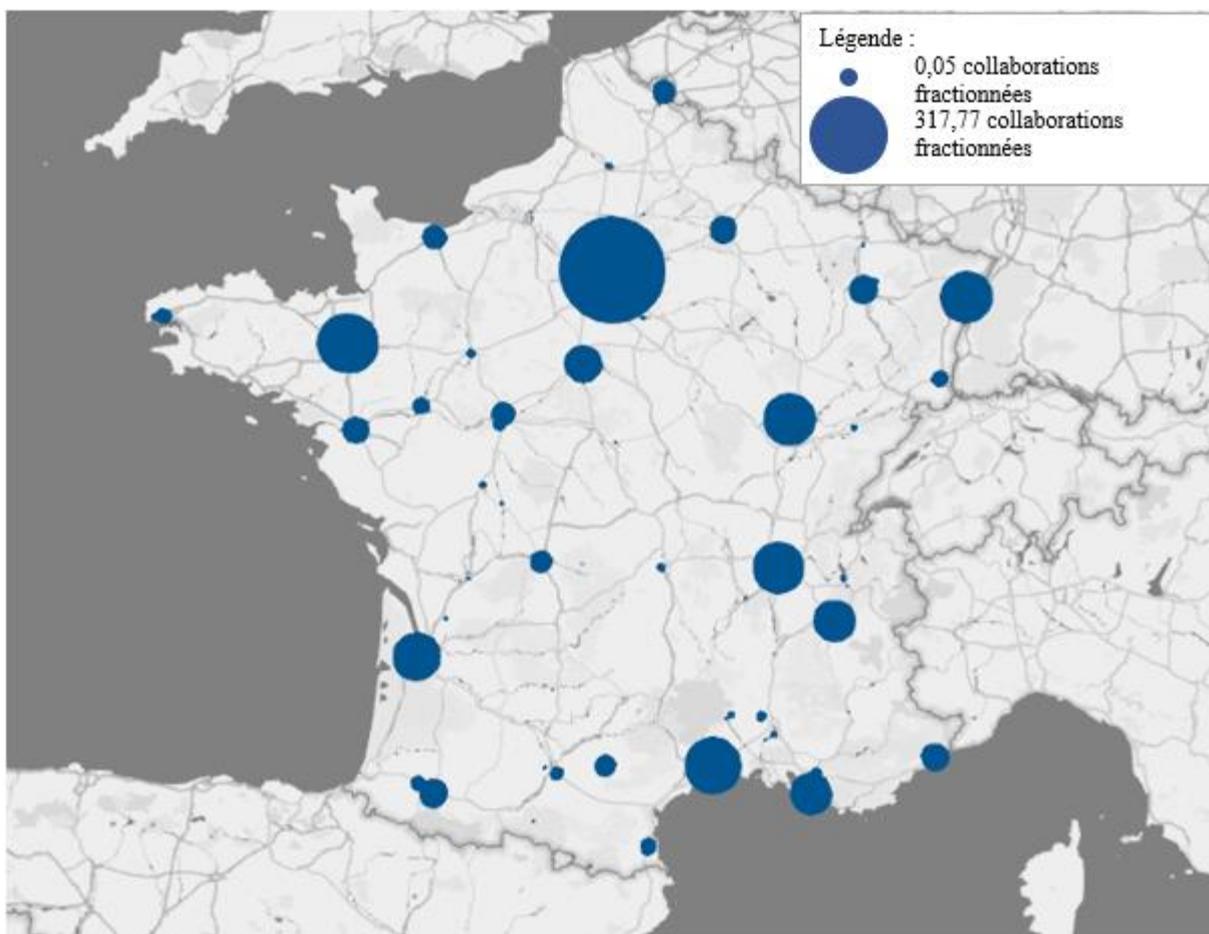
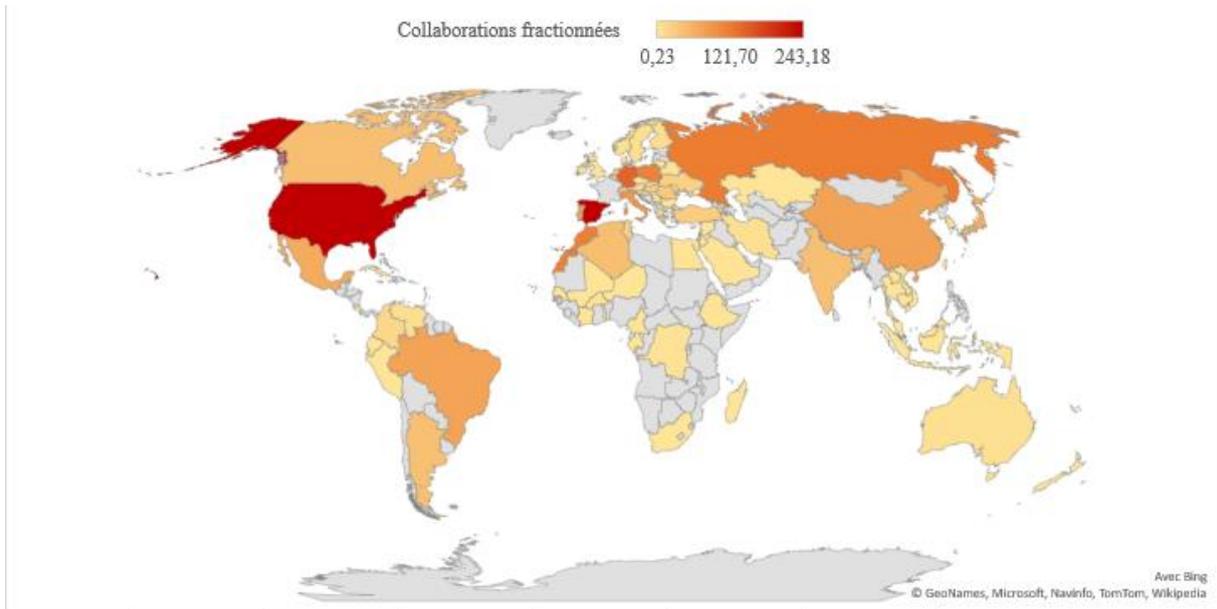


Figure 19 : Distribution des collaborations internationales (en compte fractionné)



III- Réseaux personnels et temporalité de carrière

Enfin, à partir de notre base de données bibliométrique, nous choisissons de reconstituer **le réseau de co-authorship des 80 chercheurs** sous deux dimensions temporelles : par année d'activité ou par période de carrière (segmentation des carrières en quatre périodes). L'objectif de la thèse étant de contribuer à la **compréhension du comportement collaboratif des chercheurs**, nous voudrions appréhender la capacité du chercheur à mobiliser son réseau de co-authorship sous deux temporalités différentes. Dans un premier temps, nous présenterons la démarche de création des réseaux par année d'activité dans une perspective de panel. Nous présenterons notamment les implications de cette temporalité en termes de rythme annuel de publication et la dimension collaborative des articles au fil de la carrière. Nous présenterons aussi un premier résultat concernant la composition (sociale et géographique) de l'équipe de recherche au fil de la carrière. Dans un second temps, nous présenterons la démarche de création des réseaux par période d'activité dans une perspective d'Analyse de Réseaux Sociaux. Plus précisément, nous détaillerons la méthodologie de segmentation des carrières et de création de réseaux *ego*-centrés. Nous finirons par présenter des statistiques descriptives relatives à la taille du réseau (habitudes collaboratives, hétérogénéité des tailles).

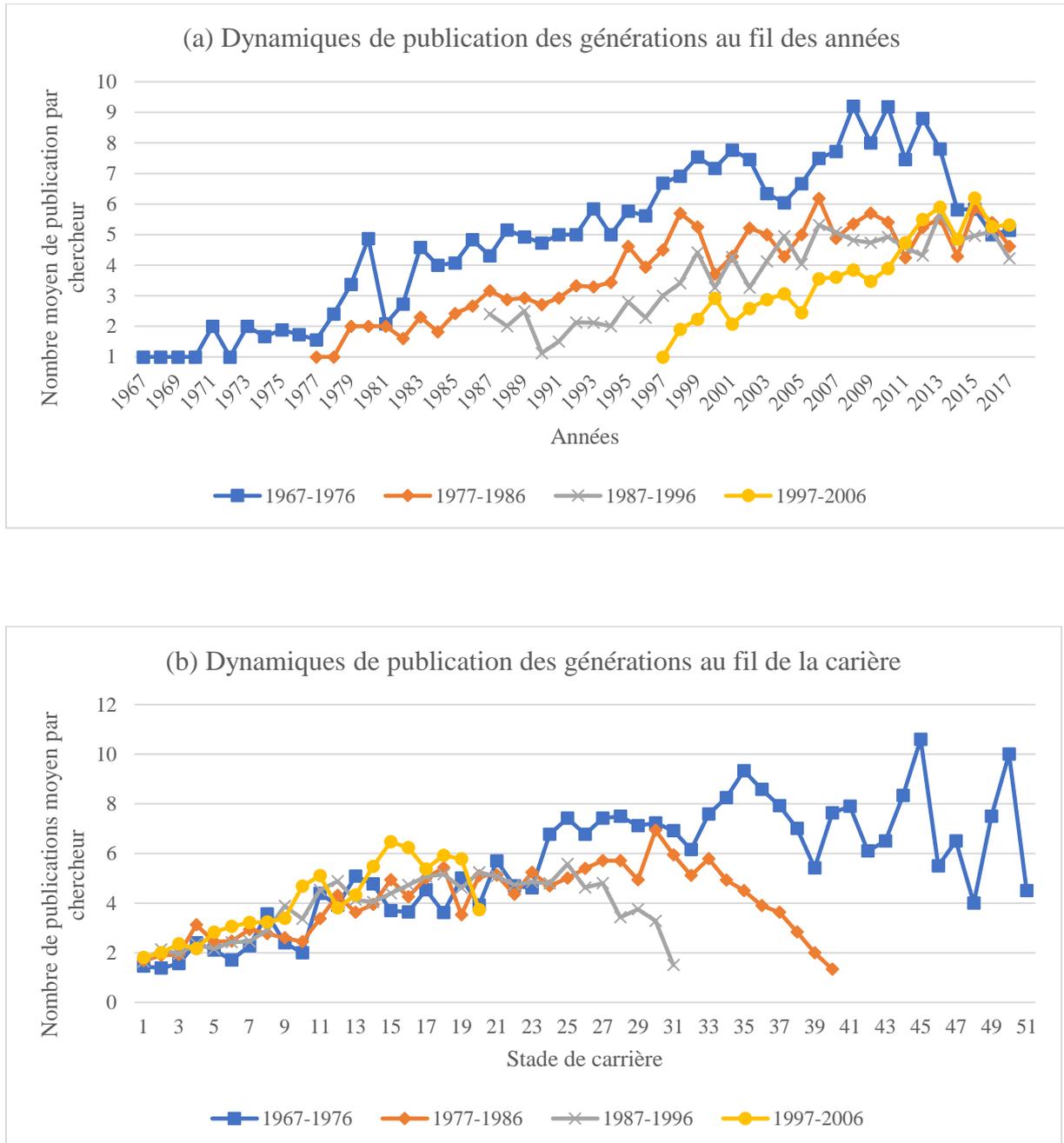
A) Réseaux de co-authorship par année d'activité : Approche de panel

Dans cette partie, nous intégrons **la dimension temporelle du stade de carrière en séquençant l'activité de publication en termes d'années**. Ce séquençage se fait à compter de l'année de la première publication. Les carrières étant de durées différentes, la variable de stade de carrière est comprise entre 12 (plus courte carrière) et 51 (plus longue carrière) : nous obtenons donc $t \in \mathbb{N}$ tel que [12 ;51].

Les graphiques des figures 20 et 21 montrent les dynamiques de publication et de collaboration sous deux échelles temporelles : par année et par stade de carrière. L'intérêt de ces deux approches graphiques est de montrer un même phénomène sous deux perspectives différentes : le contexte de recherche (au fil des années) et le contexte de carrière (au fil des stades de carrière).

Les graphiques de la figure 20 montrent comment le nombre de publications annuelles évolue par génération au fil des années et des stades de carrière. Bien que les chercheurs n'aient pas commencé leur carrière au même contexte historique, **toutes les générations suivent le même rythme croissant de publication annuel au fil des années et de la carrière**. La fréquence annuelle de publication varie en fonction du stade de carrière et non en fonction de la génération. Cependant, dans le graphique *_(a)*, nous observons dans les dernières années, la génération 1997-2006 rattrape le rythme moyen des autres générations plus anciennes : ce fait conjoncturel est dû aux exigences croissantes quant à l'activité de publication dans le système scientifique actuel.

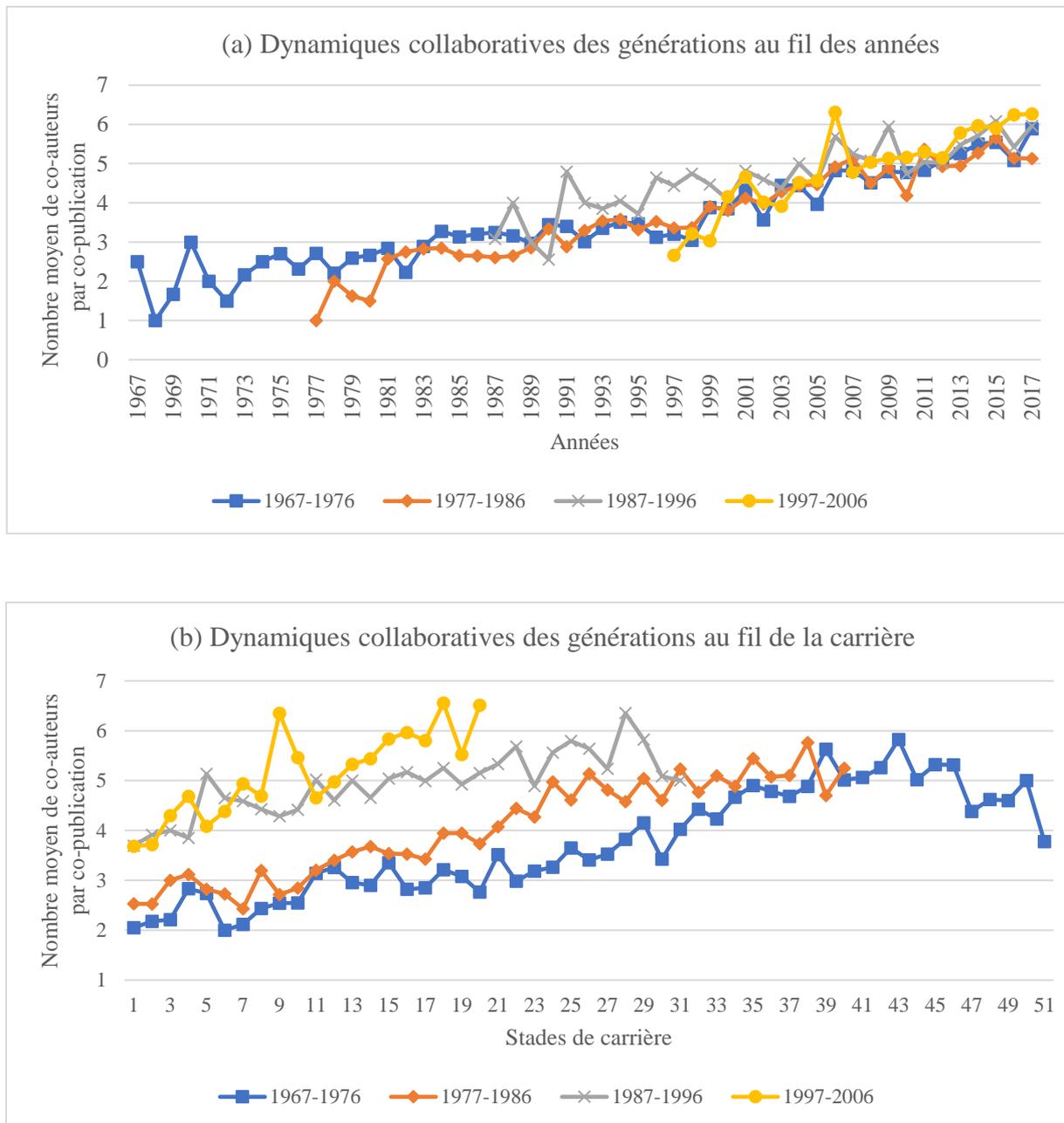
Figure 20 : Nombre moyen de publications annuelles par génération au fil des années (a) et des stades de carrière (b)



Les graphiques de la figure 21 présentent le nombre moyen de co-auteurs par publication par génération au fil des années et des stades de carrière. **La nature collaborative des publications dépend à la fois du stade de carrière et de la génération.** En effet, le nombre moyen de co-auteurs par publication s'accroît au fil de la carrière. En même temps, le contexte scientifique encourage les chercheurs à publier avec plus de co-auteurs. Par exemple, en 2000, bien que la quatrième cohorte de génération commence

juste à publier alors que la première cohorte a commencé sa carrière trente ans auparavant, les publications ont la même dimension collaborative.

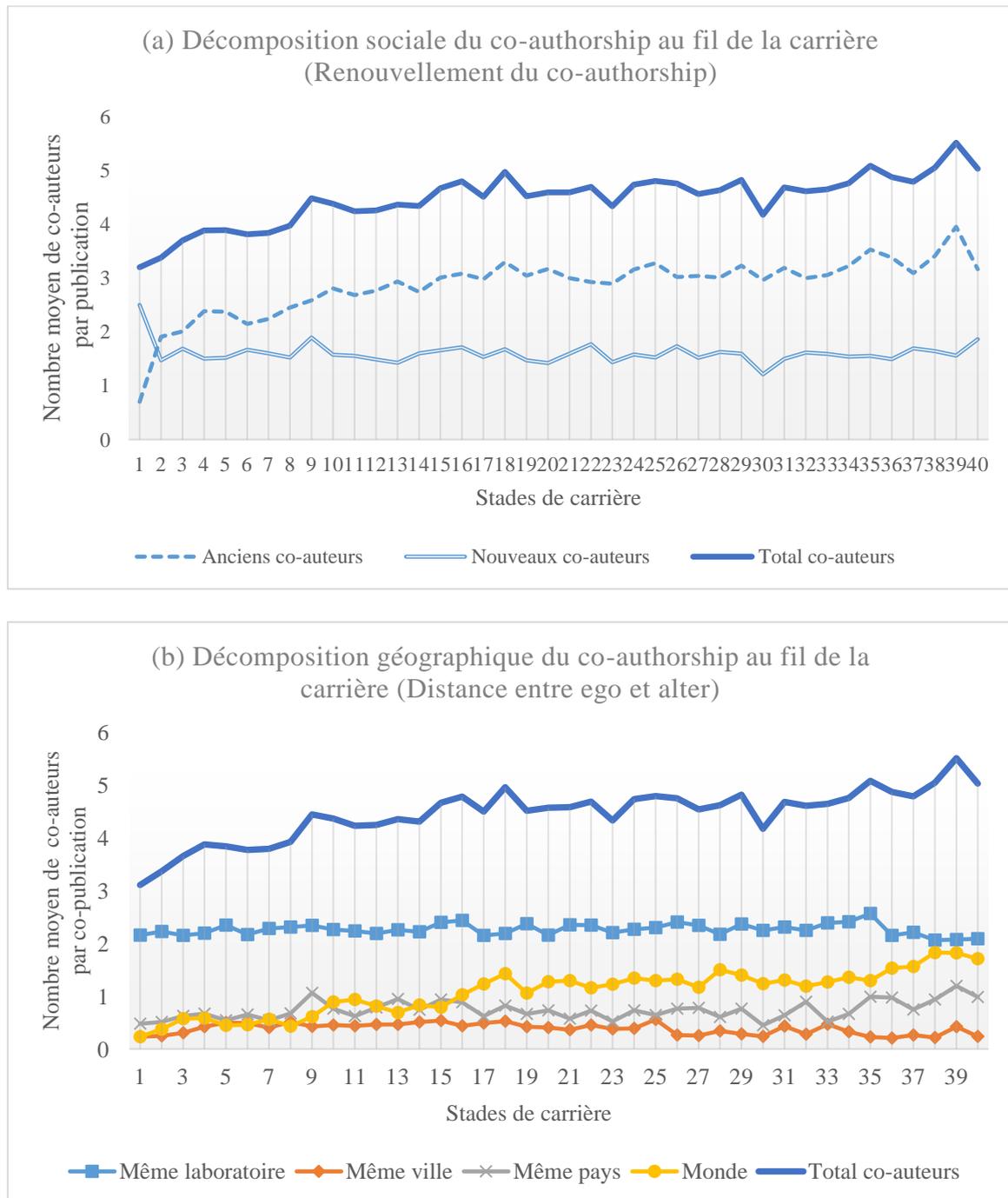
Figure 21 : Nombre moyen de co-auteurs par article par génération au fil des années (a) et des stades de carrière (b)



Enfin, au travers de ces données longitudinales, nous pouvons présenter un premier résultat préliminaire relatif au comportement collaboratif du chercheur au fil de sa carrière. Les graphiques de la figure 22 décomposent l'équipe de publication en fonction du renouvellement social et de la distance

géographique *ego-alter*. Dans chaque graphique, la courbe du haut réfère à la tendance croissante du nombre total de co-auteurs par publication au fil de la carrière (pour toute génération, la progression de carrière se traduit par l'accroissement de la dimension collaborative des articles).

Figure 22 : Décomposition sociale (a) et géographique (b) de l'équipe de publication au fil de la carrière



Le graphique 22(a) montre le poids moyen des nouveaux co-auteurs dans l'équipe de publication en comparaison au poids des anciens co-auteurs (c'est-à-dire les relations créées durant les années précédentes et renouvelées). **La croissance de la dimension collaborative des articles au fil des stades de carrière est le résultat du maintien des relations avec les anciens co-auteurs et de la stabilité du nombre de nouveaux co-auteurs par article (environ 1,5).** Par conséquent, lorsque le chercheur gravit les différents stades de carrière, les chercheurs ne collaborent pas de plus en plus avec de nouveaux co-auteurs mais plutôt avec les anciens co-auteurs comme le soulignent Cabanac *et al.* (2015).

Le graphique 22(b) illustre la distribution spatiale des relations *ego-alter* dans l'équipe de publication. Concernant la distribution spatiale du co-authorship au fil de la carrière et la distance géographique qui sépare *ego* et *alter*, nous notons l'importance et la stabilité du poids des collègues du laboratoire d'*ego* (environ 2,0 par publication) dans l'activité de publication au fil de la carrière. Les *alter* centraux apparaissent centraux au fil de la carrière. Bien que les *alter* non locaux soient minoritaires, ils s'accroissent au fil de la carrière. Nous observons une ouverture géographique progressive du co-authorship et même une internationalisation du réseau scientifique pour toutes les générations. **La densité du réseau repose à la fois sur le renforcement du poids des collègues du laboratoire et l'internationalisation du co-authorship.**

B) Réseaux de co-authorship par période d'activité : Approche de l'analyse de réseaux

Une seconde approche du réseau de co-authorship est mobilisée via la théorie sociologique de l'Analyse de Réseaux Sociaux. Cette nouvelle appréhension des réseaux repose sur une segmentation des carrières en différentes périodes et la réalisation de réseaux *ego*-centrés. Plusieurs statistiques descriptives seront ensuite présentées, relatives à la taille des réseaux de co-authorship.

1- Segmentation des carrières

Nous avons reconstitué les réseaux de co-authorship dans le but d'étudier les styles de collaboration des chercheurs au fil de leur carrière. Nous avons donc segmenté la carrière de nos chercheurs afin d'identifier la structure du réseau à des stades de carrière précis et d'opter une vision dynamique quant à son évolution.

Notre base de données bibliométrique met en lumière la carrière de chacun de nos 80 *ego* à partir de leur première publication. Ainsi, nous choisissons cette année comme référence de l'activité de publication des chercheurs. Par conséquent, nous avons segmenté la carrière des *ego* **en quatre stades professionnels** relatifs à la première année de publication d'*ego* nommée t_0 :

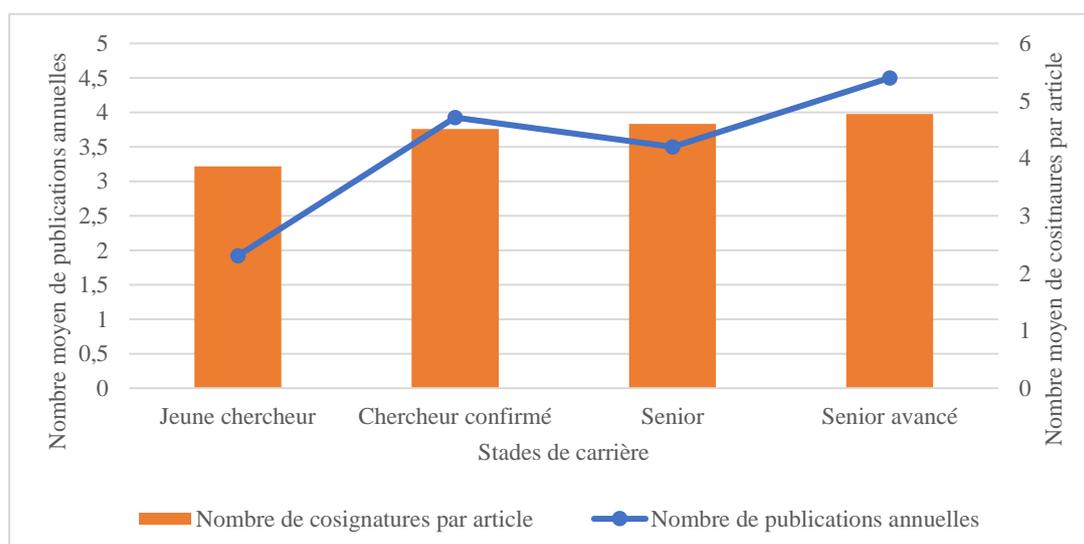
Tableau 20 : Description des stades de carrière

Stades de carrière	Critère du stade de carrière : Publications d' <i>ego</i> réalisées ...
Jeune chercheur	Entre t_0 et $(t_0 + 9)$
Chercheur confirmé	Entre $(t_0 + 10)$ et $(t_0 + 18)$
Senior	Entre $(t_0 + 19)$ et $(t_0 + 27)$
Senior avancé	Entre $(t_0 + 28)$ et $(t_0 + 50)$

Nous souhaitons nous rapprocher le plus finement possible des différentes phases qui composent la carrière d'un scientifique. Grâce l'étude des CV des chercheurs et aux entretiens semi-directifs réalisés au cours des chapitres précédents, nous avons pu capter les principaux stades professionnels.

Durant ses dix premières années de carrière, le jeune chercheur réalise ses premières publications dans le cadre de son doctorat (ou post-doctorat), puis en tant que chargé de recherche ou de maître de conférences. Son rythme annuel moyen de publication de 1,92 publications dans lesquels il collabore en moyenne avec 3,86 co-auteurs. Ensuite, durant sa seconde décennie de carrière, le chercheur confirmé est généralement promu aux postes de professeur ou de directeur de recherche. Son activité de publication s'intensifie puisqu'il publie en moyenne 3,93 publications par an avec en moyenne 4,51 co-auteurs. Durant sa troisième décennie de carrière, le chercheur dit senior poursuit son activité au sein du laboratoire en disposant d'une forte ancienneté et expérience. Enfin, le senior avancé termine sa carrière en tant que chercheur émérite, durant laquelle il participe à des projets très collaboratifs : il collabore en moyenne avec 4,5 co-auteurs par article.

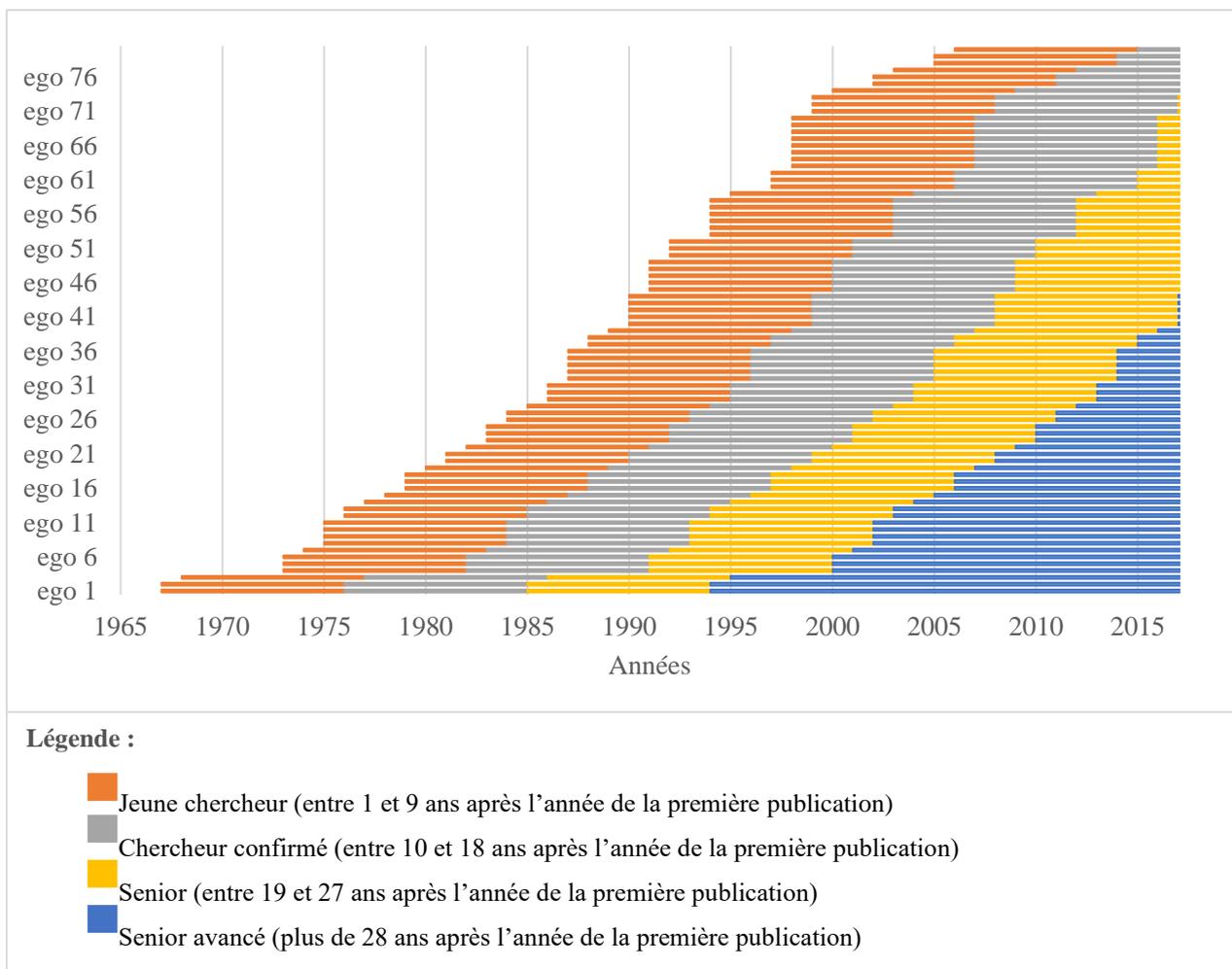
Figure 23 : Activité de publication au fil des stades de carrière



Nous avons choisi des **périodes longues d’au moins dix ans** pour deux raisons. Premièrement, nous souhaitons segmenter nos carrières en phases suffisamment longues pour identifier les collaborations émis au sein d’un nombre conséquent de publications. Notre objectif est d’avoir une vision représentative de l’activité de publication et de collaboration des chercheurs. Ensuite, ces périodes sont également assez courtes pour capter différentes phases de la trajectoire de carrière : nous disposons de données longitudinales qui nous permettent d’étudier de longues carrières (la plus longue étant de 51 ans).

Une fois la méthodologie de segmentation choisie, nous l’avons appliqué aux 80 carrières. Le graphique suivant illustre ce processus de segmentation des carrières et l’hétérogénéité des durées de carrière. Chaque barre est associée à un chercheur : le graphique dispose donc de 80 barres horizontales au total. Chacune représente la carrière d’un chercheur, allant de l’année de la première publication jusqu’à 2018. La barre est donc segmentée et colorée en fonction du stade de carrière dans lequel le chercheur se trouve, relativement à sa première année de publication.

Figure 24 : Segmentation des carrières



Trois précautions ont été prises en compte relativement aux différentes générations des chercheurs dans ce processus de création de réseau :

i) La première précaution concerne le fait que certains les chercheurs **n'aient pas publié à certains stades de carrière**. En effet, plus le chercheur a commencé à publier tard, moins il a de chance d'avoir publié en tant que « senior » ou « senior avancé ». Par exemple, les chercheurs qui ont commencé leur carrière entre 1997 et 2006 ont moins de 20 ans de carrière en 2018 : les plus jeunes d'entre eux n'ont pu publier qu'en tant que « jeune chercheur », que « chercheur confirmé ». Par conséquent, le nombre de réseaux reconstitués ne sera pas le même pour chaque stade de carrière.

ii) La seconde précaution concerne **le stade de « senior avancé »**. Alors que les trois stades précédents ont une durée régulière de 9 ans, le stade de « senior avancé » démarre à la 28^{ème} année de publication jusqu'à ce que le chercheur décide de clore sa carrière. Ainsi, l'activité de publication du « senior avancé » dure entre 2 et 25 ans, selon la génération du chercheur.

iii) Enfin, la troisième précaution concerne la création de réseaux sur **les périodes dites incomplètes**. Cela concerne notamment les chercheurs des générations « 1987-1996 » et « 1997-2006 » (qui avaient entre 12 et 31 années de carrière en 2018) et les stades de « senior » et « senior avancé ». Il peut arriver que ces chercheurs publient sur une période inférieure à celle du stade. Par exemple, un chercheur qui commence à publier en 1998 deviendra jeune chercheur en 2006 puis chercheur confirmé en 2015. Jusqu'en 2018, il n'aura donc que 3 années d'activité en tant que « senior ». Prendre en compte le réseau de ce chercheur relatif à ces trois années d'activité soulève le risque de créer une illustration non représentative des collaborations du chercheur (trop peu de publications sont étudiées). Nous avons donc suivi deux conditions pour retenir les réseaux de courte période. La première condition est que l'activité de publication soit suffisamment dense pour pouvoir capter des collaborations (le chercheur doit avoir publié au moins 10 co-publications). La seconde condition concerne les cas où les réseaux sont dotés de moins de 10 co-publications : nous ne retenons que les réseaux dont le volume de publication reste représentatif de l'activité d'*ego* (c'est-à-dire où *ego* n'avait pas beaucoup publié durant les périodes précédentes).

Tableau 21 : Distribution des réseaux selon la génération et le stade de carrière du chercheur

	Stade de carrière	<i>Jeune chercheur</i>	<i>Chercheur confirmé</i>	<i>Senior</i>	<i>Senior avancé</i>	Total
Année de la première publication	<i>1967-1976</i>	13	13	13	13	52
	<i>1977-1986</i>	18	18	18	16	70
	<i>1987-1996</i>	28	28	28	5	89
	<i>1997-2006</i>	21	21	6		48
	Total	80	80	65	34	259

Grâce à notre base de données, nous avons reconstitué 259 réseaux de co-authorship, dans lesquels les 80 chimistes ont réalisé 9146 co-publications avec 14399 co-auteurs différents (sommets).

Comme nous l'avons souligné, la longueur des trajectoires de carrière n'est pas la même selon l'année durant laquelle le chercheur a commencé à publier. Les 80 chercheurs sont représentés aux stades de « jeune chercheur » et « chercheur confirmé ». En revanche, à partir du stade de « senior », nous n'étudions plus que 65 réseaux et 34 à partir du stade de « chercheur confirmé ».

La génération « 1997-2006 » est sous-représentée aux stades de « senior » et « senior avancé » et la génération « 1987-1996 » l'est également au stade de « senior ». Dans ces deux cas particuliers, nous convenons qu'il s'agit de petits réseaux dans le sens où le chercheur n'a qu'une courte période de publication.

2- Lecture graphique des réseaux ego-centrés

Bien que nos traitements statistiques et économétriques mobilisent davantage des indicateurs de réseaux, les graphiques ont néanmoins une grande place dans notre analyse. Nous avons reconstitué les représentations graphiques des 259 réseaux. Nous avons deux possibilités quant à la méthode de construction des réseaux, en fonction de la présence d'*ego* ou non. Les deux graphiques ci-dessous représentent ces deux options pour les mêmes collaborations.

Figure 25 : Réseau personnel dans lequel *ego* est représenté

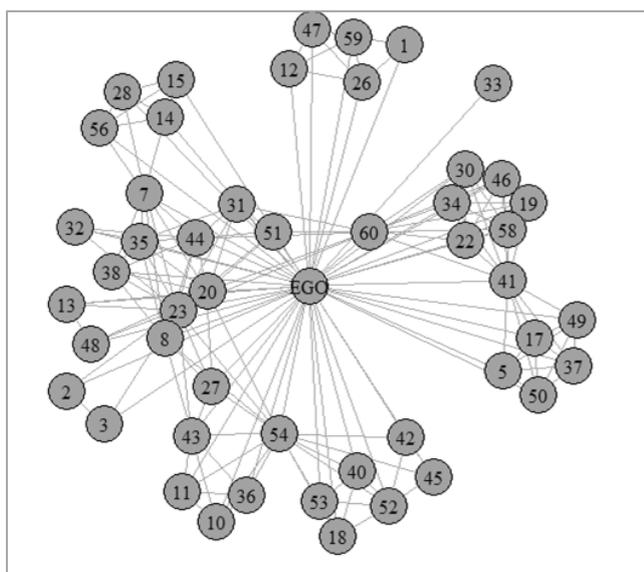
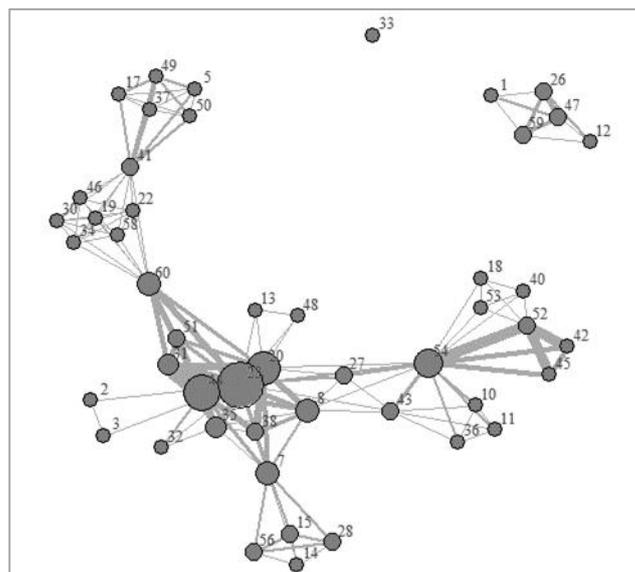


Figure 26 : Réseau personnel *ego*-centré



Dans le graphique de gauche, *ego* est représenté au même titre que ses co-auteurs (chacun étant nommé par un numéro). *Ego* est donc au centre des collaborations puisqu'il est relié à chacun de ses *alter*. Cependant, ce réseau dispose de plusieurs inconvénients. Tout d'abord, il est peu lisible : il est surchargé par les flux d'*ego* qui cachent les différents flux entre les *alter*. Ensuite, il contient de nombreuses informations superficielles et non utiles. En effet, la présence d'*alter* dans le réseau est conditionnée par la collaboration avec *ego* : ainsi, les arrêtes reliant chaque *alter* à *ego* ne sont pas nécessaires dans le sens où on sait qu'*alter* a forcément collaboré avec *ego* pour être présent dans le réseau. Ce type de réseau est davantage utilisé dans le cas de réseaux complets dans lesquels *ego* représente également le co-auteur d'un autre chercheur. Or, dans le cas de réseau personnel à l'échelle micro, la présence d'*ego* n'est pas nécessaire.

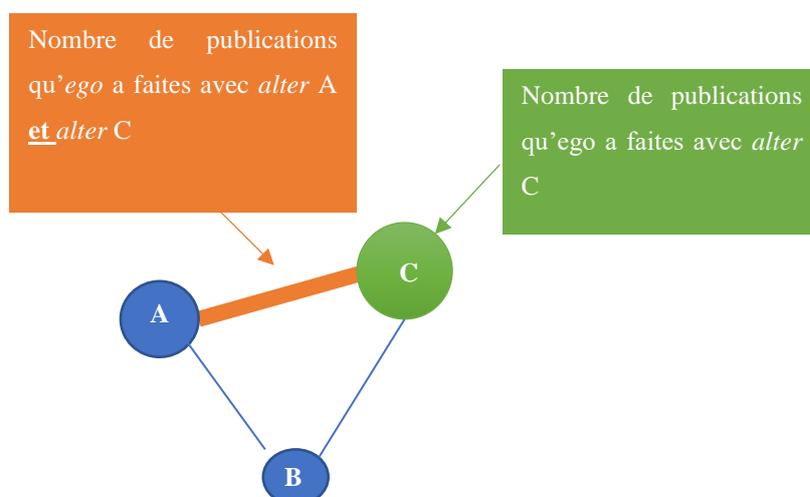
Ainsi, **le réseau *ego*-centré nous semble plus adapté à notre objectif de recherche qui est d'identifier la structure du réseau au travers des flux entre les *alter*.** Le graphique de droite en est une illustration : il s'agit du même graphique que celui de gauche soustrait de la présence d'*ego* et de ses différentes arrêtes. Le réseau *ego*-centré est donc allégé de la présence d'*ego* et permet de se concentrer sur les flux entre les autres nœuds numérotés (qui collaborent avec *ego* par définition). Bien qu'*ego* ne soit plus représenté graphiquement dans le réseau, deux indicateurs (et caractères graphiques) nous permettent de décrire ses collaborations avec chacun de ses *alter* :

- **Le diamètre de chaque nœud i** illustre le nombre de publications qu'*ego* a réalisé avec son co-auteur i . Ainsi, plus le nœud dispose d'un fort diamètre, plus cet *alter* a été récurrent dans les publications d'*ego* dans le stade de carrière étudié. La grosseur du nœud nous permet donc d'évaluer la durée de la relation collaborative entre *ego* et chacun de ses *alter*.

- **L'épaisseur de l'arrête entre deux nœuds i et j** représente le nombre de publications qu'*ego* a réalisé conjointement avec ses *alter* i et j . Plus l'arrête est épaisse, plus *ego* a collaboré simultanément avec le duo d'*alter* i et j . Ainsi, l'épaisseur de l'arrête nous permet d'évaluer la durée de la relation collaborative entre *ego* et chacun de ses couples d'*alter* i et j , dans le cas où ils ont collaboré tous les trois.

Le schéma ci-dessous illustre la lecture de ce type de graphique. Tout d'abord, nous pouvons constater que le diamètre du nœud C est plus gros que ceux des nœuds A et B. Cela signifie qu'*ego* a réalisé plus de co-publications avec C qu'avec A ou B. Ensuite, l'arrête reliant les nœuds A et C est plus épaisse que les deux autres arrêtes (A-B et B-C). Cela signifie qu'A et C ont réalisé plus de publications conjointement avec *ego* qu'A et B ou bien B et C l'ont fait.

Figure 27 : Lecture du réseau ego-centré



Après avoir expliqué la méthodologie de création et de lecture des réseaux, nous pouvons les décrire au travers de statistiques descriptives.

3- Statistiques descriptives

Cette troisième partie a pour but de décrire les 259 réseaux relativement aux habitudes de collaboration des chercheurs et de leur hétérogénéité.

Les habitudes collaboratives et à la taille des réseaux

Plusieurs résultats descriptifs nous permettent de comprendre les pratiques collaboratives des *ego* au travers de la structure des réseaux.

i) Taille du réseau et nombre de publication : Tout d'abord, la taille moyenne du réseau varie selon la description d'*ego*. En moyenne, *ego* réalise en moyenne 35,31 publications (4,03 publications par an) et collabore avec 64,58 *alter* dans chaque réseau.

Les réseaux grossissent au fil de la carrière d'*ego*. Durant ses 9 années de jeune chercheur, *ego* publie en moyenne 17,28 publications (2,30 publications par année d'activité⁴⁵) avec 29,59 co-auteurs. Le rythme de publication d'*ego* s'intensifie durant ses phases de chercheur confirmé (en moyenne 35,34 publications et 4,27 publications par année d'activité) et de senior (en moyenne 39,17 publications et 5,04 publications par année d'activité). Cela se traduit par une accentuation progressive de la taille du réseau : en moyenne, *ego* collabore avec 64,10 *alter* en tant que chercheur confirmé et 75,32 *alter* en tant que senior. Enfin, on note une forte accentuation de la taille du réseau sur la fin de carrière d'*ego*.

⁴⁵ Année durant laquelle il a réalisé au moins une publication

En moyenne, le senior avancé réalise 70,29 publications dans lesquelles il collabore avec 127,50 *alter*. Cette forte croissance est partiellement expliquée par la hausse du rythme de publication annuel (en moyenne 5,62 publications par année d'activité). Cette croissance de la taille du réseau au stade de senior avancé est surtout due au fait que ce stade de carrière est beaucoup plus long que les autres : en moyenne, ego publie durant 7,51 années distinctes au tant que jeune chercheur, chercheur confirmé et senior. En revanche, il publie en moyenne durant 11,15 années distinctes en tant que senior avancé.

La génération d'ego influence également la taille de son réseau. Par exemple, les chercheurs qui ont commencé à publier entre 1967 et 1976 ont des réseaux de 79,65 *alter* en moyenne alors que ceux qui ont commencé à publier entre 1997 et 2006 ont en moyenne un réseau de 64,79 *alter*. Ceci est dû aux différentes durées des carrières. Les anciennes générations sont davantage représentées en tant que chercheurs expérimentés (« senior » ou « senior avancé »). Les générations plus récentes ont de courtes carrières : leurs réseaux sont associés aux périodes où le rythme de publication annuel est moins dense (« jeune chercheur », « chercheur confirmé »). En moyenne, les réseaux de la génération 1967-1976 concentrent 54,46 publications et ceux de la génération 1997-2006 concentrent 28,48 publications. En revanche, cette tendance n'est pas linéaire puisque le réseau moyen de la génération 1997-2006 est plus gros que ceux des générations 1977-1986 et 1987-1996 : le volume de publications ne suffit pas à expliquer la taille du réseau selon la génération d'ego.

Les réseaux des ego hommes sont en moyenne plus gros que ceux des ego femmes. Cela peut être dû au fait que les ego femmes ont commencé leur carrière plus tardivement que les hommes : parmi nos 22 ego femmes, 14 d'entre elles ont commencé à publier entre 1987 et 1996. Ainsi, elles sont très peu représentées aux stades de « senior » et « senior avancé », qui sont les stades où l'activité de publication et la taille des réseaux sont les plus riches.

Enfin, **les réseaux des directeurs de recherche sont aussi plus gros que ceux des professeurs** : en moyenne, le directeur de recherche publie 42,77 publications avec 75,40 *alter* par réseau alors que le professeur ne publie que 29,07 publications avec 55,52 *alter*. Cela peut être expliqué par un rythme de publication différent : dans chaque réseau, les chercheurs directeurs de recherche publient en moyenne 4,68 publications par année d'activité et les professeurs 3,48 publications.

ii) Taille du réseau et nature collaborative des articles : **La taille du réseau s'explique également par la nature collaborative des articles, soit le nombre d'*alter* moyen avec qui ego collabore.** Pour un même nombre de publication, deux réseaux disposant d'équipes plus ou moins nombreuses auront une taille plus ou moins importante. En moyenne, ego collabore avec 4,5 *alter* par publication dans ses réseaux et chaque *alter* accompagne ego dans la collaboration avec 4,03 autres *alter*. D'un point de vue graphique, cela signifie que chaque réseau dispose en moyenne de 4,5 nœuds qui ont chacun environ 4,03 arrêtes.

La hausse de la nature collaborative des articles explique également la croissance et la densité des réseaux au fil de la carrière et des générations. Au fil des stades de carrière, le nombre moyen d'*alter* par article s'accroît. En moyenne, un « jeune chercheur » collabore avec 3,86 *alter* par publication et un « senior avancé » collabore avec 4,77 *alter*. Ainsi, en moyenne, *alter* accompagne *ego* « jeune chercheur » dans ses liens collaboratifs avec 3,39 *alter* différents et accompagne *ego* « senior avancé » avec 4,24 *alter*. Aussi, la nature collaborative des articles s'accroît fortement au fil des générations. En moyenne, la génération « 1967-1976 » collabore avec 3,94 *alter* par article et la génération « 1997-2006 » avec 5,27 *alter*. Ainsi, plus *ego* commence sa carrière tard, plus son réseau grossit et ses *alter* sont connectés dans son réseau (ie, l'accompagnent dans ses collaborations avec différents co-auteurs). En moyenne, l'*alter* d'un *ego* de la génération « 1967-76 » l'accompagne dans ses liens collaboratifs avec 3,57 autres *alter* alors que l'*alter* d'un *ego* de la génération « 1997-06 » l'accompagne dans ses liens collaboratifs avec 4,88 *alter*.

Aussi, **la nature collaborative des publications des *ego* femmes est plus forte que celle des hommes** : ceci est dû au fait qu'elles appartiennent aux générations récentes. Ainsi, elles sont plus propices à collaborer avec de grosses équipes. Enfin, **les directeurs de recherche ont également des articles de plus forte nature collaborative que les professeurs.** En moyenne, les directeurs de recherche collaborent 4,60 *alter* par article et les professeurs avec 4,38 *alter*. Aussi, en moyenne, *alter* accompagne un directeur de recherche dans ses liens collaboratifs avec 4,11 autres *alter* et accompagne un professeur avec 3,94 *alter*.

iii) Taille du réseau et maintien de relation collaborative : Contrairement au rythme de publication et à la nature collaborative des articles, **la durée du maintien de relations collaboratives** (nombre de cosignatures moyen avec chaque sommet, nombre de collaborations par arrêtes) **ne semble pas affecter la taille du réseau d'*ego***. En effet, en moyenne, *ego* échange 2,46 collaborations avec chaque *alter* et 1,49 collaboration avec chaque couple d'*alter*. Néanmoins, nous constatons une certaine homogénéité relative à ces indicateurs chez tous nos *ego*. Ainsi, quels que soient la nature collaborative des articles ou le nombre de publication réalisées, la durée du maintien de relation collaborative reste stable et n'affecte pas la taille du réseau.

Tableau 22 : Statistiques descriptives des réseaux

	Stade de carrière				Total
	Jeune chercheur	Chercheur confirmé	Senior	Senior avancé	
Nombre de réseaux	80	80	65	34	259
Nombre de co-publications	1383	2827	2546	2390	9146
Nombre de cosignatures d'ego	5340	12738	11708	11393	41179
Nombre moyen de publications par réseau	17,29	35,34	39,17	70,29	35,31
Nombre moyen de cosignatures par co-publication	3,86	4,51	4,60	4,77	4,50
Nombre total de sommets	2367	5128	4896	4335	16726
Nombre moyen de sommet par réseau	29,59	64,10	75,32	127,50	64,58
Nombre de cosignatures moyen avec chaque sommet	2,26	2,48	2,39	2,63	2,46
Nombre d'arrêtes	8018	21807	19203	18364	67392
Nombre total de liens collaboratifs au sein des arrêtes*	12252	32073	27991	27953	100269
Nombre de liens collaboratifs par arrêtes	1,53	1,47	1,46	1,52	1,49
Nombre d'arrêtes par sommet	3,39	4,25	3,92	4,24	4,03

	Génération				Genre		Type de carrière		
	1967-1976	1977-1986	1987-1996	1997-2006	Femmes	Hommes	Directeurs de recherche	Professeurs	Total
Nombre de réseaux	52	70	89	48	72	187	118	141	259
Nombre de co-publications	2832	2357	2590	1367	1974	7172	5047	4099	9146
Nombre de cosignatures d'ego	11157	9848	12965	7209	9598	31581	23218	17961	41179
Nombre moyen de publications par réseau	54,46	33,67	29,10	28,48	27,42	38,35	42,77	29,07	35,31
Nombre moyen de cosignatures par co-publication	3,94	4,18	5,01	5,27	4,86	4,40	4,60	4,38	4,50
Nombre total de sommets	4142	3920	5554	3110	4144	12582	8897	7829	16726
Nombre moyen de sommet par réseau	79,65	56,00	62,40	64,79	57,56	67,28	75,40	55,52	64,58
Nombre de cosignatures moyen avec chaque sommet	2,69	2,51	2,33	2,32	2,32	2,51	2,61	2,29	2,46
Nombre d'arrêtes	14797	14540	22871	15184	17277	50115	36538	30854	67392
Nombre total de liens collaboratifs au sein des arrêtes*	22719	21662	33646	22242	24530	75739	55707	44562	100269
Nombre de liens collaboratifs par arrêtes	1,54	1,49	1,47	1,46	1,42	1,51	1,52	1,44	1,49
Nombre d'arrêtes par sommet	3,57	3,71	4,12	4,88	4,17	3,98	4,11	3,94	4,03

L'hétérogénéité de la taille des réseaux

Après avoir constaté que la taille du réseau évoluait principalement selon le nombre de publications réalisées et la nature collaborative des articles, nous constatons l'hétérogénéité de la taille des réseaux.

Tableau 23 : Dispersion des tailles de réseaux (selon le nombre d'*alter*)

	Min	1er Quartile	Médiane	Moyenne	Écart type	Coefficient de variation	3 ^{ème} Quartile	Max
Total	4	25	48	64,58	65,43	1,01	79	539
Jeune chercheur	4	11	22	29,59	24,17	0,82	41,75	107
Chercheur confirmé	7	33	49	64,10	49,86	0,78	83,5	278
Senior	14	43	64	75,32	47,16	0,63	86	239
Senior avancé	13	65,5	84,5	127,50	120,54	0,95	159,25	539

Le tableau ci-dessus nous permet d'appréhender la **dispersion des tailles de réseaux** sur l'ensemble des carrières et par stade de carrière.

Nous notons la présence de réseaux extrêmes qui sont beaucoup plus gros que la moyenne. En effet, 75% des réseaux ont entre 4 et 79 *alter* et 25% concentrent entre 80 et 539 *alter*. **Cette dispersion relative à la taille des réseaux s'accroît au fil des stades de carrière.** En effet, au stade de jeune chercheur, nous notons une concentration de petits réseaux (25% des réseaux ont entre 4 et 11 *alter*) et de plus gros réseaux (25% des réseaux concentrent entre 42 et 107 *alter*). La dispersion s'atténue aux stades de chercheur confirmé et de senior, durant lesquels le coefficient de variation se réduit. En revanche, la dispersion s'accroît au stade de senior avancé : 50% des réseaux ont entre 13 et 85 *alter*, 25% des réseaux ont entre 86 et 159 *alter* et 25% des réseaux ont entre 160 et 539 *alter*. Ainsi, au stade de senior avancé, nous notons une concentration de petits réseaux et quelques réseaux extrêmement gros. Cela s'explique notamment par l'hétérogénéité du rythme de publications et de la durée de la période. Dans les petits réseaux (qui ont entre 13 et 85 *alter*), les chercheurs ont publié en moyenne 5,62 publications annuelles durant en moyenne 11,15 années alors que dans les gros réseaux (qui ont entre 160 et 539 *alter*), les chercheurs ont publié en moyenne 10,61 publications annuelles durant en moyenne 16,67 années.

IV- Conclusion

L'objectif du chapitre 2 est de présenter la base de données bibliométrique mobilisée dans les traitements empiriques de la thèse via ses aspects méthodologiques et des statistiques descriptifs.

Nous avons commencé par décrire les choix méthodologiques qui ont permis la construction de la base. Nous avons focalisé notre attention sur la chimie puisque cette discipline est un très bon champ d'étude des dynamiques collaboratives des chercheurs. Outre les enjeux économiques et scientifiques relatifs au positionnement mondial de la chimie française et de sa stratégie de spécialisation thématique, la bibliométrie des travaux de chimie offrent de belles opportunités d'étude du comportement collaboratif des chercheurs. Nous avons choisi deux laboratoires français de chimie qui partagent des points communs sur différents domaines : la spécialisation dans des thématiques précises, la mise en place de plateformes de collaboration interne, la visibilité nationale et internationale et la qualité reconnue des travaux scientifiques. En plus de cette comparabilité, nous devons néanmoins prendre en compte certaines spécificités structurelles relatives à l'historique des laboratoires, à leur statut d'unité de recherche et à leur structure interne. Enfin, nous avons sélectionné parmi ces deux structures les 80 chercheurs qui étaient directeurs de recherche CNRS ou professeurs des universités en 2018. Ce critère assure la performance scientifique des chercheurs ainsi qu'une longévité de carrière et une réputation scientifique.

A partir de cette liste de chercheurs, nous avons extrait leurs 9233 co-publications afin de reconstituer leurs réseaux personnels à l'échelle micro via la création de paires de co-authorship. Nous avons enrichi la base de données à partir de données complémentaires et de codage relatives à la dimension géographique des relations collaborations, à la description des co-auteurs (genre, année de soutenance de thèse) et des chercheurs de l'échantillon (nom du directeur de thèse et liste des doctorants sous leur supervision). Enfin, nous avons regroupé les 80 chercheurs en cohortes générationnelles. Puisque ces chercheurs ont commencé leur carrière à des contextes différents, ils n'ont pas la même durée de carrière (la plus courte étant de 12 ans et la plus longue de 51 ans). Ainsi, nous avons regroupé les chercheurs en durée de carrière comparables.

Dans un second temps, nous avons présenté des statistiques descriptives relatives aux activités de publication et de collaboration des chercheurs. Tout d'abord, une grande partie des relations *ego-alter* sont ponctuelles et dure que le temps d'une publication, quelques soient le genre, la génération ou le type de carrière du chercheur. Aussi, nous avons montré la forte fréquence de l'activité de publication et l'intense dimension collaboration de la recherche de chimistes. Seule la dimension collaborative des articles dépend de la génération du chercheur.

A partir des données complémentaires sociales collectées, nous avons observé que la relation de mentorship des chercheurs influence leurs habitudes de collaboration. En effet, les équipes de publications des hommes sont très structurées par les doctorants qu'ils supervisent (ou ont supervisé) et

les publications des chimistes femmes sont davantage structurées par le maintien de relation avec leur directeur de thèse durant leur carrière. Également, nous observons que la croissance de la nature collaborative des articles se traduit par un recours croissant aux co-auteurs diplômés d'un doctorat et non de doctorants ou de support technique ou administratif. Au fil des générations, les publications sont de plus en moins structurées par la présence du directeur de thèse.

Enfin, à partir des données relatives à la localisation géographique des *ego* et *alter* pour chaque collaboration, nous pouvons identifier la géographie des collaborations fractionnées. Nous constatons une forte dimension française du co-authorship : les cosignatures fractionnées issues de villes françaises ont contribué à 75,82% des co-publications. Ces contributions françaises viennent majoritairement de Poitiers, de Toulouse et d'autres pôles français plus ou moins spécialisés dans des thématiques précises. Nous notons également la présence de pôles internationaux, notamment en Espagne, aux Etats-Unis, en Allemagne et en Russie.

Dans un troisième temps, nous avons proposé deux approches différentes de reconstitution des réseaux de co-authorship. La première approche est une approche de panel par activité de publication. Dans cette approche longitudinale, plusieurs implications ont été soulignées relatives à l'activité de publication et de collaboration des chercheurs au fil de la carrière. En effet, nous avons montré que le rythme annuel de publication augmentait au fil de la carrière. Aussi, la dimension collaborative des articles s'accroît au fil de la carrière et des générations de chercheurs. Un premier résultat a été souligné concernant le comportement collaboratif des chercheurs au fil de la carrière via la décomposition des équipes de publication. Nous avons montré que la hausse de la nature collaborative des articles est dû au maintien de relations collaboratives (la création de nouvelles relations collaboratives par article étant stable). Aussi, l'équipe de publication étant constitué d'un noyau stable et structurant de collègues du laboratoire au fil de la carrière, la hausse de la nature collaborative s'explique par l'internationalisation progressive des collaborations. La seconde approche est une approche d'Analyse de Réseaux Sociaux. Dans cette perspective, la carrière scientifique est segmentée en quatre périodes régulières et chaque réseau est associé la période d'un chercheur. La distribution de ces périodes professionnelles varie en fonction de la génération du chercheur. A partir de cette segmentation, 259 réseaux *ego*-centrés ont été créés afin d'identifier la structure des collaborations. Enfin, via des statistiques descriptives, nous avons montré que la taille du réseau dépend à la fois du nombre de publication et de la nature collaborative des articles. Néanmoins, le maintien des relations collaboratives ne semble pas avoir d'impact sur la taille du réseau. Au vu des différentes pratiques de publication et de collaboration, les 259 réseaux sont de taille hétérogène.

A partir de cette base de données bibliométrique, nous proposons trois études empiriques à l'échelle micro qui ont pour but de mieux comprendre le comportement collaboratif des chimistes au fil de leur carrière

Tableau 24 : Description des trois études empiriques réalisées à partir de la base de données

Objectif de la thèse : Etude du comportement collaboratif des chimistes au fil de leur carrière				
<i>Type d'approche</i>	<i>Objectif</i>	<i>Données / Méthodes</i>	<i>Chapitres</i>	<i>Références de publication</i>
Econométrie de panel (Réseaux de co-authorship par année)	Etude à visée exploratoire et compréhensive : Mise en lumière de l'évolution de la dynamique relationnelle des chimistes au fil des stades de carrière et des différents contextes de recherche	<u>Données</u> : Etude de 5 chimistes émérites 893 co-publications, 1 073 co-auteurs et 3 343 cosignatures <u>Méthodologie mixte</u> : Données bibliométriques sur les relations de co-authorship des chercheurs, CV et entretiens semi-directifs	Chapitre 3 : « Carrières des chercheurs et dynamique des réseaux scientifiques Une mise en intrigue à partir des cas de cinq chimistes émérites »	Bernela, B., Bernard, M., & Ferru, M. (2018). Carrières des chercheurs et dynamique des réseaux scientifiques. Une mise en intrigue à partir des cas de cinq chimistes émérites. <i>Temporalités. Revue de sciences sociales et humaines</i> , (27).
	Evaluer l'impact de la mobilité géographique sur la formation et les dynamiques du réseau de co-authorship	<u>Données</u> : Etudes des 80 chimistes, 9233 co-publications, 14603 co-auteurs, 41720 cosignatures <u>Méthodologie</u> : modèles économétriques de panel	Chapitre 4 : « Est-ce que la mobilité géographique façonne le réseau de co-authorship ? Une approche de panel des carrières scientifiques »	Bernard, M., Bernela, B., & Ferru, M. (2020). Does the geographical mobility of scientists shape their collaboration network? A panel approach of chemists' careers. <i>Papers in Regional Science</i> .
Analyse de réseaux sociaux (Réseaux de co-authorship par période)	Création d'une typologie de collaboration à partir de la structure des réseaux personnels Explication des styles collaboratifs à partir de données de carrière	<u>Données</u> : Etude des 80 chimistes, 259 réseaux ego-centrés <u>Méthodologie</u> : Analyse de réseaux sociaux, Analyse de composantes principales, Classification hiérarchique, modèles <i>logit</i>	Chapitre 5 : « Existe-t-il plusieurs façons de collaborer au fil de la carrière ? Une analyse de réseaux sociaux de co-authorship des chimistes »	Working paper

Chapitre 3 : Carrières des chercheurs et dynamique des réseaux scientifiques : une mise en intrigue à partir des cas de cinq chimistes émérites

L'objectif de ce chapitre est de proposer une analyse de la dynamique des réseaux scientifiques, en privilégiant une entrée par l'individu, avec la carrière du chercheur comme échelle d'analyse. L'originalité de cette étude porte sur le **couplage des échelles (action et temps)** : nous construisons le réseau personnel d'un chercheur à partir de l'ensemble de ses relations de co-authorship et repérons les moments d'émergence, de renforcement ou d'affaiblissement des liens avec ses coauteurs (Suitor *et al.*, 1997) ainsi que la géographie de ces relations. Afin d'identifier dans quelle mesure le contexte historique de recherche et le stade de carrière impactent les dynamiques relationnelles des chercheurs, nous choisissons de réaliser une **étude à visée exploratoire et compréhensive**. A partir d'une étude de cas de 5 chercheurs émérites, nous recontextualisons leurs collaborations entre 1967 et 2018. Pour ce faire, nous mobilisons une **méthodologie mixte** (publications/CV/entretiens semi-directifs) pour capter le plus finement possible la dimension relationnelle de l'activité de publication et la dynamique d'émergence et de développement des réseaux scientifiques tout au long de la carrière des chercheurs.

I- Contribution méthodologique : articulation des échelles d'analyse et des matériaux empiriques

Afin de mieux comprendre l'impact des événements de carrière sur la dynamique du réseau de co-authorship, il nous semble primordial de prendre plusieurs dispositions méthodologiques. Nous commencerons par porter une attention particulière à **l'intégration de la dynamique temporelle** dans le processus de carrière. En effet, plusieurs travaux ont souligné l'importance de la temporalité dans les processus sociaux. Dans un deuxième temps, nous soulignons l'intérêt de mobiliser conjointement plusieurs **matériaux empiriques de nature quantitative et qualitative**. Cette articulation méthodologique originale va nous permettre d'affiner l'analyse micro des dynamiques collaboratives.

A) Combiner les échelles d'action et de temps : pour une analyse processuelle des carrières scientifiques

L'objectif de notre recherche est de mieux comprendre **l'impact des éléments inhérents à la carrière d'un chercheur sur sa dynamique de collaboration**. L'intégration d'une perspective dynamique constitue un enjeu essentiel dans les sciences sociales, bien que son traitement ne soit pas toujours satisfaisant (Bidart et Mendez, 2016). En effet, « *le temps est souvent considéré comme une variable additionnelle mais sans que soit pris en compte ce qu'il modifie dans ce qui est observé. Dans d'autres cas, les auteurs se concentrent sur les événements et les dates et les processus sont fragmentés : la continuité de l'histoire et son lien avec le contexte peuvent être oubliés.* » (Bidart et al., 2012, p. 743). Abbott (1990 ; 2001) souligne dans ses travaux le rôle crucial du temps dans l'analyse des phénomènes sociaux : le temps modifie au fur et à mesure les phénomènes observés, les conditions de leur émergence et les contextes de leur réalisation.

S'appuyant sur ces travaux, Grossetti (2004, 2006), Brochier *et al.* (2010), Bidart et Mendez (2016) ont proposé ces dernières années de réintégrer une perspective dynamique dans l'analyse des phénomènes sociaux à travers une conception historique des processus. Les processus renvoient à un « *ensemble de phénomènes organisés dans le temps qui combinent des ingrédients et des moteurs dont le changement éventuel permet d'identifier des séquences* » (Brochier *et al.*, 2010, p. 90). Ces auteurs montrent en effet que le contexte ne peut être extérieur au processus et qu'une décomposition en séquences distinctes est essentielle afin d'identifier ce qui a changé (continuité versus changements plus ou moins marqués d'un point de vue géographique, thématique...) ainsi que les événements clés pouvant modifier fortement l'orientation de la trajectoire.

Enfin, il convient de rendre compte des moteurs de la dynamique observée via l'existence d'interdépendances et de causalités multiples. Ces éléments permettent de dépasser une analyse centrée sur une période afin d'élargir une simple analyse chronologique qui donnerait une vision fragmentée et isolée des phénomènes sociaux. Bidard et Mendez (2016, p.223) souligne : « *la carrière [doit être analysée] comme un tout, comme une combinaison de contraintes, d'opportunités et de choix, qui comporte une cohérence interne* ». Pour développer une telle démarche, Grossetti (2006) propose d'intégrer trois échelles d'analyse :

i) *L'échelle du temps* suppose de tenir compte des **temporalités multiples** pouvant se superposer au cours d'un processus. En effet, le temps court d'une expérience peut modifier la progression d'un projet ou même d'une carrière dans un temps plus long. La prise en considération de la multiplicité des échelles de temps nous invite à ne pas limiter l'étude de la carrière à l'histoire de la carrière professionnelle du chercheur. En effet, elle invite plutôt à intégrer la dynamique de l'équipe, de

la discipline scientifique tout en tenant compte du système national ou international de la recherche plus ou moins stable dans le temps.

ii) *L'échelle des masses* renvoie aux **niveaux micro et macro de la carrière**. Elle permet de passer du niveau de l'acteur individuel à celui de l'organisation.

iii) *L'échelle de la généralité* permet de réintégrer **le rôle du contexte**. Les frontières entre les échelles des masses et de la généralité sont poreuses et les différents niveaux d'action peuvent entrer en interaction. Par exemple, la participation d'un chercheur à des instances nationales de la recherche peut avoir une incidence sur les décisions prises au niveau du laboratoire.

B) Combiner données bibliométriques et matériaux qualitatifs : une analyse relationnelle de l'activité de publication

L'adoption d'une telle perspective analytique nécessite de développer une **méthodologie adaptée**. En effet, le processus doit être conceptualisé comme une histoire (Bidart *et al.*, 2012) et nécessite du matériau narratif qui rend compte de l'articulation de séquences dans le temps (Abbott, 2001).

D'un point de vue empirique, le développement de la scientométrie (via la couverture croissante des bases de données bibliométriques) a conduit au développement rapide des techniques d'analyse des réseaux sociaux pour l'étude du monde scientifique. Bien que ces travaux aient permis de mieux décrire l'activité de publication et de collaboration scientifique, les études exclusivement bibliométriques butent sur l'information limitée que présentent les publications (listes de noms et d'affiliations, années et revues). Ces données permettent de décrire l'évolution du réseau scientifique d'un chercheur d'un point de vue quantitatif, sans pouvoir identifier les mécanismes collaboratifs. **Afin de saisir plus finement la dimension relationnelle de l'activité de publication et la dynamique d'émergence et de développement des réseaux scientifiques tout au long de la carrière des chercheurs**, nous proposons de combiner les données bibliométriques avec des matériaux plus qualitatifs : les CV et les entretiens semi-directifs. L'articulation de ces différentes sources de données est une façon originale de rendre compte de la complexité de l'effet des trajectoires sur les collaborations scientifiques (Bernela et Milard, 2016).

1- Les paires de co-authorship issues de la base de données bibliométrique

Nous mobiliserons les données de notre **base de données bibliométrique comme matériel d'étude des collaborations**. Contrairement à de nombreuses études bibliométriques, les publications ne sont pas mobilisées comme un indicateur de productivité scientifique qu'un indicateur des dynamiques relationnelles à l'œuvre dans l'activité de publication (Milard, 2013) : « *la publication est un moyen*

privilegié de diffusion des recherches et que les collaborations scientifiques donnent de plus en plus souvent lieu à des publications cosignées par plusieurs partenaires » (Maisonobe, 2015, p. 158). De plus, à partir des informations complémentaires collectées (géographie de la relation ego-alter, description d'ego et d'alter), la base de données est un outil adéquate à une première contextualisation de chaque pair de co-authorship.

2- Les Curriculum Vitae

Les CV, largement sous-utilisés dans les études sur les sciences (Cañibano et Bozeman, 2009), représentent un matériau riche pour étudier les carrières. Dietz et al., 2000 soulignent l'intérêt de ce matériau dans l'étude des carrières : « *Parallèlement aux carrières, les CV évoluent au fil du temps, en tenant compte de l'évolution des intérêts, des emplois et des collaborations* ». Également, Cañibano et al. (2008, p. 18-19) soulignent la richesse des informations : « *Les CV reflètent à la fois les trajectoires de carrière et les résultats et les caractéristiques de cette carrière* ». Les CV fournissent plusieurs types d'informations :

i) Tout d'abord, ils permettent de retracer précisément la **trajectoire professionnelle et géographique d'un chercheur**. Ainsi, il est possible de coder la nature, le lieu et la durée de chaque poste occupé depuis le doctorat.

ii) Ensuite, ils fournissent la **liste des publications scientifiques**, ce qui permet de consolider les données de la base de données bibliométrique (issue de l'extraction des publications sur le Web of Science).

iii) Enfin, ils donnent accès à des **données relationnelles très riches** dans la qualification de la nature des relations qu'ego peut entretenir avec ses alter (participation à des jurys de thèse et d'HDR, encadrement de doctorat et post-doctorat, participation à des projets de recherche collectifs, responsabilités administratives, séjours dans des laboratoires...).

Ces informations peuvent quasiment être considérées comme exhaustives dans la mesure où le CV s'est imposé comme le support normé de gestion des carrières (fourni lors des promotions, des demandes de financement, etc.) et donc actualisé en permanence par son auteur.

3- Les entretiens semi-directifs

En dépit de la richesse de leurs données, la base de données bibliométrique et les CV ne permettent pas d'expliquer les déterminants de la dynamique des réseaux scientifiques au fil de la carrière des chercheurs. Afin de **comprendre les interactions entre carrière et réseau**, il est indispensable de collecter du **matériel narratif**. Le questionnaire n'étant pas adapté pour mener une analyse processuelle des phénomènes sociaux (Pettigrew, 1990 ; Edwards, 2010), nous menons des entretiens semi-directifs avec les protagonistes, c'est-à-dire les chercheurs dont nous étudions la trajectoire. Les **entretiens**

permettent de comprendre l'enchaînement des différentes séquences de la carrière, le sens de la causalité entre les phénomènes observés et d'intégrer une large diversité de déterminants, sans préjuger de ce qui va être important. L'enquêté parle assez librement de sa carrière et de ses collaborations, en insistant sur ce qui lui paraît structurant, et les relances sont là pour mieux contextualiser les événements abordés et objectiver la nature des relations décrites. Tous les chercheurs enquêtés ont été dans leur carrière directeurs d'équipe ou de laboratoire. Ainsi, la question de la structuration locale et nationale de la chimie. Les entretiens ont duré en moyenne deux heures et ont été réalisés en automne 2017.

Nous proposons donc de développer **une méthodologique originale** afin d'identifier l'impact des événements de carrière sur la dynamique du réseau de co-authorship à une échelle très fine. Nous combinons les échelles d'action et de temps afin de prendre en compte la dimension temporelle des carrières. Aussi, nous proposons une méthode mixte fondée sur la lecture qualitative de résultats quantitatifs. Les données bibliométriques permettent de caractériser la dynamique du co-authorship étudié. L'appariement avec deux autres sources données – les CV et les entretiens – permet de redonner une épaisseur sociale aux résultats froidement livrés par l'analyse statistique des publications.

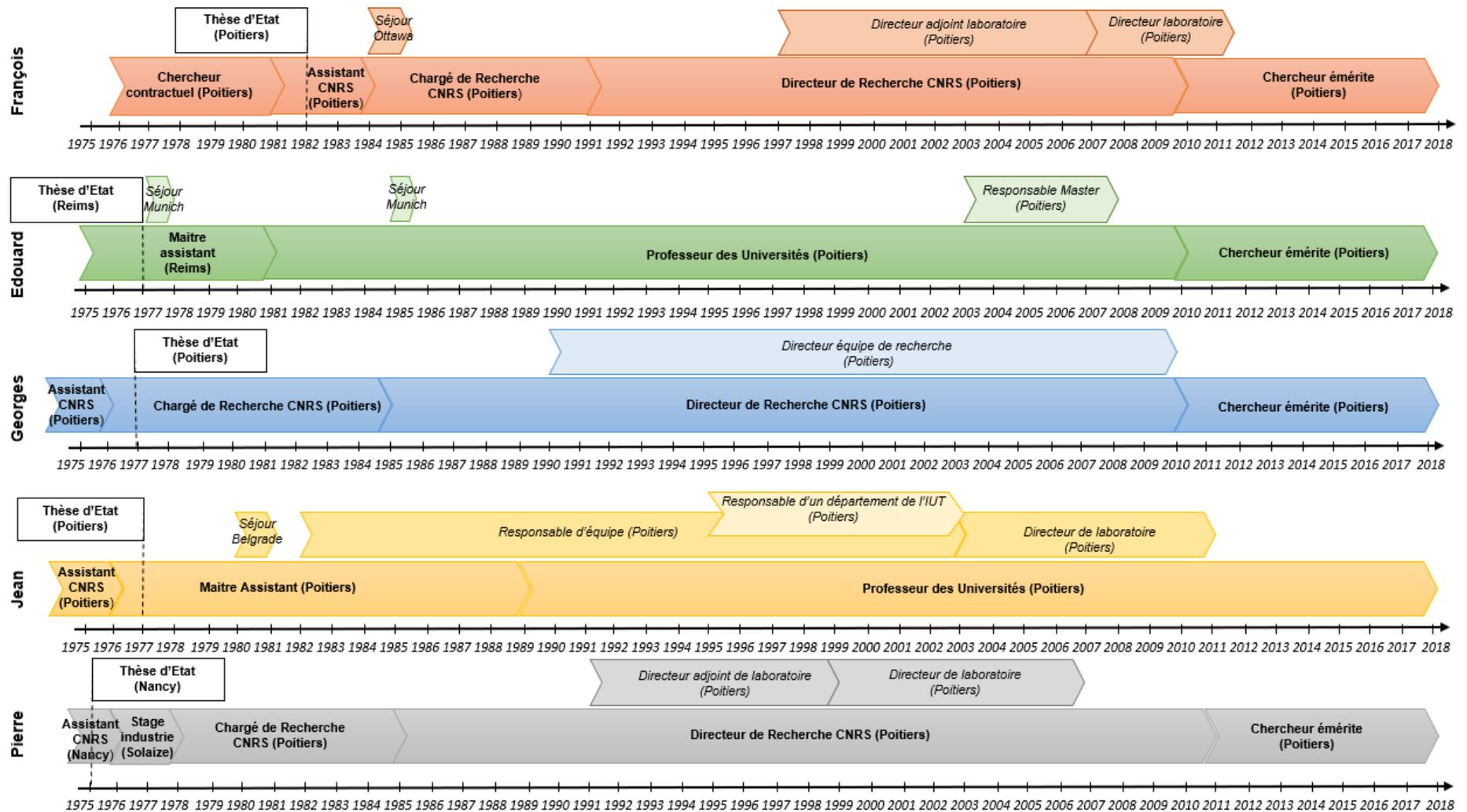
II - Le cas de cinq chimistes émérites de l'Université de Poitiers

A- Constitution de l'échantillon

Nous centrons notre étude sur une petite population de chercheurs puisque nous sommes dans une logique d'étude exploratoire. Nous sélectionnons **cinq chimistes émérites de l'Université de Poitiers**. Ce critère permet d'assurer une profondeur historique des carrières étudiées : le fait que les chimistes étudiés soient devenus Professeur des Universités ou Directeur de Recherche au CNRS assure une activité de publication — et donc relationnelle — significative. Aussi, ce critère sous-entend une homogénéité des carrières dans la mesure où ils ont fini leur carrière dans le même laboratoire de chimie (dans des spécialités de la chimie plus ou moins proches). Ils sont tous nés entre 1945 et 1950 et ils ont « grandi » scientifiquement dans le même contexte scientifique national. **L'homogénéité générationnelle** nous semble importante afin de conserver le même contexte d'émergence des réseaux scientifiques. Si les résultats obtenus n'ont pas vocation à être généralisés, l'étude détaillée de quelques cas de chercheurs prolifiques permet de dégager des déterminants récurrents de la dynamique des réseaux scientifiques tout au long de leur carrière, ou à l'inverse des événements singuliers. La figure 28 donne une représentation simplifiée des CV et permet de comparer le déroulement des carrières des cinq chercheurs⁴⁶, en datant les principales promotions et prises de responsabilité.

⁴⁶ Les prénoms choisis ne correspondent pas à l'identité réelle des chercheurs.

Figure 28 : Trajectoire professionnelle des chercheurs



B) Statistiques descriptives de l'activité de publication et de collaboration

Notre objectif est donc de réaliser une « mise en intrigue » à partir d'un petit groupe de chercheurs. Dans *Temps et Récit* publié en 1983, le philosophe Paul Ricœur définit la « mise en intrigue » comme la configuration de l'expérience temporelle à travers l'acte narratif. C'est **de l'effort de narration de la dimension temporelle que provient le récit**. L'intrigue constitue l'unité dynamique et la mise en intrigue l'opération structurante de l'œuvre.

Ainsi, les données quantitatives des tableaux et figures introduits dans cette section seront mises en perspective à travers les enseignements des narrations. En d'autres termes, ces données n'ont pas vocation à être détaillées dans cette section mais plutôt à **servir de support dans l'interprétation des récits**. La dynamique relationnelle observée à travers la publication sera « racontée » par le protagoniste, dans le cadre d'entretiens semi-directifs biographiques, à la manière d'un « récit de carrière ».

1- Statistiques descriptives statistiques

Le tableau 26 présente quelques statistiques descriptives sur la carrière des chercheurs, leur activité de publication et la dimension collaborative de cette activité scientifique. Il s'agit à ce stade d'une analyse statique de leur activité de publication et de collaboration, abordée comme un stock, une fois la carrière terminée.

Tableau 25 : Statistiques descriptives sur la carrière et le co-authorship

	Pierre	Jean	Georges	Edouard	François
Année de la première publication	1976	1975	1975	1975	1979
Nombre de publications	261	90	214	113	215
Nombre moyen de publications par an	6.37	2.14	5.10	2.69	5.66
Nombre de cosignatures	975	328	781	418	841
Nombre de co-auteurs	300	122	277	182	192
Nombre moyen de co-auteurs par publication	3.74	3.64	3.65	3.70	3.91
Nombre moyen de nouveaux co-auteurs par publication	1.15	1.36	1.29	1.61	0.89
Co-auteurs « one-shot »	142 (47.3%)	72 (59.0%)	155 (56.0%)	112 (61.5%)	88 (45.8%)

On voit une **différence assez nette concernant le volume de publication** entre les chercheurs CNRS (Pierre, Georges et François) et les enseignants-chercheurs (Jean et Édouard) : « *on ne joue pas dans la même cour* » décrit Édouard. En effet, les chercheurs CNRS ont la possibilité de consacrer significativement plus de temps à leur mission de recherche que les enseignants-chercheurs. Le nombre total de publications sur la carrière est environ deux fois plus élevé (261, 215 et 214) pour les chercheurs CNRS que pour les universitaires (113 et 90). Les **volumes de collaboration** connaissent mécaniquement les mêmes écarts dans la mesure où le nombre moyen de coauteurs par publication est

sensiblement le même pour les cinq chimistes (environ 3,7). Les deux dernières lignes du tableau 26 donnent quelques premiers éléments sur le **renouvellement du co-authorship et le poids des relations éphémères** à l'échelle de la carrière. En moyenne pour chaque publication, environ un co-auteur sur trois est nouveau. Cette part est relativement plus faible pour François et plus forte pour Édouard, indiquant des différences dans le degré de renouvellement du réseau scientifique au fil de la carrière. Enfin, environ la moitié des co-auteurs des cinq chimistes étudiés correspondent à ce qu'on appelle des co-auteurs « one-shot », c'est-à-dire des personnes qui n'apparaissent qu'une seule fois dans les publications (Cabanac *et al.*, 2014), indiquant le caractère très éphémère de certaines relations.

2- Statistiques descriptives dynamiques

Dans la perspective d'une analyse dynamique des réseaux scientifiques, nous construisons des périodes afin de mieux saisir le renouvellement du réseau (Figure 29) et la géographie du co-authorship (Figure 30). Nous choisissons de **construire des périodes de cinq ans**, à l'exception de la première et de la dernière période.

i) Les cinq chimistes étudiés ont soutenu leur thèse d'État à la même période (un en 1975, un en 1976, deux en 1977 et un en 1982). Les premières années de la carrière sont marquées par une activité de publication relativement peu soutenue, ce qui justifie le choix de faire une première période plus longue, qui fait référence à toutes les publications parues avant 1984 (année incluse).

ii) Pour tous les chercheurs, les périodes suivantes sont de cinq années : 1985-1989, 1990-1994, 1995-1999, 2000-2004, 2005-2009.

iii) Enfin, la dernière période fait référence aux publications parues depuis 2010. Les chercheurs interrogés ont obtenu leur statut d'éméritat entre 2010 et 2012 : l'activité de publication s'étant ralentie pour la plupart (excepté pour Pierre), nous avons fait le choix de rassembler les publications post-éméritat dans une même période, qui marque la sortie de carrière.

Les écarts de volume de publications entre les cinq chimistes sont évacués dans la construction des figures 29 et 30, en exprimant le renouvellement et la géographie du co-authorship en moyenne par publication et par période. Ce choix est d'autant plus naturel que notre intention n'est pas normative mais plutôt compréhensive et exploratoire : nous nous intéressons à l'évolution du réseau et non à la performance scientifique.

Avec le support de ces matériaux quantitatifs, les entretiens ont été réalisés auprès des 5 chercheurs. **Plusieurs périodes ont été décrites, tant via une dimension contextuelle (contexte historique de la recherche en chimie) que structurelle (stade de carrière).**

Figure 29 : Renouvellement du co-auteursip

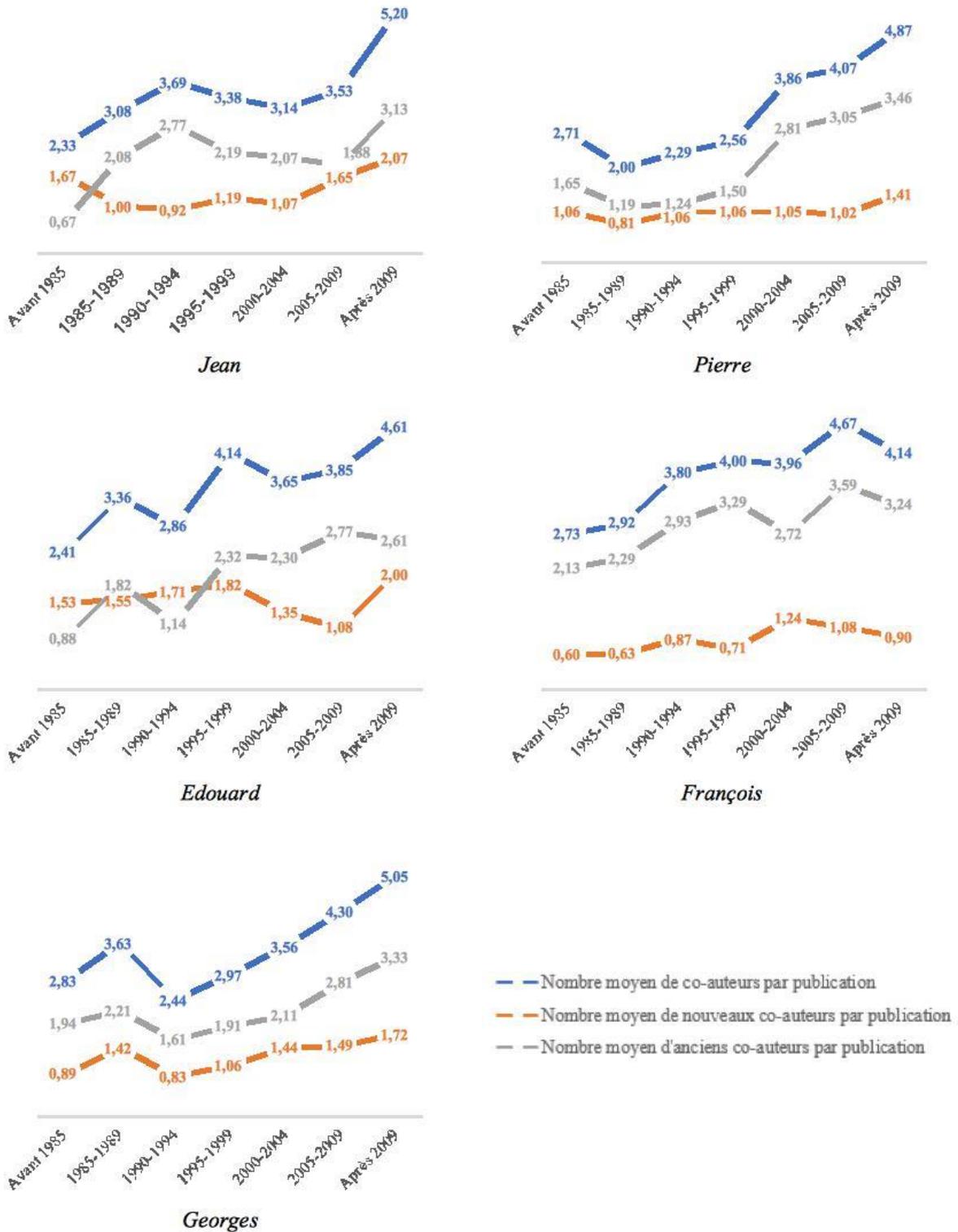
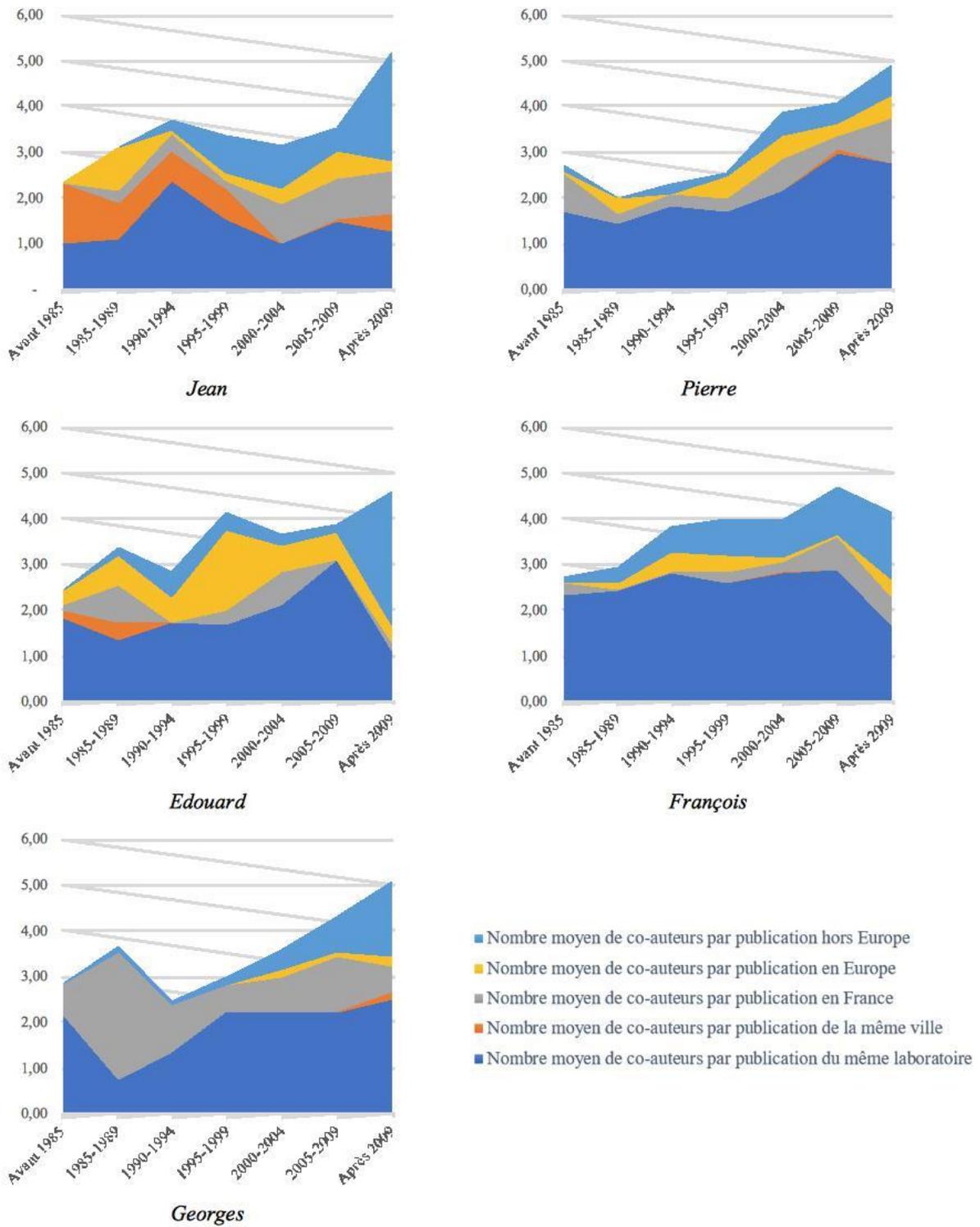


Figure 30 : Evolution de la géographie du co-authorship



III- Résultats : contextualisation des relations collaboratives

A partir des récits des cinq chercheurs, nous identifions trois **contextes scientifiques**. Le premier concerne les années 1960-1970 : il s'agit de l' « *âge d'or* » de la recherche et des premières années d'activité des chercheurs. Le second contexte concerne les années 1980-1990 : c'est un contexte d'autonomisation scientifique, multiplication des collaborations et de prises de responsabilités. Enfin, à partir des 2000, nous identifions une phase de rationalisation et de fusion des laboratoires au cours des dernières années d'activité des chercheurs.

A- Les années 1960-1970 : L'« âge d'or » de la recherche et l'insertion dans les réseaux scientifiques des « patrons »

1- L'émergence de la chimie à Poitiers

Les cinq chercheurs commencent leur carrière académique dans un **contexte d'aménagement scientifique du territoire**. En effet, les années 1960-1970 constituent selon leurs propres termes l' « *âge d'or* » de la recherche. À Poitiers, comme dans de nombreuses autres villes, le campus émergeait, ainsi que les équipes de recherche et les formations : « *il fallait tout créer* » (Jean), « *on a démarré l'équipe à trois, j'étais le seul étudiant* » (François). Les dotations des laboratoires étaient confortables et les conditions de recrutement favorables : « *honnêtement, ce n'était pas compliqué de trouver un poste à cette période* » (Jean). Les cinq chimistes étudiés ont ainsi été recrutés dans un environnement peu concurrentiel.

Si les cinq chercheurs étudiés ont fini leur carrière dans le même laboratoire (Institut de Chimie des Milieux et Matériaux de Poitiers), ils ont auparavant évolué dans des laboratoires différents, qui ont aussi connu de multiples restructurations. Quand nos cinq chimistes prennent leur fonction à Poitiers, ils sont respectivement dans des équipes de catalyse (Pierre et Georges), de chimie organique (Jean), de chimie minérale (Édouard) et d'électrochimie (François), indépendantes les unes des autres. La structuration de ces équipes de recherche s'est faite à l'initiative de personnalités très encadrées dans les instances nationales (CNRS et Ministère) et qui constituaient des figures emblématiques du domaine. Il s'agissait notamment d'individus non originaires de Poitiers qui étaient venus s'implanter avec le soutien du CNRS, « *important avec eux leurs sujets de recherche* » et qui participaient activement au développement de formations (création de l'IUT⁴⁷ en 1966) et d'équipes de recherche. L'un de ces fondateurs de la chimie à Poitiers, qui venait de Lille, connaissait parfaitement « *la carte de la chimie*

⁴⁷ Institut Universitaire de Technologie

en France » du fait de ses fonctions nationales. Il avait des relations avec de nombreux chercheurs à l'époque. Pierre raconte que cette figure « *a fait venir* » des chercheurs d'autres villes dans des conditions avantageuses, citant le cas d'un collègue recruté à Poitiers avec sa conjointe et son assistant. Jean, Pierre, Georges et François qualifient ces figures de « *patrons* » : ils se positionnent comme les chevilles ouvrières de leurs ambitions scientifiques.

2- Une carte française de la chimie très lisible : Poitiers bien encadrée dans les réseaux

A l'époque, **les relations interpersonnelles entre les quelques grandes figures de la chimie française, s'organisaient autour de sept pôles français** (Caen, Lille, Lyon, Montpellier, Paris, Poitiers, Strasbourg). Ces canaux nationaux ont structuré les réseaux scientifiques des cinq enquêtés durant leur début de carrière : **les collaborations sont encore très peu internationalisées**. La structuration de la chimie en sept pôles et les effets de réputation associés permettent aux chercheurs de « *savoir qui fait quoi en France en catalyse* ». En commentant une publication, Pierre précise que l'intégration d'un de ses co-auteurs est issue d'une ressource particulière affiliée Lyon : « *d'une technique particulière, qui n'existait encore nulle part ailleurs* ». Le co-authorship de la première période est très fortement marqué par **sa dimension locale** puisque 75 % des cosignatures en moyenne se font avec des collègues du laboratoire : les structures de recherche se développent, cherchent à se faire connaître, en travaillant essentiellement de manière autonome. Les quelques collaborations se font avec les « *villes françaises de la chimie* » citées plus haut, essentiellement Paris, Lille et Lyon. Pierre et Édouard, qui viennent respectivement de Nancy et Reims où ils ont soutenu leur thèse d'État, maintiennent des relations quelque temps après leur arrivée à Poitiers. Le réseau de Jean est un cas particulier : dans son début de carrière, les collègues issus d'un autre laboratoire de Poitiers pèsent autant dans le réseau que ceux de son propre laboratoire. Le développement de ses collaborations à l'échelle de sa ville traduit une forte collaboration interdisciplinaire. Entreprise avec ses « *patrons* » et des collègues géologues, Jean structure son réseau scientifique via le développement d'une expertise locale durant une bonne partie de sa carrière.

Pour résumé, deux niveaux de relation collaborative transparaissent dans les réseaux des cinq chercheurs : ils sont structurés autour des « *patrons* » **du laboratoire**. Le premier niveau concerne les **relations interpersonnelles des « patrons »** : les chercheurs développent ces liens dans le cadre d'une communauté scientifique nationale en structuration. Le second niveau concerne les **relations plus institutionnelles** : les chercheurs ont accès à ces relations via l'encastrement de ces figures emblématiques de la chimie dans des instances nationales de la recherche. Par conséquent, l'émergence des collaborations des cinq chimistes est essentiellement issue des orientations prises par leur chef d'équipe, qui sont impliquées quasi systématiquement dans leurs publications.

B) Les années 1980-1990 : Autonomisation scientifique et multiplication des collaborations

1- La « course à la publication »

A partir des années 1980-1990, nous identifions une **intensification du rythme de publication des chercheurs**. Les enquêtés l'associent à une phase de « *professionnalisation* » de la recherche et à un engagement à la productivité de la communauté scientifique sous la forme de compétition académique et de « *course à la publication* ». Cette intensification a généré une **densification progressive du co-authorship**, assez nette dans la figure 29 : Le nombre moyen de coauteurs par publication passe de 2,66 avant 1985 à 3,36 de 1995 à 1999 et à 4,82 après 2009. Cette densification du co-authorship est également soulignée par plusieurs études bibliométriques (Glänzel et Schubert, 2004 ; Newman, 2004).

Les enquêtés qualifient cette période comme les prémices du fameux « *publish or perish* », qui fait de l'activité de **publication dans des revues internationales à comité de lecture** la clé de voûte de l'évaluation des chercheurs (pour accéder à une promotion, décrocher des projets de recherche, etc.). Le poids de ce nouveau critère de performance scientifique dans la carrière va transformer l'activité scientifique des chercheurs.

i) Tout d'abord, la « course à la publication » va reléguer au second plan les autres fonctions du travail académique (enseignement et responsabilités collectives.).

ii) Ensuite, elle va aussi transformer l'activité de collaboration. En commentant ses copublications, un enquêté explique que la présence de certains co-auteurs sur les articles résulte davantage d'un « *arrangement* » que d'un réel investissement scientifique. Cet arrangement vise à permettre aux collègues d'attendre l'objectif de publication nécessaire dans l'évolution de la carrière. Ainsi, la réalité relationnelle et le degré d'interaction entre des paires d'acteurs impliqués dans une même collaboration est parfois très hétérogène (Bernela et Levy, 2017).

2- Construction de l'équipe de recherche en tant qu'animateur d'un collectif de recherche

C'est aux alentours de 1990 que les cinq chercheurs **deviennent Directeur de Recherche au CNRS ou Professeur des Universités**. Ces promotions interviennent souvent simultanément avec la prise de **direction d'équipe ou de laboratoire**. Georges souligne l'importance de l'entourage des bras droits scientifiques dans ce type de poste : « *prendre les rênes d'une équipe, c'est une preuve de maturité et d'envergure scientifique (...). Pour être un bon premier, il faut avoir de bons seconds* ». Il développe : « *on peut tirer une machine mais il faut que les résultats suivent, que les manip soient bien réalisées. On ne peut pas écrire les publis, chercher du financement, gérer les collaborations internationales et veiller au grain tous les jours au labo* ». Ainsi, pour que l'activité scientifique d'un chercheur puisse

monter en puissance et que sa thématique de recherche se voit pérennisée au sein du laboratoire, il est décisif de disposer à un moment dans la carrière de **sa propre équipe**. Via cette prise de responsabilités, le parole du chef d'équipe pèse davantage dans les choix de recrutements. Comme l'indiquent les figures 29 et 30, on observe que cette prise de responsabilité se traduit par un **pic des collaborations locales** chez les cinq chercheurs. Cela traduit un resserrement des activités scientifiques sur celles de l'équipe.

On observe également **la place croissante des étudiants** dans le co-authorship des chercheurs étudiés. En effet, les étudiants sont indispensables pour « *faire tourner* » la recherche dans cette discipline expérimentale : ils sont au cœur du développement continu du réseau de co-authorship des chercheurs. Enfin, compter sur son équipe permet aussi aux chercheurs de ne pas relâcher l'activité de recherche malgré le poids de lourdes responsabilités administratives (direction d'UMR, de département d'IUT, etc.).

Par conséquent, ces périodes de prise de direction de laboratoire ou d'équipe se traduisent par une **accélération du rythme de publication grâce au soutien de l'équipe de recherche**.

3- La mobilité scientifique et à l'internationalisation des réseaux

L'intensification du rythme de publication s'accompagne également **d'une internationalisation des collaborations académiques**. En effet, l'internationalisation du co-authorship est un sujet abordée fréquemment dans les narrations. La part des cosignatures avec des co-auteurs internationaux augmente fortement à partir de 1995. C'est notamment le cas d'Edouard, dont 50% des cosignatures sont internationales. Dans le cas de ce dernier, il s'agit essentiellement de nouvelles collaborations académiques et industrielles, financées par les programmes-cadres européens de la recherche. Cette dynamique porte toute l'équipe et est impulsée par les besoins d'industriels de l'aérospatial : « *confrontés à des problèmes techniques liés à la propulsion, ils sont venus nous chercher en 1994 pour notre expertise en catalyse* ». S'ensuit alors un investissement scientifique sur un sujet original, liant catalyse et propulsion, qui va internationaliser les relations d'Edouard et de son équipe, notamment vers l'Europe de l'Est.

Au-delà de ce cas particulier, le déterminant majeur de l'internationalisation du co-authorship est le rôle de **la mobilité temporaire et de l'accueil de chercheurs à Poitiers**. La mobilité entrante apparaît structurante dans la construction de leur réseau scientifique et résulte en partie de l'attractivité du pôle de chimie de Poitiers. En effet, la présence de ressources matérielles et immatérielles spécifiques attire les chercheurs localisés dans des laboratoires moins bien dotés, notamment de l'Europe de l'Est, du Maghreb et de l'Amérique du Sud. Edouard souligne : « *Les années 1990 étaient une période faste en termes de supports de chercheurs invités, qui pouvaient venir jusqu'à six ou neuf mois* ». François et

Jean confirment avoir également mobilisé plusieurs fois ce dispositif. La genèse des collaborations scientifiques entre deux équipes suit assez souvent le schéma suivant :

i) Les chefs d'équipe se rencontrent pendant les congrès, où de premières discussions permettent d'envisager un travail en commun.

ii) Si cette première collaboration est fructueuse, un membre de l'équipe partenaire est invité à Poitiers pour approfondir les discussions

iii) Une fois la collaboration pérenne, les doctorants et post-doctorants réalisent des séjours dans les laboratoires.

Édouard et François parlent de « *filières* » qui se mettent en place au fil du temps, permettant souvent un découplage⁴⁸ : « *avec Budapest, ce sont des dizaines d'étudiants qui ont été tantôt envoyés tantôt reçus pour travailler sur leur thèse chez le partenaire* ». Ces relations internationales sont maintenues dans le temps grâce aux mobilités d'étudiants et traversent les décennies : « *on a formé ce jeune brésilien, qui est venu faire une partie de sa thèse à Poitiers, et c'est maintenant lui, aujourd'hui à la tête de son équipe chez lui, qui nous envoie ses propres étudiants* » (François). Parfois, ces collaborations s'expliquent également par des déterminants plus personnels. Par exemple, après avoir passé son enfance au Maroc, Jean s'est intéressé aux sédiments de ce pays et a noué des collaborations structurantes avec des laboratoires marocains, dans lesquels il aime se rendre à l'occasion, en famille.

Malgré une faible mobilité géographique des chercheurs étudiés, certains ont réalisé des **séjours dans des laboratoires à l'étranger**. Jean a par exemple fait un séjour en Yougoslavie en 1980 pour rencontrer un chercheur spécialiste des sédiments anciens. Cette mobilité a eu une répercussion importante et durable au sein de son réseau scientifique. Entre 1985 et 1989, le poids des collaborations avec des co-auteurs yougoslaves est particulièrement élevé (20,7 %), et renvoie à des reconductions de partenariats avec les chercheurs du laboratoire de Belgrade. Jean a ensuite étendu ses partenariats à un second cercle de co-auteurs issus de ce laboratoire, grâce à des médiations interpersonnelles, et a gardé « l'habitude » de recevoir des étudiants serbes en séjour de recherche à Poitiers.

Les chercheurs indiquent que la conception de carrière comme un processus local est encline à se raréfier dans les années à venir. L'Europe mène une politique scientifique interventionniste pour gagner en compétitivité dans l'économie de la connaissance sous l'hypothèse que la mobilité géographique des chercheurs favorise une meilleure circulation des connaissances (Mahroum, 2000 ; Ackers, 2005). Ils n'ont pour leur part pas connu d'« *injonction à la mobilité* », ce qui ne les a pas empêchés de mener des

⁴⁸ Conformément à la définition de Grossetti *et al.* (2006), alors que l'encastrement correspond à un accroissement de la dépendance d'une entité sociale par rapport à d'autres, le découplage (ou désencastrement) renvoie, au contraire, au processus d'autonomisation d'une entité sociale par rapport à d'autres entités. Grâce à des allers-retours entre les niveaux individuels et organisationnels, il est possible d'observer un processus de découplage des relations personnelles et réciproquement un encastrement dans des dispositifs plus formels.

collaborations internationales. Par conséquent, ce n'est pas via une mobilité intrinsèque que leur co-authorship s'est internationalisé mais via **le positionnement de leur équipe dans le jeu des circulations scientifiques** (invitations, reconductions de collaborations avec des partenaires privilégiés).

C) A partir des années 2000 : rationalisation et fusion des laboratoires

A partir des années 2000, les chercheurs dépeignent un contexte scientifique plutôt maussade, relativement à ce qu'ils ont pu connaître dans leur carrière. En effet, le contexte de recherche a changé et s'oriente vers la **recherche de visibilité des laboratoires, la rationalisation et la réduction des moyens** (via des économies d'échelles). A Poitiers, cela s'est traduit par une restructuration locale de la chimie en phase avec le mouvement de fusions des structures de l'enseignement supérieur et de la recherche observé nationalement. Les enquêtés, qui ont succédé à la génération des « patrons créateurs » des premiers laboratoires, se retrouvent impliqués dans le dossier de fusion de quatre laboratoires de chimie et de géosciences, lequel entraîne des tensions humaines fortes. Une nouvelle UMR CNRS est ainsi créée au 1er janvier 2012, regroupant trois unités mixtes : les onze équipes de recherche émanant des quatre UMR d'origine ont été restructurées en cinq équipes.

Cette restructuration n'est pas sans effet sur les réseaux de co-authorship de fin de carrière des chercheurs étudiés. Ceux qui n'adhèrent pas à ce projet de fusion tendent à **prendre de la distance avec la stratégie du laboratoire naissant**. De plus, les thématiques de recherche de certains émérites sont faiblement représentées dans la nouvelle architecture alors que se positionnent comme le cœur de nouvelles équipes. L'activité de publication post-éméritat va donc être **plus ou moins liée aux activités locales**, selon la relation qui est maintenue avec les ex-collègues. Un des chercheurs quitte la structure plutôt en mauvais terme, ne gardant que des relations personnelles avec des chercheurs qu'il estime. Le point commun des cinq chimistes en fin de carrière est une très **forte internationalisation du co-authorship**, qui consiste essentiellement en des reconductions de collaborations, avec des personnes souvent devenues des amis.

IV-Conclusion

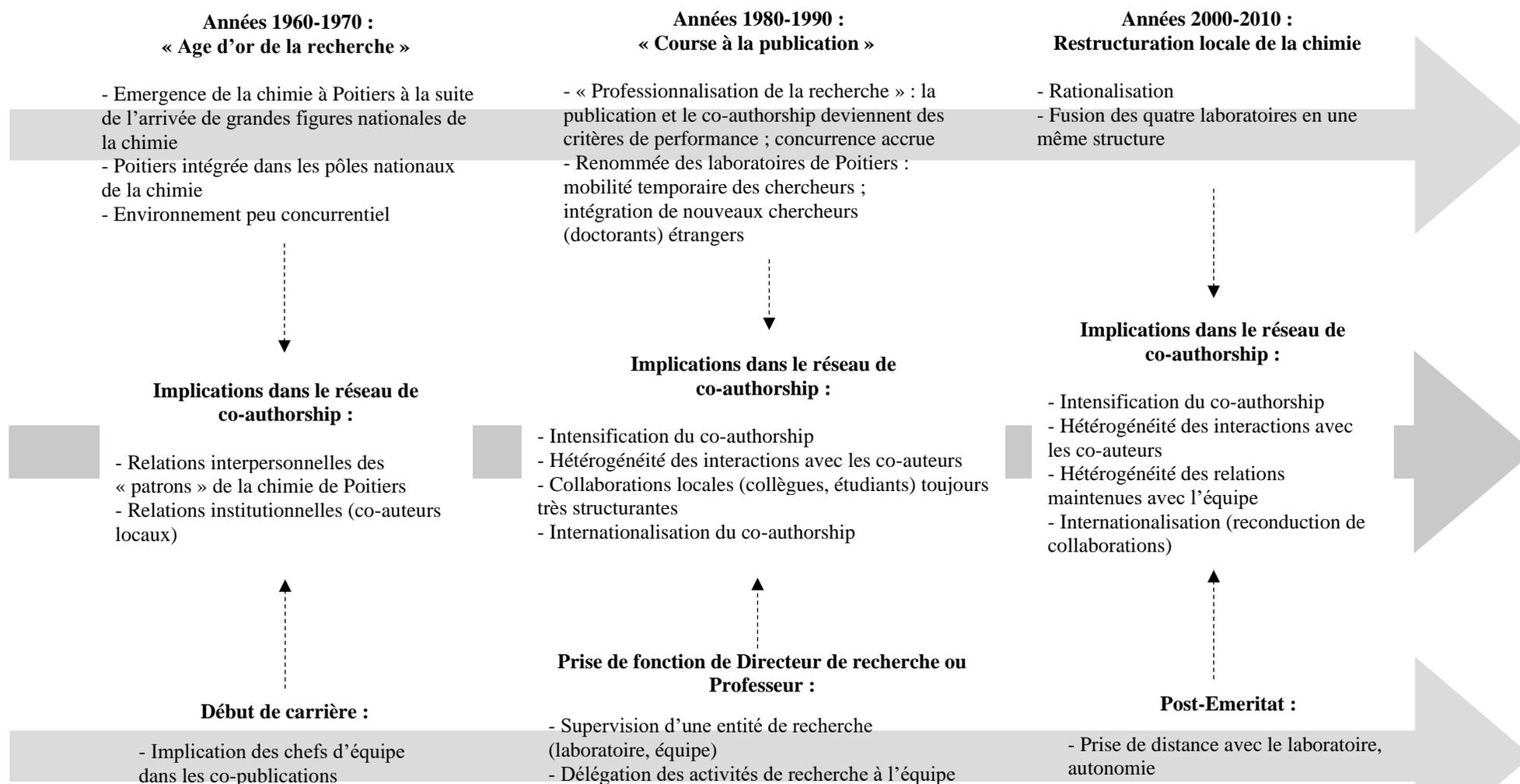
Nous avons proposé ici une **analyse dynamique des réseaux scientifiques de chercheurs, à l'échelle des carrières**. Une méthodologie cohérente avec cet objectif consiste à apparier des données longitudinales sur le réseau scientifique des chercheurs (co-authorship sur quarante ans à partir du Web of Science) avec des données qualitatives permettant de comprendre la dynamique de ce réseau au fil de la carrière. Cette méthodologie est opérationnalisée sur le cas des cinq chimistes émérites de l'Université de Poitiers.

Empiriquement, on observe une relative **homogénéité dans le séquençage des carrières**, avec un co-authorship **d'abord très local qui s'internationalise tardivement**, et une **autonomisation scientifique** suivie d'une prise de responsabilité, qui assure une forme de centralité dans le collectif. Une telle homogénéité s'explique, au moins en partie, par le **rôle particulièrement structurant du contexte**, notamment disciplinaire. En observant des chercheurs d'une même génération, d'une même discipline et d'un même territoire, nous avons non seulement pu apprécier **l'ampleur de l'impact du contexte** mais également saisir plus nettement les éléments différenciant situés à d'autres échelles. Nous avons ainsi observé une différence nette d'intensité des publications entre les chercheurs selon leur appartenance au CNRS ou à l'Université, selon leur orientation thématique, selon leurs liens avec l'industrie, influençant nécessairement les opportunités de financement et les dynamiques de réseau.

Ces résultats, au-delà de leur contribution empirique à l'analyse des dynamiques de réseaux scientifiques, révèlent également l'importance méthodologique **d'une démarche processuelle intégrant différentes échelles d'analyse**. Les narrations ont révélé en effet l'importance de l'enchevêtrement entre les niveaux individuel et collectif dans l'évolution de carrière des chercheurs, des processus de découplage dans la pérennisation des réseaux scientifiques des laboratoires, etc. Le schéma 31 synthétise dans quelle mesure l'articulation entre le contexte historique de recherche et le stade de chercheur impactent les dynamiques collaboratives pour chaque période.

Bien que cette recherche n'ait qu'une visée exploratoire, la fiabilité de nos résultats est assurée par la corroboration des résultats par les sujets interviewés et la triangulation assurée par la confrontation de trois sources de données. Les résultats issus de l'analyse des publications ainsi que le CV sont utilisés pendant l'entretien, sous les yeux du chercheur interrogé, afin de baliser la subjectivité des sujets. Il s'agit là clairement d'une forme de triangulation, reposant sur l'appariement de divers matériaux, qui permet de contrôler partiellement le discours de l'enquêté.

Figure 31 : Dans quelle mesure le contexte historique de recherche et le stade de carrière impactent les dynamiques collaboratives ?



Chapitre 4 : Est-ce que la mobilité géographique façonne le réseau de co-authorship ? Une approche de panel des carrières scientifiques

Ce chapitre se situe dans le prolongement des statistiques descriptives du chapitre 2 relatives aux dynamiques sociales et géographiques de co-authorship des chimistes au fil de la carrière. Pour rappel, nous avons montré que quelle que soit la trajectoire individuelle du chercheur, **les dynamiques de co-authorship étaient relativement stables au fil de la carrière**. En effet, bien que la dimension collaborative des articles s'accroisse, le nombre moyen de nouveaux co-auteurs (ouverture sociale du réseau) est stable et le nombre moyen d'anciennes relations par article (maintien de relations) augmente au fil de la carrière. De plus, le poids des collègues du laboratoire (co-auteurs locaux) dans l'activité de publication est central et stable au fil de la carrière. La densification du co-authorship se traduit pas l'internationalisation des relations collaboratives.

Dans ce chapitre, nous souhaiterions **comprendre dans quelle mesure l'événement de la mobilité géographique bouleverse ces dynamiques sociales et géographiques de co-authorship**. En d'autres termes, nous souhaiterions comprendre si le changement d'affiliation du chercheur favorise ou restreint la dimension sociale et géographique de son réseau de co-authorship.

Nous commencerons par présenter les motivations de cette étude relative à l'intégration de la mobilité dans la trajectoire du chercheur. La **contextualisation de l'événement de mobilité** dans la carrière soulève de nombreux enjeux scientifiques et nécessite plusieurs dispositions méthodologiques qui ont permis de soulever quelques résultats préliminaires. Dans un second temps, nous présenterons la **stratégie économétrique** suivie via la construction de modèles de panel visant à estimer les créations et les ruptures de relations collaboratives à la suite de la mobilité géographique. Nous présenterons ensuite les **résultats obtenus** ainsi que les implications politiques qu'ils soulèvent.

I- Intégration de la mobilité géographique dans la trajectoire du chercheur

Cette première partie du chapitre a pour but de décrire les **fondements scientifiques et méthodologiques** de cette étude ainsi que des **résultats préliminaires**. Nous commencerons par rappeler pour quelle raison l'étude de l'impact de la mobilité sur la diversification du réseau soulève de nombreux enjeux scientifiques. Ensuite, nous décrirons dans quelle mesure nous avons choisi

d'identifier la mobilité des chercheurs : des précautions ont été prises pour prendre en compte la richesse de cet événement et pour le réintégrer dans la carrière des chercheurs. Enfin, nous présentons deux résultats préliminaires qui tendent à souligner l'effet restreint de la mobilité scientifique dans la construction du réseau de co-authorship.

A) Enjeux scientifiques de la mobilité et de la diversification du réseau

Dans la littérature, de nombreux travaux étudient la dimension géographique de la science au travers de la distribution spatiale des activités scientifiques (Livingstone, 1995 ; Eckert et Baron, 2013). La mobilité est alors considérée soit comme un mécanisme de **réallocation des scientifiques** entre les régions et les pays, soit comme un levier d'accès aux ressources fixes (Jöns, 2007). Elle est au cœur de multiples programmes politiques européens (Morano-Foadi, 2005) pour favoriser la **diffusion des connaissances individuelles** propres aux scientifiques (Cañibano, 2006). Dans ce contexte, le **scientifique mobile est jugé comme performant** puisqu'il est amené à rencontrer, à collaborer et à partager ses connaissances avec de nouveaux chercheurs. Cette diversification scientifique du réseau de co-authorship est indispensable à la **production et à la circulation des connaissances dans le monde** : la mobilité géographique en serait donc le moteur.

L'extension du réseau peut être considérée en termes de compétences ou de co-auteurs (par exemple, nouveaux partenaires de collaboration) et en termes d'affiliation géographique (nouvelles villes de collaboration). Ainsi, les chercheurs diversifient leur réseau en collaborant avec de nouveaux partenaires (extension sociale) et en collaborant avec des partenaires issus d'une nouvelle localisation (extension spatiale). Cette approche suggère **l'étude de la formation et des dynamiques des réseaux scientifiques personnels et d'identifier la contribution de la mobilité géographique des chercheurs dans cette dynamique**. Les dynamiques géographiques des réseaux scientifiques ont été largement étudiées dans une perspective macro (au niveau des villes et des pays). En effet, de nombreux travaux allant de ceux de Bhagwati (1976) à ceux de Berschi *et al.* (2017) traitent la mobilité via le concept de « brain drain » : la mobilité est traitée via l'analyse des flux migratoires entre les territoires. Dans cette perspective macro, les nœuds du réseau scientifique ne sont pas des individus mais des organisations, des villes et des pays. En parallèle, **très peu d'études portent sur les carrières individuelles dans une perspective micro**. En effet, certains travaux évaluent l'impact de la mobilité sur la performance scientifique (Hoisl, 2007; Lawson et Shibayama, 2015). **Plus marginalement, certains auteurs ont cherché à mettre en lien la mobilité et le réseau scientifique**. Par exemple, Fontes, Videira, and Calapez (2013) ont montré que la mobilité internationale enrichissait le réseau de connaissance des chercheurs à partir d'une études de trajectoires de doctorants. Aussi, Wolley *et al.*, (2008) ont vérifié ce résultat pour la mobilité post-

doctorale des chercheurs asiatiques. Enfin, Melin (2004,2005) a nuancé les bienfaits de la mobilité (« *the dark side of mobility* ») dans une étude sur la mobilité des chercheurs suédois.

Ce chapitre se situe dans la lignée de ces précédents travaux. **Nous souhaitons tester le rôle de la mobilité dans la construction et le développement du réseau scientifique à l'échelle micro.** Pour donner suite aux résultats issus du chapitre 3 relatifs à la nécessité de contextualiser la carrière scientifique, nous considérons que la mobilité ne doit pas être étudiée de manière isolée dans la carrière : il nous semble crucial de réintégrer l'impact de la mobilité dans le processus de la carrière des chercheurs pour évaluer ses effets. Ainsi, nous souhaitons interroger l'impact de la mobilité dans une perspective de dynamique de carrière via une méthodologie appropriée.

B) Evaluation de la mobilité dans notre base de données

A partir de l'information spatiale issue de la base de données bibliométrique, l'impact de mobilité sur le réseau scientifique des chercheurs est évalué via :

i) La **trajectoire géographique** des 80 chercheurs

ii) La **distribution spatiale du réseau de co-authorship** à partir de l'information spatiale issue de la base de données bibliométrique.

Nous considérons la mobilité comme un **changement d'affiliation** d'*ego* d'une année à l'autre à l'échelle de la ville ⁴⁹. Nous identifions le volume d'évènements de mobilité parmi les carrières de chercheurs et nous distinguons les mobilités nationales et internationales. Dans la figure 32(a), nous constatons que 17 des 80 chercheurs n'ont jamais eu d'expérience de mobilité alors que les autres ont été mobiles au fil de leur carrière. Ce chiffre est supérieur à ce qui a pu être observé dans d'autres disciplines en France (Bernela, 2018).

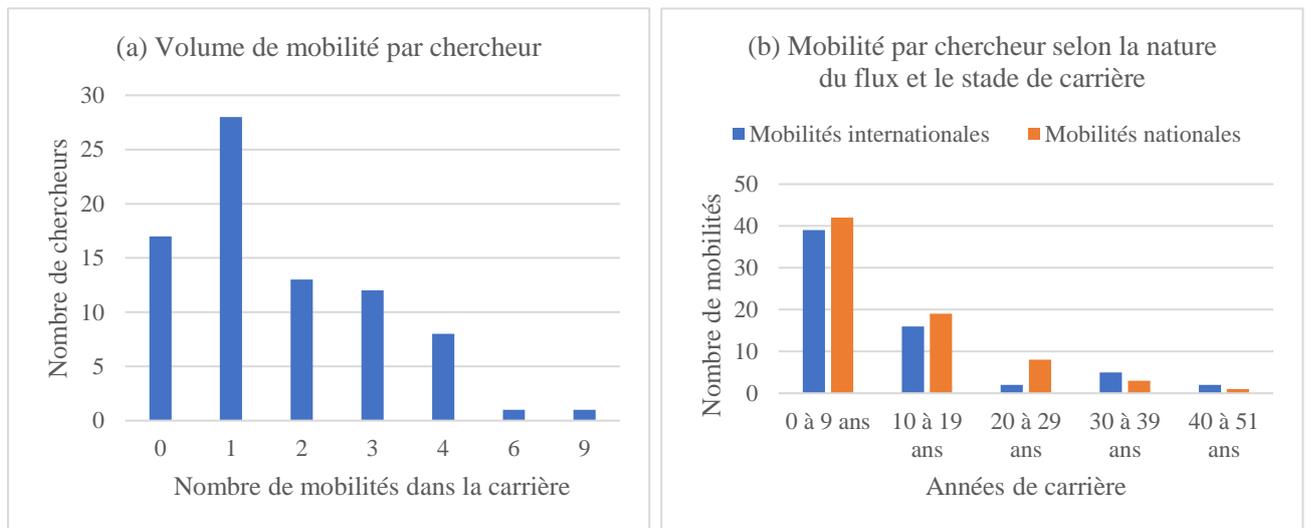
Beaucoup d'études utilisent une **caractérisation binaire de la mobilité** et de la non-mobilité (Aksnes *et al.* 2013 ; Chinchilla-Rodriguez *et al.*, 2018 ; Franzoni *et al.*, 2014). Cette conception de la mobilité **n'est pas adéquate avec la diversité des profils de mobilité**. Ainsi, capter la complexité de la mobilité au travers d'une méthodologie appropriée et de données bibliométriques est une réelle opportunité scientifique (Robinson-Garcia *et al.*, 2019).

La figure 32(b) montre qu'il y a autant de mobilités internationales que nationales dans la base de données. La mobilité apparaît principalement au **début de la carrière** : cela souligne le rôle central de la mobilité post-doctorale en chimie. Puisque même les chercheurs mobiles ont été sédentaires avant leur première

⁴⁹ Nous n'observons pas la mobilité institutionnelle au sein d'une même ville, à l'exception des changements de noms des laboratoires à Poitiers avant la fusion

mobilité, il nous semble nécessaire de contextualiser chaque évènement de mobilité au sein des carrières pour évaluer le plus justement possible l'impact de la mobilité sur l'extension du réseau de co-authorship. De plus, nous identifions et notons l'ensemble des villes fréquentées par ego dans le passé : nous sommes donc en mesure **de contextualiser chaque publication (et chaque pair de co-authorship) dans la trajectoire géographique du chercheur**. D'un point de vue dynamique, nous évaluons la continuité des relations avec chaque co-auteurs affiliés à une ville qu'ego a fréquentée auparavant.

Figure 32 : Volume des évènements de mobilité par chercheur et durant la carrière



C) Premiers résultats : la mobilité ne semble pas nécessaire à la diversification du réseau

Le tableau 27 apporte des premiers résultats relatifs l'impact de la mobilité sur l'extension du réseau en réintroduisant chaque relation collaboration *ego-alter* dans la trajectoire géographique d'*ego*. Nous adoptons deux approches de la relation collaborative : le **moment d'intégration des nouveaux co-auteurs** dans le réseau (ie les nouvelles co-signatures) et l'**ensemble des collaborations réalisés à la suite de cette intégration** (ie, l'ensemble des co-signatures). Ces deux approches sont appréhendées **en fonction de la trajectoire géographique d'ego**. Ainsi, chaque co-signature est appréhendée en fonction de la localisation d'*alter* : est-il affilié à la ville actuelle d'*ego*, à une ville qu'*ego* a fréquentée auparavant ou à une ville qu'*ego* n'a jamais fréquenté ?

Deux premiers résultats émergent de ce tableau :

i) Le premier résultat est que le **localisme des relations collaboratives empêche tout effet premium de la mobilité**. En effet, le laboratoire est une grande source de co-authorship : 32,46% des alter ont intégré le réseau d'ego alors qu'ils étaient affiliés à son laboratoire. Aussi, 48,86% des co-signatures totales ont lieu avec un collègue du laboratoire : les relations collaboratives avec les collègues sont donc maintenues et contribuent hautement à la construction du réseau. En dépit de ce localisme, les collaborations ont **également une dimension internationale** : 33,97% des alter qui ont intégré le réseau d'ego étaient affiliés à un autre pays qu'ego au moment de leur première collaboration. En revanche, **ces collaborations sont courtes et contribuent peu à la construction du réseau** : seulement un quart des cosignatures totales sont faites avec un co-auteur international.

ii) Le second résultat est que **la mobilité ne semble pas être un prérequis pour l'ouverture du réseau de co-authorship**. Ce tableau révèle que seulement 3,14% des alter sont affiliés à une ville qu'ego a fréquenté dans le passé. Cela signifie qu'ego ne maintient que très peu ses relations collaboratives avec ses villes fréquentées à la suite de la mobilité : les relations personnelles sont difficiles à maintenir à la suite d'une mobilité géographique. Les villes qu'ego n'a pas fréquentées sont très structurantes dans la construction du réseau de co-authorship. En effet, 53,74% des alter ont intégré le réseau alors qu'ils étaient affiliés à une ville qu'ego n'avait jamais fréquenté et 37,89% des collaborations sont issues de ce type de co-authorship.

Nous venons de montrer l'opportunité scientifique que représente l'évaluation de l'impact de la mobilité sur l'extension du réseau avec des données micro. Pour ce faire, nous adoptons des précautions méthodologiques afin d'identifier la richesse de l'évènement de mobilité et de le recontextualiser dans la carrière. Des premiers résultats préalables soulignent l'effet restreint de la mobilité sur l'extension du réseau, du fait du localisme des relations et qu'une grande partie des co-auteurs se trouvent en dehors de la trajectoire géographique d'ego.

Tableau 26 : Géographie des relations *ego-alter*

		Distance géographique entre <i>ego</i> et <i>alter</i> (% des cosignatures)						Localisation d' <i>alter</i> par rapport à la trajectoire géographique d' <i>ego</i> (% des cosignatures)				
		<i>Même laboratoire</i>	<i>Même ville</i>	<i>Même pays</i>	<i>International</i>	<i>Non trouvé</i>	Total	<i>Ville actuelle</i>	<i>Ville fréquentée auparavant</i>	<i>Ville non fréquentée</i>	<i>Non trouvé</i>	Total
Co-signatures avec des nouveaux co-auteurs	n = 14604	32,46%	9,97%	22,92%	33,97%	0,68%	100,00%	42,43%	3,14%	53,74%	0,68%	100,00%
Toutes les co-signatures	n = 41720	49,86%	8,69%	16,09%	25,10%	0,27%	100,00%	58,55%	3,29%	37,89%	0,27%	100,00%

II- Présentation du modèle économétrique

Pour capter l'impact de la mobilité sur la diversification du réseau, **l'approche de panel** semble être la méthode la plus appropriée dans une étude des trajectoires individuelles basée sur des données longitudinales. Bien que certaines études aient porté sur la capacité de la mobilité à nourrir le réseau, nous contribuons à la littérature en estimant **l'effet de la mobilité sur la « naissance » et la « mort » des relations collaboratives au sein du réseau d'ego**. En d'autres termes, nous souhaitons vérifier si la mobilité géographique se traduit par la création ou la rupture des relations collaboratives et dans quelle mesure.

Nous introduisons une dimension individuelle qui représente chaque ego ($i \in N [0; 80]$) et une dimension temporelle qui représente chaque année de carrière d'ego ($t \in N [0 ;51]$)⁵⁰. Nous adoptons un **modèle à effet fixe** pour deux raisons :

i) Tout d'abord, ce type de modèle est plus approprié à notre base de données : **l'hétérogénéité inobservée** pourrait biaiser les résultats dans un modèle à effet aléatoire. Notre test de Hausman (1978) est significatif : il confirme que l'estimation à effet fixe est significativement différente à l'estimation à effets aléatoires.

ii) Ensuite, cette méthode est plus adéquate à notre problématique : nous souhaitons nous **focaliser sur la variance intra-individuelle et non sur la variance inter-individuelle**. En effet, nous approchons implique d'entrer dans chaque trajectoire de carrière pour tester l'impact de la mobilité géographique et non de les comparer entre elles. Plus précisément, mobiliser un modèle « *within* » nous émancipe de l'hétérogénéité de l'activité de publication et de collaboration entre les individus : chaque variable est centrée et l'impact de la mobilité géographique est estimé relativement à l'activité moyenne de chaque individu (sans comparer le niveau d'activité entre les individus).

Le choix du modèle à effet fixe nous évite d'utiliser les variables individuelles (comme le genre d'ego par exemple) pour expliquer les dynamiques de collaboration. Notre stratégie d'économétrie de panel vise à **favoriser l'étude des dynamiques collaboratives au sein des carrières individuelles plutôt que de comparer les individus entre eux**.

Pour évaluer la diversification du réseau de co-authorship, nous **mesurons les flux annuels de nouveaux alter et d'alter sortants**⁵¹ **dans le réseaux**. Ces deux dimensions complémentaires nous

⁵⁰ La plus longue carrière est de 51 ans (entre l'année de la première publication et celle de la dernière dans la carrière).

⁵¹ Un alter est jugé sortant s'il réalise sa dernière collaboration avec ego. Nous ne comptons pas les « sorties » de réseau associées aux publications faites deux avant la fin de carrière. Ce choix vise à ne pas prendre en compte les « sorties » de réseau artificielles induites par le fait que nos données bibliométriques s'arrêtent en 2018

permettent d'évaluer les potentielles créations et ruptures de relations à la suite de l'expérience de mobilité. Deux choix méthodologiques ont été fait quant à l'appréhension de ces flux :

i) Tout d'abord, **chaque flux** d'entrées dans le réseau (ie, nombre total de nouveaux co-auteurs avec qui *ego* collabore) et de sorties (ie, nombre total de co-auteurs avec qui *ego* collabore pour la dernière fois) sera ramené **sur le nombre total de publications annuelles** par ratio. Sachant que le volume de co-authorship est très corrélé par le volume de publication annuelles, nous souhaitons affiner notre perception des flux, non pas à l'échelle du réseau mais de l'équipe de publication. Ainsi, nous souhaitons évaluer dans quelle mesure l'équipe de publication moyenne annuelle est structurée par l'intégration ou la sortie des co-auteurs.

ii) Aussi, nous choisissons de mesurer ces flux pour **l'ensemble des co-auteurs du réseau** (tous les *alter*) et plus spécifiquement pour **les alter affiliés à une ville non fréquentée par ego**. Nous souhaitons donc évaluer l'impact de la mobilité sur la dynamique relationnelle *ego-alter*, qu'*alter* fasse partie de la trajectoire géographique d'*ego* ou pas. Cette distinction va nous permettre d'identifier **l'effet direct de la mobilité** (concernant les *alter* qui sont intégrés dans la trajectoire géographique d'*ego*) et **l'effet indirect** (concernant les *alter* qui sont exclus de la trajectoire géographique d'*ego*).

Par conséquent, nous notons les variables dépendantes suivantes :

Tableau 27 : Variables dépendantes

$Y_{i,t}^{entrant}$	Le nombre de co-auteurs entrants par publication dans le réseau d' <i>ego</i> <i>i</i> en <i>t</i>
$Y_{i,t}^{entrant_nonfreq}$	Le nombre de co-auteurs entrants affiliés à une ville non fréquentée par <i>ego</i> <i>i</i> par publication dans le réseau d' <i>ego</i> <i>i</i> en <i>t</i>
$Y_{i,t}^{sortant}$	Le nombre total de co-auteurs sortants par publication dans le réseau d' <i>ego</i> <i>i</i> en <i>t</i>
$Y_{i,t}^{sortant_nonfreq}$	Le nombre de co-auteurs sortants affiliés à une ville non fréquentée par <i>ego</i> <i>i</i> par publication dans le réseau d' <i>ego</i> <i>i</i> en <i>t</i>

Nous construisons **deux variables explicatives longitudinales** pour caractériser la mobilité du chercheur au fil de sa carrière :

i) La première variable $X_{i,t}^{mob_nature}$ permet de distinguer **la nature du flux de mobilité** d'*ego* *i* en *t*, à savoir si *ego* n'est pas mobile ou s'il est mobile à l'échelle nationale ou internationale. L'objectif est donc de tester si la mobilité impacte la construction du réseau de co-authorship en fonction de l'échelle de son flux.

ii) La seconde variable $X_{i,t}^{derniere_mob}$ est construite en **gardant en mémoire la trajectoire géographique du chercheur passée**. Pour chaque année *t* d'*ego* *i*, elle exprime quand est arrivé le dernier évènement de mobilité dans la carrière. Elle teste deux effets complémentaires de la mobilité :

L'effet immédiat et l'effet retardé. Nous testons l'hypothèse selon laquelle la mobilité aurait un effet de court terme (à la suite d'une rapide dissipation) ou de long terme sur la construction du réseau de co-authorship.

Par conséquent, nous notons les deux variables explicatives suivantes :

Tableau 28 : Variables explicatives relatives à la mobilité

$X_{i,t}^{mob-nature}$	La nature de la mobilité d'ego i en t relativement à $(t-1)$ (« pas mobile », « mobilité nationale », « mobilité internationale »)
$X_{i,t}^{dernière_mob}$	Le nombre d'années entre t et la dernière mobilité d'ego i

L'équation générale du modèle estimée est la suivante :

$$(Z_{i,t} - \bar{Z}_{i,t}) = \beta^{mob} (X_{i,t}^{mob} - \bar{X}_i^{mob}) + \beta^{collab} (X_{i,t}^{collab} - \bar{X}_i^{collab}) + (\varepsilon_{i,t} - \varepsilon_i)$$

i) Dans cette équation, $Z_{i,t}$ est égale à l'une des quatre variables dépendantes $Y_{i,t}^{entrant}$, $Y_{i,t}^{entrant_nonfreq}$, $Y_{i,t}^{sortant}$, $Y_{i,t}^{sortant_nonfreq}$

ii) $X_{i,t}^{mob}$ correspond à l'une des variables de mobilité $X_{i,t}^{mob-nature}$ ou $X_{i,t}^{dernière_mob}$

iii) $X_{i,t}^{collab}$ est la matrice des variables de contrôle relatives à l'activité de publication et de collaboration de chaque ego i à chaque année t

iv) $\varepsilon_{i,t}$ est le résidu.

Finalement, en combinant les différentes variables explicatives et expliquées, nous estimons **huit modèles**, détaillés dans le tableau 30.

Tableau 29 : Synthèse des modèles estimés

		Variables dépendantes			
		Entrées dans le réseau		Sorties du réseau	
		<i>Tous les alter</i> ($Y_{i,t}^{entrant}$)	<i>Seuls les alter</i> <i>affiliés à une ville</i> <i>non fréquentée par</i> <i>ego</i> ($Y_{i,t}^{entrant_nonfreq}$)	<i>Tous les</i> <i>alter</i> ($Y_{i,t}^{sortant}$)	<i>Seuls les alter</i> <i>affiliés à une ville</i> <i>non fréquentée</i> <i>par ego</i> ($Y_{i,t}^{sortant_nonfreq}$)
Variables explicatives	<i>La nature de la mobilité</i> ($X_{i,t}^{mob-nature}$)	Modèle 1	Modèle 3	Modèle 5	Modèle 7
	<i>L'historique de la mobilité</i> ($X_{i,t}^{dernière_mob}$)	Modèle 2	Modèle 4	Modèle 6	Modèle 8

Le tableau 31 décrit les statistiques descriptives des variables dépendantes, explicatives et de contrôle.

Tableau 30 : Statistiques descriptives des variables dépendantes, explicatives et de contrôle

Variables		N	Mean	Std Error	Min	Max
Entrées et sorties du réseau (variables dépendantes)						
$Y_{i,t}^{\text{entrant}}$	Entrées par publication (tous les <i>alter</i>)	2109 ⁵²	1,737	1,419	0,00	17,00
$Y_{i,t}^{\text{entrant_nonfreq}}$	Entrées par publication (seulement les <i>alter</i> de villes non fréquentées par <i>ego</i>)	2109	0,824	1,092	0,00	10,00
$Y_{i,t}^{\text{sortant}}$	Sorties par publication (tous les <i>alter</i>)	2109	1,710	1,662	0,00	15,00
$Y_{i,t}^{\text{sortant_nonfreq}}$	Sorties par publication (seulement les <i>alter</i> de villes non fréquentées par <i>ego</i>)	2109	0,853	1,208	0,00	11,00
Expérience de mobilité (variables explicatives)						
$X_{i,t}^{\text{mob-nature}}$	<i>Pas_mobile</i>	<i>Ego i</i> est affilié à la même ville entre (<i>t-1</i>) et <i>t</i>	1972	0,935	0	1
	<i>National</i>	<i>Ego i</i> est affilié à une autre ville nationale en <i>t</i> relativement à (<i>t-1</i>)	73	0,035	0	1
	<i>Internat</i>	<i>Ego i</i> est affilié à un autre pays en <i>t</i> relativement à (<i>t-1</i>)	64	0,030	0	1
	<i>Sédentaire</i>	<i>Ego i</i> n'a jamais été mobile jusqu'en <i>t</i>	782	0,371	0	1
$X_{i,t}^{\text{dernière_mob}}$	<i>Cette_année</i>	<i>Ego i</i> est mobile en <i>t</i>	133	0,063	0	1
	<i>1_à_3_ans</i>	<i>Ego i</i> a été mobile entre (<i>t-1</i>) et (<i>t-3</i>)	275	0,130	0	1
	<i>4_à_6_ans</i>	<i>Ego i</i> a été mobile entre (<i>t-4</i>) et (<i>t-6</i>)	216	0,102	0	1
	<i>7_à_9_ans</i>	<i>Ego i</i> a été mobile entre (<i>t-7</i>) et (<i>t-9</i>)	176	0,083	0	1
	<i>Plus_de_10_ans</i>	<i>Ego i</i> a été mobile avant (<i>t-10</i>)	527	0,250	0	1
Dynamiques de publication et de collaboration (variables de contrôle)						
$Cosign_{i,t}$	Moyenne mobile du nombre moyen de cosignatures par articles d' <i>ego i</i> entre (<i>t-1</i>) et (<i>t+1</i>)	2099	4,31	1,66	1	17
$Publis_{i,t}$	Moyenne mobile du nombre annuel de publications d' <i>ego i</i> entre (<i>t-1</i>) et (<i>t+1</i>)	2099	4,27	3,67	0,33	24,67
$Geo_net_{i,t}$	Nombre de villes avec qui <i>ego i</i> a cosigné jusqu'en (<i>t-1</i>)	2099	17,26	18,16	0	116
$Social_net_{i,t}$	Nombre d' <i>alter</i> avec qui <i>ego i</i> a cosigné jusqu'en (<i>t-1</i>)	2099	81,19	93,77	0	673
$Citati_{i,t}$ ⁵³	Nombre de citations qu' <i>ego i</i> reçu jusqu'en (<i>t-1</i>)	2099	7038,19	19274,17	0	267733

⁵² Cela correspond au nombre total d'années de publication observées dans les 80 carrières

⁵³ Cette variable est basée sur l'hypothèse d'une distribution linéaire des citations annuelles, à partir du stock de citations de chaque publication reçu en 2018

III- Résultats

Le tableau 32 décrit les résultats des huit modèles. Indépendamment de l'expérience de mobilité, nous constatons que le nombre moyen de co-signatures par article a un effet significatif et positif sur la création et la rupture de relations collaboratives. En revanche, l'intensité de l'activité de publication a un effet négatif.

Concernant les variables de mobilité, nous constatons toute chose égale par ailleurs quatre résultats précis :

i) La mobilité encourage la création de nouvelles relations : Concernant le modèle 1, la mobilité accroît significativement **l'intégration de nouveaux alter dans le réseau** par rapport aux années de non-mobilité. En effet, quand le chercheur est mobile (que ce soit à l'échelle nationale ou internationale), nous observons que cette expérience a un effet significatif et positif sur l'entrée des alter. Être mobile stimule l'intégration de nouveaux alter par publication en t , et contrairement aux attentes, être mobile internationalement n'a pas de plus grand effet qu'être mobile nationalement.

ii) L'effet de création des relations est instantané : Si nous définissons la mobilité en prenant en compte si l'expérience est ancienne ($t, t-1$) ou plus ancienne ($t-2, t-3, t-4, t-5, \dots$), nous observons dans le modèle 2 que la mobilité a un effet significatif et positive sur l'intégration de nouveaux alter affilié à la ville fréquentée par *ego*. Néanmoins, **la mobilité a un effet de court terme** parce qu'il n'accroît le nombre de nouveaux co-auteurs seulement durant l'année où la mobilité a lieu. Par conséquent, nous n'observons pas d'effet premium de la mobilité comme dans des études précédentes (Jonkers et Tijssen, 2008).

iii) Les relations en dehors du laboratoire ne dépendent pas de la mobilité : Également, si nous regardons les flux d'alter affiliés à une ville qu'*ego* n'a pas fréquentée (modèles 3,4,7 et 8), nous notons que la mobilité n'a pas d'impact sur la création ou la rupture de relation. La naissance et la mort des collaboration à la suite de la mobilité ne concerne que les alter affiliés à une ville fréquentée par *ego*. Ainsi, bien que la mobilité favorise la diversification du réseau, **son impact géographique est relativement restreint aux villes fréquentées par ego**. La mobilité n'a qu'un effet direct sur la construction du réseau parce qu'elle ne permet pas le développement de co-authorship avec des alter qui sont en dehors de la trajectoire géographique d'*ego*.

iv) La mobilité est à la fois synonyme de création de relations que de rupture de relations entre ego et alter : Dans le modèle 5, nous observons que l'expérience de mobilité accroît significativement la probabilité que des co-auteurs sortent du réseau de co-authorship. En effet, **la mobilité fragilise les relations qu'ego avait durant son ancienne affiliation**. De plus, cet effet de rupture dure plus longtemps que l'effet de création de relations.

Tableau 31 : Estimations de panel (impact de la mobilité sur l'entrée et la sortie d'alter dans le réseau d'ego)

Entrées												
		Tous les alter						Alter affiliés à une ville non fréquentée				
		Modèle (1)		Modèle (2)		Modèle (3)		Modèle (4)				
		Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	
$\Delta_{i,t}^{mob-nature}$	<i>Pas_mobile</i>	REF				REF						
	<i>National</i>	0,554	0,155	***				-0,06	0,12			
	<i>Internat</i>	0,449	0,165	**				0,05	0,13			
$X_{i,t}^{derniere_mob}$	<i>Sédentaire</i>				REF					REF		
	<i>Cette_année</i>				0,49	3,92	***			-0,03	0,10	
	<i>1_à_3_ans</i>				0,13	1,27				0,03	0,08	
	<i>4_à_6_ans</i>				0,02	0,16				-0,11	0,08	
	<i>7_à_ans</i>				0,06	0,54				-0,11	0,09	
	<i>Plus_de_10_ans</i>				-0,08	-0,86				-0,05	0,08	
$Cosign_{i,t}$	0,347	0,020	***	0,35	17,15	***	0,26	0,02	***	0,26	0,02	***
$Publis_{i,t}$	-0,057	0,010	***	-0,06	-5,44	***	-0,01	0,01		-0,01	0,01	
$Geo_net_{i,t}$	0,001	0,001		0,00	0,77		0,00	0,00		0,00	0,00	
$Social_net_{i,t}$	-0,006	0,005		0,00	-0,87		0,01	0,00	*	0,01	0,00	*
$Citat_{i,t}$	0,000	0,000		0,00	-1,30		0,00	0,00	.	0,00	0,00	.

Sorties												
		Tous les <i>alter</i>						<i>Alter</i> affiliés à une ville non fréquentée				
		Modèle (5)			Modèle (6)			Modèle (7)		Modèle (8)		
		Estimate	Std.Error		Estimate	Std.Error		Estimate	Std.Error	Estimate	Std.Error	
$X_{i,t}^{mob-nature}$	<i>Pas_mobile</i>	REF			REF			REF				
	<i>National</i>	0,151	0,890					-0,105	0,126			
	<i>Internat</i>	0,096	0,527					-0,052	0,135			
$X_{i,t}^{derniere_mob}$	<i>Sédentaire</i>	REF			REF			REF				
	<i>Cette année</i>				0,275	1,998	*				-0,091	0,101
	<i>1_à_3_ans</i>				0,348	3,220	**				0,019	0,079
	<i>4_à_6_ans</i>				0,053	0,456					-0,132	0,086
	<i>7_à_9_ans</i>				0,164	1,295					-0,150	0,093
	<i>Plus_de_10_ans</i>				0,186	1,797	.				0,028	0,076
<i>Cosign</i> _{i,t}	0,412	18,416	***	0,405	17,887	***	0,274	0,017	***	0,275	0,017	***
<i>Publis</i> _{i,t}	-0,120	-10,407	***	-0,124	-10,671	***	-0,044	0,008	***	-0,043	0,009	***
<i>Geo_net</i> _{i,t}	0,004	3,451	***	0,005	3,498	***	-0,002	0,001	.	-0,002	0,001	.
<i>Social_net</i> _{i,t}	0,007	1,196		0,005	0,969		0,025	0,004	***	0,024	0,004	***
<i>Citat</i> _{i,t}	0,000	-1,713	.	0,000	-1,603		0,000	0,000		0,000	0,000	

IV- Conclusion et implications politiques

Le but de ce chapitre est **d'évaluer l'impact de la mobilité sur l'extension du réseau de co-authorship**. A partir de la base de données bibliométrique, nous avons testé l'impact du changement d'affiliation d'*ego* sur la **création et la rupture de relation avec *alter*** dans le réseau de co-authorship, au fil de la carrière et à l'échelle micro.

Cette étude se situe dans la lignée des statistiques descriptives du chapitre 2 relatives aux dynamiques sociales et géographiques au fil de la carrière. En effet, nous avons montré que quelle que soit la trajectoire du chercheur, le renforcement de la dimension collaborative des articles s'expliquait via le renouvellement des relations et l'internationalisation du co-authorship. Nous avons également montré que l'ouverture du réseau (nombre de nouveaux *alter* par publication) était stable tout au long de la carrière. Aussi, les collègues d'*ego* forment un noyau structurant et stable dans l'activité de publication tout au fil de sa carrière.

Dans les modèles de panel, nous évaluons plus précisément **l'impact de l'évènement de mobilité à l'échelle micro sur l'extension sociale et géographique du réseau**. Nous avons montré que l'expérience de mobilité augmentait **l'intégration de nouveaux co-auteurs** par article. Néanmoins, cet effet est limité pour trois raisons :

i) La mobilité a un **effet instantané sur la diversification sociale** du réseau. En effet, cet effet de diversification sociale du réseau ne se traduit que l'année où la mobilité est faite et ne s'étale pas au fil de la carrière (Jonkers et Tijssen, 2008).

ii) La mobilité n'a **pas d'impact sur la diversification géographique** du réseau en dehors de la trajectoire géographique d'*ego*. En effet, l'intégration de nouveaux co-auteurs est restreinte aux villes fréquentées par *ego* : il n'y a pas d'effet « bonus » d'ouverture du réseau lié à l'affiliation à une nouvelle ville. Chinchilla-Rodriguez *et al.* (2018) ont montré que le réseau international de mobilité est beaucoup plus petit que le réseau international de co-authorship. Ils soulignent que la mobilité internationale du chercheur n'est pas le seul facteur explicatif des co-signatures internationales.

iii) La mobilité **fragilise les anciennes relations** : cet effet de rupture est plus long dans la carrière que l'effet de création de relations. Ce résultat est cohérent avec les travaux de Bauder (2020) et de Wang *et al.* (2019). A partir d'entretiens de chercheurs mobiles canadiens et allemands, Bauder (2020) a montré que la mobilité pouvait créer du capital social mais aussi entraîner la perte de capital

social national. Aussi, à partir d'un échantillon de chercheurs très mobiles de Singapour, Wang *et al* (2019) ont montré que à la suite d'une mobilité les collaborations de recherche augmentent fortement dans le pays hôte et que les collaborations avec les anciens pays affiliés disparaissent.

Ces effets induisent plusieurs **implications politiques**, notamment concernant les incitations à la mobilité. Encourager la mobilité permanente (« job-to-job ») pour étendre le réseau de co-authorship semble contreproductif pour deux raisons : la mobilité détruit autant de relations qu'elle en crée et ses effets sont restreints à la trajectoire géographique d'*ego*. Lorsque le chercheur intègre une nouvelle ville et un nouveau laboratoire, le chercheur ne tisse de liens qu'avec ses nouveaux collègues et remplace ses anciennes relations. Ackers (2008) a montré que dans les pays où les scientifiques étaient hautement qualifiés, la mobilité n'était considérée comme un signe d'excellence seulement si le chercheur avait accès aux co-auteurs locaux et qu'il co-signait avec eux. Cependant, la capacité à collaborer avec des *alter* de villes différentes ne dépend pas de la mobilité du chercheur. **Par conséquent, la mobilité de court-terme devrait être encouragée puisqu'elle semble aussi efficace pour créer de nouvelles relations et réduit le risque de rupture de relations.**

Chapitre 5 : « Existe-t-il plusieurs façons de collaborer au fil de la carrière ? Une analyse de réseaux sociaux de co-authorship des chimistes »

L'objectif de ce chapitre est de répondre à deux questions : **Existe-t-il différentes façons de collaborer au fil de la carrière ? Si oui, quelles sont-elles ?**

A partir des 259 réseaux *ego*-centrés présentés dans le chapitre 2, nous souhaiterions identifier et décrire les différents styles collaboratifs pratiqués par nos 80 chercheurs au fil de leur carrière. Nous faisons l'hypothèse que la structure du réseau de co-authorship est une importante source d'informations relative au comportement collaboratif du chercheur qui permet de montrer que la collaboration n'est pas un processus singulier.

Notre chapitre suivra trois étapes :

i) Identification des styles collaboratifs. Nous commencerons par décrire la méthodologie suivie dans le regroupement des réseaux en groupes homogènes. Après avoir décrit 13 indicateurs de réseaux *ego*-centrés relatifs à la relation *ego-alter* (position dans le réseau) et à l'organisation des collaborations (structure du réseau), nous avons identifié deux comportements collaboratifs au travers de ces variables grâce à une analyse en composantes principales. Nous avons réalisé une classification hiérarchique ascendante pour identifier quatre classes différentes de réseaux et donc quatre styles collaboratifs différents.

ii) Caractérisation des styles collaboratifs. Nous avons ensuite caractérisé avec attention chaque classe de réseaux à partir des variables les plus contributrices. Nous avons opté pour une caractérisation statique (en considérant les styles comme indépendants les uns des autres) puis une caractérisation dynamique (reconstitution du processus de classification). L'objectif est de comprendre les implications collaboratives sous-jacentes à chaque caractéristique structurelle de la classe.

iii) Estimation des styles collaboratifs. Enfin, nous avons estimé chaque style collaboratif à partir de deux types d'information. Nous avons commencé par adopter une approche statique pour déterminer la probabilité d'opter chaque style collaboratif à partir de données relatives à la carrière et aux caractéristiques du co-authorship. Ensuite, nous avons adopté une approche dynamique pour déterminer la probabilité d'opter chaque style collaboratif en fonction du « style » adopté durant le stade de carrière précédent. L'objectif est d'identifier si la progression de carrière se traduit par des mutations structurelles du réseau de co-authorship et donc des collaborations.

I- Structures des réseaux et types de collaboration

Nous décrivons le **processus regroupement structurel des réseaux** de co-authorship au travers de trois étapes méthodologiques. Nous commencerons par décrire les treize variables choisies de réseaux *ego*-centrés mobilisées pour caractériser les relations *ego-alter* (position dans le réseau) et l'organisation des collaborations (structure du réseau). Une fois ces variables présentées, nous réaliserons une analyse en composantes principales (ACP) pour appréhender les principaux comportements collaboratifs issus de ces treize variables. Cela nous permettra d'identifier deux axes : la propension d'*ego* à entretenir des relations hétérogènes avec ses *alter* et la propension d'*ego* à segmenter ses relations collaboratives. Enfin, nous présenterons les choix méthodologiques opérés dans la classification hiérarchique ascendante (CHA). Cette dernière étape nous permettra d'identifier quatre styles de collaboration différents.

A) Indicateurs des réseaux ego-centrés

Nous avons sélectionné treize indicateurs de réseaux sociaux afin d'identifier les particularités structurelles de chacun des 259 réseaux. Habituellement utilisés dans l'étude de réseaux complets, nous avons dû réinterpréter la signification de chacun d'entre eux dans le cadre de réseaux *ego*-centrés. Le détail du calcul et de l'interprétation de chacun d'entre eux se situe en Annexe 2.

Ces indicateurs donnent deux types d'information :

i) **La position des *alter* dans le réseau** : Quelles sont les différentes relations de collaboration qu'entretient *ego* avec chacun de ses co-auteurs ?

ii) **La structure du réseau** : Comment sont organisés les liens collaboratifs ?

La description détaillée de ces indicateurs est fournie dans les tableaux 33 et 34.

Tableau 32 : Synthèse des treize indicateurs de réseaux *ego*-centrés : position des *alter* dans le réseau

Variables	Signification	Comportement collaboratif d'<i>ego</i>	Modalités	Faible valeur	Forte valeur
1 <i>coef_var_pub_ego</i>	Coefficient de variation du nombre de publication entre <i>ego</i> et ses <i>alter</i>	<i>Ego</i> a des récurrences plus marquées que d'autres dans le choix de son co-authorship	$R \in [0 ; +\infty[$	Tous les <i>alter</i> publient le même nombre d'articles avec <i>ego</i>	Certains <i>alter</i> publient plus d'articles avec <i>ego</i> que d'autres
2 <i>share_alterCliques</i>	Part des nœuds qui sont intégrés dans une clique	<i>Ego</i> restreint sa relation avec ses <i>alter</i> au cadre de l'équipe de publication	$R \in [0 ; 1]$	Réseau faiblement composé de relations restreintes au cadre d'une équipe	Réseau fortement composé de relations restreintes au cadre d'une équipe
3 <i>share_unequivalent</i>	Part des nœuds en in-équivalence automorphique	<i>Ego</i> entretient une relation collaborative unique avec ses <i>alter</i> et leur place dans le réseau est non subsitutable	$R \in [0 ; 1]$	<i>Ego</i> entretient une relation individualisée avec peu d' <i>alter</i>	<i>Ego</i> entretient une relation individualisée avec beaucoup d' <i>alter</i>
4 <i>centr_degree</i>	Centralisation de degré	Hétérogénéité de probabilité des <i>alter</i> d'accompagner <i>ego</i> dans tous ses liens collaboratifs	$R \in [0 ; 1]$	Tous les <i>alter</i> ont la même probabilité d'accompagner <i>ego</i> dans ses collaborations avec l'ensemble de son réseau	Seulement quelques <i>alter</i> accompagnent <i>ego</i> dans ses collaborations avec l'ensemble de son réseau
5 <i>centr_between</i>	Centralisation d'intermédiarité	Hétérogénéité de probabilité des <i>alter</i> d'accompagner <i>ego</i> dans ses liens collaboratifs avec des sous-espaces collaboratifs différents	$R \in [0 ; 1]$	Tous les <i>alter</i> ont la même probabilité d'accompagner <i>ego</i> dans ses collaborations avec des sous-espaces collaboratifs différents	Seulement quelques <i>alter</i> accompagnent <i>ego</i> dans ses collaborations avec des sous-espaces collaboratifs différents
6 <i>centr_eigen</i>	Centralisation de vecteur propre	Hétérogénéité de probabilité des <i>alter</i> d'accompagner <i>ego</i> dans des publications avec ses co-auteurs centraux	$R \in [0 ; 1]$	Tous les <i>alter</i> ont la même probabilité d'accompagner <i>ego</i> dans ses collaborations avec ses <i>alter</i> centraux	Seulement quelques <i>alter</i> accompagnent <i>ego</i> dans ses collaborations avec ses <i>alter</i> centraux

Tableau 33 : Synthèse des treize indicateurs de réseaux *ego*-centrés : structure des liens collaboratifs

Variables	Signification	Comportement collaboratif d'<i>ego</i>	Modalités	Faible valeur	Forte valeur
7 <i>density</i>	Densité	Conception du réseau en une ou plusieurs entités singulières	$R \in [0 ; 1]$	Le réseau est considéré comme une même entité	<i>Ego</i> segmente ses liens collaboratifs en plusieurs entités
8 <i>nb_comp</i>	Nombre de composants	Compartimentation son réseau en composants connexes	$N \in [1 ; +\infty[$	Concentration du réseau en un îlot central	Décomposition du réseau en multiples composants connexes
9 <i>share_comp_princ</i>	Part du réseau dans le composant principal		$R \in]0 ; 1]$	Segmentation très structurante dans le réseau	Segmentation peu structurante sur le réseau
10 <i>clustering_global</i>	Coefficient de clustering global	Densité du voisinage des <i>alter</i> au sein des composants	$R \in [0 ; 1]$	Tous les <i>alter</i> sont au cœur des liens collaboratifs dans le composant	Quelques <i>alter</i> sont au cœur des liens collaboratifs dans le composant
11 <i>modularity</i>	Modularité	Structuration des flux autours des sous-espaces collaboratifs	$R \in [0 ; 1]$	Les liens collaboratifs sont peu structurés au sein de sous-espaces collaboratifs	Les liens collaboratifs sont très structurés au sein de sous-espaces collaboratifs
12 <i>coef_var_taille_comm</i>	Coefficient de variation de la taille des sous-espaces	Certains sous-espaces sont plus structurants que d'autres	$R \in]0 ; +\infty[$	Sous-espaces collaboratifs sont de taille homogène	Sous-espaces collaboratifs sont de tailles hétérogènes
13 <i>intensivity_pub_comm</i>	Rapport entre le nombre de publications et de sous-espaces	Diversification des sous-espaces collaboratifs	$R \in [1 ; +\infty[$	Diversification des sous-espaces (autant de sous-espaces que de publications)	Maintien et croissance des sous-espaces collaboratifs dans lesquels <i>ego</i> a fait plusieurs publications

B) Analyse de composantes principales

Après avoir présenté ces treize indicateurs, nous avons choisi de réaliser une ACP en amont de la création de notre typologie de structure de réseaux pour trois raisons.

i) Tout d'abord, elle nous permet de rendre compte **des corrélations** entre les 13 variables mobilisées. Ainsi, nous pourrions préalablement comprendre quelles sont les principales tendances collaboratives décrites par nos variables au travers d'un nombre limité de nouvelles variables non corrélées.

ii) Ensuite, nous pourrions **introduire ces nouveaux facteurs dans la CHA**. L'objectif est d'accroître la qualité du partitionnement des classes en mobilisant des combinaisons linéaires non corrélées et dont l'information n'est pas redondante. Cela nous aidera également à interpréter les différentes classes en fonction de leur position sur les axes.

iii) Enfin, la méthode permet de **projeter les données dans un espace aux dimensions réduites**. Il est donc possible d'identifier graphiquement les groupes d'individus homogènes ainsi que les individus atypiques relativement aux axes conservés. Cette première visualisation graphique des individus est un préliminaire nécessaire pour comprendre et interpréter le processus de classification.

1- Description des variables actives

Tableau 34 : Statistiques descriptives des 13 variables actives

		Min	Premier quartile	Médiane	Moyenne	Troisième Quartile	Max
La position des alter dans le réseau	<i>coef_var_pub_ego</i>	0,00	0,85	1,10	1,18	1,43	3,74
	<i>share_alter_cliqes</i>	0,00	0,60	0,67	0,66	0,73	1,00
	<i>share_unequivalent</i>	0,00	0,06	0,09	0,10	0,13	0,57
	<i>centr_degree</i>	0,04	0,27	0,39	0,42	0,56	0,96
	<i>centr_eigen</i>	0,21	0,71	0,78	0,76	0,83	0,96
	<i>centr_between</i>	0,00	0,19	0,33	0,34	0,48	0,91
L'organisation des liens collaboratifs	<i>density</i>	0,01	0,09	0,13	0,17	0,21	0,77
	<i>nb_comp</i>	0,00	2,00	3,00	2,99	4,00	23,00
	<i>share_comp_princ</i>	0,29	0,67	0,84	0,80	0,98	1,00
	<i>clustering_global</i>	0,00	0,43	0,54	0,56	0,66	1,00
	<i>modularity</i>	0,00	0,42	0,51	0,50	0,60	0,82
	<i>coef_var_taille_comm</i>	0,00	0,42	0,57	0,56	0,71	1,49
	<i>intensity_pub_comm</i>	1,00	3,00	4,50	5,38	7,00	28,54

Le tableau 35 détaille les statistiques descriptives de nos 13 variables qui seront utilisées lors de l'ACP. Nous pouvons interpréter les valeurs moyennes de chacune d'entre elles pour avoir un premier aperçu des styles collaboratifs des chercheurs.

Concernant les relations *ego-alter* le réseau, nous notons plusieurs résultats :

i) Au sein de chaque réseau, nous notons une **légère hétérogénéité moyenne concernant le nombre de publications d'ego avec chacun de ses alter** puisque le coefficient de variation moyen est légèrement supérieur à 1 (1,18). En moyenne, certains *alter* sont plus récurrents que d'autres dans la bibliométrie d'*ego*. Nous notons cependant de fortes valeurs extrêmes pour cette variable : la valeur maximale correspond au réseau où l'écart type est presque quatre fois supérieur au nombre moyen de publication par *alter*. Ainsi, dans certains réseaux, la récurrence dans les publications d'*ego* est très hétérogène entre les *alter*.

ii) Aussi, **très peu d'alter disposent d'une position particulière dans le réseau** : 75% des réseaux ont au plus 13% d'*alter* disposant d'un degré particulier et 75% des réseaux ont plus de 60% d'*alter* compris dans une clique (relation de collaboration avec *ego* restreinte au cadre de la publication).

iii) Les **moyennes de centralisation de degré et d'intermédiarité** sont plutôt **faibles** (respectivement 0,42 et 0,34) : il y a peu d'hétérogénéité entre les *alter* concernant l'accompagnement d'*ego* dans ses liens collaboratifs avec le réseau ou des sous-espaces collaboratifs différents. En revanche, **la moyenne de la centralisation de vecteur propre est plutôt forte (0,76)** : en moyenne, tous les *alter* n'ont pas la même probabilité d'accompagner *ego* dans ses publications avec ses *alter* centraux.

Plusieurs résultats sont également remarquables concernant l'organisation des liens collaboratifs :

i) Nous notons une très **faible densité** dans la majorité de nos réseaux : en moyenne, chaque réseau dispose de 17% des arrêtes possibles au vue de son nombre de nœuds. Seule une petite partie des réseaux semblent très dense : 25% des réseaux ont entre 21% et 77% des arrêtes possibles au regard de leur nombre de nœuds. La **segmentation des relations collaboratifs est donc très pratiquée** : en moyenne, le réseau d'*ego* est segmenté en 3 composants connexes. Certains réseaux sont plus segmentés que d'autres : 25% des réseaux sont segmentés entre 4 et 23 composants. Cette décomposition du réseau en composants n'est néanmoins que très **peu structurante dans le sens où une forte majorité des alter reste concentrée dans un même composant** : 50% des réseaux ont un composant qui concentre plus de 84% des *alter*. Très peu de réseaux disposent de composants dans lesquels les *alter* sont équitablement distribués.

ii) Concernant la densité du réseau à l'échelle des *alter* dans les composants, les réseaux sont relativement bien distribués : dans 50% des réseaux, les **alter ont moins de 54% de chance d'avoir un voisinage connecté**.

iii) Enfin, concernant l'organisation des liens selon les **sous-espaces collaboratifs**, on note également que la modularité est équitablement distribuée dans les réseaux : 50% des réseaux ont une modularité inférieure à 0,51. En moyenne, les **sous-espaces sont de taille homogène** (le coefficient de variation moyen vaut 0,56) et *ego* réalise 5,38 publications avec chacun d'entre eux. On note une forte dispersion concernant la relation collaborative entre *ego* et ses sous-espaces : dans 25% des réseaux, *ego* ne réalise qu'entre 1 et 3 publications avec chaque sous-espace alors que dans 25% des réseaux, il en réalise entre 7 et 28,54 avec chaque sous-espace.

Avant d'appliquer l'ACP, nous avons choisi de **centrer et réduire les 13 variables** (Annexe 3). En effet, nos variables n'ont pas toutes la même unité⁵⁴ et certaines variables sont très dispersées⁵⁵.

2 – Application de l'Analyse de composantes principales

Nous appliquons donc l'analyse de composantes principales sur 259 réseaux (n=259) décrits par 13 variables actives centrées réduites (p=13).

Les détails de la méthodologie suivie lors de l'analyse de la décomposition de l'inertie totale, ainsi que la répartition graphique des individus et de la contribution totale des 13 variables actives font partie de l'Annexe 4. A l'issue de cette analyse, nous choisissons de retenir deux valeurs propres, chacune associée à une dimension collaborative particulière.

La première dimension concerne l'hétérogénéité des relations collaboratives (soit, 35% de l'inertie totale)

Tableau 35 : Variables contributrices à la caractérisation de la dimension 1

	Coordonnée	p.value	Contribution	Cos2
coef_var_pub_ego	0,9025	5.008E-96*	17,83	0,81
centr_between	0,8103	1.416E-61	14,37	0,66
intensity_pub_comm	0,7544	6.694E-49	12,46	0,57
centr_degree	0,7338	4.767E-45	11,79	0,54
share_comp_princ	0,7189	1.789E-42	11,31	0,52
centr_eigen	0,5293	4.200E-20	6,13	0,28
coef_var_taille_comm	0,2652	1.524E-05	1,54	0,07
nb_comp	-0,2409	9.011E-05	1,27	0,06
density	-0,5478	1.101E-21	6,57	0,30
clustering_global	-0,8655	4.120E-79	16,40	0,75

* ce tableau ne représente que les variables qui décrivent significativement la dimension (dont |p.value| < 0.05)

⁵⁴ Par exemple, le nombre de composants connexes (*nb_comp*) n'est pas dans la même unité que la densité (*density*).

⁵⁵ C'est le cas par exemple du nombre de composants (*nb_comp*) ou du nombre moyen de publications réalisées avec chaque sous-espace collaboratif (*intensity_pub_comm*).

A droite de l'axe, nous trouvons les réseaux de co-authorship fortement structurés par **l'hétérogénéité des relations collaboratives qu'entretient ego et chacun de ses alter** :

i) Des alter sont plus centraux que d'autres dans les liens collaboratifs d'ego :

Tout d'abord, nous observons une dualité entre les *alter* du réseau : ceux qui sont récurrents dans les publications d'*ego* et ceux qui ne le sont pas. En effet, l'hétérogénéité du nombre de publications partagées entre *ego* et chacun des *alter* est la variable qui a la plus grande coordonnée sur la droite de l'axe (*coef_var_pub_ego*). Ainsi, dans ces réseaux, *ego* choisit de partager une grande partie de ses publications avec certains de ses co-auteurs récurrents alors que d'autres n'auront qu'une contribution ponctuelle. La contribution de ces co-auteurs récurrents n'est pas que bibliométrique : ils sont aussi au cœur d'une grande partie des relations de co-authorship d'*ego* à l'échelle de son réseau, de sous-espaces collaboratifs différents et des co-auteurs centraux. En effet, les centralisations de degré (*centr_degree*), d'intermédiation (*centr_between*) et de vecteur propre (*centr_eigen*) disposent de coordonnées de forte valeur sur la droite de l'axe.

ii) Sous-espaces collaboratifs autour de quelques alter centraux :

Nous observons également une dualité de sous-espaces collaboratifs qui sont de taille différente (*coef_var_taille_comm*) : ceux où les *alter* centraux sont présents sont plus gros que les autres. En effet, l'activité de publication d'*ego* se traduit par la croissance progressive des sous-espaces collaboratifs (*intensity_pub_comm*) dans lesquels les *alter* centraux sont logés. Les autres sous-espaces collaboratifs, plus petits, gravitent autour des *alter* centraux (récurrents dans les liens collaboratifs qu'*ego* a avec chacun d'entre eux). Le réseau va donc être concentré dans un même composant (*share_comp_princ*).

A gauche de l'axe, nous trouvons des réseaux qui se caractérisent par une **forte homogénéité des relations collaboratives entre ego et ses alter** :

i) Des alter au cœur des liens collaboratifs :

Dans ces réseaux, tous les *alter* ont une forte probabilité d'être proche d'un voisinage connecté. En d'autres termes, dans chaque composant, les *alter* collaborent conjointement avec *ego* au sein de publications communes (*clustering_global*).

ii) Les sous-espaces collaboratifs qui ne gravitent pas autour des alter centraux :

Les réseaux situés à gauche de l'axe peuvent avoir deux types de structure. Dans un premier cas, le réseau peut être très dense à l'échelle globale : tous les *alter* du réseau font partie d'une même entité, dans laquelle ils sont mobilisés conjointement. Les liens collaboratifs ne sont pas cloisonnés au sein de sous-espaces collaboratifs hermétiques et ne gravitent pas autour de co-auteurs centraux (*density*). Dans

un second cas, le réseau peut être éclaté en composantes connexes, très denses. *Ego* collabore indépendamment avec chacun d'entre eux, sans être accompagné par des co-auteurs récurrents (*nb_comp*).

La seconde dimension concerne la segmentation du réseau en plusieurs entités (soit, 28% de l'inertie totale)

Tableau 36 : Variables contributrices à la caractérisation de la dimension 2

	Coordonnée	p.value	Contribution	Cos2
modularity	0,8595	7.416E-77*	20,26	0,74
centr_eigen	0,7297	2.501E-44	14,60	0,53
nb_comp	0,7257	1.240E-43	14,44	0,53
coef_var_taille_comm	0,5522	4.518E-22	8,36	0,30
intensity_pub_comm	0,1259	4.288E-02	0,43	0,02
centr_between	-0,3701	7.894E-10	3,76	0,14
centr_degree	-0,4583	7.487E-15	5,76	0,21
share_comp_princ	-0,4683	1.602E-15	6,01	0,22
share_unequivalent	-0,6382	4.948E-31	11,17	0,41
density	-0,7317	1.129E-44	14,68	0,54

* ce tableau ne représente que les variables qui décrivent significativement la dimension (dont $|p.value| < 0.05$)

En haut de l'axe, nous trouvons des **réseaux éclatés en sous-espaces collaboratifs et en composants**.

i) Réseaux éclatés en plusieurs entités :

Ces réseaux sont caractérisés par une forte segmentation des relations collaboratives. En effet, la modularité (*modularity*) est la variable qui a la plus forte coordonnée et qui contribue le plus à l'axe. Cela signifie que les liens collaboratifs sont cloisonnés au sein de sous-espaces collaboratifs denses en interne et hermétiques au reste du réseau. Ces sous-espaces sont de taille différente (*coef_var_taille_comm*) et grossissent au fil des publications d'*ego* (*intensity_pub_comm*). Cette segmentation peut être extrême dans le cas où le réseau est compartimenté en composants connexes (*nb_comp*).

ii) Les liens collaboratifs avec les co-auteurs centraux sont isolés dans le réseau :

Au sein de ces réseaux, tous les *alter* ont une faible probabilité d'accompagner *ego* dans ses liens collaboratifs avec ses *alter* centraux. En effet, les projets collaboratifs dans lesquels *ego* collaborent avec ses *alter* centraux se font au sein d'un composant ou d'un sous-espace collaboratif et sont isolés du reste du réseau (*centr_eigen*).

En bas de l'axe, nous trouvons des réseaux dont les **liens ne sont pas cloisonnés au sein de sous espaces collaboratifs ou de composants connexes.**

i) **Le réseau est considéré comme une même entité :**

Ego ne compartimente pas ses liens collaboratifs. En effet, la densité (*density*) est la seconde variable contributrice de l'axe. Les liens collaboratifs ne sont pas cloisonnés au sein de sous-espaces collaboratifs hermétiques ou au sein de composantes : tous les *alter* du réseau font partie d'une même entité, dans laquelle ils sont mobilisés conjointement. Ainsi, une grande partie du réseau est concentrée dans la composante principal (*share_comp_princ*) qui est dense.

ii) **Hétérogénéité des *alter* :**

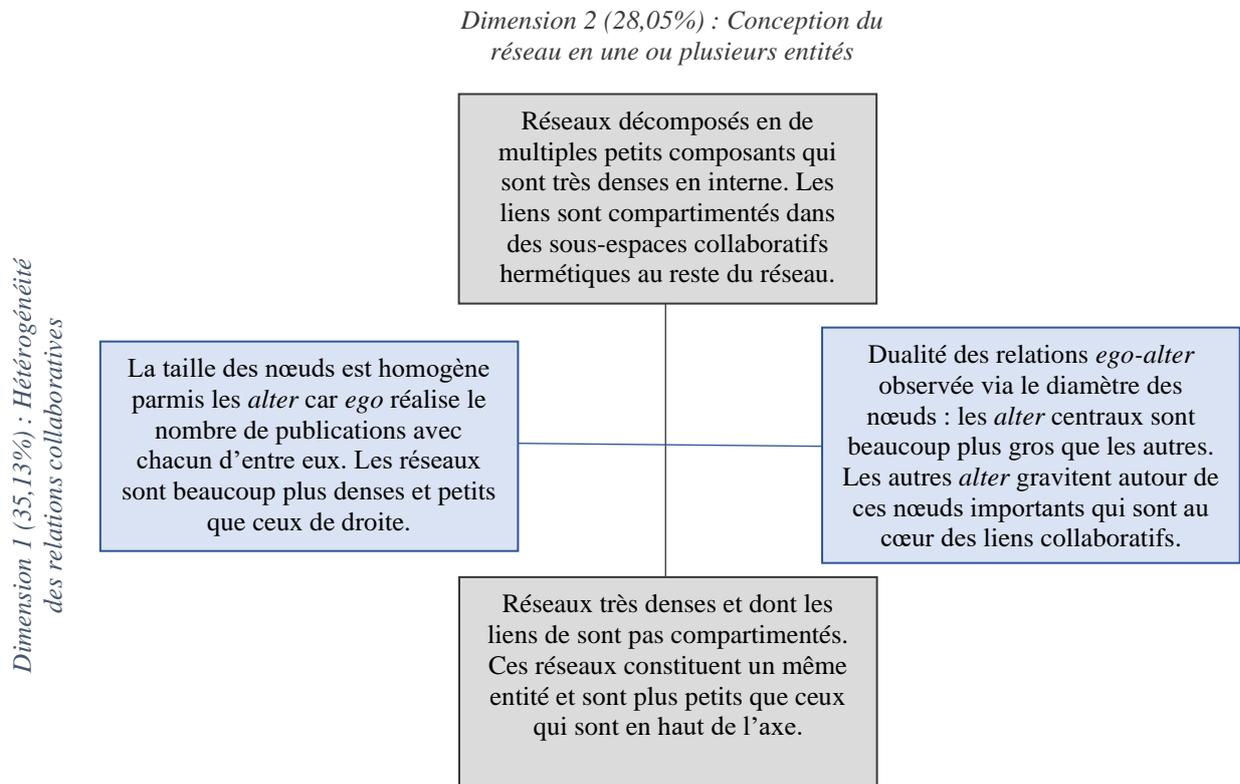
Ego entretient une relation singulière avec une grande partie de ses *alter*. En effet, la part des *alter* disposant d'une position unique dans le réseau (*share_unequivalent*) a une forte coordonnée sur le bas de l'axe. Tous les *alter* de l'équipe d'une même publication entrant dans le réseau ont la même place : une fois la publication faite, *ego* maintient le lien avec une grande partie d'entre eux et personnalise ses relations. Cette relation individuelle se traduit aussi par une centralisation de degré (*centr_degree*) et d'intermédiation (*centr_between*) élevées qui traduisent la récurrence de certains *alter* dans les collaborations d'*ego*.

Après avoir interpréter ces deux dimensions, nous pouvons terminer notre caractérisation des axes en identifiant les particularités graphiques selon leur positionnement dans le plan.

3- Représentation graphique des réseaux les plus contributifs

Ces deux axes sont donc associés à des particularités structurelles. Le schéma 33 décrit la structure des réseaux positionnés à gauche et à droite de la dimension 1 et en haut et en bas de la dimension 2. La représentation graphique des 5 réseaux les plus contributeurs de chaque dimension se trouve en Annexe 5.

Figure 33 : Structure des réseaux et axes de l'ACP



L'analyse de composantes principales a donc permis de créer deux nouvelles variables (« dimension 1 » et « dimension 2 ») qui permettent d'expliquer 63% de l'inertie totale relative à la structure de nos réseaux. **Ces deux variables décrivent deux comportements collaboratifs des chercheurs.** Elles seront intégrées dans la création de la typologie pour améliorer la qualité du partitionnement et l'interprétation des classes. La CHA dresse des liens hiérarchiques entre nos 259 réseaux afin de démarquer ceux qui se ressemblent et se distinguent le plus. La classification sera donc faite à partir de 15 variables : les 13 variables actives et les deux axes collaboratifs issus de l'analyse de composante principale.

C) Description de la méthodologie de la classification hiérarchique ascendante

Une classification est un processus qui consiste à regrouper des individus dans des classes homogènes. Une multitude de partitions sont possibles selon les outils mobilisés et les objectifs suivis. Ainsi, il convient de choisir celle qui donnera la meilleure qualité possible.

La **qualité d'une partition** est soumise aux deux critères suivants (détaillés dans l'Annexe 7) :

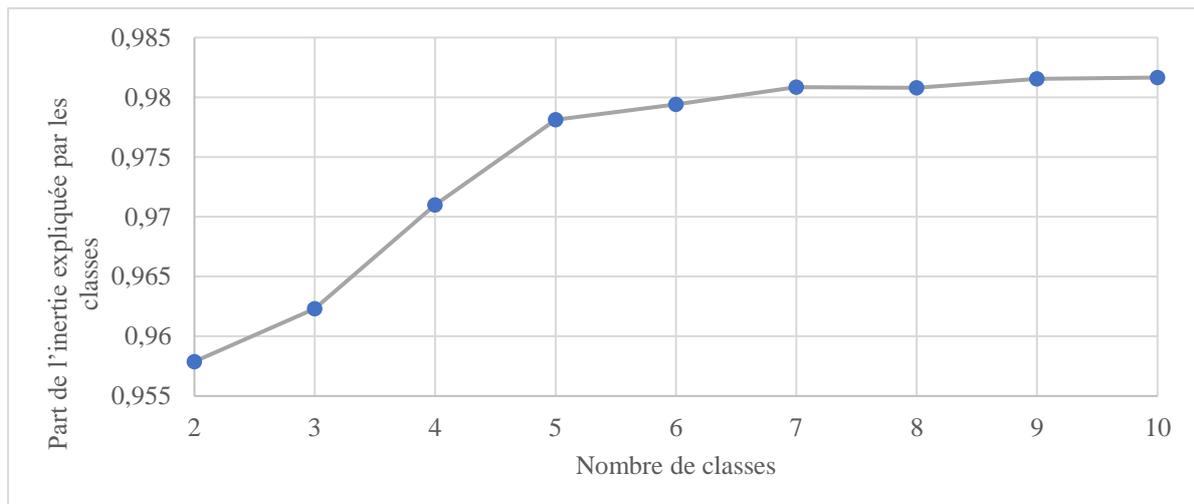
i) **La proximité des individus au sein des classes** (faible variabilité intra-classe) à partir de la matrice de distance euclidienne

ii) **La distance des individus de classes différentes** (forte variabilité inter-classe) grâce à la méthode d'agrégation de Ward.

L'algorithme de la CHA a été appliqué sur les 259 réseaux décrits par 15 variables (13 variables actives et des deux axes collaboratifs retenus à la suite de l'ACP). Une fois que l'algorithme a regroupé nos réseaux dans une même entité, nous avons choisi l'étape du partitionnement permettant la meilleure qualité de résultats.

Par définition, chaque agrégation a deux effets : elle entraîne à la fois une baisse de l'inertie interclasse (baisse de la différence entre les classes) et une augmentation de l'inertie intra-classe (augmentation de la différence au sein des classes).

Figure 34 : Perte relative d'inertie selon le nombre de classes



Sur le graphique suivant, nous observons une chute d'inertie à partir de la classification en 5 classes. Il s'agit donc du nombre maximal de classes que notre classification peut supporter. Nous observons une pente plus importante concernant la classification à 3 classes : selon le critère de Ward, c'est à ce seuil que la perte d'inertie est maximale. La seconde classification optimale est celle à 4 classes, puisqu'il s'agit de la seconde pente la plus forte. Nous choisissons donc entre ces deux classifications à partir des graphiques de nuage de points des individus sur le plan des axes de l'ACP et des dendrogrammes de chaque partitionnement (Annexe 6).

i) Nous remarquons que la **classification à 3 classes** a un risque **d’inertie intra-classe élevé**. Par exemple, dans le nuage de points, on remarque que la classe 2 (jaune) et la classe 3 (grise) disposent de réseaux très concentrés graphiquement sur le plan.

ii) La **classification à 4 classes** nous semble être un **choix optimal**. En effet, elle est un bon compromis concernant les critères d’inertie inter et intra-classe. Sur le dendrogramme, on remarque que la distance relative aux différents regroupements est suffisamment élevée. Dans le nuage de points, les quatre classes ont des frontières distinctes, ce qui sous-entend une homogénéité intra-classe suffisante.

II- Caractérisation des quatre structures de réseau

Grâce à la CHA, nous avons regroupé les 259 réseaux en quatre groupes homogènes au regard de leur structure. Nous commencerons par décrire une à une les quatre classes et nous dresserons un schéma synthétique du processus de classification.

A) Caractérisation des classes

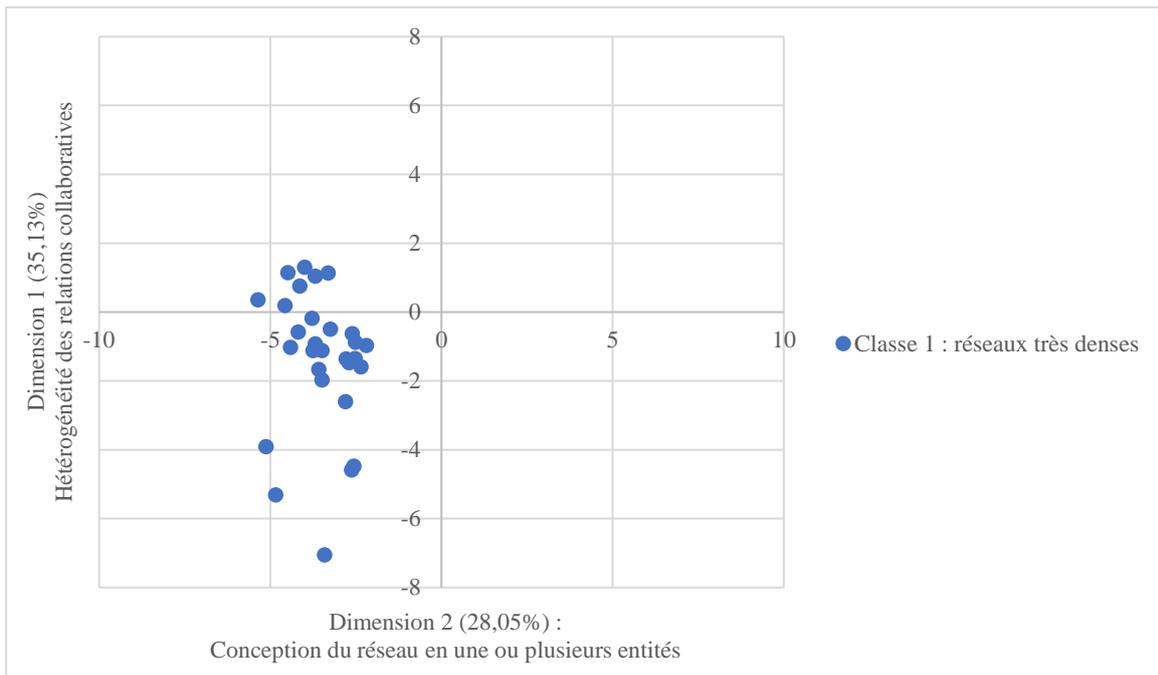
Nous pouvons à présent caractériser les quatre classes de réseau à partir des treize indicateurs de réseaux ego-centrés et de ces deux axes issus de l’ACP, qui sont extrêmement contributeurs dans la classification (Annexe 8).

1- Caractérisation de la classe 1 : Réseaux à très forte densité

Tableau 37 : Caractérisation de la classe 1

		v.test	Mean category	Overall mean	p.value
Dimensions collaboratives	<i>Dim.2</i>	-4,055	-1,358	-1,31E-16	5,01E-05
	<i>Dim.1</i>	-9,397	-3,521	-2,52E-17	5,63E-21
Variables actives	<i>density</i>	10,148	1,776	1,06E-17	3,38E-24
	<i>clustering_global</i>	9,686	1,695	-1,02E-16	3,44E-22
	<i>coef_var_taille_comm</i>	-2,980	-0,521	-2,22E-16	2,88E-03
	<i>intensivity_comm_pub</i>	-4,750	-0,831	-3,18E-17	2,03E-06
	<i>modularity</i>	-4,861	-0,851	-1,81E-16	1,17E-06
	<i>share_comp_princ</i>	-5,227	-0,915	-1,46E-16	1,72E-07
	<i>centr_degree</i>	-5,474	-0,958	5,46E-17	4,41E-08
	<i>centr_between</i>	-6,610	-1,157	-3,90E-17	3,84E-11
	<i>coef_var_pub_ego</i>	-6,980	-1,221	-9,34E-17	2,96E-12
	<i>centr_eigen</i>	-9,323	-1,631	-4,48E-16	1,13E-20

Figure 35 : Distribution graphique de la classe 1

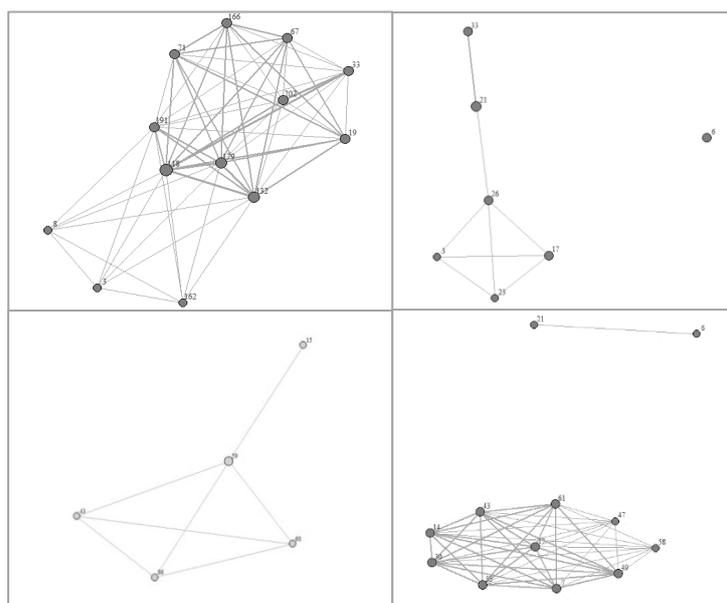


La classe 1 est caractérisée par une **forte densité**, aussi bien à l'échelle locale que globale du réseau.

A l'échelle locale, tous les *alter* sont au cœur des liens collaboratifs. En effet, ces réseaux se trouvent à gauche de l'axe 1. Dans ce type de structure, les *alter* ont un voisinage significativement plus dense que la moyenne (*clustering_global*). Les composants sont donc très **connectés en interne**. Tous les *alter* accompagnent *ego* dans ses liens collaboratifs, que ce soit à l'échelle du réseau entier ou avec des sous-espaces collaboratifs différents (*centr_degree*, *centr_between*). Cela est en partie dû au fait qu'*ego* entretienne des relations très homogènes avec ses *alter*. En effet, il réalise autant de publications avec chacun d'entre eux (*coef_var_pub_ego*) : les nœuds du graphique sont donc de même diamètre. Aussi, cette densité interne s'explique par une certaine autonomie collaborative. *Ego* collabore avec les composants de son réseau sans être accompagné par de co-auteurs récurrents : le réseau ne gravite pas autour de co-auteurs centraux. (*share_comp_princ*).

A l'échelle globale, on remarque que **les liens collaboratifs ne sont pas compartimentés**. En effet, la classe 1 est significativement caractérisée comme en dessous du second axe. Tous les co-auteurs font partie d'une même entité, mobilisés conjointement et tous complémentaires entre eux. Ces réseaux sont significativement plus denses que la moyenne (*density*) : tous les *alter* accompagnent *ego* dans ses relations avec ses co-auteurs centraux (*centr_eigen*). Le réseau n'est pas segmenté et les relations collaboratives ne sont pas cloisonnées. En effet, les sous-espaces collaboratifs sont très peu structurants : ils sont mobilisés avec le reste du réseau (*modularity*), de taille homogène (*coef_var_taille_comm*) et *ego* publie très peu avec chacun d'entre eux (*intensity_comm_pub*).

Figure 36 : Réseaux caractéristiques de la classe 1 (les plus éloignés du centre de gravité des autres classes)

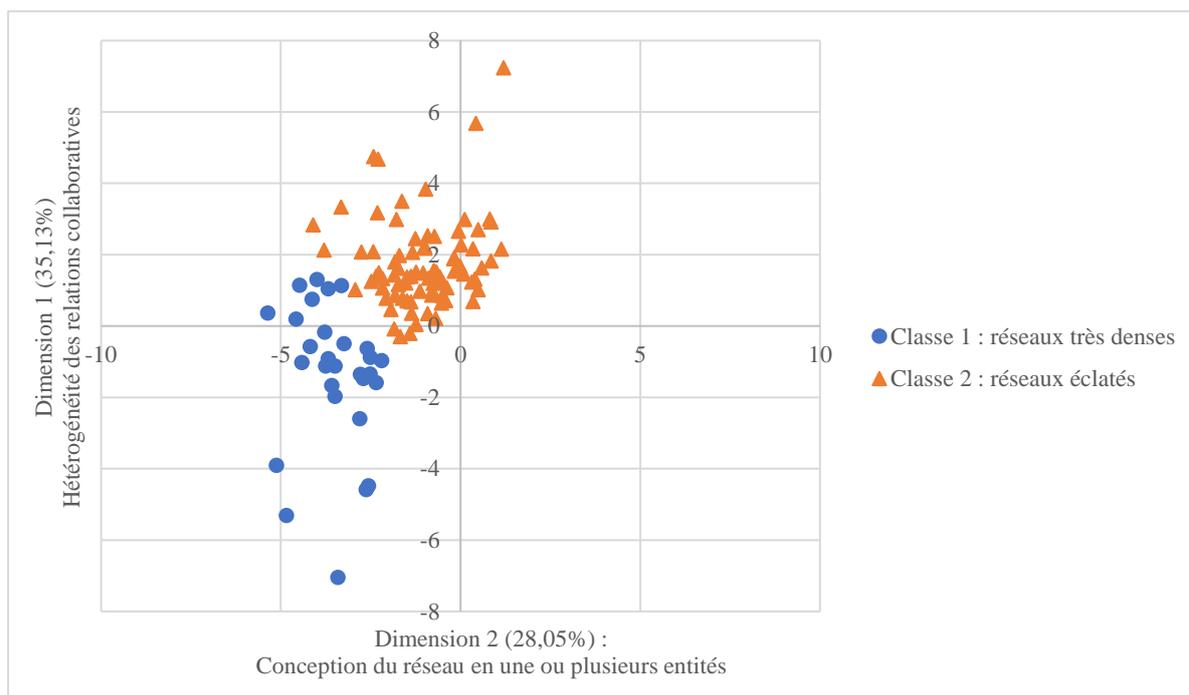


2- Caractérisation de la classe 2 : Réseaux éclatés

Tableau 38 : Caractérisation de la classe 2

		v.test	Mean in category	Overall mean	p.value
Dimensions collaboratives	<i>Dim.2</i>	9,916	1,747	-1,31E-16	3,56E-23
	<i>Dim.1</i>	-5,303	-1,046	-2,52E-17	1,14E-07
Variables actives	<i>nb_comp</i>	8,999	0,829	1,07E-18	2,27E-19
	<i>modularity</i>	8,990	0,828	-1,81E-16	2,47E-19
	<i>centr_eigen</i>	4,955	0,456	-4,48E-16	7,25E-07
	<i>clustering_global</i>	4,584	0,422	-1,02E-16	4,55E-06
	<i>coef_var_taille_comm</i>	3,626	0,334	-2,22E-16	2,88E-04
	<i>share_alterCliques</i>	2,027	0,187	-1,41E-16	4,26E-02
	<i>intensivity_comm_pub</i>	-3,985	-0,367	-3,18E-17	6,75E-05
	<i>share_unequivalent</i>	-4,329	-0,399	-2,86E-17	1,50E-05
	<i>coef_var_pub_ego</i>	-4,546	-0,419	-9,34E-17	5,48E-06
	<i>density</i>	-4,751	-0,438	1,06E-17	2,02E-06
	<i>centr_between</i>	-7,182	-0,662	-3,90E-17	6,88E-13
	<i>centr_degree</i>	-8,188	-0,754	5,46E-17	2,65E-16
	<i>share_comp_princ</i>	-8,513	-0,784	-1,46E-16	1,69E-17

Figure 37 : Distribution graphique des classes 1 et 2

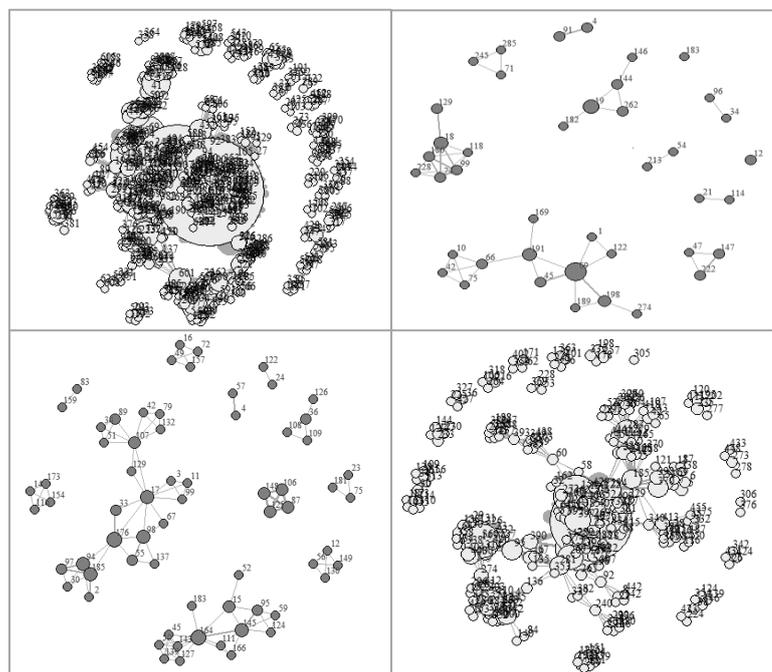


Les réseaux de la classe 2 sont **éclatés en de multiples composants** qui sont très denses en interne.

A l'échelle globale, on note une forte segmentation des relations collaboratives : les réseaux de la classe 2 sont **décomposés en plusieurs entités**. En effet, la dimension 2 est très contributrice dans la construction de cette classe, les réseaux étant au-dessus de son axe. La densité (*density*) est significativement plus faible que la moyenne : *ego* compartimente ses liens de collaboration au sein de sous-espaces collaboratifs et de composants connexes. La modularité (*modularity*) des réseaux est aussi significativement plus forte que la moyenne : les sous-espaces collaboratifs sont isolés au sein de composants différents (*nb_comp*, *share_comp_princ*). Les sous-espaces sont de taille différente (*coef_var_taille_comm*) et *ego* réalise peu de publications avec chacun d'entre eux (*intensivité_comm_pub*). Les *alter* étant répartis dans ces composants isolés, les co-auteurs centraux sont exclus du reste du réseau : seuls les *alter* issus de leur composant accompagnent *ego* dans ce type de collaboration (*centr_eigen*).

A l'échelle locale, on note **une forte densité interne**. La classe 2 se situe à gauche de la dimension 1 : *ego* collabore en toute autonomie avec chacun de ses composants au sein desquels tous les *alter* sont au cœur de ses relations collaboratifs. Ainsi, le voisinage des *alter* est significativement plus dense que la moyenne (*clustering_global*). Aucun *alter* n'accompagne *ego* dans l'ensemble de ses liens collaboratifs avec le réseau (*centr_degree*) et avec des sous-espaces collaboratifs différents (*centr_between*). Cela s'explique par l'homogénéité des relations qu'*ego* entretient des relations avec chacun de ses *alter*. En effet, *ego* réalise autant de publications avec chacun d'entre eux (*coef_var_pub_ego*). Une grande partie du réseau est composée de cliques (*share_alter_cliques*) : pour beaucoup d'*alter*, la relation avec *ego* est limitée au cadre de l'équipe de publication. Ainsi, très peu d'*alter* dispose d'une position particulière dans le réseau (*share_unequivalent*).

Figure 38 : Réseaux caractéristiques de la classe 2 (les plus éloignés du centre de gravité des autres classes)

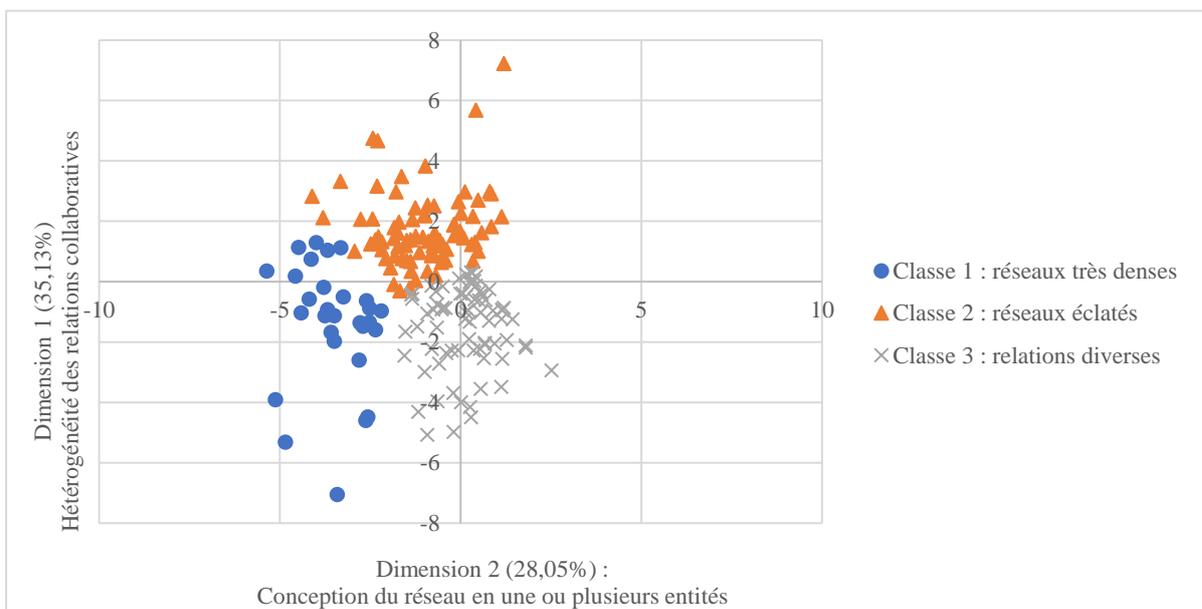


3- Caractérisation de la classe 3 : Densité et diversité des relations

Tableau 39 : Caractérisation de la classe 3

		v.test	Mean in category	Overall mean	p.value
Dimensions collaboratives Variables actives	<i>Dim.2</i>	-8,176	-1,537	-1,31E-16	2,93E-16
	<i>share_unequivalent</i>	6,919	0,680	-2,86E-17	4,56E-12
	<i>share_comp_princ</i>	5,247	0,516	-1,46E-16	1,54E-07
	<i>density</i>	4,234	0,416	1,06E-17	2,29E-05
	<i>centr_degree</i>	4,165	0,409	5,46E-17	3,11E-05
	<i>centr_between</i>	3,429	0,337	-3,90E-17	6,05E-04
	<i>share_alterCliques</i>	-2,053	-0,202	-1,41E-16	4,00E-02
	<i>clustering_global</i>	-2,585	-0,254	-1,02E-16	9,72E-03
	<i>intensity_comm_aggr</i>	-2,606	-0,256	-3,18E-17	9,16E-03
	<i>centr_eigen</i>	-4,865	-0,478	-4,48E-16	1,15E-06
	<i>coef_var_taille_comm</i>	-4,888	-0,480	-2,22E-16	1,02E-06
	<i>nb_comp</i>	-4,974	-0,489	1,07E-18	6,57E-07
	<i>modularity</i>	-7,044	-0,692	-1,81E-16	1,87E-12

Figure 39 : Distribution graphique des classes 1,2 et 3

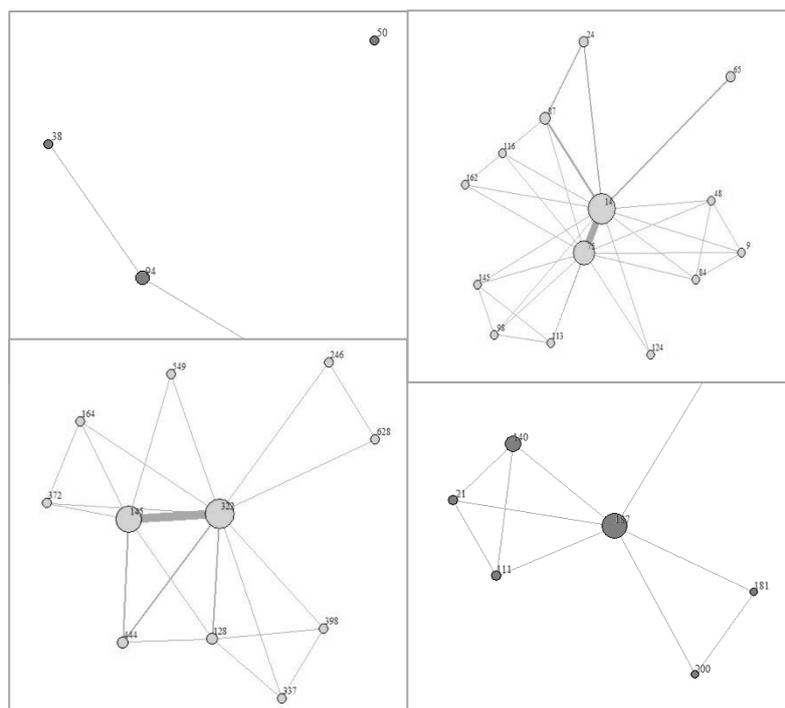


Le réseau de la classe 3 est considéré comme une seule entité collaborative dans laquelle *ego* entretient des **relations individuelles et particulières avec chacun des *alter***.

A l'échelle globale, **les liens collaboratifs ne sont pas compartimentés**. Comme la classe 1, les réseaux de la classe 3 se situent en dessous de l'axe de la dimension 2. La densité (*density*) est significativement plus forte que la moyenne et tous les *alter* ont une forte probabilité d'accompagner *ego* dans ses liens avec ses *alter* centraux (*centr_eigen*). Les liens ne sont ni compartimentés dans des composants connexes, ni dans des sous-espaces collaboratifs isolés : le nombre de composants connexes et la modularité sont inférieurs à la moyenne (*nb_comp*, *modularity*). Ces sous-espaces sont peu structurants : ils sont de taille homogène (*coef_var_taille_comm*) et *ego* réalise peu de publications avec chacun d'entre eux (*intensiviy_comm_pub*).

A l'échelle locale, contrairement à la classe 1, *ego* entretient des relations très hétérogènes avec ses *alter*. En effet, une grande partie du réseau dispose d'une position unique et non substituable (*share_unequivalent*). **Les relations de collaboration sont individualisées** : *ego* collabore différemment avec chaque *alter* et le mobilise avec différentes équipes. Le poids des cliques est inférieur à la moyenne (*share_alter_cliques*) : la relation entre *ego* et *alter* va au-delà du cadre d'une unique publication. Ainsi, certains *alter* sont plus structurants que d'autres dans les liens collaboratifs d'*ego* : le voisinage d'*alter* est moins dense que la moyenne (*clustering_global*). En effet, seulement quelques *alter* accompagnent *ego* dans l'ensemble de ses liens collaboratifs avec le réseau (*centr_degree*) et avec des sous-espaces collaboratifs différents (*centr_between*). Le réseau est donc concentré dans un même composant principal autour de ses *alter* centraux (*share_comp_princ*).

Figure 40 : Réseaux caractéristiques de la classe 3 (les plus éloignés du centre de gravité des autres classes)

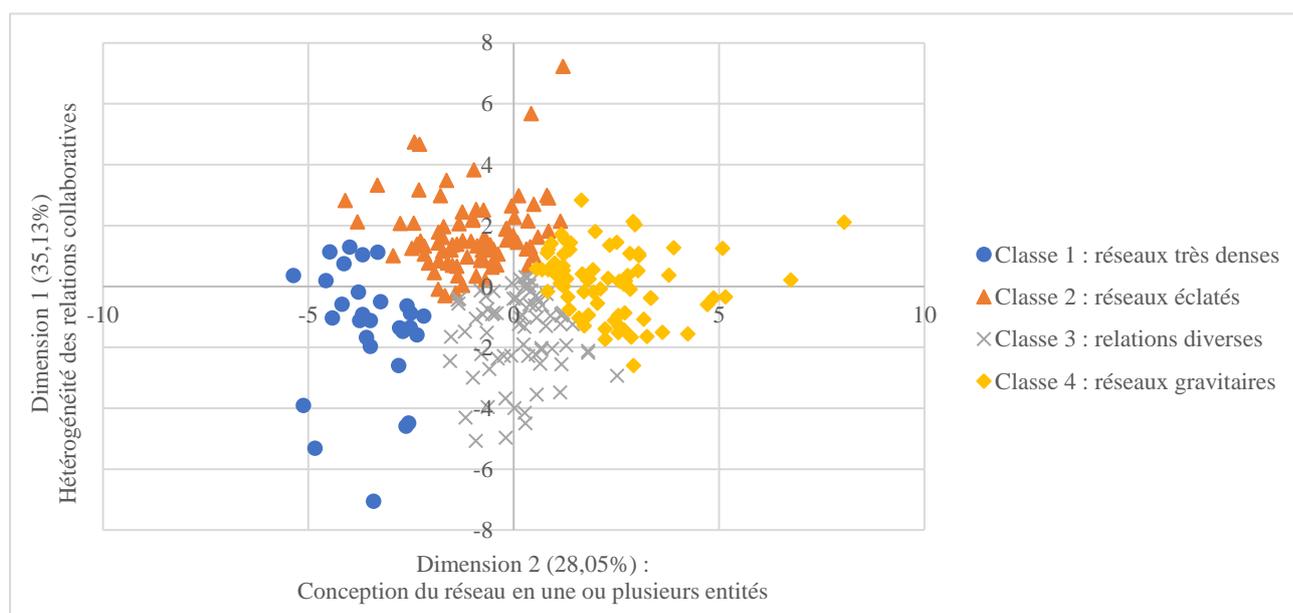


4- Caractérisation de la classe 4 : Réseaux gravitaires et dualité des relations

Tableau 40 : Caractérisation de la classe 4

Dimensions collaboratives	Dim.1	v.test	Mean in category	Overall mean	p.value
Variables actives		11,344	2,364	-2,52E-17	7,97E-30
	<i>coef_var_pub_ego</i>	10,790	1,050	-9,34E-17	3,86E-27
	<i>intensivity_comm_aggr</i>	9,971	0,970	-3,18E-17	2,04E-23
	<i>centr_between</i>	8,521	0,829	-3,90E-17	1,58E-17
	<i>centr_degree</i>	8,027	0,781	5,46E-17	1,00E-15
	<i>share_comp_princ</i>	7,110	0,692	-1,46E-16	1,16E-12
	<i>centr_eigen</i>	6,263	0,610	-4,48E-16	3,79E-10
	<i>coef_var_taille_comm</i>	3,234	0,315	-2,22E-16	1,22E-03
	<i>share_unequivalent</i>	-3,557	-0,346	-2,86E-17	3,75E-04
	<i>nb_comp</i>	-3,736	-0,364	1,07E-18	1,87E-04
	<i>density</i>	-6,416	-0,624	1,06E-17	1,40E-10
	<i>clustering_global</i>	-8,845	-0,861	-1,02E-16	9,15E-19

Figure 41 : Distribution graphique des classes 1,2,3 et 4



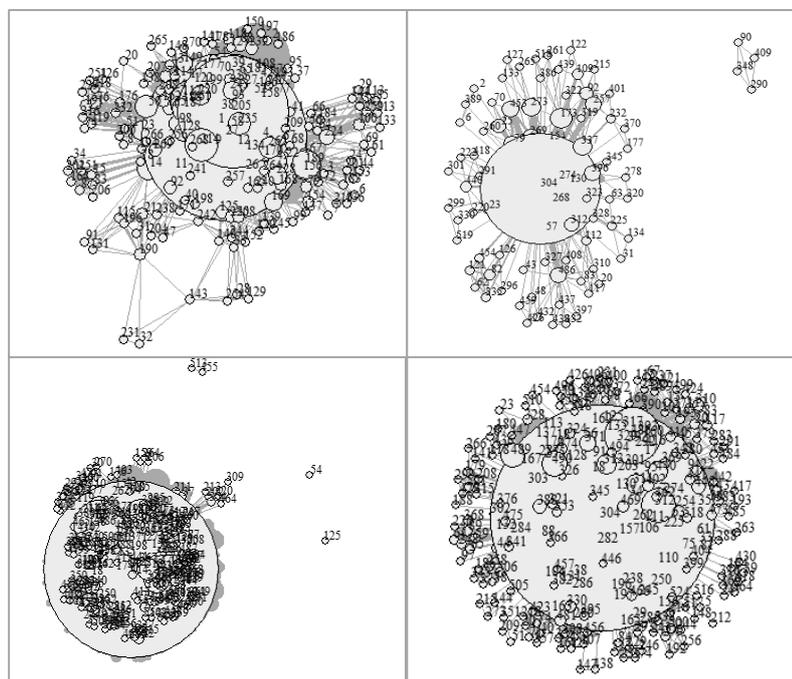
Le réseau de la classe 4 dispose d'une synergie particulière : le réseau **grave autour de quelques co-auteurs centraux**.

A l'échelle locale, le réseau est très structuré par **l'hétérogénéité des relations d'ego avec ses alter**. La classe 4 est très caractérisée par la dimension 1 : ses réseaux se trouvent à droite de l'axe. L'hétérogénéité du nombre de publications qu'ego échange avec chaque alter est très supérieure à la moyenne (*coef_var_pub_ego*). Le réseau est composé de deux types d'alter : quelques co-auteurs centraux avec qui ego collabore beaucoup (le diamètre de ces nœuds est très fort) et le reste du réseau

avec qui il ne collabore que ponctuellement (le diamètre de ces nœuds est plus petit). Les **co-auteurs centraux** accompagnent *ego* dans l'ensemble de ses liens collaboratifs dans des sous-espaces collaboratifs différents (*centr_between*) et avec l'ensemble du réseau (*centr_degree*). En parallèle, les autres **alter gravitent autour de ces alter centraux dans un même composant** (*share_comp_princ*, *nb_comp*). Contrairement aux *alter* centraux, le reste du réseau ne dispose pas de position unique et non substituable dans le réseau : ils partagent la même relation collaborative avec *ego* que les autres membres de l'équipe de publication (*share_unequivalent*). Le voisinage des *alter* est moins dense que la moyenne (*clustering_global*) : seuls les *alter* centraux sont au cœur des liens collaboratifs.

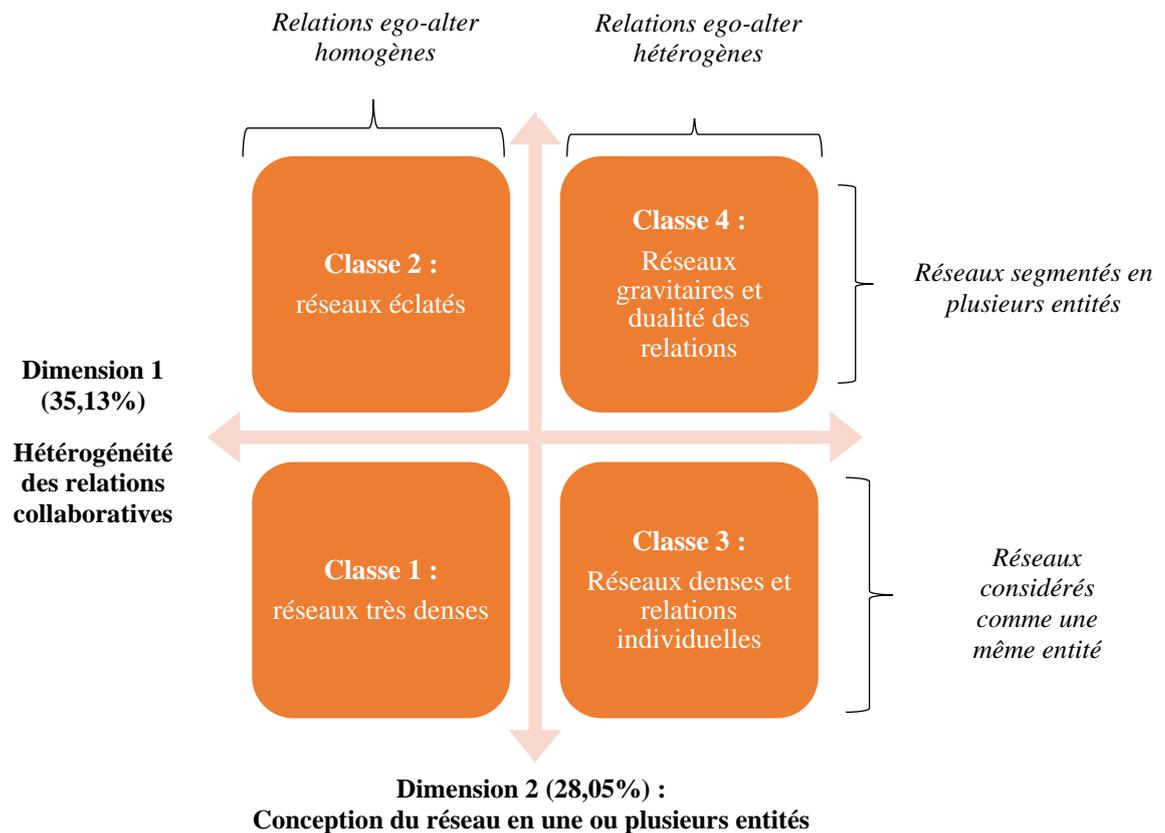
A l'échelle globale, comme la classe 2, les réseaux de la classe 4 ont une densité plus faible que la moyenne (*density*) : **ego compartimente ses relations**. Nous distinguons deux types de sous-espaces collaboratifs hétérogènes (*coef_var_taille_comm*, *intensivity_comm_pub*). Le premier concerne les sous-espaces qui incluent les *alter* centraux : ces sous-espaces sont très gros et se développent au fil de l'activité de publication d'*ego*. Les publications qui mobilisent conjointement les *alter* centraux sont limités à la frontière de ces sous-espaces-là : le reste du réseau en est exclus (*centr_eigen*). Les autres sous-espaces collaboratifs sont beaucoup plus petits et gravitent autour des co-auteurs centraux.

Figure 42 : Réseaux caractéristiques de la classe 4 (les plus éloignés du centre de gravité des autres classes)



Au vue de ces différentes caractérisations, nous pouvons synthétiser ces quatre classes relativement aux deux axes de l'ACP avec la matrice suivante :

Figure 43 : Matrice synthétique de caractérisation des classes



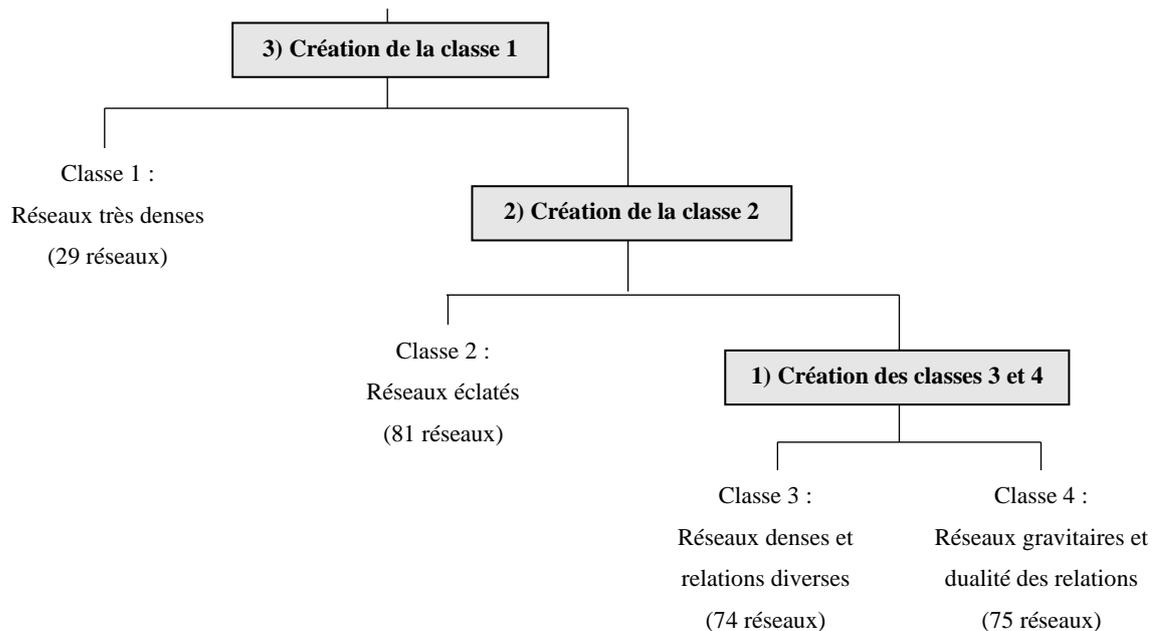
Néanmoins, la description de ces quatre classes ne peut se limiter à la caractérisation par les 15 variables mobilisées. En effet, les classes ne sont pas indépendantes les unes des autres mais résultent d'un processus progressif de regroupement. Ainsi, la caractérisation doit être complétée par une compréhension des dernières étapes de classification afin de comprendre ce qui oppose et ce qui réunit les classes entre elles.

B) Le processus de classification

L'objectif de cette dernière sous-section du chapitre est de **comprendre ce qui justifie la création de chacune des classes**. En effet, chaque étape de l'algorithme a eu pour but de favoriser l'identité de la classe (réduire l'inertie intra-classe) et d'accentuer la distance entre les classes (accroître l'inertie inter-

classe). Grâce à la condition de proximité intra-classe, il nous est possible de comparer les caractères structurels des différentes classes. En revanche, nous ne pouvons pas encore expliquer pourquoi **telle classe a été créée en opposition aux autres**. Nous allons donc reconstruire les dernières étapes de l'algorithme afin de comprendre dans quelle mesure les différents types de structure de réseau sont liés entre eux.

Figure 44 : Recontextualisation des créations de classe dans le processus de classification



Pour rappel, notre classification hiérarchique ascendante est partie de 259 réseaux distincts : l'algorithme les a réunis un à un pour ne former qu'une seule entité à la fin. Le graphique ci-dessus nous décrit les trois étapes que nous allons analyser :

i) Etape 1 : **La création de la classe 3 « Réseaux denses et relations diverses »** à la suite du regroupement de 74 réseaux et **la création de la classe 4 « Réseaux gravitaires et dualité des relations »** à la suite du regroupement de 75 réseaux. A cette étape, les 110 réseaux restants ne sont pas encore regroupés dans les classes définitives.

ii) Etape 2 : **La création de la classe 2 « Réseaux éclatés »** à la suite du regroupement de 81 réseaux parmi les 110 non classés. A cette étape, les 29 réseaux restants ne sont pas encore regroupés dans les classes définitives.

iii) Etape 3 : **La création de la classe 1 « Réseaux très denses »** à la suite au regroupement des 29 réseaux restants. A l'issus de cette étape, les 4 classes sont créés.

1- Contextualisation de la création des classes 3 et 4

Nous commencerons par contextualiser la création simultanée des classes 3 et 4. A cette étape, l'algorithme fusionne 74 réseaux distincts dans la classe 3, fusionne 75 réseaux dans la classe 4 et 110 réseaux ne sont toujours pas classés (futures classes 1 et 2). Nous voudrions comprendre en quoi ces deux classes se distinguent entre elles et avec les 110 réseaux restants non classés.

On lit plusieurs **différences structurelles qui opposent la classe 3 à la classe 4**. Ces différences concernent la part d'*alter* disposant d'une position particulière dans le réseau (*share_unequivalent*), la centralisation de vecteur propre (*centr_eigen*), la densité (*density*), l'hétérogénéité de la taille des sous-espaces collaboratifs (*coef_var_taille_comm*) et le nombre moyen de publications réalisées avec chaque sous-espace (*intensity_comm_pub*).

i) Les réseaux de la **classe 3** sont **plus denses que la moyenne** : tous les *alter* accompagnent *ego* dans ses publications où il collabore conjointement avec ses *alter* centraux. Aussi, *ego* entretient une **relation individuelle et particulière** avec ses *alter* qui ont donc des positions particulières dans le réseau. Enfin, les **sous-espaces collaboratifs sont peu structurants** : ils sont de même taille et *ego* collabore peu avec chacun d'entre eux.

ii) En comparaison, les réseaux de la **classe 4** sont **moins denses que la moyenne** : seulement quelques *alter* accompagnent *ego* dans ses publications où collabore conjointement avec ses *alter* centraux. **Très peu d'*alter* entretiennent une relation individuelle et particulière avec *ego***. Les **sous-espaces collaboratifs sont très structurants** dans le réseau : ils sont de taille différente et *ego* collabore beaucoup avec certains d'entre eux.

Néanmoins, ces **deux classes partagent des similitudes structurelles** qui permettent de les distinguer des autres réseaux non classés. Ces similitudes concernent l'hétérogénéité de récurrence dans les travaux d'*ego* (*coef_var_pub_ego*), la centralisation de degré et d'intermédierité (*centr_degree*, *centre_between*), la décomposition en composants connexes (*nb_comp*, *share_comp_princ*) et la densité du voisinage d'*alter* (*clustering_global*).

i) Dans les réseaux des **classes 3 et 4**, certains *alter* sont beaucoup plus récurrents que d'autres dans la bibliométrie d'*ego*. Ces ***alter* centraux sont au cœur des liens collaboratifs d'*ego*** : ils l'accompagnent dans les collaborations avec le réseau entier et avec des sous-espaces collaboratifs différents. C'est pour cette raison que les **réseaux sont très peu décomposés** : tous les *alter* sont concentrés autour d'un même composant autour des co-auteurs centraux.

ii) En revanche, les **110 réseaux non classés** ne partagent pas ces caractéristiques. Chaque *alter* du réseau réalise le même nombre de publication avec *ego* et **tous les *alter* sont au cœur des liens**

collaboratifs dans les composants. Le réseau n'est pas structuré autour de quelques *alter* qui seraient récurrents. Le réseau est **segmenté en composants** avec qui *ego* collabore en toute autonomie.

2- Contextualisation de la création de la classe 2

À la suite de la création des classes 3 et 4, **l'algorithme fusionne 81 des 110 réseaux restants dans la classe 2.** Nous avons déjà expliqué préalablement pourquoi les réseaux de la classe 2 ne peuvent fusionner avec les classes 3 et 4. En revanche, nous ne savons pas encore pourquoi les 29 réseaux restants non classés ne peuvent fusionner avec la classe 2. Sachant décrire le réseau « type » de la classe 2, nous pouvons le comparer avec les autres réseaux restants. Plusieurs caractéristiques opposent cette classe avec les 29 autres réseaux non classés, à savoir la centralisation de vecteur propre (*centr_eigen*), la densité (*density*), la modularité (*modularity*) et l'hétérogénéité des tailles des sous-espaces collaboratifs (*coef_var taille_comm*).

i) Les réseaux de la **classe 2** sont moins denses que la moyenne. Les **réseaux étant éclatés**, seuls quelques *alter* accompagnent *ego* dans ses publications où il collabore conjointement avec ses *alter* centraux. Les **sous-espaces collaboratifs sont très denses en interne et isolés avec le reste du réseau** : *ego* collabore indépendamment avec chacun d'entre eux. Ils sont très structurants dans les liens collaboratifs puisqu'ils sont de taille différente.

ii) Les **29 autres réseaux non classés** sont **plus denses** que la moyenne : tous les *alter* accompagnent *ego* dans les publications où il collabore conjointement avec ses *alter* centraux. Les **sous-espaces collaboratifs sont de faible qualité** puisqu'*ego* collabore simultanément avec eux. Ils sont très peu structurants car ils sont de taille homogène.

3- Contextualisation de la création de la classe 1

Enfin, la troisième étape concerne la réunion des réseaux restants dans la classe 1. En supposant que le critère de proximité au sein de la classe soit appliqué, nous voudrions comprendre la particularité structurelle de ces réseaux relativement aux autres classes : les réseaux de la classe 1 sont-ils suffisamment éloignés des autres classes ?

i) Les réseaux de la **classe 1 sont très denses** : ils sont caractérisés positivement par la densité à l'échelle du réseau (*densité*) et des *alter* (*clustering_global*). Ainsi, les réseaux de la classe 1 sont plus denses que la moyenne à la fois à **l'échelle globale et locale.**

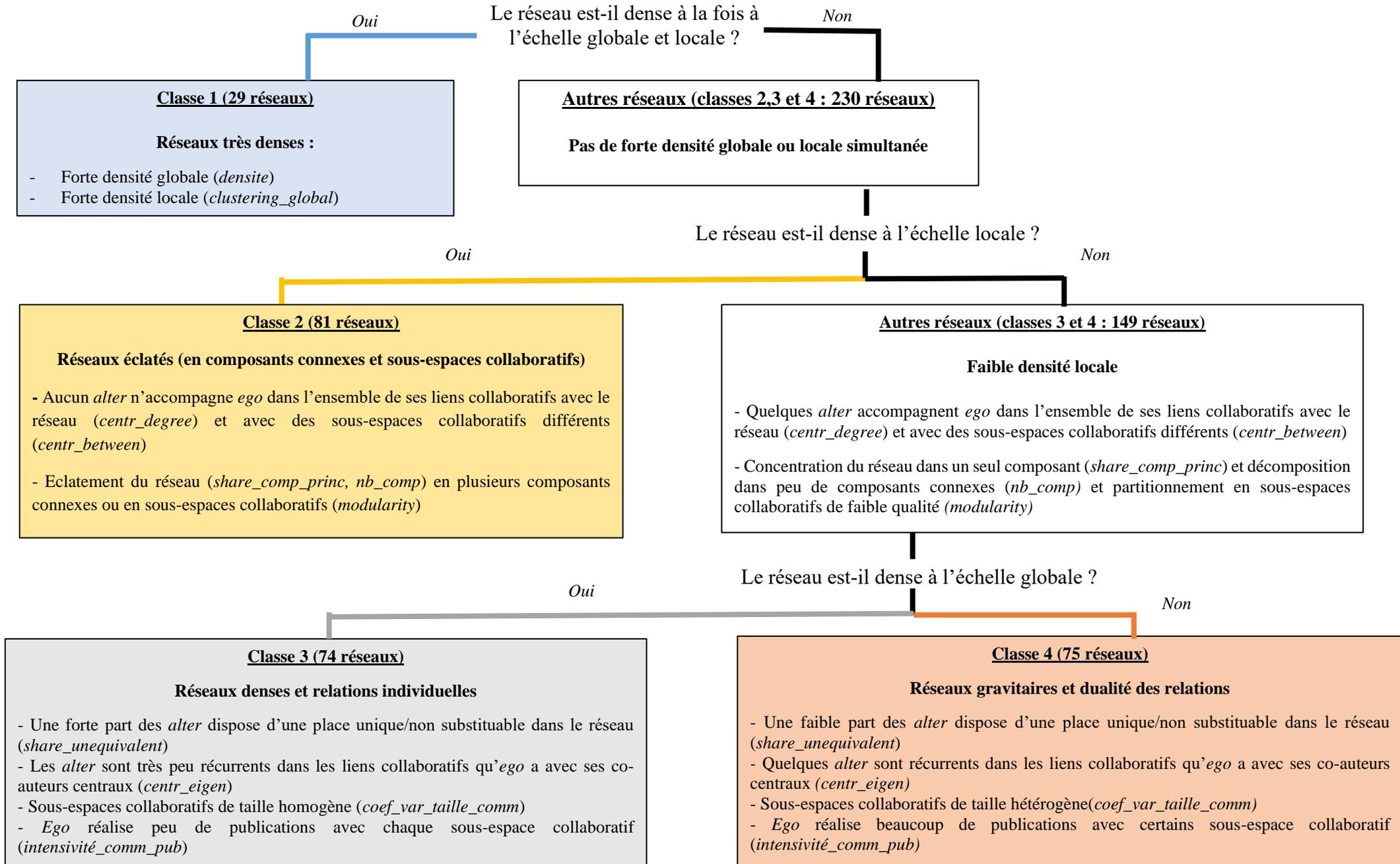
ii) Cela n'est pas le cas pour **les autres classes** : aucune n'est caractérisée à la fois par une forte **densité interne et globale.** La classe 2 est caractérisée par une forte densité locale mais faible à l'échelle globale. A l'inverse, la classe 3 est caractérisée par une faible densité locale mais forte à l'échelle

globale. Enfin, la classe 4 est caractérisée par une faible densité, aussi bien à l'échelle globale que locale. La classe 1 réunit donc un cas particulier de réseaux : ceux qui sont très denses.

Le schéma ci-dessous synthétise les différentes étapes du processus de création de classes.

La prochaine partie aura pour but d'expliquer la probabilité d'avoir chacune de ces structures de réseau en fonction de données de carrière d'*ego*. Nous testerons des variables explicatives relatives à la temporalité de la carrière, à la description d'*alter* et à la dynamique de carrière.

Figure 45 : Reconstruction du dendrogramme et du processus de classification



III- Styles de collaboration et carrière scientifique

L'objectif de cette dernière partie est d'expliquer la typologie de structure des réseaux à partir de données relatives à la carrière d'*ego*. Nous souhaiterions comprendre **dans quelle mesure la carrière d'*ego* influence la structure de son réseau et donc son style collaboratif**. Nous procéderons en deux étapes :

i) Dans un premier temps, nous estimerons la classe de chacun des 259 réseaux en fonction de la description du type d'*alter* avec qui *ego* collabore. **L'objectif est de tester si chaque type de réseau est dû à un co-authorship particulier.**

ii) Dans un second temps, nous recontextualiserons chaque réseau dans la dynamique de carrière d'*ego*. L'objectif ici n'est plus de tester les 259 réseaux comme des entités indépendantes les unes des autres mais seulement les 179 réseaux à compter du stade de « chercheur confirmé » qui serviront à reconstruire les trajectoires de carrière des 80 chercheurs. Pour chacun des 179 réseaux, **nous estimerons la classe de chacun d'entre eux en fonction sa dynamique structurelle et l'évolution de l'activité de publication d'*ego*.**

Dans ces deux approches, nous testerons également des variables explicatives relatives à la temporalité de la carrière d'*ego* (stade de carrière et génération) et des variables de contrôle (relatives au genre et au type de carrière suivie par *ego*).

Nous commencerons par dresser les statistiques descriptives de chaque classe afin d'avoir une première appréhension de la structure des réseaux en fonction des caractéristiques d'*ego* et de la diversification du co-authorship. Ensuite, nous présenterons la méthodologie utilisée et les résultats pour chaque approche économétrique.

A) Statistiques descriptives des quatre classes

Ces statistiques descriptives servent de résultats préliminaires à la caractérisation des styles de collaboration. Ainsi, nous montrerons les quatre styles collaboratifs sont caractérisés par des **caractéristiques d'*ego*** et par des comportements de **diversification du co-authorship** particuliers.

1- Structure du réseau et caractéristiques d'*ego*

Tableau 41 : Répartition des 259 réseaux en fonction de leur structure et des caractéristiques d'*ego*

		Classe 1	Classe 2	Classe 3	Classe 4	Total
Stade de carrière d'<i>ego</i>	Jeune chercheur	19	21	31	9	80
	Chercheur confirmé	6	21	20	33	80
	Senior	2	26	18	19	65
	Senior avancé	2	13	5	14	34
Génération d'<i>ego</i>	1967-1976	5	21	10	16	52
	1977-1986	8	16	20	26	70
	1987-1996	12	24	31	22	89
	1997-2006	4	20	13	11	48
Genre d'<i>ego</i>	Femme	7	18	30	17	72
	Homme	22	63	44	58	187
Type de carrière d'<i>ego</i>	Directeur de recherche	7	30	35	46	118
	Professeur	22	51	39	29	141
	Total	29	81	74	75	259

Tout d'abord, nous notons que **chaque stade de carrière a des particularités structurelles.**

i) La classe 1 est surreprésentée au stade de « **jeune chercheur** » : en effet, parmi les 29 de cette structure, 19 sont associés à la première décennie de carrière d'*ego*. Ainsi, avoir un réseau très dense doté de **relations collaboratives homogènes** concerne plutôt le début de carrière durant lequel *ego* a peu d'expérience professionnelle et un faible rythme de publication. D'autres structures sont également caractéristiques de ce stade de carrière : parmi les 80 jeunes chercheurs, 31 ont un réseau dense aux relations individuelles (classe 3) et 21 ont un réseau éclaté (classe 2).

ii) Ensuite, à partir de la seconde décennie de carrière, la structure des réseaux est différente. Parmi les 80 **chercheurs confirmés**, 33 disposent d'un réseau gravitant autour d'*alter* centraux dans la classe 4. Ainsi, à partir de ce stade, le réseau de co-authorship se structure davantage autour de **l'hétérogénéité des relations** : en plus des réseaux de la classe 4, 20 des 80 chercheurs confirmés ont un réseau dense aux relations individuelles (classe 3).

iii) Par la suite, la classe 2 et la classe 4 sont les structures majoritaires au sein des réseaux de « **senior** » et « **senior avancé** » : la **fin de carrière d'*ego*** se traduit donc par une **forte segmentation des liens collaboratives** avec des relations collaboratives plus ou moins hétérogènes.

Nous notons également des **particularités structurelles entre les différentes générations**.

i) Bien que leurs durées de carrière soient très différentes et qu'ils publient durant différents contextes, les chercheurs des **générations « 1967-1976 » et « 1997-2006 »** ont plus de réseaux éclatés de classe 2. Ces générations sont plus à même à **segmenter leurs liens collaboratifs** et à entretenir des **relations homogènes** avec leurs *alter*.

ii) Ensuite, parmi les 70 réseaux issus de la **génération « 1977-1986 »**, 26 sont issus de la classe 4 et 20 de la classe 3 : cette génération semble entretenir des **relations hétérogènes** avec leurs *alter*, avec une densité globale plus ou moins forte.

iii) Enfin, parmi les 89 réseaux de la **génération « 1987-1996 »**, 31 sont issus de la classe 3 : ils sont considérés comme une **même entité dotée de relations ego-alter particulières**.

La structure du réseau évolue également en fonction du **genre d'ego**. En effet, parmi les 72 réseaux de femmes, 30 appartiennent à la classe 3. En revanche, parmi les 187 réseaux d'hommes, 63 d'entre eux sont éclatés (classe 2) et 58 sont gravitaires (classe 4). Ainsi, les **femmes** ont tendance à avoir des **réseaux denses** alors que les hommes sont davantage impactés par la segmentation des liens collaboratifs, quel que soit l'hétérogénéité des relations collaboratives.

Enfin, nous notons une forte concentration de **réseaux éclatés** au sein des **carrières de professeur** : parmi les 141 réseaux de ce type de carrière, 51 appartiennent à la classe 2. Alors que nous ne notons pas de spécialisation structurelle particulière auprès des carrières de directeur de recherche, nous notons une certaine appétence des professeurs à segmenter leurs liens collaboratifs et à entretenir des **relations homogènes avec leurs alter**.

2- Structure du réseau et diversité des *alter*

Nous pouvons caractériser chaque classe de réseau en fonction des habitudes collaboratives du chercheur en termes de **diversification sociale et géographique du co-authorship**.

Structure du réseau et diversification sociale : Nous constatons que la **diversification sociale** du réseau (la collaboration avec de nouveaux co-auteurs) est liée à **l'hétérogénéité des relations sociales ego-alter** dans le réseau.

i) En effet, les classes caractérisées par des **relations homogènes** (à gauche de l'axe 1) sont caractérisées par **une forte diversification des relations collaboratives** : en moyenne, 55,19% des collaborations de la classe 1 sont issues de nouvelles relations et c'est le cas pour 46,20% des collaborations de la classe 2. Ainsi, l'homogénéité des relations semble être induite par des **relations collaboratives très courtes et ponctuelles** : plus *ego* collabore avec de nouvelles relations, moins son réseau est structuré autour de co-auteurs centraux.

Tableau 42 : Diversification sociale et géographique des collaborations au sein de chaque réseau

	Renouvellement des collaborations		Proximité géographique et organisationnelle des collaborations		Internationalisation des collaborations	
	Part des nouvelles collaborations	Part des collaborations renouvelées	Part des collaborations avec le laboratoire	Part des collaborations en dehors du laboratoire	Part des collaborations internationales	Part des collaborations avec le pays <i>d'ego</i>
Classe 1	55,19%	44,81%	63,04%	36,96%	8,81%	91,19%
Classe 2	46,20%	53,80%	50,69%	49,31%	23,56%	76,44%
Classe 3	37,81%	62,19%	62,86%	37,14%	14,98%	85,02%
Classe 4	33,22%	66,78%	54,50%	45,50%	21,73%	78,27%
Total	41,05%	58,95%	56,66%	43,34%	18,93%	81,07%

ii) En revanche, les classes caractérisées par **l'hétérogénéité des relations collaboratives** (à droite de l'axe 1) sont caractérisées par **le maintien des relations collaboratives** : en moyenne, 62,19% des collaborations de la classe 3 sont issues de relations renouvelées et c'est le cas pour 66,78 % des collaborations de la classe 4. Ainsi, l'hétérogénéité des relations collaboratives est dû à la poursuite des relations collaboratives, **certaines étant plus longues que d'autres** : plus *ego* collabore avec d'anciennes relations, plus son réseau est structuré autour de co-auteurs centraux.

Structure du réseau et diversification géographique : Nous constatons que la **diversification géographique** du réseau (la collaboration avec de nouvelles villes) est liée à **la densité** dans le réseau.

i) Les classes où les **réseaux sont caractérisés comme très denses** (en bas de l'axe 2) sont caractérisées par des collaborations à la **dimension géographique restreinte**. En effet, en moyenne, 63,04% des collaborations de la classe 1 sont issues du laboratoire d'*ego* et seulement 8,81% des collaborations sont internationales. Aussi, en moyenne, 62,86% des collaborations de la classe 3 sont issues du laboratoire d'*ego* et seulement 14,98% des collaborations sont internationales. Ainsi, plus ***ego* collabore localement, plus son réseau est dense et considéré comme une même entité.**

ii) En revanche, les classes dont les réseaux sont considérés **comme segmentés en plusieurs entités** (en haut de l'axe 2) sont caractérisées par des collaborations à la **dimension géographique plus étendue**. En effet, en moyenne, 49,31% des collaborations de la classe 2 sont issues en dehors du laboratoire d'*ego* et 23,56% d'entre elles sont internationales. Aussi, en moyenne, 45,50% des collaborations de la classe 4 sont issues en dehors du laboratoire d'*ego* et 21,73% d'entre elles sont internationales. Ainsi, **plus *ego* collaborera avec des *alter* éloignés géographiquement de lui-même, plus son réseau est segmenté en sous-entités.**

Ces premiers résultats soulignent que la **structure du réseau** dépend des caractéristiques d'*ego* et de la diversification du réseau. Ils soulèvent l'intérêt d'évaluer le style collaboratif du chercheur (en termes de classe de réseau) en fonction de données relatives à sa carrière.

B) Estimation du style collaboratif en fonction de la temporalité de carrière et de la description du co-authorship

Cette première approche économétrique a pour but **d'estimer le comportement collaboratif du chercheur** en fonction de données relatives à **la temporalité de sa carrière** et à **la composition de son co-authorship**. Nous présentons la méthodologie utilisée et les résultats qui en résultent.

1- Présentation des modèles

L'objectif est d'identifier les variables qui favorisent significativement la probabilité d'avoir une structure de réseau particulière. Nous avons réalisé plusieurs modèles de régression logistique binaire dont nous avons estimé, toute chose égale par ailleurs, les effets marginaux.

Nous testons quatre variables dépendantes nommées $Classe_j$ pour tout $j \in \{1;4\}$:

$$Classe_j \begin{cases} = 1 \text{ si } Classe_j^* \geq 0 \\ = 0 \text{ sinon} \end{cases}$$

$$\text{Dont : } Classe_j^* = \alpha + Z_j \beta + \mu_j$$

Dans ces modèles de régression, chaque variable expliquée $Classe_j$ est dichotomique : elle prend la valeur 1 si le réseau appartient à la classe j et 0 si ce n'est pas le cas. Également, la variable $Classe_j^*$ est la variable latente : elle est décrite en fonction de la matrice des variable explicatives Z_j et du vecteur des paramètres associés β . La probabilité d'appartenance à la classe j est donc :

$$P (Classe_j = 1 | Z_j) = \frac{\exp (\alpha + Z_j \beta)}{1 + \exp (\alpha + Z_j \beta)}$$

Tableau 43 : Variables dépendantes : Classej

	<i>Classe_j = 1</i>	<i>Classe_j = 0</i>	<i>Total</i>
<i>Classe₁</i>	29	230	259
<i>Classe₂</i>	81	178	259
<i>Classe₃</i>	74	185	259
<i>Classe₄</i>	75	184	259

Nous remarquons que les quatre variables $Classe_j$ n'ont pas la même probabilité d'avoir 1 comme valeur. En effet, seulement 29 des 259 réseaux sont très denses et appartiennent à la classe 1 ($Classe_1 = 1$) : 230 réseaux ne cumulent pas une forte densité globale et locale ($Classe_1 = 0$). Les réseaux éclatés de la classe 2 pèsent 31,3% de nos réseaux ($Classe_2 = 1$). Aussi 149 réseaux disposent de relations *ego-alter* hétérogènes : 74 d'entre eux sont des réseaux denses aux relations hétérogènes ($Classe_3 = 1$) et 75 sont des réseaux gravitaires et d'une dualité de relations ($Classe_4 = 1$).

Chacune de ces quatre variables dépendantes sera testée par trois sous-modèles en fonction des variables explicatives utilisées.

i) Le **premier sous-modèle** se focalise sur les variables explicatives relatives à la **temporalité de la carrière d'ego**. Ainsi, nous testerons donc l'effet **du stade de carrière** (*stade_de_carriere*) pour vérifier si certaines structures de réseaux dépendent du niveau d'expérience professionnelle du chercheur. Aussi, nous testerons l'effet de la **génération** du chercheur (*generation*) pour voir si la structure du réseau est déterminée par la période durant laquelle *ego* a commencé sa carrière. Ces deux variables ont pour objectif de contextualiser le réseau dans la carrière d'*ego* sans décrire la trajectoire qu'il a eu auparavant. Nous nous limiterons à **identifier le contexte structurel et conjoncturel dans lequel chaque réseau est issu**.

ii) Le **second sous-modèle** se focalise sur les variables relatives à la **description d'alter**. Elles contextualisent la structure du réseau en fonction de **données individuelles** relatives à chacun des *alter*. Trois types de variables sont mobilisés :

- Les premières variables décrivent à quel point les **caractéristiques d'alter structurent les collaborations d'ego**. Nous évaluons le poids **des hommes** (*collab_hommes*) et des **one-shot** (*collab_one_shot*) dans l'ensemble des collaborations d'*ego*. Respectivement, une faible valeur de ces variables indique une forte proportion d'*alter* femmes et de relations durables dans les collaborations. Aussi, nous évaluons le poids des **doctorants en formation** dans les collaborations (*collab_doctorants*) : un co-auteur est jugé en formation de doctorat dans le sens où la collaboration a lieu au maximum durant son année de soutenance. Enfin, nous évaluons à quel point **les relations hiérarchiques entre ego et alter** structurent les collaborations d'*ego*. La variable *collab_doctorants_ego* évalue le poids des **alter qui sont ou ont été sous la supervision d'ego durant leur doctorat** et la variable *collab_directeur* évalue le poids des publications dans lesquels *ego* a collaboré avec **son ou ses anciens directeurs de thèse**. L'objectif est d'identifier à quel point la proximité durant la réalisation d'une thèse (par *ego* ou *alter*) détermine la structure actuelle et future du réseau de co-authorship d'*ego*.

- Les secondes variables **décomposent l'équipe moyenne de publication** du réseau en fonction de la géographie des *alter*. L'objectif est d'évaluer l'impact de la **dimension géographique moyenne des projets collaboratifs** auprès de la structure du réseau. La variable *equipe_locaux* décrit le nombre moyen d'*alter* issus du **laboratoire ou de la ville d'ego par publication**. Concernant les *alter* exclus de la ville d'*ego*, la variable *equipe_nationaux* décrit le nombre **moyen d'alter du pays d'ego** par

publication et *equipe_internationaux* décrit le nombre **moyen d'alter hors du pays d'ego** par publication.

- Enfin, les troisièmes variables concernent la **durée des relations** qu'*ego* entretient avec chacun de ses *alter* **en fonction de la géographie d'alter**. L'objectif est de tester si la structure du réseau dépend de la force de la relation collaborative qu'*ego* entretient avec certains *alter*. La variable *publis_locaux* décrit le nombre moyen de publications qu'*ego* réalise avec chacun des *alter* de son laboratoire ou de sa ville. Les variables *publis_nationaux* et *publis_internationaux* décrivent le nombre moyen de publications qu'*ego* réalise avec chacun de ses *alter* nationaux ou internationaux.

iii) Le **troisième sous-modèle** capte **l'effet simultané** des variables explicatives relatives à la **carrière et aux caractéristiques des alter**. Après avoir testé séparément ces deux types de variables, nous voudrions tester leur impact dans un même modèle. Nous avons séparé ces variables dans les deux premiers sous-modèles puisque certaines variables relatives à la description du co-authorship sont corrélées à la temporalité de la carrière d'*ego*. Par exemple, le poids des publications dans lesquelles le directeur de thèse est présent décroît au cours des carrières : il est de 16,7% dans les réseaux des « jeunes chercheurs » et inférieur à 4% à partir du stade de « chercheur confirmé » : *ego* s'émancipe peu à peu de la présence de son directeur de thèse dans ses publications. Aussi, le poids des collaborations faites avec des hommes décroît également au fil de la carrière : les hommes pèsent en moyenne 80,3% des collaborations des « jeunes chercheurs » et ne pèsent que 64,5% dans les réseaux de « seniors avancés ». Ceci est notamment dû au fait que les femmes sont de plus en plus présentes dans le co-authorship des chimistes ce qui incite les chercheurs à diversifier le genre de leur alter au fil de leur carrière.

Tableau 44 : Rapports de corrélation élevés entre la description du co-authorship et le stade de carrière

	Rapport de corrélation avec le stade de carrière d'ego
<i>collab_hommes</i>	0,383
<i>collab_directeur</i>	0,503
<i>collab_doctorants_ego</i>	0,190

Par conséquent, en croisant les différentes variables explicatives et expliquées, nous estimons 12 modèles synthétisés dans le tableau suivant. En plus des variables explicatives, des variables de contrôle seront également testées dans chacun des modèles : il s'agit du genre d'*ego* (*genre*) et du type de carrière qu'il suit (*type_carriere*), à savoir une carrière de directeur de recherche ou de professeur.

Tableau 45 : Synthèse des modèles estimés

Variables expliquées :	<i>Classe₁</i>	<i>Classe₂</i>	<i>Classe₃</i>	<i>Classe₄</i>
Variables explicatives :				
<i>Temporalité de la carrière</i>	Modèle 1a.	Modèle 2a.	Modèle 3a.	Modèle 4a.
<i>Description des co-auteurs</i>	Modèle 1b.	Modèle 2b.	Modèle 3b.	Modèle 4b.
<i>Temporalité de la carrière et description des co-auteurs</i>	Modèle 1c.	Modèle 2c.	Modèle 3c.	Modèle 4c.

Le tableau suivant présente les statistiques descriptives relatives aux variables expliquées, explicatives et de contrôle.

Tableau 46 : Statistiques descriptives des variables

			N	Moyenne	Écart type	Min	Max
Variables expliquées							
	Classe ₁	Le réseau appartient à la classe 1 « réseaux très denses »	259	0,112		0	1
	Classe ₂	Le réseau appartient à la classe 2 « réseaux éclatés »	259	0,313		0	1
	Classe ₃	Le réseau appartient à la classe 3 « réseaux denses et relations individuelles »	259	0,286		0	1
	Classe ₄	Le réseau appartient à la classe 4 « réseaux gravitaires et dualité des relations »	259	0,290		0	1
Variables explicatives							
1- Temporalité des carrières							
<i>Stade_de_carriere_</i>	<i>Jeune_chercheur</i>	Le réseau est celui d'un jeune chercheur	80	0,309		0	1
	<i>Chercheur_confirme</i>	Le réseau est celui d'un chercheur confirmé	80	0,309		0	1
	<i>Senior</i>	Le réseau est celui d'un senior	65	0,251		0	1
	<i>Senior_avance</i>	Le réseau est celui d'un senior avancé	34	0,131		0	1
<i>Generation</i>	<i>1967_76</i>	Le réseau appartient à la génération "1967-1976"	52	0,201		0	1
	<i>1977_86</i>	Le réseau appartient à la génération "1977-1986"	70	0,270		0	1
	<i>1987_96</i>	Le réseau appartient à la génération "1987-1996"	89	0,344		0	1
	<i>1997_06</i>	Le réseau appartient à la génération "1997-2006"	48	0,185		0	1
2- Description du co-authorship							
Décomposition des collaborations	<i>collab_hommes</i>	Part des hommes dans les collaborations	259	0,333	0,731	0,122	1,000
	<i>collab_one_shot</i>	Part des one-shot dans les collaborations	259	0,000	0,223	0,107	0,778
	<i>collab_doctorants</i>	Part des doctorants dans les collaborations	259	0,000	0,144	0,077	0,529
	<i>collab_doctorants_ego</i>	Part des doctorants sous la supervision d' <i>ego</i> dans les collaborations	259	0,000	0,077	0,104	0,588
	<i>collab_directeur</i>	Part des publications où le directeur de thèse d' <i>ego</i> est présent	259	0,000	0,239	0,345	1,000
Membres de l'équipe de publication	<i>equipe_locaux</i>	Nombre moyen d' <i>alter</i> locaux par publication	259	0,000	0,739	0,879	9,143
	<i>equipe_nationaux</i>	Nombre moyen d' <i>alter</i> nationaux par publication	259	0,240	2,637	1,017	6,652
	<i>equipe_internationaux</i>	Nombre moyen d' <i>alter</i> internationaux par publication	259	0,000	0,877	0,811	4,303
Durée des relations collaboratives	<i>publis_locaux</i>	Nombre moyen publications réalisées par <i>ego</i> et chaque de ses <i>alter</i> locaux	259	1,000	2,795	1,144	7,839
	<i>publis_nationaux</i>	Nombre moyen publications réalisées par <i>ego</i> et chaque de ses <i>alter</i> nationaux	259	0,000	1,372	0,654	4,000
	<i>publis_internationaux</i>	Nombre moyen publications réalisées par <i>ego</i> et chaque de ses <i>alter</i> internationaux	259	0,000	1,354	0,741	6,000
Variables de contrôles							
<i>Genre</i>	<i>Femme</i>	Le réseau est celui d'une femme	72	0,278		0	1
	<i>Homme</i>	Le réseau est celui d'un homme	187	0,722		0	1
<i>Type_carriere</i>	<i>Directeur de recherche</i>	Le réseau est celui d'un directeur de recherche	118	0,456		0	1
	<i>Professeur</i>	Le réseau est celui d'un professeur	141	0,544		0	1

2- Les résultats

Nous pouvons à présent interpréter les résultats pour identifier les déterminants de la structure du réseau et donc le style collaboratif. Pour chaque modèle, nous hiérarchisons les facteurs en commençant par décrire ceux qui ont la plus forte significativité et le plus fort effet marginal.

Facteurs qui déterminent la probabilité d'avoir un réseau très dense (classe 1)

Pour rappel, la classe 1 est caractérisée par sa **forte densité locale et globale**. Dans ce type de structure, tous les *alter* sont au cœur des liens collaboratifs et sont donc très connectés en interne. Tous les *alter* ont une forte probabilité d'accompagner *ego* dans ses liens collaboratifs, que ce soit à l'échelle du réseau entier, avec des sous-espaces collaboratifs différents ou avec les *alter* centraux. *Ego* entretient des **relations très homogènes** avec eux. Il dispose d'une certaine autonomie dans la collaboration avec ses sous-espaces.

Concernant les modèles 1a, 1b et 1c, plusieurs facteurs sont significativement identifiables.

i) Le premier est le **poids de one-shot** dans les collaborations d'*ego* : l'effet marginal est de -4 points dans le modèle 1b et de -3,5 points dans le modèle 1c. La collaboration avec ces co-auteurs ponctuels réduit la probabilité d'avoir un réseau aux relations homogènes et considéré comme une même entité : les one-shot se **retrouvent isolés du reste du réseau**.

ii) Ensuite, la probabilité d'avoir ce type de réseau **décroit également au fil de la carrière**. Parmi les 29 réseaux inclus dans la classe 1, 19 concernent le stade de « jeune chercheur ». En effet, dans le modèle 1a, être à un autre stade de carrière que « jeune chercheur » réduit la probabilité d'au moins 1,7 points : lorsqu'*ego* est jeune, son **rythme de publication est faible** et ses **relations collaboratives sont encore homogènes**.

iii) Nous remarquons que le **maintien de relation** avec *alter* réduit également la probabilité d'avoir cette structure de réseau, **quelle que soit la géographie d'*alter***. *Ego* réalise donc peu de publications avec chacun de ses *alter* (effet marginal de -1.2 pour les *alter* locaux dans le modèle 1b).

iv) Enfin, avoir un réseau très dense est favorisé par la collaboration avec beaucoup **d'*alter* nationaux dans l'équipe de publication**. En effet, nous supposons que cette distance géographique est suffisamment forte pour que les **liens collaboratifs ne soient pas trop personnels** (ce qui pourrait être le cas les *alter* locaux) et suffisamment faible que les *alter* en question **ne soient pas isolés dans le réseau**. *Ego* collabore donc avec des sous-espaces collaboratifs nationaux qu'il mobilise dans plusieurs de ses publications avec différentes équipes.

Facteurs qui déterminent la probabilité d'avoir un réseau éclaté (classe 2)

Dans les réseaux de la classe 2, les liens collaboratifs sont **compartimentés en de multiples composants** qui sont très **denses et isolés les uns des autres**. Ces sous-espaces collaboratifs sont de taille différente et beaucoup n'ont qu'une contribution ponctuelle dans les travaux d'*ego*. *Ego* collabore en toute autonomie avec chacun de ses composants au sein desquels **tous les alter sont au cœur des liens collaboratifs**. *Ego* entretient des **relations homogènes** avec ses *alter* : une grande partie d'entre eux étant au sein de cliques.

Plusieurs facteurs favorisent d'avoir ce type de réseau.

i) Tout d'abord, le **poids des one-shot** dans les collaborations est un facteur très significatif et dont l'effet marginal est très fort (9,2 points dans le modèle 2b et 8,2 points dans le modèle 2c.). En effet, les one-shot pèsent en moyenne 27,40% des collaborations dans la classe 2 contre 19,92% dans les autres classes. Ces collaborations ponctuelles expliquent la **forte proportion de cliques isolées** dans la structure du réseau.

ii) Ensuite, ce type de réseau concerne davantage les chercheurs **en fin de carrière** : dans le modèle 2a, être au stade de « senior » plutôt que « jeune chercheur » augmente la probabilité d'avoir ce type de structure de 1,8 points. En effet, ce type de collaboration nécessite une **certaine autonomie collaborative** et une **certaine ouverture** qui s'acquièrent au fil de l'expérience.

iii) Aussi, avoir **commencé à publier entre 1977 et 1986** (relativement à la période de 1967-1976) réduirait la probabilité d'avoir ce type de réseaux d'1,7 points (modèle 2c).

iv) Enfin, ce type de structure requiert une **certaine émancipation du laboratoire** dans le choix de ses co-auteurs. Dans le modèle 2c, le poids des *alter* locaux dans l'équipe de publication a un effet marginal de -1,1 points et le nombre moyen de publication avec chaque *alter* local a un effet marginal de -0,9 points. Il s'agit d'une particularité géographique : seuls les locaux ont un effet marginal significatif dans les modèles 2b et 2c. Ainsi, la dimension géographique des collaborations est très importante de ce type de structure : *ego élargit son co-authorship au-delà de sa ville*.

Facteurs qui déterminent la probabilité d'avoir un réseau dense aux relations individuelles (classe 3)

Dans les réseaux de la classe 3, les liens collaboratifs ne sont pas compartimentés et sont **très denses** : tous les *alter* ont une forte probabilité d'accompagner *ego* dans ses collaborations avec ses *alter* centraux. Le réseau n'est **pas segmenté en composants connexes** et *ego* mobilise ses sous-espaces-collaboratifs dans des publications conjointes. Aussi, *ego* entretient des relations hétérogènes avec ses *alter* : une grande partie du réseau dispose d'une position unique et non substituable. **Les relations de collaboration sont individualisées** : *ego* collabore différemment avec chaque *alter* et le mobilise avec

différentes équipes. Le réseau s'articule autour des co-auteurs centraux, qui sont très présents dans les liens collaboratifs d'*ego*.

Ce type de réseau est très fortement lié au type d'*alter* avec qui *ego* collabore.

i) Le **poids des one-shot** dans l'ensemble des collaborations décroît très fortement la probabilité d'avoir cette structure de réseau (dans les modèles 3b et 3c, l'effet marginal est de -9,8 points). En effet, la collaboration avec les one-shot peut avoir tendance à « dualiser » les relations collaboratives *ego-alter* entre eux (relations ponctuelles) et le reste du réseau avec qui *ego* poursuit sa relation. Or, dans cette classe, *ego* tisse une relation collaborative singulière avec une grande partie de ses réseaux, au-delà de l'équipe de publication initiale.

ii) Aussi, la collaboration avec des **doctorants en formation** réduit la probabilité d'avoir ce type de structure (l'effet marginal de cette variable est de -8,7 points dans le modèle 3c). En effet, le doctorant suit souvent son directeur de thèse de ses collaborations : il est rare qu'il entretienne une relation individuelle et singulière avec *ego* avant d'obtenir son doctorat.

iii) En revanche, plusieurs types d'*alter* favorise ce type de structure. Plus *ego* collabore avec des **hommes**, plus il a de chance d'avoir un réseau de classe 3 (effet marginal de 5,9 dans le modèle 3b). Ce résultat est intéressant dans la mesure où la probabilité d'avoir ce type de réseau s'accroît si ***ego* est une femme** : en effet, 30 de nos 72 réseaux de femmes sont de classe 3.

iv) La collaboration avec des **chercheurs qui sont ou ont été sous la supervision d'*ego*** favorise ce type de structure (l'effet marginal est de 5,8 dans le modèle 3c). Il semblerait que la proximité durant la réalisation de la thèse d'*alter* favorise une relation singulière avec *ego* et une position non substituable dans son réseau. .

Facteurs qui déterminent la probabilité d'avoir un réseau gravitaire avec une dualité de relations (classe 4) ?

Le réseau de la classe 4 est très structuré par l'hétérogénéité des relations d'*ego* avec ses *alter*. Il est composé de **deux types d'*alter*** : quelques **co-auteurs centraux** qui sont très structurants dans la bibliométrie et les liens collaboratifs d'*ego* et **le reste du réseau** avec qui *ego* ne collabore que ponctuellement. *Ego* compartimente ses relations collaboratives autour de deux types de sous-espaces collaboratifs : ceux qui incluent les *alter* centraux (ils sont très gros et se développent au fil de l'activité de publication d'*ego*) et les autres sous-espaces collaboratifs sont beaucoup plus petits et gravitent autour des co-auteurs centraux.

Seules deux types de variables favorisent significativement ce style de réseau : le stade de carrière d'*ego* et le maintien de relation avec les *alter* locaux et nationaux. En effet, le fait de **ne plus être au stade de « jeune chercheur »** augmente très fortement la probabilité d'avoir ce type de réseau : dans le modèle

4a, être un « chercheur confirmé » augmente de 3,1 points cette probabilité. Le réseau gravitaire s'acquière avec **l'expérience professionnelle** du fait de l'intensification du rythme de publication, de l'élargissement et de l'intensification des relations collaboratives avec les locaux et les nationaux. En moyenne, dans la classe 4, *ego* réalise 3,96 publications avec chaque *alter* issus de son laboratoire contre 2,54 publications dans les autres classes. Ainsi, cette **dualité des relations est avant tout géographique** : quelques ***alter* locaux et nationaux structurent les liens collaboratifs** alors que les autres et les **internationaux gravitent autour d'eux**. En effet, les effets marginaux relatifs au nombre de publications échangées avec chaque *alter* local ou national sont de 1,7 points et 1,3 points dans le modèle 4b.

Tableau 47 : Effets marginaux associés aux probabilités d’avoir un réseau de la classe 1 ou 2 en %

		Modèle 1a.	Modèle 1b.	Modèle 1c.	Modèle 2a.	Modèle 2b.	Modèle 2c.				
		dy/dx	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx				
Genre	<i>Femme</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>				
	<i>Homme</i>	0,01	0,02	0,06	0,00	0,08	0,05				
Type de carrière	<i>Directeur de recherche</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>				
	<i>Professeur</i>	0,09	*	0,01	-0,01	0,10	0,01				
Temporalité des carrières	Stade de carrière	<i>Jeune_chercheur</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>				
		<i>Chercheur_confirme</i>	-0,17	**	-0,14	0,00	-0,02				
		<i>Senior</i>	-0,21	***	-0,19	*	0,18	*			
		<i>Senior_avance</i>	-0,19	**	-0,16	0,17	0,10				
Génération	<i>1967_76</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>	<i>ref</i>				
	<i>1977_86</i>	0,03		0,05	-0,15	-0,17	*				
	<i>1987_96</i>	0,04		0,06	-0,10	-0,03					
	<i>1997_06</i>	-0,04		-0,07	0,09	0,11					
Description des alter	Décomposition des collaborations	<i>collab_hommes</i>		0,12	-0,03	-0,44	-0,36				
		<i>collab_one_shot</i>		-0,40	**	-0,35	*	0,92	**	0,82	**
		<i>collab_doctorants</i>		0,22	0,09	0,14	0,20				
		<i>collab_doctorants_ego</i>		-0,14	0,18	-0,22	-0,41				
		<i>collab_directeur</i>		0,01	-0,07	-0,23	*	-0,18			
	Membres de l'équipe de publication	<i>equipe_locaux</i>		0,00	0,01	-0,09	**	-0,11	***		
		<i>equipe_nationaux</i>		0,04	0,07	**	-0,02	-0,04			
		<i>equipe_internationaux</i>		-0,01	0,00	-0,02	-0,04				
	Durée des relations collaboratives	<i>publis_locaux</i>		-0,12	***	-0,11	***	-0,09	*	-0,09	*
		<i>publis_nationaux</i>		-0,06	*	-0,07	*	0,04	0,05		
<i>publis_internationaux</i>			-0,07	*	-0,07	*	0,08	0,09			
Pseudo-R² McFadden		0.15	0.35	0.42	0.06	0.19	0.23				
Critère d'Akaike		172.29	146.56	145.43	321.36	288.26	287.53				

Significatif au seuil de * : 10 % ; ** : 5 % ; *** : 1 %. Les écarts-types sont disponibles auprès de l’auteur.

Tableau 48 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 3 ou 4 en %

		Modèle 3a.		Modèle 3b.		Modèle 3c.		Modèle 4a.		Modèle 4b.		Modèle 4c.			
		dy/dx		dy/dx		dy/dx		dy/dx		dy/dx		dy/dx			
	Genre	<i>Femme</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>			
		<i>Homme</i>		-0,17	*	-0,15	*	-0,14	*	0,15	**	0,07		0,08	
	Type de carrière	<i>Directeur de recherche</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>			
		<i>Professeur</i>		0,02		0,01		0,01		-0,22	***	-0,08		-0,09	
Temporalité des carrières	Stade de carrière	<i>Jeune_chercheur</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>				<i>ref</i>			
		<i>Chercheur_confirme</i>		-0,14				-0,09		0,31	***			0,25	***
		<i>Senior</i>		-0,12				-0,05		0,16	*			0,11	
		<i>Senior_avance</i>		-0,24	**			-0,14		0,26	**			0,16	
	Génération	<i>1967_76</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>				<i>ref</i>			
		<i>1977_86</i>		0,05				0,07		0,09				0,06	
		<i>1987_96</i>		0,05				0,04		0,01				-0,03	
		<i>1997_06</i>		0,02				0,03		-0,04				-0,04	
Description des alter	Décomposition des collaborations	<i>collab_hommes</i>				0,59	**	0,50	*			-0,40		-0,19	
		<i>collab_one_shot</i>				-0,98	**	-0,98	**			0,40		0,45	
		<i>collab_doctorants</i>				-0,86	*	-0,87	*			0,66		0,58	
		<i>collab_doctorants_ego</i>				0,49		0,58	*			-0,08		-0,17	
		<i>collab_directeur</i>				0,12		0,07				0,01		0,11	
	Membres de l'équipe de publication	<i>equipe_locaux</i>				0,08	**	0,08	**			0,02		0,04	
		<i>equipe_nationaux</i>				-0,03		-0,03				-0,01		-0,02	
		<i>equipe_internationaux</i>				0,00		0,01				0,02		0,03	
	Durée des relations collaboratives	<i>publis_locaux</i>				-0,04		-0,04				0,17	***	0,15	***
		<i>publis_nationaux</i>				-0,07		-0,07				0,13	**	0,13	**
		<i>publis_internationaux</i>				-0,01		-0,01				0,07		0,05	
			Pseudo-R² McFadden		0,05		0,17		0,17		0,13		0,27		0,32
		Critère d'Akaike		311.16		287.15		296.47		287.77		254.29		251.05	

Significatif au seuil de * : 10 % ; ** : 5 % ; *** : 1 %. Les écarts-types sont disponibles auprès de l'auteur.

C) Estimation du style collaboratif en fonction de la dynamique de carrière

Après avoir testé les facteurs relatifs à la carrière d'*ego* et aux caractéristiques des *alter*, nous analysons **la dimension dynamique structurelle du réseau de co-authorship**. Dans cette partie, nous n'évaluons plus les réseaux comme des entités indépendantes les unes des autres mais nous **les recontextualisons dans la carrière de chacun des ego**. Nous avons retiré les 80 réseaux de « jeunes chercheurs » parmi les 259 réseaux de notre base de données puisque nous évaluons l'évolution de la structure du réseau relativement à la période précédente : ainsi, nous nous focalisons sur 179 réseaux à compter du stade de « chercheur confirmé ». L'objectif est donc **d'identifier les facteurs relatifs à la dynamique de carrière qui déterminent la structure du réseau de co-authorship**. Nous commencerons par dresser quelques statistiques descriptives afin de comprendre les principales dynamiques structurelles des réseaux au fil de la carrière. Ensuite, nous détaillerons les modèles économétriques et les variables mobilisées. Nous finirons par présenter les principaux résultats.

1- Statistiques descriptives des dynamiques relationnelles des réseaux

Nous souhaitons décrire les principales tendances de la dynamique structurelle des réseaux au fil de la carrière. En d'autres termes, nous ne décrivons plus la dispersion des quatre structures de réseaux mais plutôt comment le passage d'un stade de carrière à l'autre se traduit structurellement dans le réseau. Nous commencerons par dresser la matrice des différentes dynamiques structurelles possibles pour l'ensemble des 179 réseaux étudiés. Ensuite, nous détaillerons plus précisément les principales dynamiques structurelles relatives à chaque stade de carrière.

Tableau 49 : Distribution des 179 réseaux au sein des 16 dynamiques structurelles possibles

		Classe durant le stade de carrière t				Total
		Classe 1 : « Réseaux très denses »	Classe 2 : « Réseaux éclatés »	Classe 3 : « Réseaux denses et relations individuelles »	Classe 4 : « Réseaux gravitaires et dualité des relations »	
Classe durant le stade de carrière (t-1)	Classe 1 : « Réseaux très denses »	4	7	9	5	25
	Classe 2 : « Réseaux éclatés »	3	26	8	14	51
	Classe 3 : « Réseaux denses et relations individuelles »	3	17	16	19	55
	Classe 4 : « Réseaux gravitaires et dualité des relations »		10	10	28	48
	Total	10	60	43	66	179

Cette matrice décrit les **16 différentes dynamiques structurelles** possibles au sein de nos 80 trajectoires de carrière. Les colonnes décrivent la répartition des 179 réseaux dans les 4 classes structurelles : 10 réseaux sont de classe 1, 60 sont de classe 2, 43 sont de classe 3 et 66 sont de classe 4. Ces colonnes sont décomposées en fonction de la structure que chaque réseau avait au stade de carrière précédent. Ainsi, chaque case correspond à une dynamique structurelle, à savoir l'expression d'une structure de réseau en t à partir de la structure de la période précédente ($t-1$)⁵⁶. Par exemple, on peut lire que parmi les 10 réseaux de classe 1 « Réseaux très denses », 4 d'entre eux avaient déjà cette structure au stade de carrière précédent. L'intérêt de cette matrice est donc de pouvoir **contextualiser chaque réseau dans la carrière d'ego en fonction de sa structure au stade de carrière précédent et d'évaluer s'il a connu un fort changement structurel ou non.**

Nous remarquons que la **stabilité de la structure du réseau entre deux stades de carrière** se lit sur la diagonale de la matrice : à chaque case, le réseau d'ego en t est de même structure qu'en ($t-1$). Ainsi, parmi les 179 réseaux à partir du stade de « chercheur confirmé », 74 d'entre eux (41,43% des réseaux) sont de même structure que la période précédente et 105 réseaux ont donc connu un **bouleversement structurel**. Ce résultats confirme l'intérêt d'identifier les facteurs de la dynamique de carrière dans le sens où les 179 réseaux ont connu de nombreux bouleversements structurels au sein 80 trajectoires de carrière.

Il convient à présent d'identifier et d'interpréter ces différentes dynamiques structurelles à **chaque stade de carrière**. Nous étudierons trois phases transitoires au fil de la carrière d'ego : le passage de « jeune chercheur » à « chercheur confirmé », le passage de « chercheur confirmé » à « senior » et le passage de « senior » à « senior avancé ». A chaque étape, nous mobiliserons deux outils.

i) Le premier est le **nuage de points des réseaux dans le plan des axes de l'ACP** : il permet de visualiser la dispersion des réseaux de chaque classe en fonction de s'ils sont issus d'un bouleversement structurel ou non. L'objectif est **d'identifier quelles sont les coordonnées des réseaux qui sont les plus susceptibles de conserver leur classe de structure.**

ii) Le second outil est la **matrice des différentes dynamiques structurelles** : elle permet d'évaluer le poids du bouleversement structurel et de **décrire les principales trajectoires structurelles** des réseaux au stade de carrière étudié.

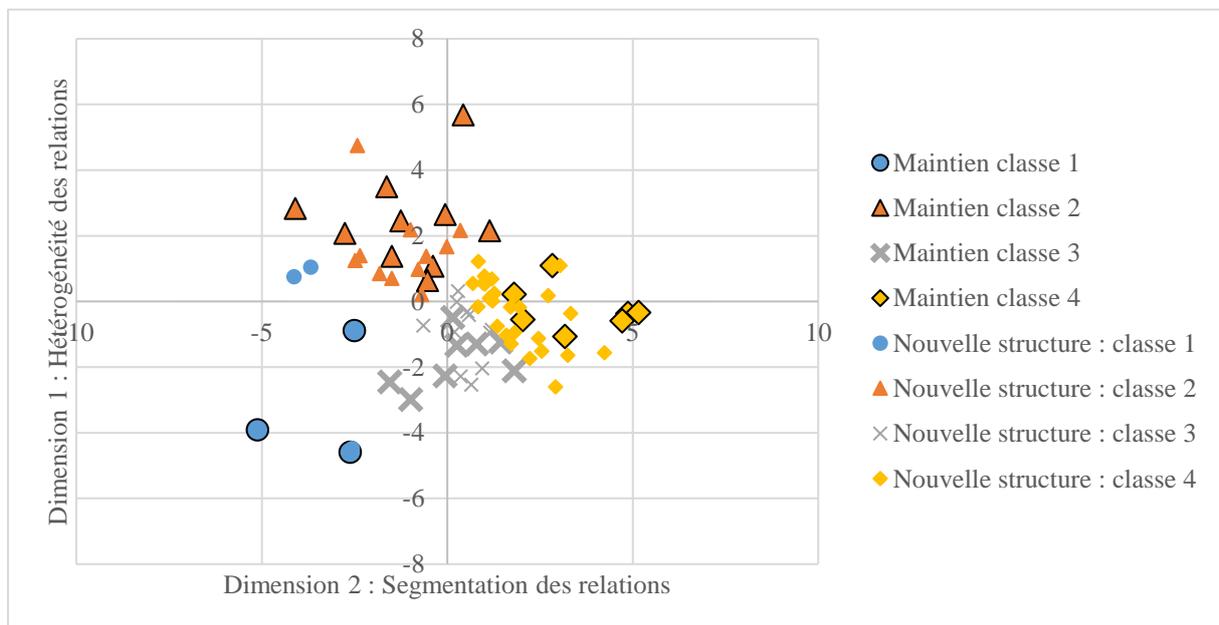
⁵⁶ pour tout t correspondant à un stade de carrière et définit tel que $N \in [2 ; 4]$

i. Le passage de « jeune chercheur » à « chercheur confirmé » se traduit-il par une mutation spécifique du réseau ?

Tableau 50 : Structure du réseau au stade de chercheur confirmé

	<i>Classe 1</i>	<i>Classe 2</i>	<i>Classe 3</i>	<i>Classe 4</i>	Total	
Structure du réseau au stade de jeune chercheur	<i>Classe 1 : « Réseaux très denses »</i>	3	4	7	5	19
	<i>Classe 2 : « Réseaux éclatés »</i>		10	4	7	21
	<i>Classe 3 : « Réseaux denses et relations individuelles »</i>	3	6	8	14	31
	<i>Classe 4 : « Réseaux gravitaires et dualité des relations »</i>		1	1	7	9
Total	6	21	20	33	80	

Figure 46 : Dynamiques structurelles des réseaux au stade de chercheur confirmé



Le passage de « jeune chercheur » à « chercheur confirmé » **se traduit par un fort bouleversement structurel du réseau**. En effet, parmi les 80 réseaux de « chercheur confirmé », seuls 28 d’entre eux ont la même structure que durant le stade de « jeune chercheur ». Ainsi, 52 des 80 chercheurs ont **changé leur style de collaboration et de diffusion de connaissance** à ce stade de carrière, concernant l’hétérogénéité de leurs relations ou à la segmentation de leurs collaborations.

i) **Les « réseaux éclatés » de la classe 2 montrent une certaine stabilité**. En effet, cette structure pèse le même poids parmi les réseaux de « jeunes chercheurs » et ceux des « chercheurs confirmés » : parmi les 21 réseaux éclatés de « chercheurs confirmés », 10 d’entre eux avaient déjà cette structure durant la période. Dans le nuage de points, on remarque que les réseaux de classe 2 qui ont connu une stabilité structurelle partagent les mêmes coordonnées que les réseaux issus d’un

bouleversement structurel : la stabilité structurelle **ne correspond pas à un positionnement particulier dans le plan.**

ii) **Beaucoup de réseaux denses à l'échelle globale (classes 1 et 3) de « jeunes chercheurs » changent de structure lorsqu'ego devient chercheur confirmé.** En effet, cette structure de réseau est surreprésentée au stade de « jeune chercheur » : elle correspond à 50 des 80 réseaux. En revanche, une fois qu'ego devient « chercheur confirmé », elle ne pèse que 26 des 80 réseaux. Cela signifie que les **réseaux denses à l'échelle globale** des « jeunes chercheurs » connaissent **un changement structurel.** Dans le nuage de points, on remarque notamment que les réseaux de classe 1 qui ont maintenu ce type de structure disposent de coordonnées extrêmement basses sur l'axe 2 : il s'agit de réseaux atypiques et particulièrement denses.

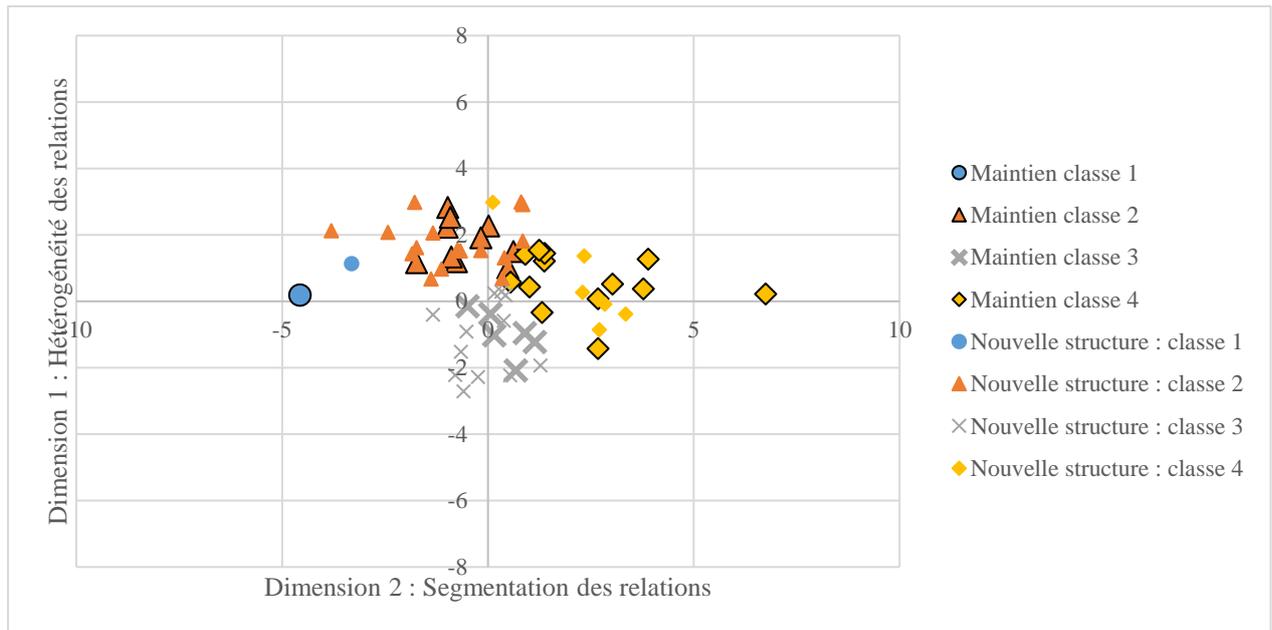
iii) Le passage au stade de « chercheur confirmé » se traduit par **l'explosion des « réseaux structurés autour d'une dualité de relation » de la classe 4.** Alors que seulement 9 des 80 réseaux de jeunes chercheurs disposaient de cette structure, elle est caractérisée par 33 réseaux de chercheurs confirmés. Ainsi, beaucoup de réseaux ont connu un **bouleversement structurel** pour intégrer cette classe au stade de chercheur confirmé. En effet, 14 des 31 réseaux de la classe 3 des jeunes chercheurs se sont mus en classe 4 au stade de chercheur confirmé : ces réseaux ont conservé leurs relations hétérogènes mais les collaborations se sont davantage compartimentées. Aussi, 7 des 33 réseaux de classe 4 avaient déjà cette structure durant la période précédente : dans le nuage de points on remarque que ces réseaux ont des coordonnées extrêmement fortes sur l'axe 1 et sont très atypiques.

ii. Le passage de « chercheur confirmé » à « senior » se traduit-il par une mutation spécifique du réseau ?

Tableau 51 : Structure du réseau au stade de senior

	<i>Classe 1</i>	<i>Classe 2</i>	<i>Classe 3</i>	<i>Classe 4</i>	Total	
Structure du réseau au stade de chercheur confirmé	<i>Classe 1 : « Réseaux très denses »</i>	1	1	2	4	
	<i>Classe 2 : « Réseaux éclatés »</i>	1	10	3	2	16
	<i>Classe 3 : « Réseaux denses et relations individuelles »</i>		8	6	4	18
	<i>Classe 4 : « Réseaux gravitaires et dualité des relations »</i>		7	7	13	27
Total	2	26	18	19	65	

Figure 47 : Dynamiques structurelles des réseaux au stade de senior



Le passage de « chercheur confirmé » à « senior » **se traduit par certaine stabilité structurelle du réseau**. En effet, parmi les 65 réseaux de « senior », 30 d’entre eux conservent la même structure qu’ils avaient à la période précédente. Ainsi, le style de collaboration et de diffusion de connaissances ne varie que pour un réseau sur deux à ce stade de carrière.

i) **Les réseaux segmentés des classes 2 et 4 montrent une certaine stabilité structurelle**. En effet, parmi les réseaux de « senior », 10 des 26 « réseaux éclatés » (classe 2) et 13 des 19 « réseaux gravitaires dotés d’une dualité de relations » (classe 4) avaient la même structure durant la période précédente. Ainsi, la **segmentation du réseau** en composants ou en sous-espaces collaboratifs est un **caractère structurel stable entre ces deux périodes**. On remarque notamment dans le nuage de points que les réseaux qui ont conservé une structure de classe 2 ou 4 ont une fois encore des coordonnées extrêmes sur les axes (réseaux atypiques).

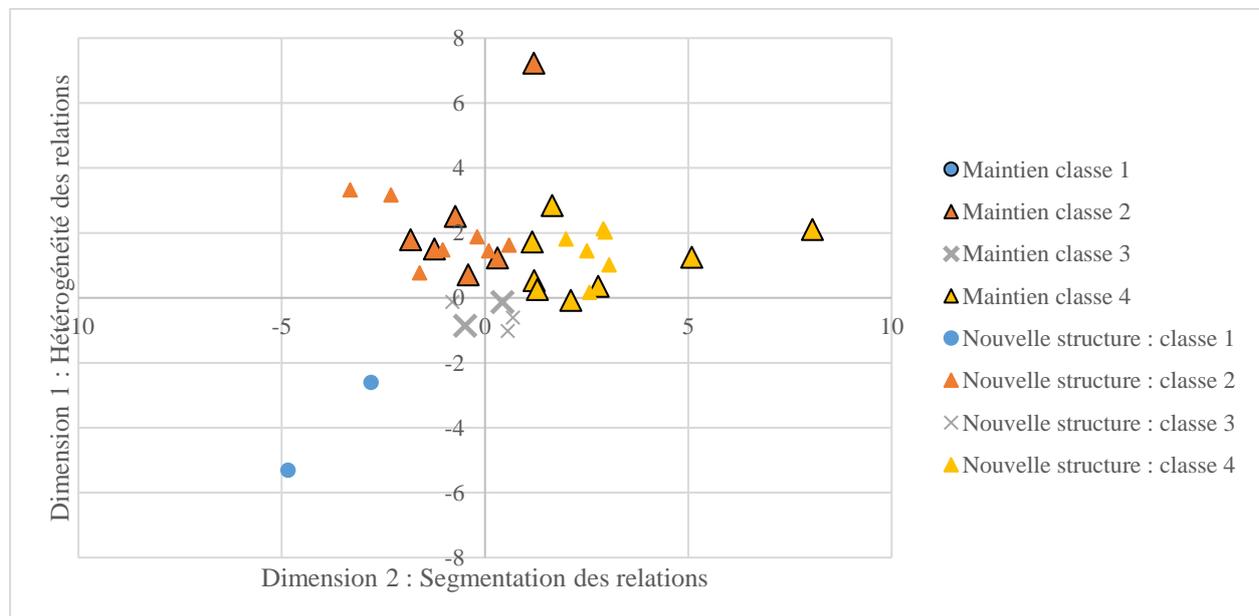
ii) Les réseaux de **classe 3** sont principalement **issus d’une densification du réseau** par rapport à la période précédente. En effet, 7 des 18 réseaux de la classe 3 étaient de classe 4 au stade de « chercheur confirmé ». La transition de carrière s’est donc faite via **le maintien de relation hétérogènes** mais **par une concentration du réseau en même entité** (et donc une baisse de la segmentation des relations).

iii. Le passage de « senior » à « senior avancé » se traduit-il par une mutation spécifique du réseau ?

Tableau 52 : Structure du réseau au stade de senior avancé

	<i>Classe 1</i>	<i>Classe 2</i>	<i>Classe 3</i>	<i>Classe 4</i>	Total	
Structure du réseau au stade de senior	<i>Classe 1 : « Réseaux très denses »</i>		2		2	
	<i>Classe 2 : « Réseaux éclatés »</i>	2	6	1	5	14
	<i>Classe 3 : « Réseaux denses et relations individuelles »</i>		3	2	1	6
	<i>Classe 4 : « Réseaux gravitaires et dualité des relations »</i>		2	2	8	12
Total	2	13	5	14	34	

Figure 48 : Dynamiques structurelles des réseaux au stade de senior avancé



Le passage de « senior » à « senior avancé » se traduit par **franche stabilité structurelle, relative aux relations hétérogènes**. En effet, parmi les 34 réseaux de « senior avancé », 16 d'entre eux conservent la même structure qu'ils avaient à la période précédente. **Les classes aux relations hétérogènes (classes 2 et 4) sont surreprésentées** : elles concernent 27 des 34 réseaux. La majorité de ces réseaux conservent la même structure qu'ils avaient auparavant.

Nous venons de décrire les principales dynamiques structurelles au fil de la carrière d'*ego*. Nous avons montré que **59% des réseaux ont une structure différente** de celle de la période précédente à partir du stade de « chercheur confirmé ». Ces bouleversements sont de plusieurs ordres et concernent davantage le passage de « jeune chercheur » à « chercheur confirmé ». En effet, la structure du réseau d'*ego* est relativement stable une fois qu'il devient « senior ». Nous allons à présent dresser des modèles économétriques qui permettent de déterminer la structure du réseau à partir de sa dynamique structurelle.

2- Présentation des modèles

L'objectif est une nouvelle fois d'identifier les variables qui favorisent systématiquement la probabilité d'avoir un réseau de telle structure ou non. Nous avons réalisé plusieurs modèles de régression logistique binaire dont nous avons estimé toute chose égale les effets marginaux.

Nous testons trois variables dépendantes nommées $Classe_dynamique_j$ pour tout $j \in \{1;4\}$:

$$Classe_dynamique_j \begin{cases} = 1 \text{ si } Classe_dynamique_j^* \geq 0 \\ = 0 \text{ sinon} \end{cases}$$

$$\text{Dont : } Classe_dynamique_j^* = \alpha + D_j \beta' + \mu_j$$

Dans ces modèles de régression, chaque variable expliquée $Classe_dynamique_j$ est dichotomique : elle prend la valeur 1 si le réseau appartient à la classe j et 0 si ce n'est pas le cas. Également, la variable $Classe_dynamique_j^*$ est la variable latente : elle est décrite en fonction de la matrice des variable explicatives D_j et du vecteur des paramètres associés β' . La probabilité d'appartenance à la classe j est donc :

$$P(Classe_dynamique_j = 1 | D_j) = \frac{\exp(\alpha + D_j \beta')}{1 + \exp(\alpha + D_j \beta')}$$

Tableau 53 : Variables dépendantes Classe-dynamiquej

	$Classe_dynamique_j = 1$	$Classe_dynamique_j = 0$	Total
$Classe_dynamique_1$	10	169	179
$Classe_dynamique_2$	60	119	179
$Classe_dynamique_3$	43	136	179
$Classe_dynamique_4$	66	113	179

Ces quatre variables dépendantes sont testées par trois sous-modèles : chacun d'entre eux disposent de variables complémentaires relatives à la dynamique structurelle du réseau.

i) Le **premier sous-modèle** décrit la dynamique structurelle de chaque réseau **en fonction de la classe à laquelle il appartenait durant la période précédente**. La variable $classe_precedente$ dispose de quatre modalités {classe 1 ; classe 2 ; classe 3 ; classe 4}. Chaque modèle qui teste la probabilité d'appartenir à la classe j prend la modalité « classe j » comme référence : par exemple, le modèle testant l'appartenance à la classe 1 prendra la modalité « classe 1 » comme référence de la *variable classe_precedente*.

ii) Le **second sous-modèle** décrit la **trajectoire structurelle du réseau** à partir de son positionnement dans le **plan des axes issus de l'ACP**. L'intérêt n'est plus d'identifier si le réseau a conservé la même classe que la période précédente mais plutôt **dans quel sens il s'est déplacé dans le plan**. Pour rappel,

les deux axes de l'ACP sont très significativement contributifs dans la construction de la classification des réseaux. Ainsi, **tout déplacement dans le plan inclue une mutation structurelle du réseau**. L'objectif est de tester si la probabilité d'avoir telle structure de réseau dépend de la trajectoire que le réseau a opéré dans le plan et donc de la mutation structurelle qu'il a subie.

- La variable *evol_dim_1* calcule la différence entre les coordonnées du réseau sur l'axe 1 de la période étudiée et la période précédente.

Si *evol_dim_1* est positive, cela signifie que le réseau s'est déplacé entre les deux périodes vers la droite de l'axe 1 dans le plan : le réseau est davantage structuré par l'hétérogénéité des relations collaboratives *ego-alter* que durant la période précédente.

Si *evol_dim_1* est négative, cela signifie que le réseau s'est déplacé entre les deux périodes vers la gauche de l'axe 1 dans le plan : le réseau est davantage structuré par l'homogénéité des relations collaboratives *ego-alter* que durant la période précédente.

- Aussi, la variable *evol_dim_2* calcule la différence entre les coordonnées du réseau sur l'axe 2 de la période étudiée et la période précédente.

Si *evol_dim_2* est positive, cela signifie que le réseau s'est déplacé entre les deux périodes vers le haut de l'axe 2 dans le plan : les liens collaboratifs sont davantage segmentés que durant la période précédente.

Si *evol_dim_2* est négative, cela signifie que le réseau s'est déplacé entre les deux périodes vers le bas de l'axe 2 dans le plan : les liens collaboratifs sont moins segmentés que durant la période précédente.

iii) Le **troisième sous-modèle** concerne également la trajectoire structurelle du réseau dans le plan. Néanmoins, il ne se concentre plus sur la direction suivie par le réseau mais **par la distance du déplacement dans le réseau**. L'objectif est de tester si **telle structure de réseau** se traduit par **un fort bouleversement structurel ou non**.

- La variable *val_abs_dim_1* correspond à la valeur absolue de *evol_dim_1* : elle ne décrit pas le sens du déplacement sur l'axe mais plutôt de combien d'unités le réseau s'est déplacé sur l'axe entre les deux périodes.

Ainsi, si *val_abs_dim_1* est élevée, cela signifie que le réseau a connu une forte mutation structurelle relative à l'hétérogénéité ou à l'homogénéité de ses relations collaboratives.

En revanche, si *val_abs_dim_1* tend vers zéro, cela signifie que le réseau a conservé le même positionnement sur l'axe 1 et que sa structure n'est pas issue d'un fort bouleversement structurel sur cet axe.

- La variable *val_abs_dim_2* correspond à la valeur absolue de *evol_dim_2*.

Si *val_abs_dim_2* est élevée, cela signifie que le réseau a connu une forte mutation structurelle relative à segmentation de ses relations collaboratives.

En revanche, si *val_abs_dim_2* tend vers zéro, cela signifie que le réseau a conservé le même positionnement sur l'axe 2 et que sa structure n'est pas issue d'un fort bouleversement structurel sur cet axe.

L'intérêt de ces trois types de variables explicatives est **d'enrichir l'appréhension de la trajectoire structurelle du réseau** en allant au-delà du changement de classe. Un exemple d'application de ces trois types de variable est disponible en Annexe 9.

Par conséquent, en croisant les différentes variables explicatives et expliquées relatives à la dynamique structurelle du réseau, nous estimons 12 modèles synthétisés dans le tableau suivant.

Tableau 54 : Synthèse des modèles estimés

	<i>Classe_dynamique</i> <i>1</i>	<i>Classe_dynamique</i> <i>2</i>	<i>Classe_dynamique</i> <i>3</i>	<i>Classe_dynamique</i> <i>4</i>
Classe de la période précédente <i>carrière_precedente</i>	Modèle 1'a.	Modèle 2'a.	Modèle 3'a.	Modèle 4'a.
Trajectoire dans le plan de l'ACP <i>evol_dim_1</i> <i>evol_dim_2</i>	Modèle 2'b.	Modèle 2'b.	Modèle 3'b.	Modèle 4'b.
Distance parcourue dans le plan de l'ACP <i>val_abs_dim_1</i> <i>val_abs_dim_2</i>	Modèle 3'c.	Modèle 2'c.	Modèle 3'c.	Modèle 4'c.

Ces douze sous-modèles économétriques testeront également l'impact de deux autres types de variables explicatives : l'évolution de l'activité de publication d'*ego* et sa temporalité de carrière.

i) La **variable *coef_mult_pub*** correspond au coefficient multiplicateur entre le nombre de publications du réseau avec celui de la période précédente. L'objectif de cette variable est d'évaluer si **telle structure dépend de l'évolution du nombre de publications réalisées et donc de l'évolution de la taille du réseau.**

ii) Également, la variable *mob* évalue l'évolution des flux de mobilité géographique d'*ego* entre les deux périodes. L'objectif est de tester si **la structure du réseau dépend également de la trajectoire géographique d'*ego*.**

- Deux modalités décrivent une stabilité des flux de mobilité : pour chaque réseau, la modalité « *mob_a_mob* » signifie qu'*ego* est affilié à plusieurs villes comme durant la période précédente et la modalité « *sed_a_sed* » signifie plutôt qu'*ego* est sédentaire comme durant la période précédente.

- En revanche, la modalité « *sed_a_mob* » décrit qu'*ego* est mobile dans le réseau alors qu'il était sédentaire auparavant durant la période précédente et la modalité « *mob_a_sed* » décrit qu'*ego* est sédentaire dans le réseau alors qu'il était mobile durant la période précédente.

Deux autres variables décrivent la **temporalité de la carrière d'*ego***. La variable *evol_carriere* décrit la transition de stade de carrière à chaque réseau, à savoir le passage de jeune chercheur à chercheur confirmé « *jeune_a_chercheur_conf* », le passage de chercheur confirmé à senior « *chercheur_conf_a_senior* » et de senior à senior avancé « *senior_a_senior_av* ». L'objectif est de tester si le stade de carrière impacte la structure du réseau. Aussi la variable *generation* recontextualise la période durant laquelle *ego* a commencé sa carrière.

Enfin, nous ajouterons deux variables de contrôle : le sexe d'*ego* (*genre*) et le type de carrière suivie (*type_carriere*) entre directeur de recherche et professeur.

Le tableau suivant décrit les statistiques descriptives des variables expliquées, explicatives et de contrôle.

Tableau 55 : Statistiques descriptives des variables

			N	Moyenne	Écart type	Min	Max
Variables expliquées							
Classe_dynamique ₁		Le réseau appartient à la classe 1 « réseaux très denses »	179	0,056		0	1
Classe_dynamique ₂		Le réseau appartient à la classe 2 « réseaux éclatés »	179	0,335		0	1
Classe_dynamique ₃		Le réseau appartient à la classe 3 « réseaux denses et relations individuelles »	179	0,240		0	1
Classe_dynamique ₄		Le réseau appartient à la classe 4 « réseaux gravitaires et dualité des relations »	179	0,369		0	1
Variables explicatives							
1- Dynamique structurelle du réseau							
Classe_precedente	<i>Classe 1</i>	Le réseau était de classe 1 en (t-1)	25	0,140		0	1
	<i>Classe 2</i>	Le réseau était de classe 2 en (t-1)	51	0,285		0	1
	<i>Classe 3</i>	Le réseau était de classe 3 en (t-1)	55	0,307		0	1
	<i>Classe 4</i>	Le réseau était de classe 4 en (t-1)	48	0,268		0	1
evol_dim_1			179	0,525	2,213	-5,750	6,730
evol_dim_2			179	0,597	2,056	-6,830	7,260
val_abs_dim_1			179	1,785	1,403	0,010	6,730
val_abs_dim_2			179	1,609	1,408	0,000	7,260
2- Dynamique de l'activité de publication et temporalités de carrière							
coef_mult_pub		Rapport entre le nombre de publications en t et (t-1)	179	2,000	0,894	1,000	4,000
mob	<i>sed_à_sed</i>	Ego était sédentaire en (t-1) et en t	179	0,263		0	1
	<i>mob_à_mob</i>	Ego était mobile en (t-1) et en t	179	0,257		0	1
	<i>sed_à_mob</i>	Ego était sédentaire en (t-1) mais a été mobile en t	179	0,291		0	1
	<i>mob_à_sed</i>	Ego était mobile en (t-1) mais sédentaire en t	179	0,190		0	1
evol_carriere	<i>jeune_a_chercheur_conf</i>	Le réseau est celui d'un chercheur confirmé	179	1,000		0	1
	<i>chercheur_conf_a_senior</i>	Le réseau est celui d'un senior	65	0,363		0	1
	<i>senior_a_senior_av</i>	Le réseau est celui d'un senior avancé	34	0,190		0	1
Generation	<i>1967_76</i>	Le réseau appartient à la génération "1967-1976"	39	0,218		0	1
	<i>1977_86</i>	Le réseau appartient à la génération "1977-1986"	52	0,291		0	1
	<i>1987_96</i>	Le réseau appartient à la génération "1987-1996"	61	0,341		0	1
	<i>1997_06</i>	Le réseau appartient à la génération "1997-2006"	27	0,151		0	1
Variables de contrôle							
Genre	<i>Femme</i>	Le réseau est celui d'une femme	72	0,278		0	1
	<i>Homme</i>	Le réseau est celui d'un homme	187	0,722		0	1
Type_carriere	<i>Directeur de recherche</i>	Le réseau est celui d'un directeur de recherche	118	0,456		0	1
	<i>Professeur</i>	Le réseau est celui d'un professeur	141	0,544		0	1

3- Les résultats

Quels sont les facteurs dynamiques qui influencent la probabilité d'avoir un réseau très dense (classe 1) ?

Bien que la classe 1 soit sous-représentée à partir du stade de « chercheur confirmé », plusieurs facteurs dynamiques expliquent ce type de structure après le stade de « jeune chercheur ». En effet, nos trois sous-modèles expliquent plus de 60% de la variance de nos données. Seules les variables relatives à la temporalité de la carrière d'*ego* ne caractérisent pas significativement cette structure.

i) Tout d'abord, ce type de réseau a plus de chance d'être issu d'une **certaine stabilité structurelle** plutôt que d'une forte mutation structurelle. En effet, le fait que le réseau soit durant la période précédente gravitaire (classe 4) plutôt que très dense (classe 1) baisse de 1.14 points la probabilité que le réseau soit de classe 1 : il est très peu probable qu'un réseau dense aux relations homogènes soit issu d'un réseau segmenté et aux relations hétérogènes. Les autres modalités de la variable *classe_precedente* ne sont pas significatives : à partir du stade de « chercheur confirmé », un réseau de **classe 1 est aussi bien issu de stabilité structurelle (maintien de la classe 1) que de mutations structurelles mineures**. Ces mutations se traduisent par **une homogénéisation des relations ego-alter d'une ampleur non significative et d'une forte densification des liens** (effet marginal de 0.4 points dans le modèle 1'c).

ii) Ensuite, la dynamique de publication influence également la probabilité de ce type de structure. Elle s'accompagne généralement par une **contraction du volume de publications** par rapport à la période précédente (effet marginal de -0.7 points dans le modèle 1'c). A partir du stade de « chercheur confirmé », il s'agit donc **de réseaux plus petits que la période précédente et de courte période**.

iii) Aussi, l'affiliation géographique impacte ce type de structure. En effet, **être mobile après une période de sédentarité** plutôt que de continuer à être sédentaire augmenterait la probabilité d'avoir ce type de classe de 0.8 points. Dans le chapitre précédent, nous avons prouvé que la mobilité générerait de **nouvelles relations collaboratives au sein du nouveau laboratoire affilié : ces relations ego-alter sont très homogènes puisqu'elles sont récentes**. En revanche, il s'avère qu'une mobilité continue (durant la période du réseau et avant celle-ci) réduirait la probabilité d'avoir cette structure de réseau de 0.6 points par rapport à une sédentarité continue. En effet, si *ego* est habitué à être affilié à plusieurs villes, son réseau a moins de chance d'être dense et d'être considéré comme une même entité.

iv) Enfin, certaines caractéristiques d'*ego* expliquent ce type de structure. Suivre une carrière de professeur plutôt que de directeur de recherche augmenterait de 1,4 points la probabilité d'avoir un réseau « très dense ». Aussi, être un homme réduirait de 0.8 points la probabilité d'avoir un réseau de classe 1. **Les professeurs et les femmes** seraient donc plus à même de mobiliser un **réseau dense aux relations homogènes** à partir du stade de « chercheur confirmé ».

Quels sont les facteurs dynamiques qui influencent la probabilité d'avoir un réseau éclaté en composants connexes (classe 2) ?

La classe 2 est une **structure surreprésentée** dans les réseaux à partir du stade de « chercheur confirmé » : elle concerne 33,5% des réseaux étudiés. Ces trois sous-modèles expliquent entre 11% et 25% de la variance de nos données. Seules les variables relatives à la dynamique structurelle du réseau et à la temporalité de la carrière semblent expliquer significativement la probabilité d'avoir un réseau dit « éclaté ».

i) Tout d'abord, si le réseau de la période précédente est de type gravitaire de classe 4 plutôt que de classe 2, la probabilité qu'il devienne un réseau éclaté baisse de 2.9 points. Ainsi, bien que ces deux réseaux partagent une forte segmentation de leurs relations collaboratives, **l'homogénéité des relations dans la classe 2 a peu de chance d'être issue d'un réseau aux relations hétérogènes de classe 4**. Ainsi, ce type de réseau résulte de **deux mutations structurelles : une homogénéisation des relations collaboratives de petite ampleur** (effet marginal de -0.8 points) et **une segmentation des relations collaboratives de très forte ampleur** (effet marginal de 0.8 points).

ii) Également, ce type de réseau se manifeste plutôt en **fin de carrière**, une fois qu'*ego* ait acquis une forte expérience professionnelle. En effet, être « senior » plutôt que « chercheur confirmé » augmente la probabilité d'avoir ce type de réseau de 2.2 points (modèle 2'a) et être « senior avancé » plutôt que « chercheur confirmé » augmente la probabilité d'avoir ce type de réseau de 2.0 points (modèle 2'b).

Quels sont les facteurs dynamiques qui influencent la probabilité d'avoir un réseau dense de relations individuelles (classe 3) ?

La classe 3 est une structure **très peu représentée** à partir du stade de « chercheur confirmé » : elle ne concerne que 24% des réseaux. Ces trois sous-modèles ne décrivent qu'entre 9% et 13% de la variance des données : seules les variables relatives à la dynamique structurelle du réseau et à l'évolution du volume de publication expliquent significativement cette structure.

i) Il n'y a pas d'unique trajectoire structurelle qui explique la classe 3 : les déplacements sur les deux axes sont d'ampleur non significative. D'une part, pour certains réseaux, être issu d'un réseau très dense (de classe 1) augmente de 2.6 points la probabilité d'être un réseau de classe 3 : ces réseaux conservent **une forte densité à l'échelle globale**. D'autre part, certains réseaux de classe 3 sont issus d'un déplacement vers le bas de l'axe 2 (effet marginal de 0.5 points) et donc d'une **densification interne**.

ii) Également, **une contraction du nombre de publication** par rapport à la période précédente augmente la probabilité d'avoir cette structure de réseau. Tout comme la classe 1, la réduction du volume de publication réduit la segmentation des liens collaboratifs et augmente la densité globale du réseau.

Quels sont les facteurs dynamiques qui influencent la probabilité d'avoir un réseau gravitaire structuré autour d'une dualité de relations (classe 4) ?

Enfin, la classe 4 est **la structure la plus représentée** à partir du stade de « chercheur confirmé » : elle représente 37% des réseaux étudiés. Les trois sous-modèles expliquent entre 21 et 25% de la variance de nos données. Toutes les variables dynamiques expliquent ce type de structure.

i) Tout d'abord, les réseaux de classe 4 sont issus d'une **forte stabilité structurelle** : être issu d'une autre structure réduit la probabilité d'être de classe 4 de plus de 3.3 points. Ainsi, **à partir du stade de « chercheur confirmé », une fois qu'un réseau est de classe 4, il a de forte chance de conserver cette structure au fil de la carrière.** Concernant les mutations structurelles, les réseaux de classe 4 sont souvent issues de déplacement vers la droite de l'axe 1, d'une portée non significative : la classe 4 provient d'une **intensification de l'hétérogénéité des ego-alter** quel que soit le positionnement sur l'axe initial. En revanche, les réseaux de la classe 4 sont issus de petits déplacements sur l'axe 2 : cela signifie qu'ils sont issus de réseau **déjà segmentés.**

ii) Ensuite, cette structure provient de dynamiques de publication précises. Tout d'abord, elle se manifeste par une **hausse du volume de publication** par rapport à la période précédente (effet marginal de 1.2 points dans le modèle 4'a). Aussi, **une mobilité continue** (durant la période du réseau et la période précédente) plutôt qu'une sédentarité continue accroît de 2.0 points (modèle 4'c) la probabilité d'avoir cette structure de réseau. En effet, la multi-affiliation accroît la segmentation des relations collaboratives contrairement à la sédentarité qui densifie le réseau composé d'*alter* locaux.

iii) La temporalité de la carrière explique également cette structure. En effet, être au stade de « senior » plutôt que de « **chercheur confirmé** » réduit la probabilité d'avoir cette structure de 2.00 points (modèle 4'a) : nous avons montré dans les statistiques descriptives l'exposition de la classe 4 au stade de « chercheur confirmé ». Aussi, appartenir à la **génération « 1977-1986 »** plutôt que « 1967-1976 » augmenterait la probabilité d'avoir ce type de structure de 1.9 points (modèle 4'b).

iv) Enfin, le profil d'*ego* détermine également la probabilité d'avoir un réseau de classe 4. Que le réseau soit celui d'un **homme** plutôt que celui d'une femme augmente de 1.9 points la probabilité d'avoir cette structure (modèle 4'c). Aussi, le fait que le réseau soit celui d'un professeur plutôt que celui d'un **directeur de recherche** réduit de 2.2 points la probabilité qu'il s'agisse d'un réseau de classe 4.

Tableau 56 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 1 en %

			Modèle 1'a.	Modèle 1'b.	Modèle 1'c.		
			dy/dx	dy/dx	dy/dx		
Dynamique structurelle du réseau	classe_precedente	<i>Classe 1</i>	REF				
		<i>Classe 2</i>	-0,065				
		<i>Classe 3</i>	-0,053				
		<i>Classe 4</i>	-0,114	*			
	evol_dim1			-0,065	**		
	evol_dim2			-0,023	**		
	valeur_abs_evol_dim1				0,012		
valeur_abs_evol_dim2				0,040	*		
Dynamiques de publication	coef_mult_pub		-0,057	**	-0,021	-0,075	**
	mob	<i>sed_a_sed</i>	REF		REF	REF	
		<i>sed_a_mob</i>	0,038		0,115	0,083	*
		<i>mob_a_sed</i>	-0,070		-0,058	-0,077	
		<i>mob_a_mob</i>	-0,045		-0,007	-0,063	*
Temporalité de la carrière d'ego	evol_carriere	<i>jeune_chercheur_a_chercheur_conf</i>	REF		REF	REF	
		<i>chercheur_conf_a_senior</i>	-0,019		-0,020	-0,010	
		<i>senior_a_senior_av</i>	-0,007		-0,009	-0,012	
	generation	<i>1967-1976</i>	REF		REF	REF	
		<i>1977-1986</i>	0,075	*	0,092	0,101	
		<i>1987-1996</i>	0,079		0,028	0,034	
		<i>1997-2006</i>	0,045		0,077	0,053	
Description d'ego	genre	<i>Femme</i>	REF		REF	REF	
		<i>Homme</i>	0,006		-0,086	**	-0,072
	type_carriere	<i>directeur_de_recherche</i>	REF		REF	REF	
		<i>professeur</i>	0,082	**	0,141	*	0,121
Pseudo-R² McFadden			0.61		0.79	0.74	
Critère d'Akaike			59.66		43.844	48.374	

Tableau 57 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 2 en %

			Modèle 2'a.	Modèle 2'b.	Modèle 2'c.		
			dy/dx	dy/dx	dy/dx		
Dynamique structurelle du réseau	classe_precedente	<i>Classe 1</i>	-0,178				
		<i>Classe 2</i>	REF				
		<i>Classe 3</i>	-0,138				
		<i>Classe 4</i>	-0,289	**			
	evol_dim1			-0,062	**		
	evol_dim2			0,068	***		
	valeur_abs_evol_dim1				-0,085	**	
	valeur_abs_evol_dim2				0,083	***	
Dynamiques de publication	coef_mult_pub		0,020	0,052	0,046		
	mob	<i>sed_a_sed</i>	REF	REF	REF		
		<i>sed_a_mob</i>	-0,104	-0,042	-0,114		
		<i>mob_a_sed</i>	0,072	-0,049	0,015		
		<i>mob_a_mob</i>	-0,048	-0,080	-0,085		
Temporalité de la carrière d'ego	evol_carriere	<i>jeune_chercheur_a_chercheur_conf</i>	REF	REF	REF		
		<i>chercheur_conf_a_senior</i>	0,218	**	0,139	0,179	*
		<i>senior_a_senior_av</i>	0,141		0,200	*	0,198
	generation	<i>1967-1976</i>	REF	REF	REF		
		<i>1977-1986</i>	-0,110	-0,159	-0,142		
		<i>1987-1996</i>	-0,066	-0,059	-0,067		
		<i>1997-2006</i>	0,022	0,099	0,108		
	Description d'ego	genre	<i>Femme</i>	REF	REF	REF	
<i>Homme</i>			0,007	-0,013	-0,017		
type_carriere		<i>directeur_de_recherche</i>	REF	REF	REF		
		<i>professeur</i>	0,092	0,109	0,091		
Pseudo-R² McFadden			0.11	0.25	0.16		
Critère d'Akaike			233.66	198.89	219.23		

Tableau 58 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 3 en %

			Modèle 3'a.	Modèle 3'b.	Modèle 3'c.
			dy/dx	dy/dx	dy/dx
Dynamique structurale du réseau	classe_precedente	<i>Classe 1</i>	0,264 *		
		<i>Classe 2</i>	-0,101		
		<i>Classe 3</i>	REF		
		<i>Classe 4</i>	-0,091		
	evol_dim1			0,010	
	evol_dim2			-0,051 **	
	valeur_abs_evol_dim1				0,026
valeur_abs_evol_dim2				-0,046	
Dynamiques de publication	coef_mult_pub		-0,075 *	-0,040	-0,058
	mob	<i>sed_a_sed</i>	REF	REF	REF
		<i>sed_a_mob</i>	-0,099	-0,088	-0,057
		<i>mob_a_sed</i>	0,005	0,067	0,042
		<i>mob_a_mob</i>	-0,052	-0,014	-0,019
Temporalité de la carrière d'ego	evol_carriere	<i>jeune_chercheur_a_chercheur_conf</i>	REF	REF	REF
		<i>chercheur_conf_a_senior</i>	0,001	-0,012	-0,036
		<i>senior_a_senior_av</i>	-0,081	-0,150	-0,143
	generation	<i>1967-1976</i>	REF	REF	REF
		<i>1977-1986</i>	-0,033	-0,019	-0,020
		<i>1987-1996</i>	0,030	0,026	0,044
		<i>1997-2006</i>	-0,026	-0,072	-0,048
Description d'ego	genre	<i>Femme</i>	REF	REF	REF
		<i>Homme</i>	-0,135	-0,150	-0,142
	type_carriere	<i>directeur_de_recherche</i>	REF	REF	REF
		<i>professeur</i>	-0,046	-0,004	0,009
	Pseudo-R² McFadden			0.12	0.13
Critère d' Akaike			204.35	199.74	207.08

Tableau 59 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 4 en %

			Modèle 4'a.	Modèle 4'b.	Modèle 4'c.
			dy/dx	dy/dx	dy/dx
Dynamique structurelle du réseau	classe_precedente	<i>Classe 1</i>	-0,619 ***		
		<i>Classe 2</i>	-0,341 ***		
		<i>Classe 3</i>	-0,331 ***		
		<i>Classe 4</i>	REF		
		evol_dim1		0,128 ***	
		evol_dim2		0,026	
	valeur_abs_evol_dim1			0,031	
	valeur_abs_evol_dim2			-0,133 ***	
Dynamiques de publication	coef_mult_pub		0,123 ***	-0,070 *	0,029
	mob	<i>sed_a_sed</i>	REF	REF	REF
		<i>sed_a_mob</i>	0,161	0,035	0,115
		<i>mob_a_sed</i>	0,013	-0,017	0,038
		<i>mob_a_mob</i>	0,194 *	0,123	0,200 *
Temporalité de la carrière d'ego	evol_carriere	<i>jeune_chercheur_a_chercheur_conf</i>	REF	REF	REF
		<i>chercheur_conf_a_senior</i>	-0,199 *	-0,012	-0,140
		<i>senior_a_senior_av</i>	-0,121	0,070	-0,113
	generation	<i>1967-1976</i>	REF	REF	REF
		<i>1977-1986</i>	0,108	0,192 *	0,077
		<i>1987-1996</i>	-0,024	0,042	-0,086
<i>1997-2006</i>		0,029	0,059	-0,074	
Description d'ego	genre	<i>Femme</i>	REF	REF	REF
		<i>Homme</i>	0,116	0,178 *	0,192 **
	type_carriere	<i>directeur_de_recherche</i>	REF	REF	REF
		<i>professeur</i>	-0,134	-0,221 ***	-0,220 ***
Pseudo-R² McFadden			0.25	0.23	0.21
Critère d'Akaike			207.02	209.96	213.68

IV-Conclusion

L'objectif de ce chapitre est **d'identifier et de décrire plusieurs styles de collaboration des chimistes**. Nous partons de l'hypothèse que la structure du réseau de co-authorship est source d'informations relatives au comportement collaboratif du chercheur. Ce chapitre est composé de plusieurs étapes qui ont permis l'identification des styles collaboratifs, leur caractérisation et leur estimation à partir de données de carrière et de dynamiques structurelles.

A partir des 259 réseaux personnels ego-centrés présentés dans le chapitre 2, nous avons mobilisé 13 indicateurs d'analyse de réseaux sociaux qui décrivent les relations collaboratives *ego-alter* (indicateurs de position) et l'organisation des liens collaboratifs (indicateurs de structure) au sein de chaque réseau. Cette caractérisation structurelle nous permet d'identifier les particularités de chaque style collaboratif. Grâce à une ACP, nous avons enrichi ces treize variables de deux variables supplémentaires décrivant deux comportements collaboratifs relatifs à **l'hétérogénéité des relations *ego-alter* dans le réseau** (dimension 1) et la **segmentation des relations collaboratives dans le réseau** (dimension 2). Nous avons intégré ces quinze variables dans l'algorithme de classification hiérarchique afin de regrouper nos 259 réseaux en groupes homogènes. Nous avons identifié **quatre types de structures de réseaux de co-authorship différents** : « les réseaux très denses », « les réseaux éclatés », « les réseaux denses aux relations individuelles » et « les réseaux gravitaires avec une dualité des relations ».

Nous avons réalisé un travail approfondi de caractérisation structurelle approfondie de chaque classe :

i) Nous avons commencé par **décrire chaque classe à partir d'indicateurs structurels** qui ont le plus contribué à la classification. L'objectif était de bien comprendre les implications collaboratives derrière les caractéristiques structurelles de chaque classe de réseaux. Cette caractérisation s'est faite d'un point de vue statique (caractériser chaque classe par rapport aux autres) et dynamique (caractériser le processus de classification et réintégrer chaque classe dedans).

ii) Nous avons ensuite **testé la probabilité d'obtenir un réseau d'un type particulier** à partir de modèles économétriques *logit* binaires. L'objectif était d'enrichir la caractérisation à partir de **données relatives à la carrière et aux caractéristiques du co-authorship**. Ces tests économétriques ont permis de comprendre dans quelles conditions chaque style collaboratif avait plus de chance d'apparaître.

iii) Enfin, nous avons développé une **perspective dynamique** afin de caractériser, pour chaque *ego*, **un style collaboratif en fonction de celui opéré lors du stade de carrière précédent**. L'objectif était d'expliquer chaque style collaboratif à partir de sa dynamique structurelle, du rythme de publication et de la mobilité géographique du chercheur. Une attention particulière a été apportée pour caractériser la dynamique structurelle du réseau.

Nous sommes donc en mesure de dresser la typologie de quatre styles collaboratifs :

Tableau 60 : Classe 1 : Réseaux « très denses »

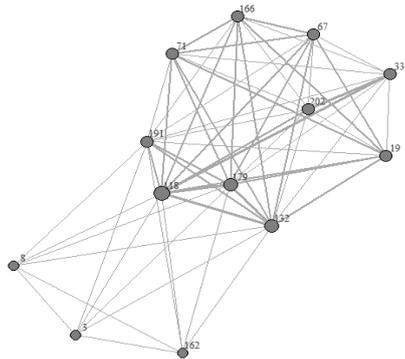
<i>Facteurs significatifs :</i>	<i>Exemple de représentation graphique :</i>
<ul style="list-style-type: none"> - Ce style collaboratif concerne davantage les chercheurs durant leurs dix premières années de carrière, lorsque leur rythme de publication est faible et que leurs relations collaboratives sont encore homogènes - Le poids des one-shot dans les collaborations d'<i>ego</i> réduit la probabilité d'avoir un réseau de cette structure. -Le maintien de relation avec <i>alter</i> réduit également la probabilité d'avoir cette structure de réseau, quelle que soit la géographie d'<i>alter</i> : <i>ego</i> réalise donc peu de publications avec chacun de ses <i>alter</i>. - Collaborer avec des équipes à dimension nationale augmente la probabilité d'avoir ce type de structure : <i>ego</i> collabore avec des sous-espaces collaboratifs nationaux qu'il mobilise dans plusieurs de ses publications avec différentes équipes - Etre une femme ou suivre une carrière de professeur accroît la probabilité de suivre ce style collaboratif 	
<i>Dynamique structurelle :</i>	
<ul style="list-style-type: none"> - A partir du stade de « chercheur confirmé », il est plus probable d'avoir un réseau de classe 1 lorsque le réseau se densifie quel que soit le niveau d'hétérogénéité entre les relations <i>ego-alter</i> - Ce bouleversement structurel s'explique par la chute du volume de publication via la réduction et la contraction du réseau - Il est aussi significativement induit par à un bouleversement dans la trajectoire géographique : notamment par une mobilité à la suite d'une période de sédentarité géographique 	

Tableau 61 : Classe 2 : Réseaux « éclatés »

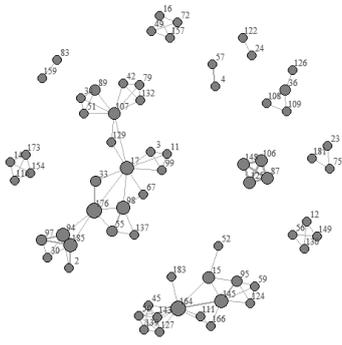
<i>Facteurs significatifs :</i>	<i>Exemple de représentation graphique :</i>
<ul style="list-style-type: none"> - La collaboration avec des one-shot augmente la probabilité d'avoir un réseau éclaté en cliques de relations ponctuelles. - Ensuite, cette structure concerne davantage les chercheurs en fin de carrière (autonomie collaborative, ouverture internationale et émancipation du laboratoire). 	
<i>Dynamique structurelle :</i>	
<ul style="list-style-type: none"> - Ce type de réseau provient de dynamiques structurelles particulières : une homogénéisation des relations collaboratives de petite ampleur et une segmentation des relations collaboratives de très forte ampleur. 	

Tableau 62 : Classe 3 : Réseaux « denses aux relations individuelles »

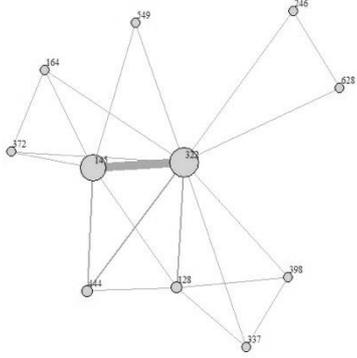
<i>Facteurs significatifs :</i>	<i>Exemple de représentation graphique :</i>
<ul style="list-style-type: none"> - La collaboration avec des relations ponctuelles telles que les one-shot et les doctorants en formation réduit la probabilité d'avoir ce type de réseau puisqu'<i>ego</i> ne tisse pas de relation individuelle et singulière avec ce type d'<i>alter</i>. - La collaboration avec des hommes et avec des <i>alter</i> qui réalisent ou ont réalisé leur thèse sous la direction d'<i>ego</i> augmente la probabilité d'avoir ce type de réseau. <i>Ego</i> manifeste plus de proximité et de relations individuelles avec ce type de co-auteurs. - Ces réseaux concernent davantage les chercheurs durant leur début ou leur milieu de carrière : la densité du réseau étant conditionnée par un faible rythme de publications. - Etre une femme augmente la probabilité d'opter à ce style collaboratif 	
<i>Dynamique structurelle de ce type de réseau</i>	
<ul style="list-style-type: none"> - Il n'existe pas de trajectoire structurelle propre à cette classe : elle se manifeste par une densification du réseau et une hétérogénéité des relations collaboration d'ampleurs non significatives. - La réduction du volume de publication augmente la densité globale du réseau et donc la probabilité d'avoir un réseau de classe 3 	

Tableau 63 : Classe 4 : Réseaux « gravitaires »

<i>Facteurs significatifs :</i>	<i>Exemple de représentation graphique :</i>
<p>- Cette structure de réseau concerne davantage les chercheurs en fin de carrière : elle s'explique par l'intensification du rythme de publication, l'élargissement et de l'intensification des relations collaboratives avec les locaux et les nationaux.</p> <p>- La dualité des relations est avant tout géographique : quelques <i>alter</i> locaux et nationaux structurent les liens collaboratifs alors que les autres et les internationaux gravitent autour d'eux.</p> <p>- Etre un homme, suivre une carrière de directeur de recherche et avoir commencé sa carrière entre 1977 et 1987 accroissent la probabilité d'opter ce style collaboratif.</p>	
<i>Dynamique structurelle de ce type de réseau :</i>	
<p>- Forte stabilité structurelle : le réseau se déplace vers la droite de l'axe 1 d'une portée non significative et très faiblement sur l'axe 2.</p> <p>- Cette structure se traduit par une hausse du volume de publication par rapport à la période précédente ainsi qu'une mobilité continue : la multi-affiliation accroît la segmentation des relations collaboratives contrairement à la sédentarité qui densifie le réseau.</p>	

Conclusion générale

Cette thèse avait pour but de contribuer à la **compréhension de la dynamique sociale et spatiale des collaborations des chimistes**. Beaucoup de travaux cherchent à évaluer l'impact des collaborations en termes de productivité à l'échelle d'un territoire ou d'un individu. L'enjeu de cette thèse était de permettre une meilleure compréhension des collaborations selon une perspective dynamique en y intégrant les dimensions sociales et spatiales. Nous avons pour cela développé notre analyse selon une démarche compréhensive basée sur des données individuelles et longitudinales.

Cette thèse fournit deux types de contributions :

i) **La première est d'ordre méthodologique**. Notre thèse a permis une meilleure caractérisation des collaborations scientifiques à l'échelle individuelle au travers de données longitudinales. Cette contribution repose sur la prise en compte du contexte des collaborations au travers de l'analyse de la carrière des chercheurs. Nous avons pour cela créé une base de données originale et réalisé différents traitements complémentaires : économétrie sur données de panel, analyse de réseaux sociaux, analyse qualitative.

ii) **La seconde contribution est empirique**, les données mobilisées permettant une meilleure compréhension du comportement collaboratif des chercheurs, des dynamiques sociale et géographique du co-authorship.

Après avoir détaillé ces deux contributions, nous présenterons les perspectives de recherches qui s'ouvrent face aux limites de cette thèse.

I- Caractériser la dynamique des collaborations scientifiques grâce à des données individuelles et longitudinales

A) Contextualiser l'analyse des collaborations scientifiques

Dans la revue de la littérature, nous avons montré que la dynamique des collaborations est fortement influencée par la carrière du chercheur, processus complexe et multi-dimensionnel qui opère sous différentes contraintes.

i) Le chercheur, considéré comme un agent autonome maximisateur de revenu, agit en de ses objectifs et de ses attentes personnelles.

ii) La seconde contrainte concerne l'environnement scientifique dans lequel le chercheur évolue, composé d'une communauté nationale et internationale de pairs : nous avons notamment montré que la mobilité scientifique est un évènement issu en partie de la pression de cet environnement.

iii) La troisième contrainte concerne le contexte évolutif qui structure et définit le métier d'enseignants-chercheurs : le monde de la science a particulièrement évolué et influence la carrière des chercheurs.

Il nous semble nécessaire de prendre en compte ces différentes dimensions de la carrière pour comprendre le comportement collaboratif du chercheur. Il nous semble primordiale de les intégrer afin d'apporter une meilleure interprétation du comportement collaboratif. **L'originalité de la thèse est donc de mobiliser des données individuelles et longitudinales relatives à la carrière scientifique permettant de contextualiser l'activité de co-authorship des chercheurs.**

B) La création d'une base de données bibliométrique originale

Afin de comprendre le comportement collaboratif des chercheurs, nous avons privilégié une entrée par l'individu avec la carrière du chercheur comme échelle d'analyse. Nous nous focalisons sur la chimie qui constitue selon nous une discipline pertinente pour l'étude des dynamiques collaboratives. Nous avons sélectionné 80 chercheurs affiliés aux laboratoires de Poitiers et Toulouse et ayant un poste de directeur de recherche CNRS ou professeur en 2018. A partir de l'extraction de l'ensemble de leurs publications sur le Web of Science et de la création de paires de co-authorship (Melin and Persson, 1996), nous avons reconstitué les dynamiques de publication et de collaboration des 80 chercheurs tout au long de leur carrière académique⁵⁷. La base de données rassemble des données relatives à 9233 co-publications réalisées par les 80 chercheurs (*ego*) dans lesquelles ils ont collaboré 41720 fois avec 14603 co-auteurs différents (*alter*). **L'originalité de la thèse repose à la fois sur la reconstitution de dynamiques collaboratives de chercheurs durant leur carrière et sur la mobilisation d'une base de données individuelles originale.** Cette base intègre des données relatives à la localisation d'*ego* et de chacun de ses *alter* à (ville, pays). Cette opération a nécessité de coder manuellement les localisations géographiques de chaque couple *ego-alter* grâce à la recherche des notices des articles réalisés avant 2008. Ces informations ont permis de reconstituer la trajectoire géographique d'*ego* et de contextualiser chaque collaboration en fonction de la distance géographique de chaque couple *ego-alter*. Nous avons également recherché des informations relatives à la description de chaque *alter* (genre, possession éventuelle de doctorat et année du diplôme) et la recontextualisation de chaque *alter* dans la carrière d'*ego* en tant qu'éventuel doctorant ou directeur de thèse d'*ego*.

⁵⁷ La carrière la plus courte étant de 12 ans et la plus longue de 51 ans. La durée de carrière moyenne de notre échantillon est 30 ans.

C) Des méthodes d'analyse variées et complémentaires

La principale contribution méthodologique de la thèse est d'expliquer le phénomène de collaboration au fil de la carrière en mobilisant plusieurs méthodes d'analyse qui ont comme point commun d'intégrer l'échelle temps (Bidart et Mendez, 2016).

Nous avons commencé par mobiliser une méthodologie mixte d'analyse grâce à la combinaison d'approches qualitatives et quantitatives (base de données bibliométriques) pour comprendre la dynamique du réseau de co-authorship au fil de la carrière. Nous avons combiné une variété de données qualitatives issues de matériaux différents (CV, entretiens semi-directifs, articles) pour réaliser une analyse relationnelle de l'activité de publication. La collecte de ces matériaux permet de comprendre l'enchevêtrement des séquences de carrière, le sens de causalité des phénomènes observés et d'intégrer de nouveaux déterminants non identifiés dans les base de données disponibles (i.e. Vos). **Ces données qualitatives complémentaires permettent de donner une épaisseur sociale aux résultats bibliométriques issus des bases de données existantes : ce type d'analyse a rarement été réalisé jusqu'alors.**

Dans un deuxième temps, nous avons réalisé une analyse exclusivement quantitative de données grâce à la réalisation de modèles de panel. L'objectif était d'évaluer l'impact de la mobilité géographique sur la diversification sociale et spatiale du réseau de co-authorship. Cette méthode apparaît particulièrement appropriée pour l'étude des trajectoires individuelles sur données longitudinales. La mobilisation de modèles à effet fixe via un modèle « within » est apparu pertinent **pour rendre compte des dynamiques collaboratives au sein des carrières individuelles et a permis d'intégrer les mobilités des chercheurs en tant qu'évènement de la carrière.**

Enfin, nous avons réalisé une analyse de réseaux pour analyser plus finement le comportement collaboratif des chercheurs en tenant compte des caractéristiques structurelles des réseaux *ego*-centrés. L'objectif de cette étude est la mise en lumière d'une pluralité des styles collaboratifs via la création d'une typologie des structures de réseau. Pour ce faire, nous avons segmenté les carrières en périodes régulières (jeune chercheur, chercheur confirmé senior, senior avancé) afin de pouvoir appréhender l'évolution du style collaboratif. **L'originalité de cette étude tient en la reconstitution de réseaux personnels sur données longitudinales à l'échelle individuelle.** En effet, de nombreuses études mobilisent des réseaux complets sur des courtes périodes dans un souci d'évaluation de performance.

Finalement, la mobilisation dans cette thèse de différentes méthodes d'analyse de données diversifiées nous a permis de contextualiser la dynamique collaborative des chercheurs.

II- Compréhension du comportement collaboratif des chercheurs au fil de la carrière

L'étude de l'activité de collaboration du chercheur a été réalisée à deux échelles temporelles : i) une échelle longitudinale (par année d'activité), ii) une échelle par période d'activité.

A) Appréhension de la carrière par année de publication : Contextualisation des dynamiques sociales et géographiques collaboratives

Les différents résultats issus de cette thèse montrent l'importance de **la mise en contexte des dynamiques collaboratives**. Nous avons montré que la nature collaborative des articles dépend de la progression de la carrière du chercheur et du contexte historique scientifique de publication. Le rythme de publication annuel augmente au fil de la carrière pour toutes les générations de chercheurs. Au travers des différents chapitres, nous avons été en mesure d'expliquer cette dynamique collaborative sous différents aspects.

1- Explication de la dynamique collaborative en fonction la composition de l'équipe de publication

Nous avons commencé par expliquer la croissance de la nature collaborative des articles en fonction de la composition sociale et géographique du co-authorship. Cette augmentation progressive de la taille du co-authorship ne s'explique pas par la création de relations : le nombre moyen de nouveaux co-auteurs par publication est stable au fil de la carrière (environ 1,5 co-auteurs par article). **Ainsi, l'intensification des collaborations provient du maintien des relations collaboratives et non de la création de nouvelles relations**. Aussi, d'un point de vue géographique, l'équipe de publication est constituée d'un noyau stable et structurant de collègues du laboratoire au fil de la carrière (environ 2 co-auteurs par article). **La hausse des collaborations s'explique par l'internationalisation progressive du co-authorship**.

2- Explication de la dynamique collaborative via la contextualisation des collaborations

A partir d'une étude comparative de cinq chercheurs d'une même génération issus d'un même territoire (Poitiers), nous avons souligné l'homogénéité de la dynamique collaborative. **Cette homogénéité de la dynamique sociale et géographique du co-authorship résulte du rôle prépondérant du contexte de carrière (stade de carrière) et du contexte historique de la recherche**.

i) La première période identifiée concerne les collaborations réalisées durant les années 1960 et 1970 : il s'agit de « l'âge d'or de la recherche ». A cette époque, les laboratoires de chimie émergeaient à Poitiers sous la coupe de grandes figures nationales. Les chercheurs ont donc débuté leur carrière dans

un environnement peu concurrentiel au sein d'équipes où les chefs d'équipe étaient très impliqués. De ces deux dimensions conjoncturels résultent des **collaborations fortes localement** et reposant sur des **relations interpersonnelles** avec les grandes figures de la chimie.

ii) La seconde période identifiée concerne les collaborations réalisées durant les années 1980 et 1990 : il s'agit de l'amorçage de la « course à la publication ». En effet, la recherche se professionnalise : publier et collaborer deviennent des critères de performance et la concurrence s'intensifient sur le marché de l'emploi scientifique. De plus, les laboratoires de Poitiers s'ouvrent à la mobilité géographique du fait de leur renommée croissante. En parallèle, les chercheurs prennent des postes de supervision d'entité de recherche et assurent une forme de centralité dans le collectif. A l'intersection de ces deux contextes, **les collaborations s'intensifient et le chercheur entretient des relations hétérogènes avec ses co-auteurs**. Aussi, les **collaborations locales restent très structurantes** (collègues, étudiants) et nous notons une **internationalisation progressive** du co-authorship.

iii) La troisième période identifiée concerne les collaborations réalisées durant les années 2000-2010 : il s'agit d'une phase de « restructuration locale de la chimie ». Au fil des différentes réformes nationales, la rationalisation s'accroît dans la recherche et les laboratoires de Poitiers fusionnent en une même structure. Les chercheurs terminent leur carrière et adoptent un statut de post-éméritat : ils réalisent leur activité avec une certaine autonomie vis-à-vis du laboratoire où ils sont affiliés. Les implications collaboratives sont les suivantes : des **relations très hétérogènes avec les membres de l'équipe du laboratoire** (en fonction des relations interpersonnelles) et le **renforcement de relations collaboratives internationales** (reconduction des collaborations).

3- Explication de la dynamique collaborative via la mobilité géographique du chercheur

Enfin, nous avons testé l'impact de la mobilité géographique dans la dynamique du réseau de co-authorship. Alors de nombreux travaux évaluent l'effet de la mobilité sur performance individuelle (productivité scientifique, profession de carrière) ou territoriale, très peu d'auteurs évaluent l'effet direct des flux migratoires sur les dynamiques de co-authorship à l'échelle individuelle. Nous avons donc évalué l'impact du changement d'affiliation du chercheur sur la construction de son réseau de co-authorship au cours du temps. Cette construction est appréhendée sous le prisme de la diversification sociale (créations et ruptures de collaborations) et de la diversification géographique (villes collaboratives incluses ou exclues dans la trajectoire géographique du chercheur). **Nous avons montré que l'expérience de mobilité augmentait la probabilité d'intégration de nouveaux co-auteurs mais qu'elle n'était que très partiellement gage de diversification du réseau** pour trois raisons :

i) La mobilité a un effet instantané sur la diversification sociale du réseau. En effet, cet effet de diversification sociale du réseau ne se traduit que l'année où la mobilité a été réalisée et ne perdure pas au fil de la carrière.

ii) La mobilité n'a pas d'impact sur la diversification géographique du réseau en dehors de la trajectoire géographique d'*ego* : l'intégration de nouveaux co-auteurs est restreinte aux villes fréquentées par *ego*. Dit autrement, on n'observe pas d'effet de « contagion de la mobilité » à d'autres villes.

iii) La mobilité fragilise le réseau de co-authorship : on observe une rupture des collaborations avec les anciens co-auteurs à la suite d'une mobilité et cet effet est plus long dans la carrière que l'effet positif de création de relations.

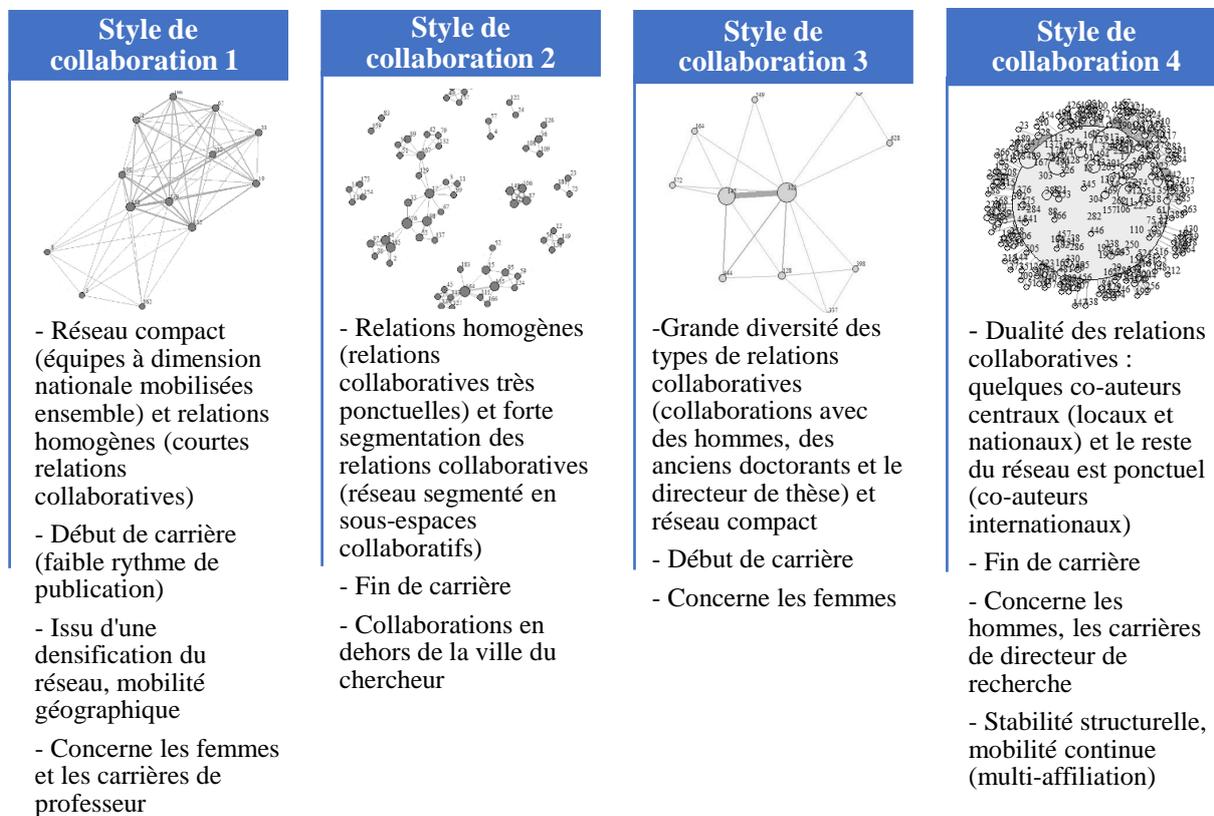
B) Création d'une typologie de styles collaboratifs

Nous avons poursuivi notre objectif de compréhension des dynamiques collaboratives des chercheurs via la création d'une typologie de différents styles collaboratifs. Dans un premier temps, nous avons choisi d'identifier le comportement collaboratif des chercheurs en fonction des caractéristiques structurelles de réseaux. A partir de 259 réseaux *ego*-centrés, nous avons identifié **quatre styles de collaboration** qui sont caractérisables sous deux dimensions. La première dimension concerne **l'hétérogénéité des relations *ego-alter***, à savoir dans quelle mesure chaque chercheur entretient des relations différentes avec ses co-auteurs. La seconde dimension concerne la **segmentation des relations collaboratives**, à savoir si *ego* mobilise tous ses co-auteurs dans des équipes conjointes (réseau considéré comme une entité) ou s'il segmente ses **collaborations** (réseau compartimenté en plusieurs composants). Ces quatre styles collaboratifs regroupent donc des réseaux disposant de similitudes structurelles relatives.

Dans un second temps, **nous avons cherché à expliquer ces quatre styles collaboratifs à partir de plusieurs facteurs relatifs à la carrière scientifique du chercheur** : en fonction de la caractérisation des co-auteurs (hommes, one-shot, doctorants en formation, anciens/actuels doctorants d'*ego*, directeur de thèse d'*ego*), de la composition géographique du co-authorship et de la durée de ces relations. Ensuite, nous avons **réintroduit l'analyse des styles collaboratifs dans la trajectoire des chercheurs** afin de comprendre dans quelle mesure le passage d'un stade de carrière à l'autre se traduit par un changement du style collaboratif. Pour ce faire, nous introduisons des facteurs relatifs à la dynamique structurelle du réseau et à la dynamique de l'activité de publication.

Ainsi, l'originalité de cette contribution tient dans **création d'une typologie de styles collaboratifs** à l'échelle microéconomique à partir de la méthode sociologique de l'analyse des réseaux sociaux :

Figure 49 : Typologie des styles de collaboration



III- Limites et perspectives de recherche possible

A) Elargir l'échantillon

L'échantillon ne concerne qu'une seule discipline et deux laboratoires français. L'homogénéité des pratiques scientifiques en chimie et dans les laboratoires IC2MP et LCC limite les possibilités de généralisation des résultats. Il nous semblerait intéressant, à partir de ces premiers résultats, d'élargir l'échantillon d'étude afin de tester les spécificités du comportement collaboratif des chimistes. En effet, **augmenter la taille de l'échantillon permettrait de consolider nos résultats et d'apporter des précisions sur les collaborations en prenant en compte d'autres particularités des carrières académiques.**

L'élargissement à d'autres disciplines pourrait permettre d'identifier les **dynamiques collaboratives communes au métier de chercheur et spécifiques aux disciplines.** Néanmoins, des précautions doivent être prises concernant le repérage des relations via l'activité de publication. En effet, la tradition

de valorisation en langue nationale en sciences humaines et sociales peut échapper aux bases de données de type Web of Science : il y a un risque de moins bonne couverture des revues dans ces disciplines.

Également, l'élargissement géographique de l'échantillon (au-delà de Poitiers et Toulouse) pourrait enrichir notre appréhension de la carrière scientifique en intégrant **le rôle de l'établissement**. En effet, nous avons choisi de ne pas intégrer le lieu d'affiliation en 2018 dans nos facteurs explicatifs des dynamiques collaboratives pour deux raisons : i) tous les chercheurs n'ont pas commencé leur carrière dans ces laboratoires, ii) le nombre de laboratoires est trop faible pour être représentatif d'une structure interne particulière. Ainsi, il pourrait être intéressant d'intégrer de nouveaux laboratoires disposant de statuts différents (UMR, UPR, EA, UMI, LRC⁵⁸,...) afin d'identifier l'effet de la structure et des réglementations internes de chaque entité sur l'activité collaborative des chercheurs. Il serait intéressant d'évaluer dans quelle mesure faire sa carrière en étant affilié à tel type de laboratoire influence le comportement collaboratif du chercheur. Aussi, l'intégration à l'analyse d'autres laboratoires réputés en chimie (ex : Strasbourg, Paris, Lyon) serait pertinente puisque ces territoires sont pleinement intégrés dans les collaborations de chimie nationales et mondiales.

B) Multiplier la mise en intrigue de carrière sur différents groupes de l'échantillon

Il serait également pertinent de répéter l'étude analytique des carrières basée sur la méthodologie mixte (bibliométrie, CV, entretiens semi-directifs) sur **différents groupes de chercheurs de l'échantillon**. Dans la thèse, nous avons appliqué cette méthodologie sur cinq chercheurs qui partagent la même discipline, la même génération et le même laboratoire source (Poitiers). Il serait intéressant de répéter notre méthodologie sur un groupe de chimistes, par exemple issu d'une génération plus jeune. Ainsi, la mise en intrigue des dynamiques collaboratives de jeunes chercheurs permettrait de mieux comprendre la croissance de la nature collaborative au fil des générations, l'influence des pressions exercées par l'environnement de recherche et la progression de carrière.

Également, il pourrait être intéressant de **multiplier les cas en croisant les contextes**. Par exemple, nous pourrions réaliser ce type d'étude sur une même discipline auprès de plusieurs générations ou bien sur une même génération auprès de différentes disciplines. Ainsi, nous ne serions plus dans la description de trajectoires individuelles homogènes mais dans la **confrontation de trajectoires différentes** : cela nous permettrait d'identifier les particularités relatives à la génération ou à la discipline.

C) Intégrer des données individuelles complémentaires

⁵⁸ Unités mixtes de recherche, Unité propre de recherche, Equipe d'accueil, Unité mixte internationale, Laboratoire de recherche commun

L'originalité de la thèse tient en partie dans la mobilisation de données individuelles relatives aux carrières de chercheurs pour interpréter et expliquer les dynamiques collaboratives. Nous avons intégré des données relatives à la géographie des couples *ego-alter* pour chaque collaboration, à la description d'*alter* (genre, année d'obtention du doctorat) et de la relation « hiérarchique » entre *ego* et *alter* (directeur de thèse et liste des doctorants d'*ego*). Nous pensons que des données complémentaires pourraient enrichir notre appréhension des dynamiques collaboratives. Notamment, des nouvelles données pourraient réduire le risque de biais d'endogénéité dans l'explication du comportement collaboratif, par exemple, dans l'étude de l'impact de la mobilité géographique sur la construction du réseau de co-authorship. **Ces données supplémentaires concerneraient aussi bien la dimension professionnelle que personnelle de la vie des chercheurs.** A partir du CV, nous pouvons introduire des variables relatives à l'emploi, la promotion, l'activité d'enseignement et les responsabilités administratives qui composent la trajectoire professionnelle du chercheur et peuvent influencer indirectement le comportement collaboratif. Également, à partir d'entretiens, nous pourrions intégrer des éléments relatifs à la trajectoire personnelle du chercheur potentiellement déterminants (lieu de vie, situation matrimoniale, nombre d'enfants...).

D) Qualifier les thématiques scientifiques de l'activité de publication

Enfin, il serait intéressant d'identifier les thématiques scientifiques suivies dans l'activité des chercheurs pour voir si elles déterminent leur comportement collaboratif. **L'intérêt serait de vérifier si les dynamiques de publication et de collaboration varient en fonction de la trajectoire thématique du chercheur.** Nous pourrions mobiliser la thématique de l'équipe de recherche dans laquelle chaque chercheur est intégré. Néanmoins, ces informations sont difficiles à obtenir notamment dans une approche longitudinale puisque tous les chercheurs n'ont pas passé leur carrière dans l'équipe et ont très bien pu changer de spécialité au fil de sa carrière. Ainsi, deux autres approches pourraient être intéressantes. La première consiste à repérer les thématiques de recherche suivies via l'étude du CV durant les différents événements de carrière (formations suivies, matières enseignées, colloques, participation à des jurys...). La seconde approche concerne la mobilisation des mots clefs des articles. Via l'extraction de ces données, il est possible d'identifier à quel moment dans la carrière le chercheur a commencé à travailler sur une thématique particulière. Néanmoins, des entretiens semblent nécessaires dans cette démarche pour compléter la compréhension de la dynamique thématique des publications compte tenu des limites inhérentes à l'analyse des articles par mots clés (Milard, 2010, 2013, 2014). En effet, de nombreux mots clefs en chimie sont utilisés sous forme d'abréviation et de symbole (par exemple le terme « HPLC » signifie « chromatographie en phase liquide à haute performance ») : leur interprétation n'est pas évidente pour un chercheur exerçant en dehors de cette discipline. Aussi, le matériel narratif permettrait de comprendre et de contextualiser la trajectoire thématique de recherche et d'identifier quels mots clefs peuvent être regroupés ensemble dans une même thématique.

Annexes

Annexe 1 : Sections CNRS de la chimie

Tableau 64 : Sections CNRS de la chimie

Sections du CNRS	Description :
11- Systèmes et matériaux supra et macromoléculaires : élaboration, propriétés, fonctions	<ul style="list-style-type: none"> - Physico-chimie et physique de la matière molle (systèmes auto-assemblés, colloïdes, fondus, fluides complexes, films minces, interfaces, systèmes bio-inspirés, ...) - Conception, synthèse et propriétés d'objets moléculaires, supramoléculaires et macromoléculaires et leurs assemblages, systèmes stimulables - Chimie des polymères synthétiques et bio-sourcés et procédés de polymérisation - Catalyses de polymérisation - Elaboration et propriétés des matériaux polymères de fonction (composites, nanocomposites, hybrides, biomatériaux, membranes, ...) - Procédés de mise en forme des polymères - Durée et cycle de vie des systèmes supra et macromoléculaires - Etude chimique et physique de systèmes biologiques, de leurs propriétés et fonctions ; approches inspirées de la matière molle, bio-mimétisme
12- Architectures moléculaires : synthèses, mécanismes et propriétés	<ul style="list-style-type: none"> - Méthodes et concepts de synthèse organique - Synthèses multi-étapes, produits naturels et molécules bioactives - Synthèse éco-compatible - Hétérochimie, chimie organométallique - Catalyse organométallique, organique, supramoléculaire, multiple, duale pour la synthèse organique - Biocatalyse et catalyse biomimétique pour la synthèse organique - Valorisation des ressources naturelles - Technologies innovantes (photochimie, mécanochemie, haute pression, flux continu, microréacteurs, ...) - Synthèse, caractérisation et contrôle des propriétés de matériaux moléculaires et hybrides pour l'optique, l'électronique et la biologie - Chimie supramoléculaire (encapsulation, autoassemblage) - Physicochimie organique, méthodologies analytiques optiques et magnétiques, spectrométrie de masse, méthodes pour l'imagerie analytique - Etude de la réactivité et des mécanismes réactionnels : spectroscopie optique, vibratoire, modélisation moléculaire
13- Chimie physique, théorique et analytique	<ul style="list-style-type: none"> - Mécanismes, interfaces, réactivités, transport et énergie - Chimie Théorique et Chemo-informatique : théories, méthodes, méthodologies, modélisations et simulations numériques - Electrochimie moléculaire, biomoléculaire, localisée, nano-électrochimie - Radiochimie - Thermodynamique - Photochimie et photophysique moléculaires et biomoléculaires - Spectroscopies et imageries : méthodologies, théories et applications - Chimie Analytique - Chimie de l'environnement : dynamique des polluants, multi-comparti, Hommes
14- Chimie de coordination, catalyse, interfaces et procédés	<ul style="list-style-type: none"> - Complexes organométalliques de coordination - Systèmes hybrides - Modèles bio-inorganiques

-
- Matériaux moléculaires à base de métaux
 - Physicochimie et réactivité des surfaces et des Interfaces, corrosion, traitement, Hommes de surfaces
 - Nanostructures et nano-chimie
 - Electrochimie interfaciale, électro-catalyse
 - Catalyseurs, catalyse homogène et hétérogène
 - Procédés catalytiques, ingénierie et modélisation cinétique, dépollution, conversion et stockage de l'énergie, traitement et valorisation de la biomasse

15- Chimie des matériaux, nanomatériaux et procédés

- Chimie du solide, Chimie de la matière condensée
- Modélisation, synthèse et caractérisation des matériaux
- Cristaux, céramiques, amorphes, multi-matériaux
- Revêtement, Hommes et couches fonctionnelles
- Matériaux hybrides et bio-inspirés
- Science et génie métallurgiques
- Thermodynamique métallurgique, procédés d'élaboration et de traitement
- Matériaux à échelle multiple ou hiérarchisée
- Matériaux pour l'énergie (batteries, photovoltaïque, piles à combustible, thermoélectrique...)
- Matériaux pour l'optique, le stockage et la transmission de l'information
- Matériaux pour la santé, biomatériaux

16- Chimie du vivant

- Conception, synthèse, analyse et propriétés de molécules d'intérêt biologique
- Chemo-biologie : Chimie pour explorer le vivant
- Concepts et outils moléculaires pour la chimie médicinale
- Chimie bio-organique, bio-inorganique et bio-inspirée
- Méthodologies pour la synthèse de molécules bioactives et de biomolécules
- Chimie in vivo
- Substances naturelles : extraction, caractérisation et synthèse
- Enzymologie, biocatalyse, bio-ingénierie
- Systèmes biomoléculaires organisés, vectorisation
- Spectroscopies optiques et magnétiques, spectrométrie de masse, méthodes de diffraction, imagerie
- Structure, dynamique et mécanismes moléculaires des biomolécules
- Bio-informatique et modélisation, appliquées aux molécules bioactives
- Aspects moléculaires de l'écotoxicologie et de l'écologie chimique

Source : <https://www.cnrs.fr/comitenational/sections/intitsec.php>

Annexe 2 : Indicateurs des réseaux ego-centrés

A) La position des alter dans le réseau : Quelles sont les différentes relations de collaboration qu'ego entretient avec ses co-auteurs ?

Avant de pouvoir déterminer les différents styles de collaboration, il convient de rappeler deux natures de relation de co-authorship représentées graphiquement dans le réseau d'ego :

La relation en « duo » : Relation de collaboration qu'entretient *ego* avec chacun de ses *alter*. Elle est représentée par le diamètre du nœud. Il s'agit du nombre de publications qu'*ego* partage avec chacun de ses *alter*.

La relation en « trio » : Relation de collaboration qu'entretient simultanément *ego* avec deux de ses *alter*. Elle est représentée par l'épaisseur de l'arrête entre le couple de nœuds en question. Il s'agit du nombre de publications dans lesquelles *ego* collabore simultanément avec ce couple d'*alter*.

1- Relations « duos » entre *ego* et *alter* : Quelles sont les différentes relations d'*ego* avec chacun de ses *alter* ?

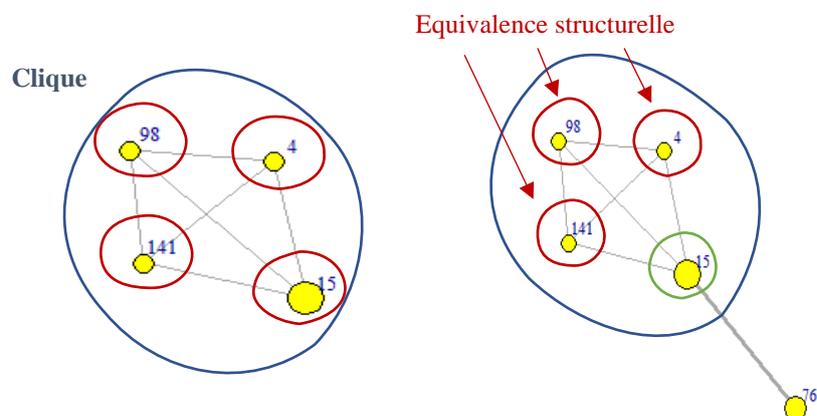
Choix de co-auteurs réguliers : *Ego a-t-il des récurrences plus marquées que d'autres dans le choix de son co-authorship ?*

Chaque nœud (*alter*) est par définition présent dans le réseau pour avoir copublié avec *ego*. Chacun se voit attribué d'un nombre de publications qu'il a partagées avec *ego*. Le coefficient de variation de cette variable indique sa dispersion dans le réseau : est-ce qu'*ego* voit certains *alter* comme des ressources de collaboration beaucoup plus récurrentes que d'autres ? L'écart de récurrence entre ses *alter* est-il très fort ?

Choix de l'équivalence structurelle : *Avec quelle part du réseau ego choisit-il ne pas poursuivre la relation de collaboration en dehors de l'équipe ?*

Dans l'analyse de réseaux, deux nœuds sont jugés équivalents d'un point de vue structurel s'ils partagent les mêmes liens avec les mêmes nœuds du réseau. En d'autres termes, deux co-auteurs structurellement équivalents pourraient être substituables l'un à l'autre sans changer la structure du réseau.

Figure 50 : Représentation graphique de l'équivalence structurelle



Dans le cadre de nos 259 réseaux centrés sur la bibliométrie de nos 80 *ego*, cela correspondrait aux *alter* inclus dans une clique. Une clique est un sous-ensemble du graphe dans lequel tous les nœuds sont reliés entre eux et qui partagent une équivalence structurelle. Le co-authorship d'une publication peut illustrer ce type de relations : chaque *alter* entretient une relation de collaboration simultanée avec *ego* et avec chacun des membres du co-authorship.

Dans le premier schéma de la figure 50, nous voyons que les quatre *alter* ont tous une équivalence structurelle car ils partagent les mêmes liens de collaboration avec les mêmes nœuds. Cependant, si l'un des *alter* de ce réseau accompagne *ego* dans la collaboration avec l'*alter* d'une autre équipe (comme le fait *alter* 15 dans le second schéma), celui-ci n'est plus en situation d'équivalence structurelle comme le reste de son équipe (*alter* 141, 98 et 4). Ainsi, retirer *alter* 15 du réseau de co-authorship bouleverserait la structure globale du réseau, contrairement aux trois autres *alter*.

La relation de collaboration entre *ego* et un *alter* impliqué dans une clique nous semble relativement précaire ou ponctuelle pour plusieurs raisons. Tout d'abord, cela signifie que la relation *ego-alter* est restreinte au cadre de l'équipe d'*alter*. Aux yeux d'*ego*, *alter* n'est qu'un composant de cette équipe : il choisit de ne pas mobiliser cette relation dans d'autres contextes. *Alter* est donc une ressource dont le potentiel collaboratif avec le reste du réseau est jugé nul pour *ego*. Ensuite, l'équivalence structurelle des *alter* en clique sous-entend qu'*ego* a adopté la même stratégie de collaboration pour tous ces nœuds, quelle que soit la contribution réelle d'*alter* dans la publication ou bien la taille de la clique dans laquelle il est présent.

Nous souhaitons donc évaluer le poids des nœuds en équivalence structurelle dans le réseau : nous les identifions comme les nœuds dont la transitivité locale vaut 1.

Choix de l'in-équivalence automorphique : Quelle part du réseau détient une position unique dans le réseau d'ego et qui n'est donc pas substituable ?

Dans l'analyse de réseau, on dit que deux nœuds sont équivalents d'un point de vue automorphique s'ils disposent de mêmes caractéristiques structurelles (et donc si leur position dans le réseau peut être échangée sans bouleverser la structure du réseau). L'équivalence structurelle est donc une application possible de l'équivalence automorphique puisqu'elle concerne ici tous les sommets appartenant à une clique. Cette fois, nous souhaitons plutôt évaluer les *alter* qui ont des degrés singuliers dans le réseau. L'intérêt ici n'est pas d'évaluer la valeur du degré d'*alter* (fort ou faible) mais plutôt de vérifier si *alter* est le seul dans le réseau à avoir cette valeur. Si c'est le cas, *alter* a une position particulière dans le réseau et aucun autre co-auteur ne peut prendre sa place sans bouleverser la structure du réseau. *Ego* a donc une relation de collaboration singulière avec lui.

2-Relations « trios » entre *ego* et chaque couple d'*alter* : Dans quelle mesure certains *alter* sont-ils plus structurants que d'autres dans les relations collaboratives qu'*ego* a avec son réseau ?

L'objectif de ces indicateurs n'est plus d'évaluer l'hétérogénéité des duos de relations *ego-alter* mais plutôt d'évaluer dans quelle mesure une petite partie des nœuds structure l'ensemble des arrêtes du réseau. En d'autres termes, nous voulons vérifier si seulement certains *alter* sont impliqués dans l'ensemble des flux de collaboration entre *ego* et ses *alter*. Nous allons donc étudier l'hétérogénéité de dispersion de trois centralités locales : la centralité de degré, d'intermédiarité et de vecteur propre.

Pour ce faire, nous utilisons à chaque fois la formule de centralisation de Freeman (1979) pour évaluer l'hétérogénéité de la dispersion de la centralité locale des *alter*.

La formule de Freeman est la suivante :

$$C = \frac{\sum_i [c^* - c_i]}{\max_y \sum_i [c^* - c_i]}$$

La formule se décompose en deux parties : le numérateur concerne le réseau étudié et le dénominateur un réseau théorique.

- Dans la première partie, nous trouvons la valeur maximale de centralité (c^*) observée dans les *alter* du réseau. Nous allons donc faire la somme des différences entre la centralité du co-auteur le plus central (c^*) avec les autres membres du réseau (c_i) .

- Dans la seconde partie, nous allons prendre un réseau de même taille que celui étudié mais disposant d'une structure la plus hétérogène possible au niveau des centralités locales. Par exemple, le réseau en « étoile » est la structure qui sera mobilisée pour la centralité de degré (le nœud central a un degré de $N-1$ alors que tous les autres nœuds ont un degré de 1) car c'est la structure la plus hétérogène possible.

La comparaison de la dispersion de ces deux réseaux varie entre 0 et 1. Si la formule vaut 0, alors la centralité des *alter* est très homogène (les relations de collaboration dans le réseau ne sont pas structurées par un petit nombre de nœuds). Si la formule vaut 1, alors la centralité des *alter* est très hétérogène (les relations de collaboration dans le réseau sont très structurées par quelques *alter*).

Centralisation de degré : Est-ce certains alter sont plus récurrents que d'autres dans les relations de collaboration qu'ego a avec le reste de son réseau ?

Degré du nœud j dans le réseau i : $c_{i,j}^d =$ somme des arrêtes incidentes à un noeuf j dans le réseau i

Le degré évalue le niveau de connectivité d'un nœud dans le réseau (le nombre de voisins avec qui il est connecté). Dans nos réseaux, cela revient à évaluer la capacité d'*alter* à être impliqué dans les relations de collaboration qu'*ego* a avec le reste de son réseau. L'objectif de la centralisation des degrés est de vérifier si la dispersion des degrés des *alter* est hétérogène ou pas. Si c'est le cas (forte centralisation de degré), cela signifierait que certains *alter* auraient beaucoup plus d'arrêtes de « trios » de collaboration avec *ego* que d'autres. En d'autres termes, certains *alter* seraient beaucoup plus impliqués que d'autres dans les relations qu'*ego* a avec le reste de son réseau. *Ego* aurait donc des récurrences de collaboration qui structurerait une grande partie de ses relations.

Centralisation d'intermédiarité : Est-ce certains alter sont plus récurrents que d'autres dans les relations de collaboration d'ego avec plusieurs équipes ?

$$\text{Centralité d'intermédiarité d'un nœud } j \text{ dans un réseau } i : c_{i,j}^b = \sum_{s \neq j \neq t} \frac{\sigma_{st}(j)}{\sigma_{st}}$$

Avec :

σ_{st} = le nombre de chemins les plus courts entre le noeud s et le noeud t

$\sigma_{st}(j)$ = le nombre de tels chemins passant par le noeud j

L'objectif de la centralisation de l'intermédiarité d'un réseau est de vérifier si la dispersion des centralités d'intermédiarité des *alter* est hétérogène ou pas. Alors que le degré correspond au niveau d'intégration d'*alter* dans le réseau, la centralité d'intermédiarité traite plutôt de sa position en tant qu'intermédiaire entre plusieurs groupes de sommets. En effet, même un *alter* disposant d'un faible degré peut se situer comme pont entre deux groupes de sommets. Ainsi, dans nos réseaux, la centralité d'intermédiarité va se traduire comme la capacité d'*alter* à être présent non plus dans un grand nombre de relations de collaboration d'*ego* (degré) mais plutôt dans différents groupes de sommets (équipes de recherche). La centralisation de l'intermédiarité va donc évaluer si ce type de position dans le réseau est présente.

Si la centralisation tend vers 0, cela signifie soit qu'*ego* choisit collaborer simultanément avec plusieurs membres de différentes équipes de collaboration, soit qu'il collabore indépendamment les unes des autres. Si elle tend vers 1, cela signifie qu'*ego* est accompagné par quelques rares *alter* (intermédiaires) dans la collaboration avec deux groupes de sommets différents.

Centralisation de vecteur propre : Est-ce certains alter sont plus récurrents que d'autres dans les relations de collaboration d'ego avec ses alter centraux ?

Centralité de vecteur propre d'un nœud j dans un réseau i : $c_{i,j}^e = \frac{1}{\lambda} \sum_{u \neq j} A_{i,j,u} c_{i,u}^e$

Avec :

λ = la plus grande valeur du vecteur propre

$A_{i,j,u}$ = la matrice adjacente du réseau i définie par $a_{i,j,u}$

= 1 si le nœud j est lié avec le nœud u ($a_{i,j,u} = 0$ dans la cas contraire)

$c_{i,u}^e$ = centralité de vecteur propre du nœud j dans le réseau i

Cette centralité a été définie par Bonacich (1987). Alors que la centralité de degré se limite à définir un nœud central comme connecté avec beaucoup de voisins, la centralité de vecteurs propres complétera cette définition en conditionnant la centralité du voisinage. En effet, pour deux nœuds ayant le même degré, celui qui sera connecté à un voisinage lui-même central dans le réseau aura une centralité de vecteur propre plus forte que l'autre. Ainsi, dans nos réseaux, la centralité de vecteur propre ne va pas traduire la capacité d'alter à être présent dans un grand nombre de relations de collaboration d'ego (degrés) mais plutôt dans les relations d'ego les plus structurantes.

Si elle tend vers 0, cela signifie que les projets collaboratifs entre ego et ses alter centraux ne sont pas limités à certains co-auteurs. Dans le premier scénario, tous les alter du réseau collaborent à ce type de projet. Dans le second scénario, aucun alter n'accompagne ego dans ces publications. En revanche, si la centralisation tend vers 1, cela signifie que certains alter sont plus récurrents que d'autres dans les relations de collaboration d'ego avec ses alter centraux.

B) L'organisation des collaborations

Cette seconde série d'indicateurs a pour but d'évaluer comment ego mobilise ses ressources de collaboration. Nous qualifierons l'organisation des collaborations au niveau macro (à l'échelle du réseau) et au niveau micro (à l'échelle des alter). Ensuite, il conviendra de comprendre l'organisation des collaborations à l'échelle méso (à l'échelle des sous-espaces collaboratifs).

1- Comment ego organise-t-il ses collaborations à l'échelle globale et locale ?

L'organisation des collaborations est un sujet qui a longtemps fait débat au sein de la littérature du capital social, notamment concernant les avantages et les handicaps qui découlent d'un réseau complet

interconnecté (Coleman, 1988 ; Burt, 1992, 1995). Au regard de ces différentes perspectives, nous allons évaluer dans quelle mesure l'interconnexion des flux est forte, au niveau global (à l'échelle du réseau) et local (à l'échelle des *alter*) dans nos réseaux personnels.

La densité globale nous permet d'évaluer dans quelle mesure le réseau est considéré comme une seule ou plusieurs entités connexes. Nous avons choisi de l'évaluer au travers de trois indicateurs : la densité, le nombre de composants connexes du réseau et la part des *alter* concentrée dans le plus gros composant du réseau.

La densité locale nous permet d'évaluer la probabilité que chaque *alter* soit proche d'un voisinage dense et interconnecté : elle évalue la densité au sein de chaque composant. Ainsi, elle exprime si certains *alter* sont au cœur des collaborations d'*ego* ou si l'ensemble des *alter* appartient à une entité connectée. Nous avons choisi de l'évaluer grâce au coefficient de clustering local.

La densité : La conception du réseau en une ou plusieurs entités

$$\text{densité du réseau } i = \frac{\text{nombre d'arrêtes dans le réseau } i}{\text{nombre d'arrêtes possibles dans le réseau } i}$$

La densité permet de situer le niveau de connexion d'un réseau par rapport à son niveau maximal possible. Dans le cadre de nos réseaux, cela revient à nous demander si *ego* a profité de toutes les relations de collaboration possibles ou pas. La densité évolue entre deux valeurs extrêmes : elle vaut 0 dans le cas où *ego* collabore indépendamment avec chacun de ses *alter* (pas d'arrêtes entre les nœuds) et 1 si le réseau est le plus dense possible (tous les nœuds sont reliés par une arrête).

Si la densité est forte, cela signifie qu'*ego* assiste au maximum de relations collaboratives possibles dans son réseau. Le réseau n'est pas segmenté en sous-espaces collaboratifs dans lesquels les relations collaboratives seraient cloisonnées. Tous les co-auteurs font partie d'une même entité, mobilisés conjointement et donc tous complémentaires entre eux.

Si la densité est faible, *ego* n'assiste pas à toutes les relations collaboratives possibles dans son réseau. Il segmente et compartimente ses relations collaboratives au sein de composants ou de sous-espaces collaboratifs. Les co-auteurs du réseau ne sont pas tous mobilisés ensemble.

Les composants : A quel point *ego* segmente-t-il ses relations ?

La compartimentation des relations d'*ego* se traduit lorsqu'il opère une certaine sélection dans les collaborations auxquelles il assiste : dans cette optique, toute relation de collaboration entre les membres de son réseau n'est pas forcément vertueuse pour lui.

Un premier indicateur évaluera le *nombre de composants connexes* dans le réseau, indépendants les uns des autres. Cette franche séparation du réseau signifie qu'*ego* ne collabore pas simultanément avec les membres de différents composants. Un réseau est au minimum composé d'un composant et peut être fragmenté en plusieurs (le maximum correspondant au nombre de nœuds du réseau).

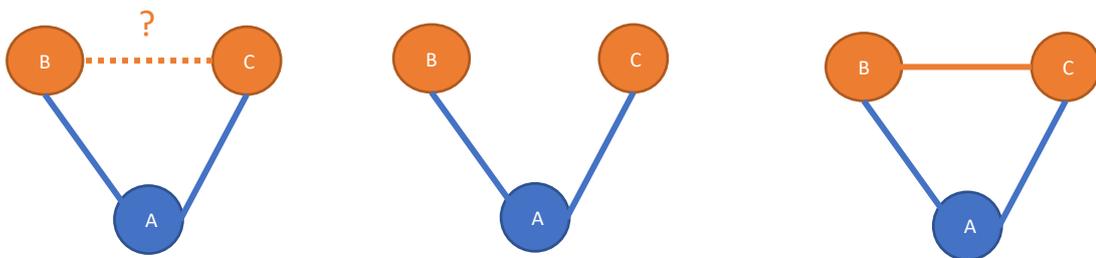
Le second indicateur évaluera la *part des nœuds du réseau incluse dans la composante principale*. L'objectif est d'évaluer l'impact de la segmentation. En effet, un réseau segmenté en de nombreux petits composants principaux peut tout de même avoir une forte partie de ses *alter* concentrée dans un même composant. La segmentation ne sera pas très structurante dans ce cas. En revanche, un réseau fragmenté en quelques composants dont chacun est doté d'une forte part d'*alter* peut être très structurante.

Le coefficient clustering global : La densité du voisinage d'alter

$$\text{coefficient clustering global du réseau } i = \frac{\text{nombre de triangle} \times 3}{\text{nombre de triades connexes}}$$

Cet indicateur évalue la compartimentation des relations du réseau, à l'échelle des relations entre les *alter* au sein des composants. En effet, nous évaluons la probabilité qu'un *alter* du réseau soit connecté à deux autres *alter* qui seront à leur tour connectés entre eux. Prenons l'exemple suivant pour comprendre l'enjeu de la transitivité à l'échelle locale :

Figure 51 : Représentation graphique de la transitivité locale



1) Dans ce réseau, *ego* collabore avec les *alter* A, B et C. Nous remarquons qu'il collabore conjointement avec les *alter* A et B, puis avec les *alter* A et C. L'*alter* A est donc connecté à deux membres du réseau. Nous évaluerons donc la probabilité qu'*ego* ait souhaité également collaborer conjointement avec les *alter* B et C.

2) Dans le cas où *ego* n'a pas souhaité collaborer conjointement avec les *alter* B et C, ces deux nœuds ne sont pas reliés entre eux. Nous pouvons donc dire que l'*alter* A est témoin d'une volonté de segmentation d'*ego* dans ses relations. En effet, il accompagne *ego* dans la collaboration avec deux nœuds qu'*ego* ne souhaite pas réunir dans un même projet. A est donc central dans les flux collaboratif d'*ego*.

3) Dans le cas où *ego* souhaite collaborer conjointement avec les *alter* B et C, ces deux nœuds sont reliés entre eux. Nous pouvons donc dire que l'*alter* A n'est pas témoin d'une quelconque volonté de segmentation d'*ego* dans ses relations. En effet, il accompagne *ego* dans la collaboration avec deux nœuds dont *ego* juge la collaboration conjointe vertueuse pour lui. A n'est donc pas plus central dans les liens collaboratifs d'*ego* que B et C.

Le coefficient clustering global synthétise donc la transitivité locale de l'ensemble des nœuds du réseau. Il est calculé à partir de la notion de triade qui une relation entre trois nœuds connectés⁵⁹.

Ainsi, si le coefficient tend vers 1, cela signifie que la probabilité que l'entourage de chaque nœud du réseau soit connecté est forte. En d'autres termes, les *alter* sont connectés avec d'autres nœuds qu'*ego* a choisi de réunir dans des projets collaboratifs communs. En somme, les *alter* récurrents ne sont pas considérés comme au cœur des liens collaboratifs d'*ego*, puisque tous les *alter* sont intégrés dans les collaborations. Ainsi, dans cette optique, deux scénarii sont possibles : soit *ego* considère son réseau comme une seule entité et assiste au maximum de relations collaboratives possibles, soit *alter* est proche d'une clique et il accompagne *ego* dans la collaboration avec des *alter* dont la relation de co-authorship se limitera au cadre de l'équipe.

En revanche, si le coefficient tend vers 0, cela signifie qu'il y a une faible probabilité que le voisinage d'*alter* soit dense. Il est très peu probable qu'*alter* accompagne *ego* dans la collaboration avec des *alter* avec qui *ego* collaborera conjointement. On observe donc une dualité entre les *alter* : ceux qui sont très peu intégrés dans le réseau (dans l'exemple précédent : B et C) et ceux qui sont centraux dans les liens de collaboration d'*ego* (dans l'exemple précédent : A).

2- Comment *ego* organise-t-il ses liens collaboratifs autour de ses sous-espaces collaboratifs ?

Ces derniers indicateurs traitent du partitionnement du réseau d'*ego*. Une partition de graphe est un sous-ensemble de nœuds qui sont plus connectés en interne qu'avec le reste du réseau. Le partitionnement du réseau de co-authorship révèle beaucoup sur l'organisation des liens collaboratifs d'*ego* puisqu'il indique les différents sous-espaces dans lesquels *ego* collabore.

⁵⁹ Par exemple, dans le 1er graphique, la triade est B-A-C. Le triangle est un cas particulier de triade fermée, composé de trois triades connexes : par exemple, dans le 3ème graphique, le triangle ABC est composé des triades A-B-C, B-A-C, A-C-B. L'indicateur évalue donc le poids des triades connexes issues de triangles sur l'ensemble des triades du réseau.

Encadré : Méthodologie de partitionnement

Le partitionnement d'un réseau est une affaire délicate puisqu'il n'existe pas de méthode officielle : il y a donc autant de partitions possibles que d'algorithmes spécialisés en la matière. Nous devons donc trouver la méthode la plus adéquate à notre base de données.

Tout d'abord, il existe les méthodes dites classiques : la méthode de bipartition des graphes, les méthodes hiérarchiques (regroupement des nœuds selon leur niveau d'équivalence, par exemple le blockmodeling), et les méthodes de clustering (pour un nombre de classes déterminé à l'avance).

La méthode divisive est très utilisée (Newman et al (2004)), notamment en physique. Elle consiste à supprimer progressivement les *alter* à forte centralité intermédiaire pour identifier les parties du graphe qui deviennent connexes (« communautés »). L'objectif est donc de partir du réseau initial et d'isoler le plus possible les nœuds au fur et à mesure. Cette méthode ne nous semble pas adaptée dans notre étude pour plusieurs raisons. Tout d'abord, elle est souvent utilisée pour des réseaux complets dans lesquels l'intégralité de la bibliométrie des *alter* est connue. Or, ce n'est pas le cas dans notre base : ne connaissant pas les relations entre les *alter* en dehors du cadre de l'activité d'*ego*, il nous semble difficile d'affirmer qu'ils puissent appartenir ou pas à telle communauté (avec la notion d'identité et les frontières que ce terme implique). Aussi, la notion d'intermédiarité ne traduit ici pas la mise en connexion de sous-groupes de nœuds via un *alter* intermédiaire mais plutôt par l'accompagnement d'un *alter* dans la collaboration d'*ego* dans plusieurs sous-groupes. Supprimer un *alter* intermédiaire ne reviendrait donc pas à dire que les sous-graphes ne collaboreraient plus ensemble (sous-graphes connexes) puisqu'ils le font déjà via la collaboration avec *ego*. Enfin, cette méthode ayant pour but de diviser le réseau en sous-groupes les plus petits possibles risque de n'identifier que les équipes de co-authorship de chaque publication. Cela nous détournerait de notre but : identifier les sous-espaces collaboration tissés par les différentes publications d'*ego*

Nous avons donc choisi la méthode de *Clauset et al* (2004), qui est une méthode agrégative et d'optimisation de la modularité. La démarche est opposée à celle de la méthode divisive puisque nous ne partons pas du réseau pour le diviser mais plutôt d'un ensemble de sommets pour les agréger entre eux étape par étape.

Le point de départ est de considérer que chaque sommet est une unique communauté. La théorie du capital social éclaire cette hypothèse de base dans le sens où chaque carrière est unique et dépend du capital social accumulé par le chercheur (Merton (1957), Bozeman et al (2001,2005)). Ainsi, nous faisons l'hypothèse que chaque *alter* dispose de compétences uniques (scientifiques, techniques, politiques) : *ego* va donc choisir comment les mobiliser dans ses différents liens de collaboration.

La première étape de l'algorithme sera de tester quels sont les couples d'*alter* (reliés par une arête) qui peuvent faire partie d'un même sous-graphe. Ce test sera conditionné par la variation de modularité occasionnée par ce regroupement. Si la fusion du couple en un même sous-espace accroît la qualité globale du partitionnement du graphe, l'algorithme les réunira. En revanche, si cette fusion détériore la qualité des partitions, l'algorithme ne les réunira pas dans un même espace. Le conditionnement de la partition par la modularité sécurise notre processus dans le sens où l'algorithme ne retient que les fusions cohérentes vis-à-vis de la stratégie d'*ego* de mobilisation des ressources collaboratives.

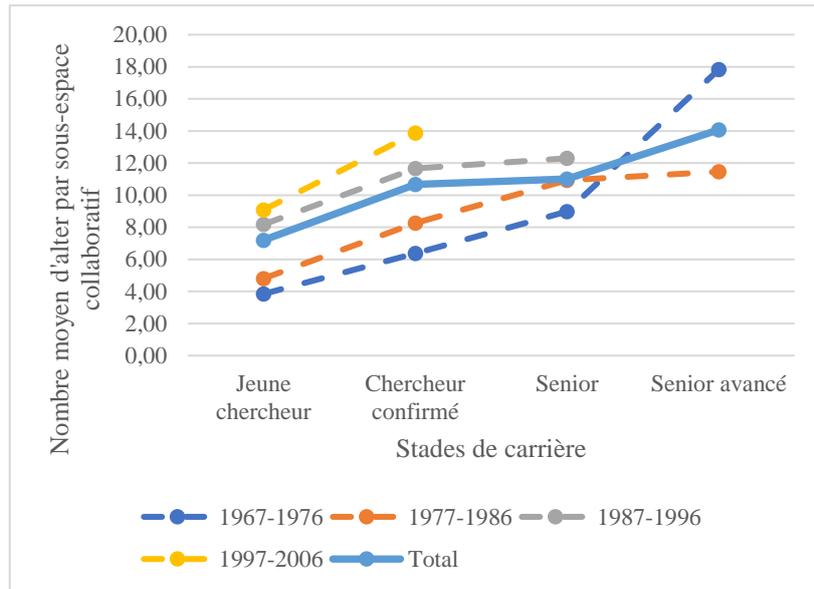
L'algorithme répétera cette même étape jusqu'à ce que tous les nœuds du graphe n'appartiennent plus qu'au même sous-ensemble. Finalement, l'algorithme ne retiendra que l'étape durant laquelle la modularité était maximum et donc durant laquelle le partitionnement des *alter* en sous-espace de collaboration est de qualité optimale.

Tableau 65 : Statistiques descriptives des sous-espaces collaboratifs

Stades de carrière	Jeune chercheur	Chercheur confirmé	Senior	Senior avancé	Total
Nombre de réseaux	80	80	65	34	259
Nombre total de sous communautés	329	481	445	308	1563
Nombre moyen de sous espace par réseaux	4,11	6,01	6,85	9,06	6,03
Taille moyenne des sous-espaces	7,20	10,66	11,00	14,07	10,70

Au total, nous avons identifié 1563 sous-espaces collaboratifs différents. En moyenne, chaque réseau est composé d'environ 6 sous-espaces collaboratifs composés chacun de 11 *alter*. Au fil de la carrière, les réseaux sont en plus en plus composés de sous-espaces collaboratifs dont la taille moyenne est croissante.

Figure 52 : Taille moyenne des sous-espaces collaboratifs



Nous remarquons également qu’au fil des générations de chercheurs, les sous-espaces collaboratifs sont de plus en plus gros dans les réseaux. Cela est en partie expliqué par la hausse de la nature collaborative des articles, dans lesquels *ego* mobilise conjointement de plus en plus d’*alter*.

Une fois le partitionnement fait, nous allons pouvoir analyser comment *ego* structure ses liens de collaboration vis-à-vis de ces sous-espaces de collaboration.

La modularité : Dans quelle mesure les sous-espaces collaboratifs structurent les liens collaboratifs d’*ego* ?

$$\text{La modularité d'un réseau } i = Q(P) = \sum_{j=1}^p (e_{c_{i,j}} - a^2_{c_{i,j}})$$

Avec :

$P = \text{la partition du réseau en } p \text{ clusters} = \{c_1, c_2, \dots, c_p\}$

$e_{c_{i,j}}$ = la part des arrêtes du cluster $c_{i,j}$ sur l'ensemble des arrêtes du réseau

$a^2_{c_{i,j}}$ = la probabilité que deux sommets reliés par une arrête se trouvent dans le cluster $c_{i,j}$

La modularité est un indicateur qui permet d’évaluer la qualité du partitionnement du réseau en différents clusters. Elle exprime la différence entre la part des arrêtes du réseau incluses dans chaque cluster avec la part des liens d’un graphe aléatoire (de même distribution de degrés) qui pourraient appartenir à ce cluster. Elle varie entre 0 et 1 : par convention, on dit qu’une modularité inférieure à 0,3

indique un partitionnement peu significatif. Un partitionnement de bonne qualité répond à deux critères : la partition doit être dense en interne (les liens collaboratifs doivent être denses au sein du sous-espaces collaboratif) et les partitions doivent être très peu connectées entre elles (*ego* mobilise très peu ses sous-espaces collaboratifs conjointement). Ainsi, une faible modularité sous-entend que les sous-espaces collaboratifs structurent peu les liens collaboratifs d'*ego*. En revanche, une forte modularité exprime un réseau communautaire, dans lequel *ego* ne mobilise ses *alter* que dans le cadre de ces sous-espaces collaboratifs.

L'hétérogénéité de taille des sous-espaces : *Certains sous-espaces collaboratifs sont-ils plus structurants que d'autres ?*

L'hétérogénéité de la taille des clusters du réseau i
 = *Coefficient de variation de la taille des clusters*

L'objectif de cet indicateur est de vérifier si les différents sous-espaces de collaboration ont une contribution homogène dans l'activité de collaboration d'*ego*. En effet, si le coefficient de variation est faible, cela signifie que le réseau d'*ego* est équitablement réparti dans les sous-espaces et que chacun d'entre eux offre le même nombre de ressources collaboratives à *ego*. *Ego* a donc la possibilité de diversifier son activité de collaboration puisqu'il a des structures collaboratives homogènes. En revanche, si le coefficient de variation est plus fort, cela signifie que les sous-espaces de collaboration du réseau d'*ego* sont de taille très différente. Par exemple, il peut avoir un sous-espace central très riche en *alter* et plusieurs sous-espaces qui seront composés de très peu d'*alter*. Cela signifie donc qu'*ego* dispose de collaborations structurantes (le sous-espace central serait alors une sorte de noyau) et diversifie son réseau en développant plusieurs autres sous-espaces ponctuels.

L'intensité moyenne de collaboration des sous-espaces : *En moyenne, combien de publications ego réalise-t-il avec chacun de ses sous-espaces collaboratifs ?*

L'intensité moyenne de collaboration des clusters du réseau i
 =
$$\frac{\text{Nombre de publications faites avec le réseau } i}{\text{Nombre de cluster dans le réseau } i}$$

L'objectif de cet indicateur sera d'évaluer la diversification de ses sous-espaces collaboratifs. En effet, une intensité faible (proche de 1) laisse supposer qu'*ego* aurait autant de sous-espaces collaboratifs que de publications faites : ainsi, chaque nouveau projet collaboratif serait l'occasion de développer un nouvel espace de collaboration. L'activité de publication d'*ego* se traduirait donc par une diversification dynamique de ses structures de collaboration. En revanche, une intensité plus forte sous-entendrait que les sous-espaces collaboratifs accompagneraient *ego* dans plusieurs publications et deviendraient donc

structurants dans son activité. Ainsi, l'activité de publication d'*ego* reposerait sur des sous-espaces récurrents qu'*ego* accroîtrait et nourrirait petit à petit.

Annexe 3 : Statistiques descriptives des 13 variables actives centrées et réduites

Tableau 66 : Statistiques descriptives des 13 variables actives centrées et réduites

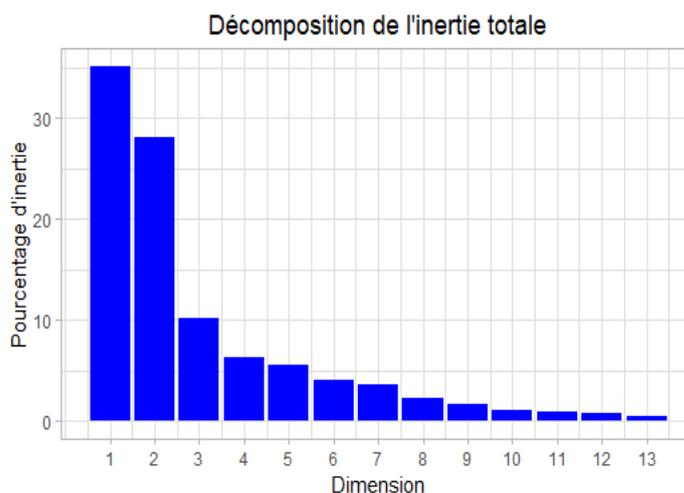
		Min	Premier quartile	Médiane	Moyenne	Troisième Quartile	Max
La position des alter dans le réseau	<i>coef_var_pub_ego</i>	-2,33	-0,65	-0,15	0,00	0,49	5,04
	<i>share_alterCliques</i>	-5,22	-0,42	0,09	0,00	0,61	2,74
	<i>share_unequivalent</i>	-1,33	-0,59	-0,19	0,00	0,33	6,08
	<i>centr_degree</i>	-1,87	-0,77	-0,17	0,00	0,70	2,66
	<i>centr_eigen</i>	-4,84	-0,40	0,19	0,00	0,63	1,72
	<i>centr_between</i>	-1,66	-0,73	-0,06	0,00	0,65	2,70
L'organisation des liens collaboratifs	<i>density</i>	-1,32	-0,66	-0,28	0,00	0,36	5,14
	<i>nb_comp</i>	-1,27	-0,42	0,00	0,00	0,43	8,47
	<i>share_comp_princ</i>	-2,63	-0,67	0,23	0,00	0,95	1,05
	<i>clustering_global</i>	-2,99	-0,70	-0,09	0,00	0,52	2,35
	<i>modularity</i>	-3,44	-0,55	0,06	0,00	0,75	2,24
	<i>coef_var_taille_comm</i>	-2,25	-0,58	0,02	0,00	0,58	3,72
	<i>intensity_pub_comm</i>	-1,36	-0,74	-0,27	0,00	0,50	7,20

Annexe 4 : Choix du nombre de valeurs propres

Tableau 67 : Part de l'inertie expliquée par chaque axe

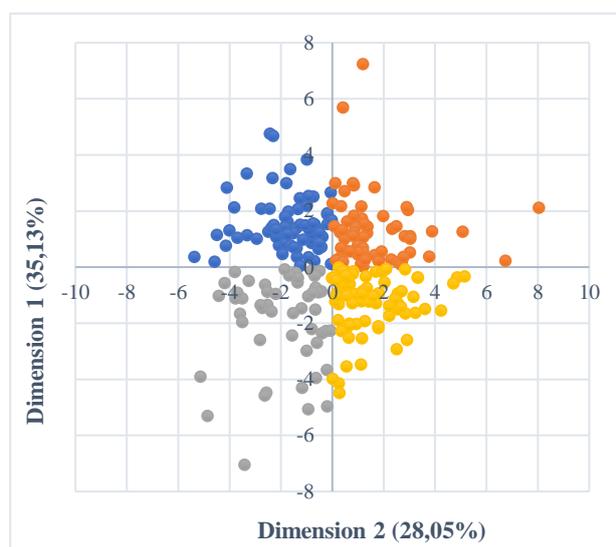
Dimension	Valeur propre	Pourcentage de variance	Pourcentage de variance cumulatif
1	4.55	35.13	35.13
2	3.63	28.05	63.18
3	1.31	10.12	73.30
4	0.81	6.28	79.58
5	0.71	5.51	85.09
6	0.52	4.03	89.12
7	0.48	3.68	92.80
8	0.30	2.29	95.09
9	0.21	1.64	96.73
10	0.15	1.14	97.87

Figure 53 : Décomposition de l'inertie totale



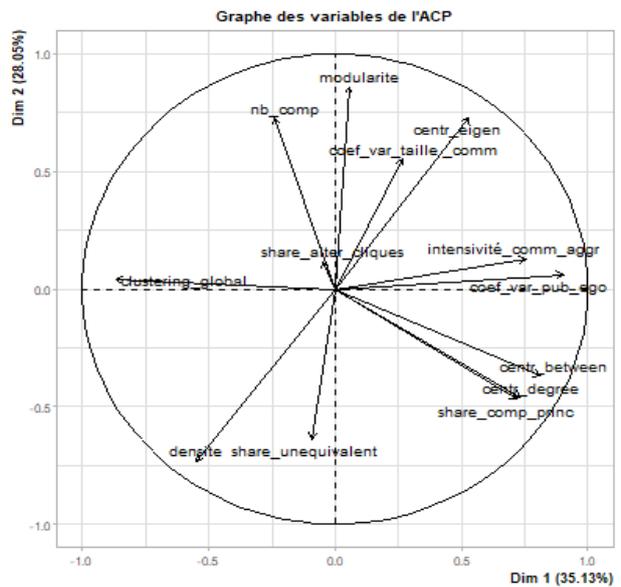
Nous estimons qu'il serait pertinent de restreindre notre analyse à l'interprétation des deux premiers axes, puisqu'ils expliquent 63,18% de la variance de nos données. Bien que le critère du coude et le critère de Kaiser nous invitent à conserver le troisième axe, nous choisissons de ne pas le retenir puisqu'il y a une réelle chute de la variance expliquée à partir de cet axe. Notre objectif étant de cerner les grandes tendances collaboratives en amont de la classification, nous pensons que les deux premiers axes expliquent une part variance suffisante.

Figure 54 : Individus de l'ACP



Le graphique ci-dessus détaille la répartition des individus dans le plan des deux dimensions collaboratives. On constate que quatre groupes d'individus se forment de chaque côté des axes. Nous notons également la présence d'individus atypiques entre la gauche de l'axe 1 et au-dessus de l'axe 2 puis entre la droite de l'axe 1 et le dessous de l'axe 2.

Figure 55 : Variables de l'ACP



Le graphique ci-dessous décrit la contribution de chacun des variables dans la caractérisation des deux axes. On remarque qu'une seule variable est très peu contributrice : la dotation de cliques du réseau (*share_alter_cliques*).

En revanche, certaines contribuent très fortement à caractériser les deux axes : l'hétérogénéité de récurrence des *alter* dans la bibliométrie d'*ego* (*coef_var_pub_ego*), le coefficient de clustering global (*clustering_global*), la densité (*density*) et les centralisations de vecteur propre (*centr_eigen*) et d'intermédiarité (*centr_between*). Aussi, la part des *alter* dans le composant principal (*share_comp_princ*) et la centralisation de degré (*centr_degre*) partagent les mêmes coordonnées dans le plan : il semble donc qu'elles caractérisent de la même manière ces deux axes.

Annexe 5 : Illustration graphique des deux axes de l'ACP

Figure 56 : Illustration graphique des 5 réseaux qui contribuent le plus à la dimension 1

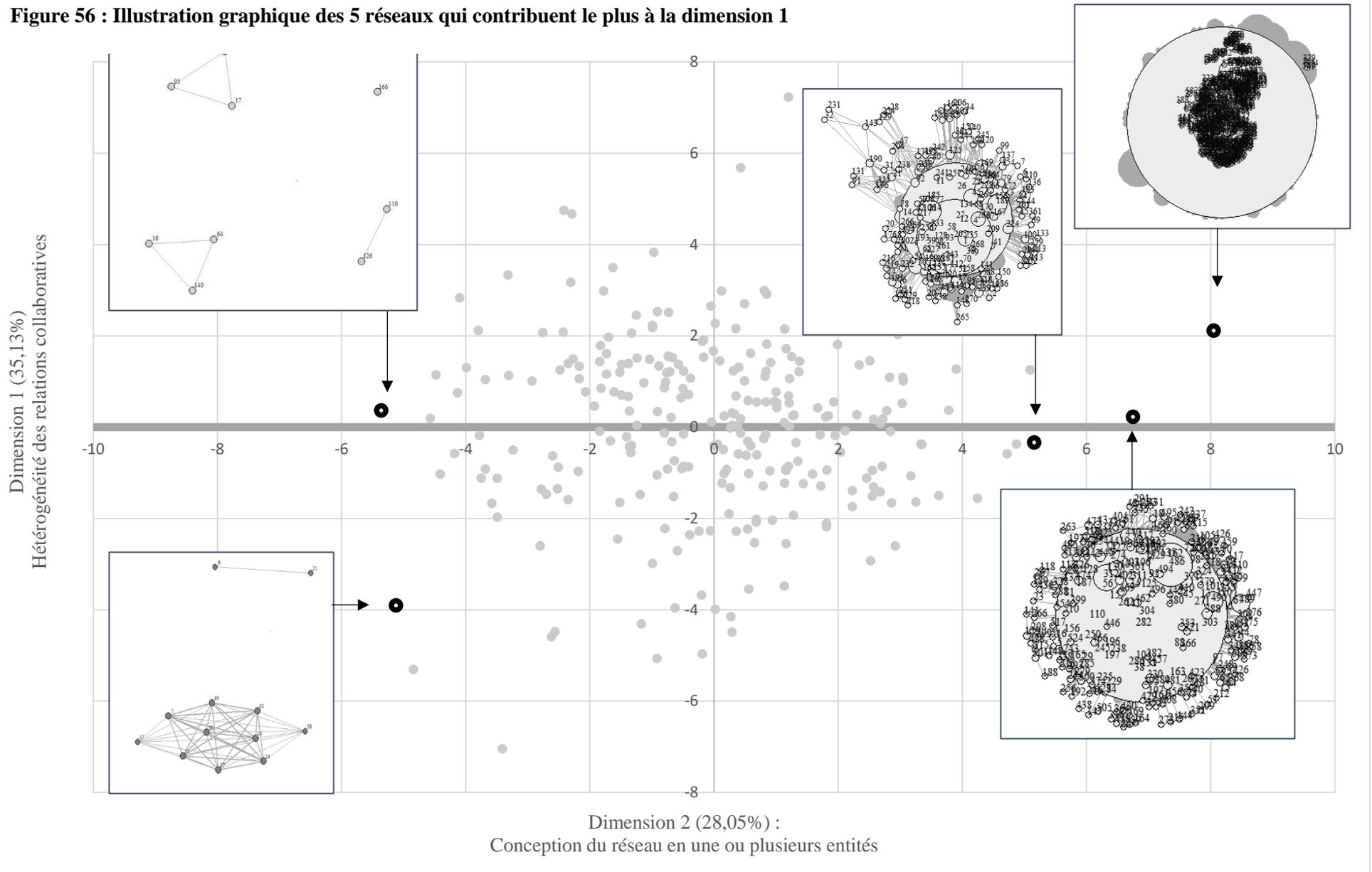
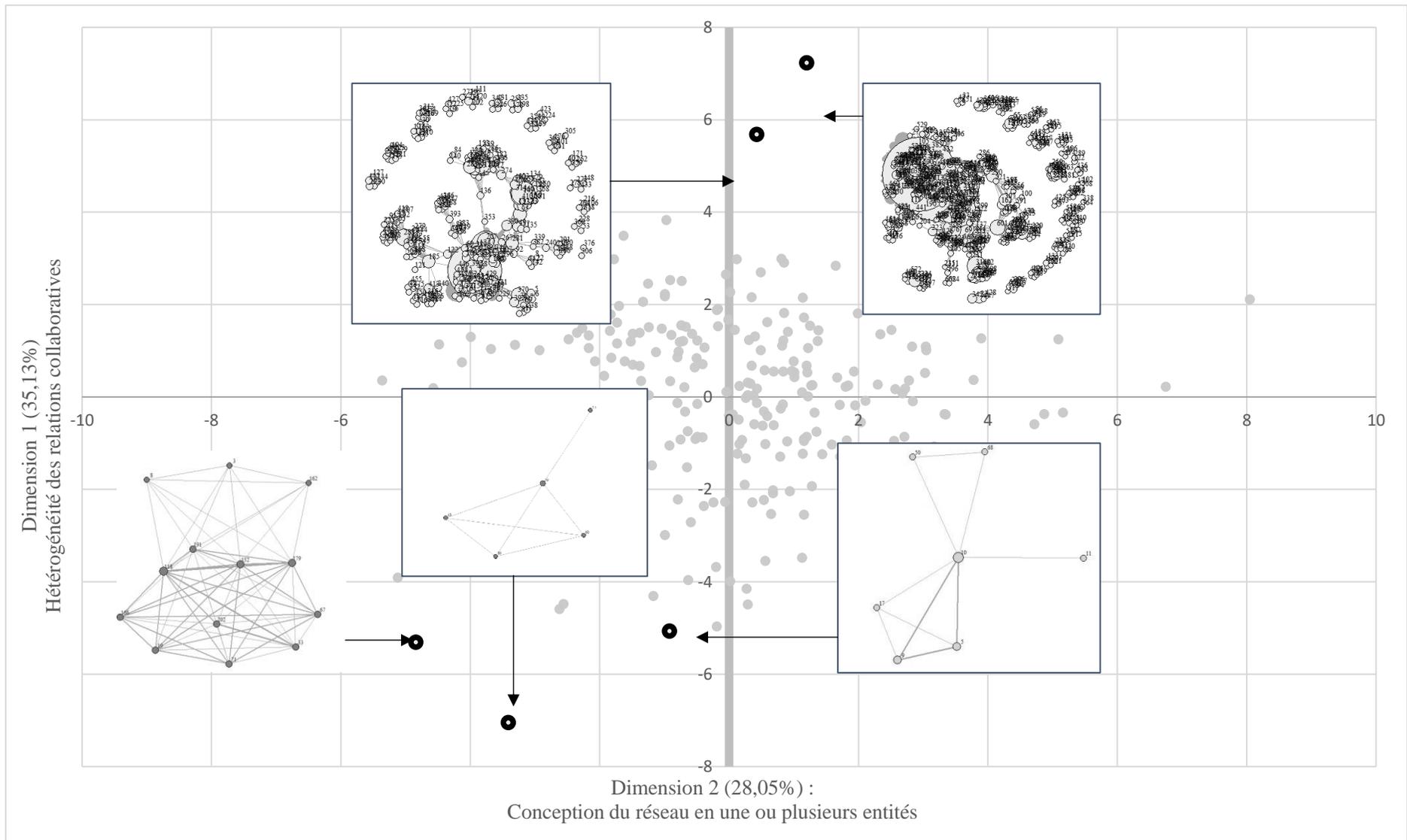
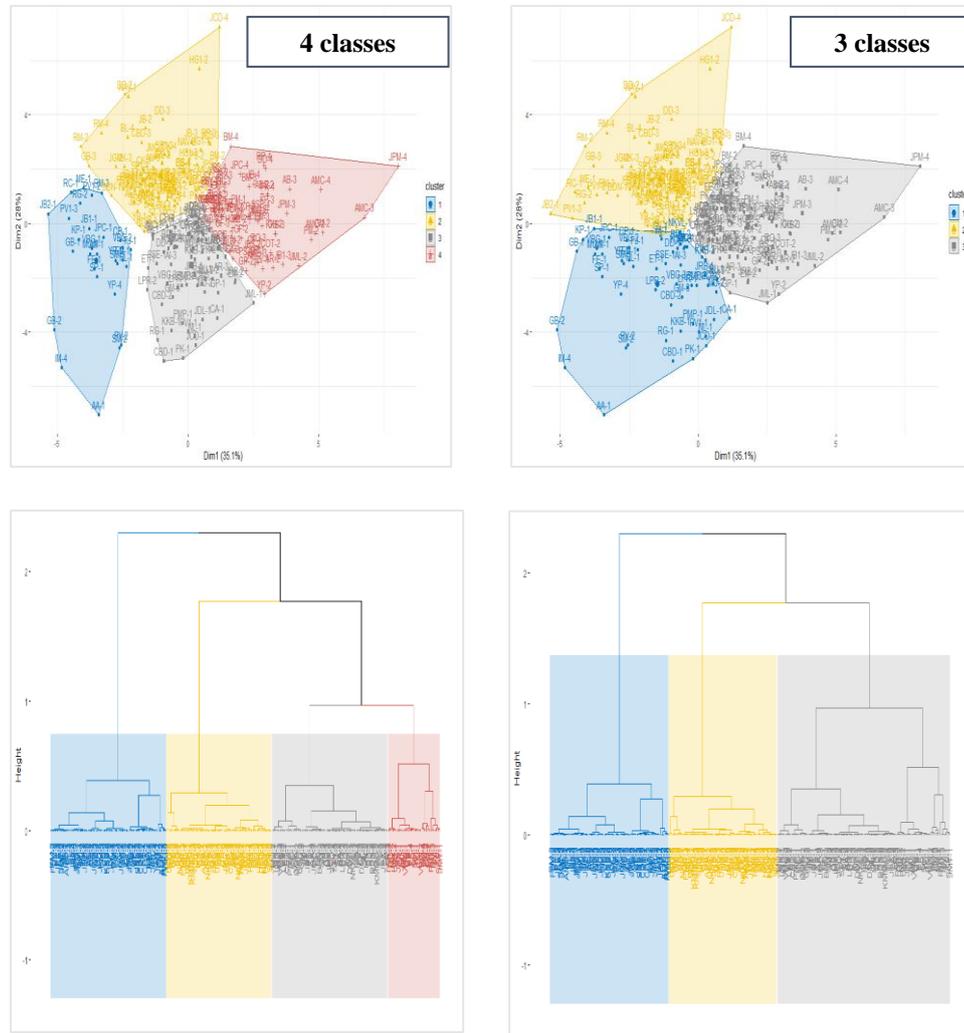


Figure 57 : Illustration graphique des 5 réseaux qui contribuent le plus à la dimension 2



Annexe 6 : Processus de classification

Figure 58 : Dendrogrammes et nuages de points : choix de la classification



Annexe 7 : Critères de classification

A) 1^{er} critère : Favoriser la proximité des individus au sein des classes

Le premier critère d'une partition réussie concerne l'homogénéité des individus au sein de chaque classe. La classification a pour but de regrouper les individus dans des classes où ils se ressemblent le plus.

La matrice de distance euclidienne est un outil que nous avons utilisé pour évaluer la proximité entre chaque réseau et pour distinguer ceux qui peuvent être regroupés ensemble. Pour un effectif de n individus, elle est de dimension $n \times n$ et représente la distance euclidienne entre chaque couple d'individus dans un

espace euclidien. La réalisation et l'actualisation de cette matrice est nécessaire à chaque étape du processus de classification.

En amont de la classification, nous disposons de 259 réseaux distincts. Nous nommons la matrice de distance euclidienne A de dimension 259×259 (pour chaque réseaux notés $x_1, x_2, \dots, x_{258}, x_{259}$) :

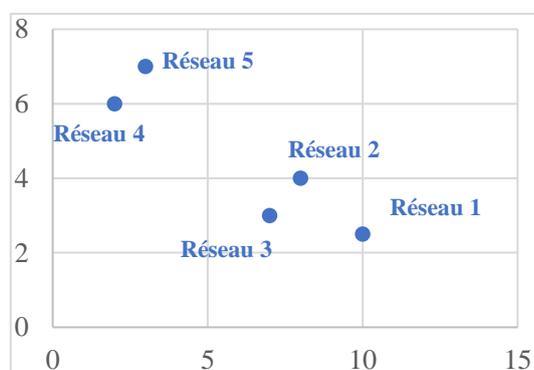
$$A = \begin{pmatrix} 0 & d_{1,2}^2 & d_{1,3}^2 & \dots & d_{1,259}^2 \\ d_{2,1}^2 & 0 & d_{2,3}^2 & \dots & d_{2,259}^2 \\ d_{3,1}^2 & d_{3,2}^2 & 0 & \dots & d_{3,259}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{259,1}^2 & d_{259,2}^2 & d_{259,3}^2 & \dots & 0 \end{pmatrix}$$

Avec $d_{i,j}^2 = \|x_i - x_j\|_2^2$ correspondant à la distance euclidienne de chaque couple de réseaux i et j (pour tout i et $j \in N / [1 ; 259]$).

Une fois la matrice A dressée pour nos 259 réseaux, nous la parcourons afin de trouver le couple de réseaux (x_i, x_j) disposant de la plus petite distance euclidienne ($d_{i,j}^2$). Nous les regroupons ensemble pour qu'il ne forme plus qu'un individu. Cela correspond à la première étape de classification. La seconde étape consiste à reproduire la matrice pour les 258 individus (257 réseaux + le couple que nous avons réuni dans la première étape), à rechercher le nouveau couple ayant la plus petite distance et à le fusionner. Nous reproduisons cette opération jusqu'à ce que tous les individus soient regroupés dans un même ensemble.

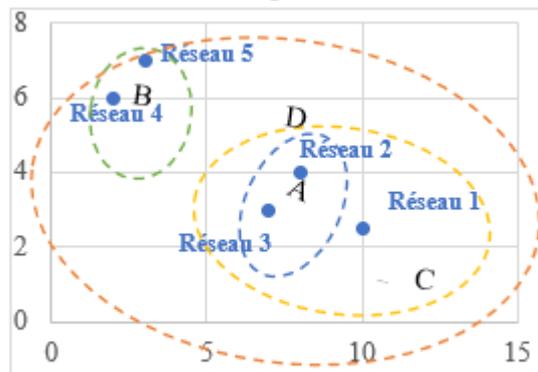
L'exemple suivant décrit graphiquement le processus de classification hiérarchique ascendante sur cinq réseaux distincts en suivant ce critère de proximité.

Figure 59 : Représentation graphique de cinq réseaux distincts avant la classification



Le plan suivant décrit cinq réseaux distincts que nous souhaitons répartir en classes homogènes. L'algorithme de classification hiérarchique va donc être composé de quatre étapes dans lesquelles il va regrouper progressivement ces réseaux en une même entité. A chaque étape, l'algorithme nous propose une partition possible.

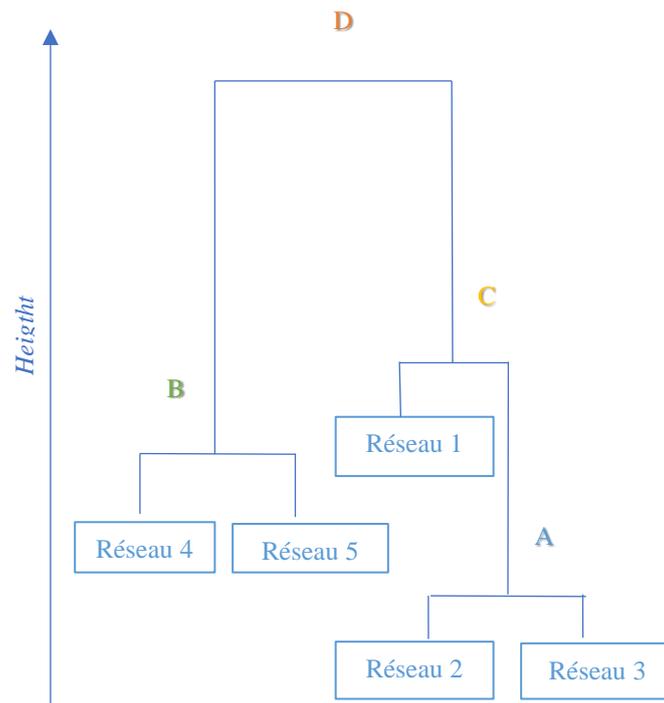
Figure 60 : Représentation graphique de cinq réseaux distincts après la classification



Dans cet exemple, la première étape du processus est l'étape A : elle consiste à regrouper les réseaux 2 et 3 dans une même classe, puisqu'il s'agit du couple dont la distance euclidienne est la plus faible. A l'issue de cette étape, nous disposons donc de 4 entités distinctes : les réseaux 1, 4, 5 et le regroupement des réseaux 2 et 3. La seconde étape du processus est l'étape B : elle consiste à regrouper les réseaux 4 et 5 dans une même classe, puisqu'il s'agit du couple dont la distance euclidienne est la plus faible. A l'issue de cette étape, nous disposons de 3 entités distinctes : le regroupement des réseaux 4 et 5, le réseau 1 et le regroupement des réseaux 2 et 3. La troisième étape du processus est l'étape C : elle consiste à intégrer le réseau 1 dans le regroupement des réseaux 2 et 3 puisqu'il s'agit du couple d'entités dont la distance euclidienne est la plus faible. Nous n'avons plus que deux entités : le regroupement des réseaux 4 et 5 et le regroupement des réseaux 1,2 et 3. Enfin, l'ultime étape D réunit ces deux entités dans une même classe, qui comprendra les cinq réseaux.

Le dendrogramme est un graphique qui décompose les différentes étapes de l'algorithme. Sur l'axe horizontal, nous voyons les individus regroupés ensemble. Par exemple, on voit que l'étape C a regroupé ensemble des réseaux 1, 2 et 3. Sur l'axe vertical, nous pouvons estimer la distance relative aux différents regroupements de couples. Ainsi, ce graphique est un outil qui permet de comprendre les étapes de classification et de trouver la partition optimale, à savoir celle qui se fera en coupant l'arbre au plus grand saut de réseaux.

Figure 61 : Exemple : dendrogramme de la classification hiérarchique des cinq réseaux



B) 2^{ème} critère : Favoriser une distance entre les réseaux les différentes classes différentes

Pour assurer la qualité du partitionnement, le critère de distance entre les réseaux doit être ajouté à celui de proximité intra-classe (garanti grâce à la matrice de distance euclidienne). Pour ce faire, nous avons utilisé la méthode d'agrégation de Ward pour contrôler la pertinence de chaque étape de regroupement de l'algorithme. En effet, cette méthode restreint les regroupements proposés par l'algorithme à ceux qui limitent la perte d'inertie interclasse.

Encadré : La méthode d'agrégation de Ward

Nous appelons le nuage de points N composé de nos 259 réseaux décrits par les 15 variables utilisées pour la classification.

Nous nommons la moyenne de chaque variable $j \in \{1, \dots, 15\}$ comme $\bar{x}_j = \frac{1}{259} \sum_{i=1}^{259} x_{j,i}$.

Nous appellerons g le centre de gravité de ce nuage de points aux coordonnées $(\bar{x}_1, \dots, \bar{x}_{15})$. Ce point représente la position globale des 259 réseaux relativement aux 15 variables. Nous utiliserons cette donnée comme référence dans le calcul de trois indicateurs de dispersion.

Supposons que notre nuage de points N soit partitionné en h sous-graphes, chacun nommé N_l (avec $l \in \{1 \dots h\}$). Dans chacune de ces partition N_l , nous avons un effectif de n_l réseaux et un centre de gravité g_l , dont les coordonnées sont $(\bar{x}_{1,l}, \dots, \bar{x}_{15,l})$. Rappelons que la somme des effectifs de chaque partition correspond à l'effectif total ($\sum_{l=1}^h n_l = 259$).

Trois indicateurs sont alors mobilisés pour caractériser la dispersion des 259 réseaux.

i) Le premier est l'inertie totale. Il évalue la dispersion totale des 259 points du nuage N relativement au centre de gravité g . Pour chaque variable j , il mesure la distance au carré de chaque réseau i , compris dans la partition l relativement à la coordonnée \bar{x}_j de g . En d'autres termes, il évalue dans quelle mesure les 259 réseaux sont dispersés entre eux.

$$I_{totale} = \frac{1}{n_l} \sum_{i=1}^{259} d^2(x_i, g_l) = \sum_{j=1}^{15} \sum_{l=1}^h \sum_{i=1}^{n_l} \frac{1}{n_l} (x_{i,l,j} - \bar{x}_j)^2$$

ii) Le second indicateur est l'inertie intra-classe. Il évalue la dispersion des réseaux au sein de chaque partition l autour de chaque centre de gravité g_l . Il mesure la distance pondérée au carré de chaque réseau i relativement à la coordonnée $\bar{x}_{j,l}$ de g_l dans la partition l . En d'autres termes, il évalue dans quelle mesure les réseaux sont dispersés au sein de chaque classe.

$$I_{intra}(N_l) = \sum_{l=1}^h \frac{n_l}{n} I_{totale} = \sum_{j=1}^{15} \sum_{l=1}^h \sum_{i=1}^{n_l} \frac{1}{n} (x_{i,l,j} - \bar{x}_{j,l})^2$$

iii) Le troisième indicateur est l'inertie inter-classe. Il évalue l'écart entre les partitions N_l (dispersion des centres de gravité g_l relativement au centre de gravité global g). Il mesure la distance au carré de chaque centre de gravité g_l relativement au centre de gravité global g . En d'autres termes, il évalue dans quelle mesure les classes sont éloignées entre elles.

$$I_{inter}(N_l) = \sum_{l=1}^h \frac{n_l}{n} d^2(g_l, g) = \sum_{j=1}^{15} \sum_{l=1}^h \sum_{i=1}^{n_l} \frac{1}{n} (\bar{x}_{j,l} - \bar{x}_j)^2$$

Grâce à la décomposition de Huygens, nous pouvons faire le lien entre ces trois indicateurs de la dispersion au sein du réseau :

$$I_{totale} = I_{intra}(N_l) + I_{inter}(N_l) \Leftrightarrow 1 = \frac{I_{intra}(N_l)}{I_{totale}} + \frac{I_{inter}(N_l)}{I_{totale}}$$

Cette décomposition nous permet d'évaluer la qualité du partitionnement :

Si $\frac{\text{Inertie inter-classe}}{\text{Inertie totale}}$ tend vers 0, cela signifie que l'inertie **inter-classe pèse très peu** dans l'inertie totale et donc que les partitions ont les mêmes centres de gravité. Le **partitionnement est donc de mauvaise qualité** car les partitions sont trop homogènes entre elles.

Si $\frac{\text{Inertie inter-classe}}{\text{Inertie totale}}$ tend vers 1, cela signifie que l'inertie **intra-classe pèse très peu** dans l'inertie totale et donc que les réseaux se concentrent autour du même centre de gravité au sein de chaque partition. Le **partitionnement est donc de très bonne qualité** car les réseaux sont très homogènes au sein de chaque partition.

Grâce à cette décomposition, la formule de l'écart de Ward estime, à chaque étape de l'algorithme, si le regroupement doit être fait ou non. Supposons, l'existence de deux groupes de réseaux A et B dans notre nuage de points N, disposants respectivement des effectifs n_a et n_b .

$$Inertie(A \cup B) = Inertie(A) + Inertie(B) + Inertie(A \cap B)$$

$$Inertie(A) + Inertie(B) = Inertie(A \cup B) - I_{inter}(A, B)$$

$$Inertie(A) + Inertie(B) = Inertie(A \cup B) - \frac{n_a n_b}{n_a + n_b} d^2(A, B)$$

Puisque : $g_{A \cup B} = \frac{n_A g_A + n_B g_B}{n_A + n_B}$

Ainsi, l'agrégation des sous-groupes se fait donc sous condition de minimisation de $\frac{n_a n_b}{n_a + n_b} d^2(A, B)$.

Le but est d'éviter une trop forte chute de l'inertie interclasse et donc de l'inertie totale. La méthode de Ward permet donc une optimisation progressive du processus de classification en appliquant cette condition à chaque étape de l'algorithme.

Annexe 8 : Variables les plus contributrices de la CAH

Tableau 68 : Contribution des variables dans la classification hiérarchique ascendante

Les cinq variables actives qui contribuent le plus à la classification en quatre classes			
	Rang	Eta2	P-value
<i>clustering_global</i>	1	0.613	2.839e-52
<i>density</i>	2	0.578	1.857e-47
<i>coef_var_pub_ego</i>	3	0.548	1.039e-43
<i>centr_eigen</i>	4	0.538	1.610e-42
<i>centr_between</i>	5	0.520	1.971e-40
Contribution des axes de l'ACP dans la classification en quatre classes			
	Rang	Eta2	P-value
<i>Dim.1 : « Hétérogénéité des relations collaboratives »</i>	1	0,734	4,67E-73
<i>Dim.2 : « Conception du réseau en une ou plusieurs entités »</i>	2	0,505	9,32E-39

Avant de décrire chacune des classes, nous pouvons voir les variables qui permettent de distinguer les quatre classes de la classification.

Au regard du premier axe, on observe que le poids des co-auteurs centraux dans l'ensemble des liens collaboratifs d'*ego* est déterminant dans la classification. En effet, trois indicateurs de centralité apparaissent structurant : le coefficient de clustering global, soit l'indicateur qui évalue la densité du réseau à l'échelle des *alter*, la variable *coef_var_pub_ego* qui évalue si certains *alter* sont plus récurrents

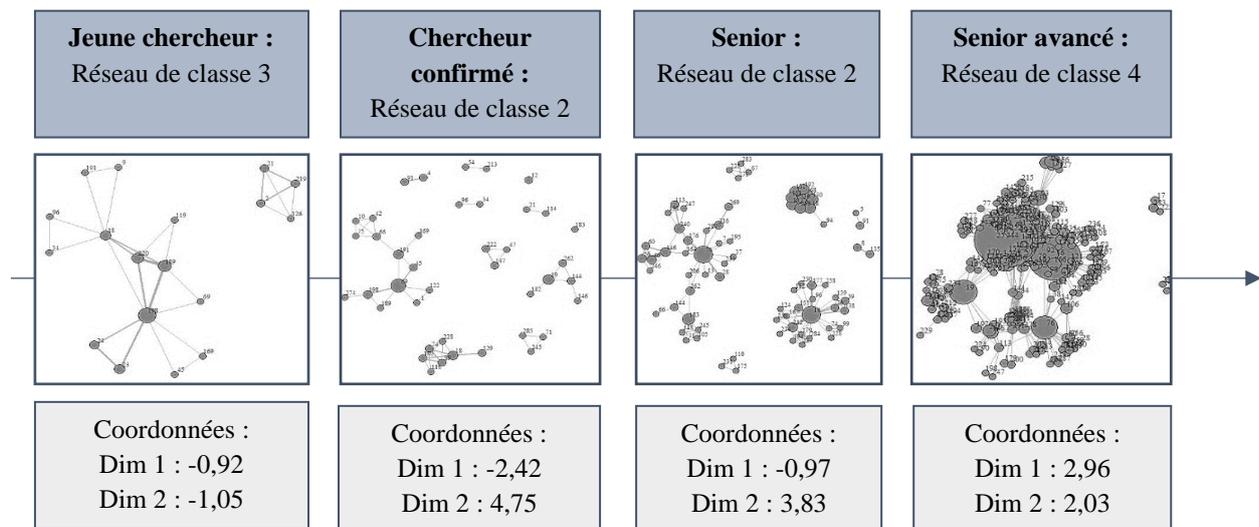
que d'autres dans la bibliométrie d'*ego* et la variable de centralisation d'intermédierité (*centr_between*) qui évalue dans quelle mesure certains alter centraux accompagne ego dans ses collaborations avec plusieurs sous-espaces collaboratifs.

La densité (*density*) est la seconde variable qui contribue le plus à la classification. Aussi, la centralité de vecteur propre (*centr_eigen*) permet d'évaluer si *ego* mobilise ses *alter* centraux conjointement avec le reste de son réseau ou non.

Annexe 9 : Exemple de calculs des indicateurs de trajectoire structurelle du réseau

L'exemple suivant illustre comment ces cinq variables sont calculées et leur intérêt. Nous choisissons l'un de nos 80 chercheurs que nous nommerons « Paul ». La trajectoire structurelle de son réseau est décrite par le schéma suivant :

Figure 62 : Exemple de la trajectoire structurelle du réseau de co-authorship de l'individu "Paul"



Les variables relatives à la dynamique structurelle prendront donc les valeurs suivantes :

Tableau 69 : Dynamique structurelle du réseau de Paul au fil de sa carrière

	Réseau de chercheur confirmé de Paul	Réseau de senior de Paul	Réseau de senior avancé de Paul
<i>Variables dépendantes</i>			
<i>Classe_dynamique₁</i>	0	0	0
<i>Classe_dynamique₂</i>	1	1	0
<i>Classe_dynamique₃</i>	0	0	0
<i>Classe_dynamique₄</i>	0	0	1
<i>Variables explicatives</i>			
<i>classe_precedente</i>	Classe 3	Classe 2	Classe 2
<i>evol_dim_1</i>	$(-2.42) - (-0.92) = -1.5$	$(-0.97) - (-2.42) = 1.45$	$2.96 - (-0.97) = 3.93$
<i>evol_dim_2</i>	$4.75 - (-1.05) = 5.8$	$3.83 - 4.75 = (-0.92)$	$2.03 - 3.83 = (-1.8)$
<i>val_abs_dim_1</i>	$ -1.5 = 1.5$	$ 1.45 = 1.45$	$ 3.93 = 3.93$
<i>val_abs_dim_2</i>	$ 5.8 = 5.8$	$ -0.92 = 0.92$	$ -1.8 = 1.8$

On constate que la structure du réseau de Paul évolue au fil de sa carrière : il est de classe 3 lorsqu'*ego-1* est « jeune chercheur », puis de classe 2 lorsqu'il est « chercheur confirmé » et « senior » puis de classe 4 lorsqu'il publie en tant que « senior avancé ».

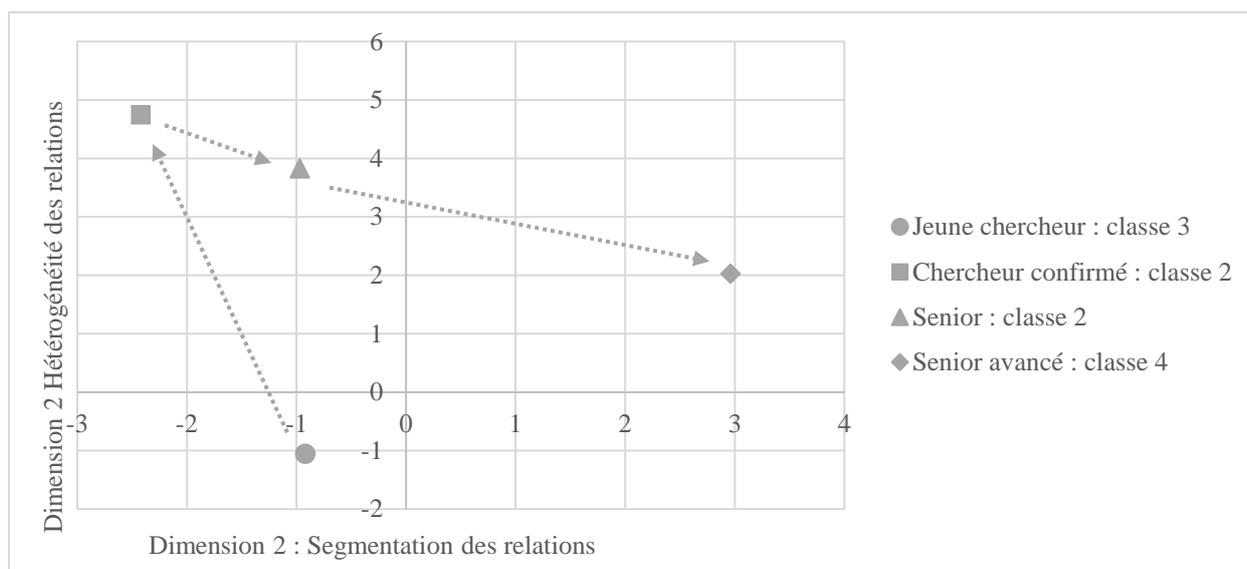
Lorsque Paul devient « chercheur confirmé », son réseau connaît un fort bouleversement structurel. Tout d'abord, il change radicalement de structure : il passe d'une structure dense dotées de relations *ego-alter* singulières à une structure éclatée en composants denses et homogènes en interne. Le réseau s'est déplacé sur la gauche de l'axe 1 sur 1,5 unités : les relations *ego-alter* se sont homogénéisées dans une faible mesure. En revanche, le réseau s'est déplacé vers le haut de l'axe 2 sur 5,8 unités : ses liens collaboratifs se sont très fortement segmentés par rapport à la structure d'origine.

Lorsque Paul devient « senior », son réseau connaît très peu de bouleversement structurel. En effet, il conserve une structure éclatée en composants denses et ne se déplace que très peu dans le plan.

Enfin, lorsque Paul devient « senior-avancé », son réseau connaît à nouveau un bouleversement structurel. Il passe d'une structure éclatée en composants denses en interne à une structure gravitaire structurée autour de quelques co-auteurs centraux. Ce bouleversement structurel se traduit par un fort déplacement vers la droite de l'axe 1 de 3,93 unités : l'hétérogénéité des relations *ego-alter* est beaucoup plus structurante que durant la période précédente. En revanche, le réseau ne se déplace que dans un faible ampleur sur l'axe 2 : les collaborations du réseau étaient déjà segmentées durant la période précédente.

Le graphique suivant illustre la trajectoire structurelle du réseau de co-authorship de Paul suivie au fil de sa carrière.

Figure 63 : Dynamiques structurelles des réseaux de Paul



Bibliographie

Abbasi, A., & Altmann, J. (2011, January). On the correlation between research performance and social network analysis measures applied to research collaboration networks. In *2011 44th Hawaii International Conference on System Sciences* (pp. 1-10). IEEE.

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594-607.

Abbasi, A., Chung, K. S. K., & Hossain, L. (2012). Egocentric analysis of co-authorship network structure, position and performance. *Information Processing & Management*, 48(4), 671-679.

Abbasi, A., Wigand, R. T., & Hossain, L. (2014). Measuring social capital through network analysis and its influence on individual performance. *Library & Information Science Research*, 36(1), 66-73.

Abbott, A., & Abbott, A. D. (2001). *Time matters: On theory and method*. University of Chicago Press.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). Research productivity: Are higher academic ranks more productive than lower ones ?. *Scientometrics*, 88(3), 915-928.

Ackers, L. (2005). Moving people and knowledge: Scientific mobility in the European Union1. *International migration*, 43(5), 99-131.

Ackers, L. (2008). Internationalisation, mobility and metrics: A new form of indirect discrimination? *Minerva*, 46(4), 411-435

Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but not forgotten : knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6(5), 571-591.

Akrich, M. C. M., & Latour, B. (2006). *Sociologie de la traduction*. Presses des MINES.

Aksnes, D. W., Rørstad, K., Piro, F. N., & Sivertsen, G. (2013). Are mobile researchers more productive and cited than non-mobile researchers? A large-scale study of Norwegian scientists. *Research Evaluation*, 22(4), 215-223.

Albarrán, P., Carrasco, R., & Ruiz-Castillo, J. (2017). Geographic mobility and research productivity in a selection of top world economics departments. *Scientometrics*, 111(1), 241-265.

Almeida, P., & Kogut, B. (1997). The exploration of technological diversity and geographic localization in innovation: Start-up firms in the semiconductor industry. *Small Business Economics*, 9(1), 21-31.

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science*, 45(7), 905-917.

Aman, V. (2018). A new bibliometric approach to measure knowledge transfer of internationally mobile scientists. *Scientometrics*, 117(1), 227-247.

Angel, D. P. (1991). High-technology agglomeration and the labor market: the case of Silicon Valley. *Environment and Planning A*, 23(10), 1501-1516.

- Arnaboldi, V., Conti, M., La Gala, M., Passarella, A., & Pezzoni, F. (2016). Ego network structure in online social networks and its impact on information diffusion. *Computer Communications*, 76, 26-41.
- Arnaboldi, V., Dunbar, R. I., Passarella, A., & Conti, M. (2016, January). Analysis of co-authorship ego networks. In *International Conference and School on Network Science* (pp. 82-96). Springer, Cham.
- Arrieta, O. A. D., Pammolli, F., & Petersen, A. M. (2017). Quantifying the negative impact of brain drain on the integration of European science. *Science advances*, 3(4), e1602232.
- Arrow K.J., (1962). The Economic Implications of Learning by Doing , *Review of Economic Studies*, vol. 29, 1962
- Arthur, M. B., & Rousseau, D. M. (1996). A career lexicon for the 21st century. *Academy of Management Perspectives*, 10(4), 28-39.
- Aschenbruck, N., Gerhards-Padilla, E., & Martini, P. (2008). A survey on mobility models for performance analysis in tactical mobile networks. *Journal of Telecommunications and Information Technology*, 54-61.
- Attali, J. (1998), Pour un modèle européen d'enseignement supérieur : rapport à M. le ministre de l'éducation nationale, de la recherche et de la technologie.
- B. Latour, Ces réseaux que la raison ignore, L'Harmattan, Paris, 1992
- Badar, K., Frantz, T. L., & Jabeen, M. (2016). Research performance and degree centrality in co-authorship networks. *Aslib Journal of Information Management*.
- Badar, K., Hite, J. M., & Ashraf, N. (2015). Knowledge network centrality, formal rank and research performance: evidence for curvilinear and interaction effects. *Scientometrics*, 105(3), 1553-1576.
- Badar, K., Hite, J. M., & Badir, Y. F. (2013). Examining the relationship of co-authorship network centrality and gender on academic research performance: the case of chemistry researchers in Pakistan. *Scientometrics*, 94(2), 755-775.
- Badar, K., Hite, J. M., & Badir, Y. F. (2014). The moderating roles of academic age and institutional sector on the relationship between co-authorship network centrality and academic research performance. *Aslib Journal of Information Management*.
- Badar, M., Javid, A. Y., & Zulfiquar, S. (2013). Impact of macroeconomic forces on nonperforming loans: An empirical study of commercial banks in Pakistan. *wseas Transactions on Business and Economics*, 10(1), 40-48.
- Badaracco, J., & Badaracco, J. L. (1991). *The knowledge link: How firms compete through strategic alliances*. Harvard Business Press.
- Bagla-Gökalp, L. (2000). Quelques approches sociologiques de réseaux sociaux. *ASp. La revue du GERAS*, (27-30), 201-229.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.
- Baret, C., Huault, I., & Picq, T. (2006). Management et réseaux sociaux Jeux d'ombres et de lumières sur les organisations. *Revue française de gestion*, (4), 93-106.

- Barnes, J. A. (1954). Class and committees in a Norwegian island parish. *Human relations*, 7(1), 39-58.
- Baruch, Y. (2013). Careers in academe: The academic labour market as an eco-system. *Career Development International*.
- Baruch, Y., & Hall, D. T. (2004). The academic career: a model for future careers in other sectors ?. *Journal of Vocational Behavior*, 64(2), 241-262.
- Baruffaldi, S. H., & Landoni, P. (2012). Return mobility and scientific productivity of researchers working abroad: The role of home country linkages. *Research policy*, 41(9), 1655-1665.
- Baruffaldi, S. H., & Landoni, P. (2016). Mobility intentions of foreign researchers: the role of non-economic motivations. *Industry and innovation*, 23(1), 87-111.
- Bauder, H. (2020). International Mobility and Social Capital in the Academic Field. *Minerva*, 1-21.
- Baudet-Michel, S., Berroir, S., Grasland, C., Guérois, M., Madelin, M., Maisonobe, M., ... & Zanin, C. (2020). Vers une désertification scientifique et universitaire du territoire français ?.
- Becker, G. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. Chicago, *University of Chicago Press*.
- Becquet, V., & Musselin, C. (2004, January). Variations autour du travail des universitaires. In *convention MENRT* (Vol. 20022007).
- Beine, M., Docquier, F., & Rapoport, H. (2008). Brain drain and human capital formation in developing countries: winners and losers. *The Economic Journal*, 118(528), 631-652.
- Bernela, B. (2016). Géographie des carrières universitaires et construction des collaborations scientifiques : une étude de cas en mathématiques. *Géographie, économie, société*, 18(2), 235-256.
- Bernela, B., & Levy, R. (2017). Collaboration networks within a French cluster: Do partners really interact with each other ?. *Papers in Regional Science*, 96(1), 115-138.
- Bernela, B., & Milard, B. (2016). Co-authorship Network Dynamics and Geographical Trajectories-What Part Does Mobility Play? *Bulletin of Sociological Methodology*, 131(1), 5-24.
- Bernela, B., Bouba-Olga, O., & Ferru, M. (2018). Spatial patterns of PhDs' internal migration in France, 1970-2000. *Journal of Innovation Economics Management*, 1, 33-56.
- Berthelot, J.M., Ponthieux, S. (1992). Les enseignants-chercheurs de l'enseignement supérieur : revenus professionnels et conditions d'activités. *Documents du CERC (Centre d'études des revenus et des coûts)*. Paris, *La Documentation Française*.
- Bhagwati, J. N. (1976). Taxing the brain drain. *Challenge*, 19(3), 34-38.
- Bhagwati, J. N. (1979). International migration of the highly skilled: economics, ethics and taxes. *Third world quarterly*, 1(3), 17-30.
- Bhagwati, J., & Dellalfar, W. (1973). The brain drain and income taxation. *World Development*, 1(1-2), 94-101.
- Bidart, C., & Mendez, A. (2016). Un système d'analyse qualitative des processus dans les sciences sociales : l'exemple de la mutation d'une organisation.

- Bidart, C., Longo, M. E., & Mendez, A. (2013). Time and process: An operational framework for processual analysis. *European Sociological Review*, 29(4), 743-751.
- Biddle, J., & Roberts, K. (1994). Private sector scientists and engineers and the transition to management. *Journal of Human Resources*, 82-107.
- Biggart, N. W., & Hamilton, G. G. (1992). On the limits of a firm-based theory to explain business networks. *Networks and Organizations*. Harvard Business School Press, Boston, 471-490.
- Biswal, A. K. (2013). An absolute index (Ab-index) to measure a researcher's useful contributions and productivity. *PloS one*, 8(12), e84334.
- Bollig, M. (1998). Moral economy and self-interest: Kinship, friendship, and exchange among the Pokot (NW Kenya). *Cambridge University Press*.
- Bordons, M., Aparicio, J., González-Albo, B., & Díaz-Faes, A. A. (2015). The relationship between the research performance of scientists and their position in co-authorship networks in three fields. *Journal of Informetrics*, 9(1), 135-144.
- Bordons, M., Zulueta, M., Romero, F., & Barrigón, S. (1999). Measuring interdisciplinary collaboration within a university: The effects of the multidisciplinary research programme. *Scientometrics*, 46(3), 383-398.
- Bouba-Olga, O., & Ferru, M. (2011). La dimension spatiale des collaborations pour l'innovation : une analyse sur données CIFRE (1981-2006). *Revue d'Economie Regionale Urbaine*, (3), 449-468.
- Bouba-Olga, O., Ferru, M., & Pépin, D. (2012). Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science*, 91(2), 355-375.
- Boudreaux, E. D., Higgins Jr, S. E., Reznik-Zellen, R., Wang, B., & Volturo, G. (2019). Scholarly productivity and impact: developing a quantifiable, norm-based benchmarking methodology for academic emergency medicine. *Academic Emergency Medicine*, 26(6), 594-604.
- Boulding, K.E. "The 'National' importance of human capital", in W. Adams (Ed.), *The Brain Drain*, Macmillan, New York, Chapter 7: 109-119
- Bourdieu, P. (1980). Le capital social : notes provisoires. *Actes de la recherche en sciences sociales*, 31(1), 2-3.
- Bourdieu, P. (1985). The levels of social capital. *Handbook of theory of research for the sociology of education*, 241-258.
- Bourdieu, P. (1988). *Homo academicus*. Stanford University Press.
- Bourdieu, P., & Wacquant, L. (1992). *Réponses* (Vol. 4). Paris : Seuil.
- Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research policy*, 33(4), 599-616.
- Bozeman, B., & Gaughan, M. (2011). How do men and women differ in research collaborations? An analysis of the collaborative motives and strategies of academic researchers. *Research policy*, 40(10), 1393-1402.
- Bozeman, B., Dietz, J. S., & Gaughan, M. (2001). Scientific and technical human capital: an alternative model for research evaluation. *International Journal of Technology Management*, 22(7-8), 716-740.

- Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of economic geography*, 9(4), 439-468.
- Breschi, S., Lissoni, F., & Miguelez, E. (2017). Foreign-origin inventors in the USA: testing for diaspora and brain gain effects. *Journal of Economic Geography*, 17(5), 1009–1038.
- Briscoe, J. P., & Hall, D. T. (2006). The interplay of boundaryless and protean careers: Combinations and implications. *Journal of vocational behavior*, 69(1), 4-18.
- Brocard, M. (1991). La science et les régions. *Paris, La Documentation française*.
- Brochier, D., Garnier, J., Gilson, A., Longo, M. E., Mendez, A., Mercier, D., ... & Bidart, C. (2010). Propositions pour un cadre théorique unifié et une méthodologie d'analyse des trajectoires des projets dans les organisations. *Management Avenir*, (6), 84-107.
- Burt, R. S. (1995). Le capital social, les trous structuraux et l'entrepreneur. *Revue française de sociologie*, 599-628.
- Burt, R. S. (2009). Structural holes: The social structure of competition. *Harvard university press*.
- Burt, R. S., & Knez, M. (1995). Kinds of third-party effects on trust. *Rationality and society*, 7(3), 255-292.
- Burt, Ronald S. (1992). Structural Holes. The Social Structure of Competition. Cambridge, Mass.: *Harvard University Press*.
- Cabanac, G., Hubert, G., & Milard, B. (2015). Academic careers in Computer Science: Continuance and transience of lifetime co-authorships. *Scientometrics*, 102(1), 135-150.
- Callon, M. (1986). Éléments pour une sociologie de la traduction : la domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc. *L'Année sociologique (1940/1948-)*, 36, 169-208.
- Callon, M. (1988). Faut-il croire en la recherche industrielle ?. *Culture technique*.
- Callon, M. (1991). Réseaux technico-économiques et irréversibilités. *Figures de l'irréversibilité en économie*, 38(1), 195-230.
- Callon, M., & Latour, B. (1985). Pour une Sociologie relativement exacte. *Ecole Nationale Supérieure des Mines de Paris, mimeo*.
- Cameron, C., Lee, H. Y., Anderson, C. B., Trachtenberg, J., & Chang, S. (2020). The role of scientific communication in predicting science identity and research career intention. *PloS one*, 15(2), e0228197.
- Cañibano, C. (2006). La gestion de la mobilité professionnelle des chercheurs: un défi pour les politiques de recherche et d'innovation. *La Revue pour l'Histoire du CNRS*, (14).
- Cañibano, C. (2017). Scientific mobility and economic assumptions: From the allocation of scientists to the socioeconomics of network transformation. *Science as Culture*, 26(4), 505-519.
- Cañibano, C., & Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: the state-of-the-art. *Research Evaluation*, 18(2), 86-94.

- Cañibano, C., & Woolley, R. (2015). Towards a socio-economics of the brain drain and distributed human capital. *International Migration*, 53(1), 115-130.
- Cañibano, C., D'Este, P., Otamendi, F. J., & Woolley, R. (2020). Scientific careers and the mobility of European researchers: an analysis of international mobility by career stage. *Higher Education*, 1-19.
- Cañibano, C., Fox, M. F., & Otamendi, F. J. (2016). Gender and patterns of temporary mobility among researchers. *Science and Public Policy*, 43(3), 320-331.
- Cañibano, C., Otamendi, F. J., & Solís, F. (2011). International temporary mobility of researchers: a cross-discipline study. *Scientometrics*, 89(2), 653-675.
- Cañibano, C., Otamendi, J., & Andújar, I. (2008). Measuring and assessing researcher mobility from CV analysis: the case of the Ramón y Cajal programme in Spain. *Research Evaluation*, 17(1), 17-31.
- Cattaneo, M., Malighetti, P., & Paleari, S. (2019). The Italian brain drain: cream and milk. *Higher Education*, 77(4), 603-622.
- Chan, H. F., Mixon, F. G., & Torgler, B. (2018). Relation of early career performance and recognition to the probability of winning the Nobel Prize in economics. *Scientometrics*, 114(3), 1069-1086.
- Chinchilla-Rodríguez, Z., Miao, L., Murray, D., Robinson-García, N., Costas, R., & Sugimoto, C. R. (2017). Networks of international collaboration and mobility: a comparative study.
- Chinchilla-Rodríguez, Z., Miao, L., Murray, D., Robinson-García, N., Costas, R., & Sugimoto, C. R. (2018). A global comparison of scientific mobility and collaboration according to national scientific capacities. *Frontiers in Research Metrics and Analytics*, 3, 17.
- Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and impacts: the influence of public research on industrial R&D. *Management science*, 48(1), 1-23.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, 94, S95-S120.
- Coleman, J. S. (1990). Commentary: Social institutions and social theory. *American Sociological Review*, 55(3), 333-339.
- Collins, H. M. (1974). The TEA set: Tacit knowledge and scientific networks. *Science studies*, 4(2), 165-185.
- Commeiras, N., & Georgescu, I. (2012). La pression à l'activité: une source de stress et d'épuisement professionnel? un regard croisé entre le contrôle de gestion et la gestion des ressources humaines: le cas des enseignants-chercheurs en sciences de gestion. *Actes du XXIIème Congrès de l'AGRH, Nancy, septembre*.
- Conseil Européen Lisbonne (23 et 24 Mars 2000). Conclusions de la présidence. Récupéré du Parlement Européen.
- Cope, S. R., Holliday, E. B., & Somerson, J. S. (2017). Factors Associated With Publication Impact at Shoulder and Elbow Surgery Fellowships. *Journal of Surgical Education*, 74(5), 883-888.
- Costas, R., & Bordons, M. (2005). Bibliometric indicators at the micro-level: some results in the area of natural resources at the Spanish CSIC. *Research evaluation*, 14(2), 110-120.
- Cour des Comptes (2005), Le rapport public annuel 2005.

Courpasson, D., & Guedri, Z. (2007). Les professeurs-chercheurs en management face à la performance scientifique. *Revue française de gestion*, (9), 173-194.

Cruz-Castro, L., & Sanz-Menéndez, L. (2010). Mobility versus job stability: Assessing tenure and productivity outcomes. *Research policy*, 39(1), 27-38.

Dabney, K. P., & Tai, R. H. (2014). Comparative analysis of female physicists in the physical sciences: Motivation and background variables. *Physical Review Special Topics-Physics Education Research*, 10(1), 010104.

Dahan, A., & Mangematin, V. (2010). Recherche, ou temps perdu ? Vers une intégration des tâches administratives au métier d'enseignant-chercheur. In *Annales des Mines-Gérer et comprendre* (No. 4, pp. 14-24). ESKA.

Dalton, G. W., Thompson, P. H., & Price, R. L. (1977). The four stages of professional career: A new look at performance by professional Organization Dynamics. *New york: American Management Association, Summer*.

de Baecque, F. (Ed.). (1974). *La situation des personnels enseignants des universités: éléments de réflexion pour une réforme*. Documentation française.

de Bologne, D. (1999). L'espace européen de l'enseignement supérieur. *Déclaration commune des*, 31.

DeFillippi, R. J., & Arthur, M. B. (1994). The boundaryless career: A competency-based perspective. *Journal of organizational behavior*, 15(4), 307-324.

Delicado, A. (2010). Going abroad to do science: Mobility trends and motivations of Portuguese researchers. *Science & Technology Studies*.

Département d'évaluation des établissements et Département d'évaluation des formations (2019) "Aide à la rédaction du dossier d'autoévaluation des écoles relevant d'une évaluation conjointe établissement/formations". HCERES

Deville, P., Wang, D., Sinatra, R., Song, C., Blondel, V. D., & Barabási, A. L. (2014). Career on the move: Geography, stratification, and scientific impact. *Scientific reports*, 4, 4770.

Di Lorenzo, F., & Tartari, V. (2014). Should i stay or should i go? How mobility explains individual scientific performance. In *Academy of Management Proceedings* (Vol. 2014, No. 1, p. 15422). Briarcliff Manor, NY 10510: Academy of Management.

Di Maria, C., & Strykowski, P. (2009). Migration, human capital accumulation and economic development. *Journal of Development Economics*, 90(2), 306-313.

Diamond, A. M. (1986). The life-cycle research productivity of mathematicians and scientists. *Journal of gerontology*, 41(4), 520-525.

Diem, A., & Wolter, S. C. (2013). The use of bibliometrics to measure research performance in education sciences. *Research in higher education*, 54(1), 86-114.

Dietz, J., Chompalov, I., Bozeman, B., Lane, E., & Park, J. (2000). Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment. *Scientometrics*, 49(3), 419-442.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, 147-160.

- Doja, A., Eady, K., Horsley, T., Bould, M. D., Victor, J. C., & Sampson, M. (2014). The h-index in medical education: an analysis of medical education journal editorial boards. *BMC medical education*, 14(1), 251.
- Dowd, K. O., & Kaplan, D. M. (2005). The career life of academics: Boundaried or boundaryless?. *Human Relations*, 58(6), 699-721.
- Duberley, J., Cohen, L., & Leeson, E. (2007). Entrepreneurial academics: developing scientific careers in changing university settings. *Higher Education Quarterly*, 61(4), 479-497.
- Durkheim, E. (1897). *Le suicide: étude de sociologie*. Alcan.
- Eckert, D., & Baron, M. (2013). Construire une géographie de la science. *M@Ppemonde*, 110(2)
- Edwards, G. (2010). Mixed-method approaches to social network analysis.
- Ejermo, O., Fassio, C., & Källström, J. (2020). Does Mobility across Universities Raise Scientific Productivity?. *Oxford Bulletin of Economics and Statistics*, 82(3), 603-624.
- Elias, N. (1991). *La société des individus*, Paris, Fayard, coll. *Agora*, trad. Etoré, J. [publié pour la première fois en 1987 sous le titre *Die Gesellschaft der Individuen*].
- Elias, N., & Hoffmann, Y. (1991). *Qu'est-ce que la sociologie*. Presses Pocket.
- Enders, J., & Kaulisch, M. (2006). The binding and unbinding of academic careers. *Wenner Gren International Series*, 83, 85.
- Eurydice (2018) France fr : Accords bilatéraux et coopération multilatérale. Commission Européenne.
- Eurydice (2018) France fr : Mobilité dans l'enseignement supérieur. Commission Européenne.
- Eurydice (2018) France fr: Mobilité dans l'enseignement supérieur. Commission Européenne.
- Faure, S., Millet, M., & Soulié, C. (2005). Enquête exploratoire sur le travail des enseignants chercheurs. Vers un bouleversement de la "table des valeurs académiques"?.
- Feldman, M., & Lichtenberg, F. R. (1998). The interaction between public and private R&D investment: cross-country evidence from European Community's R&D information service. *Annales d'Economie et de Statistique*, 49, 199-222.
- Fernandez-Zubieta, A., Geuna, A., & Lawson, C. (2013). Researchers' mobility and its impact on scientific productivity.
- Ferru, M. (2010). Formation process and geography of science–industry partnerships: the case of the University of Poitiers. *Industry and Innovation*, 17(6), 531-549.
- Fitzenberger, B., & Schulze, U. (2014). Up or out: Research incentives and career prospects of postdocs in Germany. *German Economic Review*, 15(2), 287-328.
- Foegle, J. P. (2013). Précarité du statut des jeunes chercheurs et doctorants étrangers. *Plein droit*, (2), I-VIII.
- Fontes, M. (2007). Scientific mobility policies: how Portuguese scientists envisage the return home. *Science and Public Policy*, 34(4), 284-298.

- Fontes, M., Videira, P., & Calapez, T. (2013). The impact of long-term scientific mobility on the creation of persistent knowledge networks. *Mobilities*, 8(3), 440-465.
- Fouad, N. A., Singh, R., Cappaert, K., Chang, W. H., & Wan, M. (2016). Comparison of women engineers who persist in or depart from engineering. *Journal of Vocational Behavior*, 92, 79-93.
- Fox, M. F., & Stephan, P. E. (2001). Careers of young scientists: Preferences, prospects and realities by gender and field. *Social studies of Science*, 31(1), 109-122.
- Franzoni, C., Scellato, G., & Stephan, P. (2014). The mover's advantage: The superior performance of migrant scientists. *Economics Letters*, 122(1), 89-93.
- Freeman, L. C., Roeder, D., & Mulholland, R. R. (1979). Centrality in social networks: II. Experimental results. *Social networks*, 2(2), 119-141.
- Gatignol, C. (2014). L'environnement professionnel des enseignants-chercheurs français explique-t-il leurs parcours de carrière?. @ *GRH*, (2), 51-80.
- Gaufrriau, M., Larsen, P., Maye, I., Roulin-Perriard, A., & von Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147-176.
- Gëdeshi, I., & King, R. (2019). The Albanian scientific diaspora: can the brain drain be reversed?. *Migration and Development*, 1-23.
- Gibbs Jr, K. D., & Griffin, K. A. (2013). What do I want to be with my PhD? The roles of personal values and structural dynamics in shaping the career interests of recent biomedical science PhD graduates. *CBE—Life Sciences Education*, 12(4), 711-723.
- Gibson, J., & McKenzie, D. (2012). The economic consequences of 'brain drain' of the best and brightest: Microeconomic evidence from five countries. *The Economic Journal*, 122(560), 339-375.
- Gibson, J., & McKenzie, D. (2014). Scientific mobility and knowledge networks in high emigration countries: Evidence from the Pacific. *Research policy*, 43(9), 1486-1495.
- Gingras, Y. (2003). Idées d'universités. *Actes de la recherche en sciences sociales*, (3), 3-7.
- Gingras, Y. (2008). La Fièvre de l'évaluation de la recherche du mauvais usage de faux indicateurs. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 100(1), 41-44.
- Glänzel, W., & Schubert, A. (2004). Analysing scientific networks through co-authorship. In *Handbook of quantitative science and technology research* (pp. 257-276). Springer, Dordrecht.
- Glaser, J. (2001). Macrostructures, careers and knowledge production: a neoinstitutionalist approach. *International Journal of Technology Management*, 22(7-8), 698-715.
- Gläser, J., & Laudel, G. (2007). The social construction of bibliometric evaluations. In *The changing governance of the sciences* (pp. 101-123). Springer, Dordrecht.
- Grandori, A. (1997). An organizational assessment of interfirm coordination modes. *Organization Studies*, 18(6), 897-925.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of sociology*, 91(3), 481-510.

- Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6), 1360-1380.
- Granrose, C. S., & Baccili, P. A. (2006). Do psychological contracts include boundaryless or protean careers?. *Career Development International*.
- Grossetti, M. (2004). *Sociologie de l'imprévisible*.
- Grossetti, M. (2006). Trois échelles d'action et d'analyse. *L'année Sociologique*, 56(2), 285-307.
- Grossetti, M., & Bès, M. P. (2002). Proximité spatiale et relations science-industrie: savoirs tacites ou encastrement (Polanyi ou Polanyi)?. *Revue d'Economie Regionale Urbaine*, (5), 777-787.
- Grossetti, M., Barthe, J. F., & Beslay, C. (2006). La mobilisation des relations sociales dans les processus de création d'entreprises. Aperçus à partir d'une enquête en cours. *Sociologies pratiques*, (2), 47-59.
- Grossetti, M., Eckert, D., Gingras, Y., Jegou, L., Lariviere, V., & Milard, B. (2014). Cities and the geographical deconcentration of scientific activity: A multilevel analysis of publications (1987–2007). *Urban Studies*, 51(10), 2219-2234.
- Grossetti, M., Maisonobe, M., Jégou, L., Milard, B., & Cabanac, G. (2020). L'organisation spatiale de la recherche française à travers les publications savantes: régularité des tendances de long terme et désordre des politiques publiques (1999-2017).
- Grubel, H. B., & Scott, A. D. (1966). The international flow of human capital. *The American Economic Review*, 56(1/2), 268-274.
- Györfy, B., Nagy, A. M., Herman, P., & Török, Á. (2018). Factors influencing the scientific performance of Momentum grant holders: an evaluation of the first 117 research groups. *Scientometrics*, 117(1), 409-426.
- Halevi, G., Moed, H. F., & Bar-Ilan, J. (2016). Researchers' mobility, productivity and impact: Case of top producing authors in seven disciplines. *Publishing Research Quarterly*, 32(1), 22-37.
- Hall, D. T. (1996). Protean careers of the 21st century. *Academy of Management Perspectives*, 10(4), 8-16.
- Hall, D. T. (2004). The protean career: A quarter-century journey. *Journal of vocational behavior*, 65(1), 1-13.
- Hall, D. T., & Moss, J. E. (1998). The new protean career contract: Helping organizations and employees adapt. *Organizational dynamics*, 26(3), 22-37.
- Hanifan, L. J. (1916). The rural school community center. *The Annals of the American Academy of Political and Social Science*, 67(1), 130-138.
- Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative science quarterly*, 44(1), 82-111.
- Harrison, M., Dunbar, D., Ratmansky, L., Boyd, K., & Lopatto, D. (2011). Classroom-based science research at the introductory level: changes in career choices and attitude. *CBE—Life Sciences Education*, 10(3), 279-286.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.

Haut conseil de l'évaluation de la recherche et de l'enseignement supérieur (2015, 29 Septembre). Evaluation de l'unité : Laboratoire de Chimie de Coordination LCC.

Haut conseil de l'évaluation de la recherche et de l'enseignement supérieur (2017, 18 Août). Evaluation de l'unité : Institut de chimie des milieux et matériaux de Poitiers IC2MP.

Hayat, T., & Lyons, K. (2017). A typology of collaborative research networks. *Online Information Review*.

Hayter, C. S., & Parker, M. A. (2019). Factors that influence the transition of university postdocs to non-academic scientific careers: An exploratory study. *Research Policy*, 48(3), 556-570.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431.

Hoisl, K. (2007). Tracing mobile inventors—the causality between inventor mobility and inventor productivity. In *A Study of Inventors* (pp. 65–119). Wiesbaden: Deutscher Universitätsverlag.

Horizon 2020 Le Portail Français Du Programme Européen Pour La Recherche Et l'Innovation (2013), Charte européenne du chercheur & Code de conduite pour le recrutement des chercheurs, Ministère de l'enseignement supérieur et de la recherche.

Horizon 2020 Le Portail Français Du Programme Européen Pour La Recherche Et l'Innovation (2013), Les Actions Marie Sklodowska-Curie (A.M.S.C.), Ministère de l'enseignement supérieur et de la recherche.

Horta et Yudvich (2016). The role of academic inbreeding in developing higher education systems: Challenges and possible solutions. *Technological Forecasting and Social Change*, Volume 113, Part B, 363-372

Horta, H. (2013). Deepening our understanding of academic inbreeding effects on research information exchange and scientific output: new insights for academic based research. *Higher education*, 65(4), 487-510.

Horta, H., & Yonezawa, A. (2013). Going places: exploring the impact of intra-sectoral mobility on research productivity and communication behaviors in Japanese academia. *Asia Pacific Education Review*, 14(4), 537-547.

Hossain, L., & Fazio, D. (2009). The social networks of collaborative process. *The Journal of High Technology Management Research*, 20(2), 119-130.

Hou, H., Kretschmer, H., & Liu, Z. (2008). The structure of scientific collaboration networks in Scientometrics. *Scientometrics*, 75(2), 189-202.

Hubbard, K. E., & Dunbar, S. D. (2017). Perceptions of scientific research literature and strategies for reading papers depend on academic career stage. *PloS one*, 12(12), e0189753.

Hussler, C., & Ronde, P. (2005). Les réseaux d'innovation des chercheurs universitaires sont-ils géographiquement circonscrits?. *Revue d'économie industrielle*, 109(1), 69-90.

IC2MP UMR CNRS 7285, Institut de Chimie des Milieux et des Matériaux de Poitiers (<http://ic2mp.labo.univ-poitiers.fr/>)

- Inda, M., Rodríguez, C., & Peña, J. V. (2013). Gender differences in applying social cognitive career theory in engineering students. *Journal of vocational behavior*, 83(3), 346-355.
- J. Nettelbeck, Le recrutement des professeurs d'université, Paris, *Maison des sciences de l'homme*, 1979, pp. 80 sq.
- Jean-Jacques, S. (2013). Analyse textuelle avec IRaMuTeQ et interprétations référentielles des programmes officiels de mathématiques en quatrième.
- Jensen, P., Rouquier, J. B., & Croissant, Y. (2009). Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics*, 78(3), 467-479.
- Johnson, H. (1965). The economics of the "brain drain": The Canadian case. *Minerva*, 3(3), 299-311.
- Johnson, H. (1968). An 'internationalist' model. *The brain drain*, 5, 69-91.
- Jonkers, K., & Cruz-Castro, L. (2013). Research upon return: The effect of international mobility on scientific ties, production and impact. *Research Policy*, 42(8), 1366-1377.
- Jonkers, K., & Tijssen, R. (2008). Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. *Scientometrics*, 77(2), 309-333.
- Jöns, H. (2007). Transnational mobility and the spaces of knowledge production: a comparison of global patterns, motivations and collaborations in different academic fields. *Social geography*, 2(2), 97-114.
- Jöns, H. (2009). 'Brain circulation' and transnational knowledge networks: studying long-term effects of academic mobility to Germany, 1954-2000. *Global networks*, 9(3), 315-338.
- Kalén, S., Stenfors-Hayes, T., Hylén, U., Larm, M. F., Hindbeck, H., & Ponzer, S. (2010). Mentoring medical students during clinical courses: a way to enhance professional development. *Medical teacher*, 32(8), e315-e321.
- Kang, J., Hense, J., Scheersoi, A., & Keinonen, T. (2019). Gender study on the relationships between science interest and future career perspectives. *International Journal of Science Education*, 41(1), 80-101.
- Kannappanavar, B. U., & Swamy, C. (2004). Publishing trends of Indian chemical scientists: A bibliometric study.
- Kidd, C. V. (1965). The economics of the "brain drain". *Minerva*, 4(1), 105-107.
- Kim, E. M. (1996). The industrial organization and growth of the Korean chaebol: integrating development and organizational theories. *Asian business networks*, 64, 231.
- Kim, T. (2010). Transnational academic mobility, knowledge, and identity capital. *Discourse: Studies in the cultural politics of education*, 31(5), 577-591.
- Kim, T. (2017). Academic mobility, transnational identity capital, and stratification under conditions of academic capitalism. *Higher Education*, 73(6), 981-997.
- Knight, F. H. (1921). Cost of production and price over long and short periods. *Journal of Political Economy*, 29(4), 304-335.

Kossi, Y., Lesueur, J. Y., & Sabatier, M. (2016). Publish or teach? The role of the scientific environment on academics' multitasking. *Industrial and Corporate Change*, 25(3), 487-506.

Koul, R., Lerdpornkulrat, T., & Chantara, S. (2011). Relationship between career aspirations and measures of motivation toward biology and physics, and the influence of gender. *Journal of Science Education and Technology*, 20(6), 761-770.

Krackhardt, D. (1992). The Strength of Strong Ties: The Importance of Philos in Networks and Organizations. *Networks and Organizations..* (1998). Super strong and sticky. *Power Influ. Organ*, 21.

Kühn, P. J. (1983). Analysis of Busy Period and Response Time Distributions in Queuing Networks. In *Messung, Modellierung und Bewertung von Rechensystemen* (pp. 135-149). Springer, Berlin, Heidelberg.

Kwiek, M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers?. *Scientometrics*, 115(1), 415-462.

L. Sanz et L. Cruz, « Doctorate Holders' Careers: Research and Policy Questions », working party of national experts on science and technology indicators, workshop on user needs for indicators on careers of doctorate holders, Paris, 27 septembre, DSTI/EAS/STP/NESTI/RD(2004)18, 2004.

Landes, D. S. 1993 [1958]. *Banquiers et pachas*. Paris : Albin Michel.

Langfeldt, L., Benner, M., Sivertsen, G., Kristiansen, E. H., Aksnes, D. W., Borlaug, S. B., ... & Pelkonen, A. (2015). Excellence and growth dynamics: A comparative study of the Matthew effect. *Science and Public Policy*, 42(5), 661-675.

Latour, B. (1987). Les "vues" de l'esprit. *Réseaux. Communication-Technologie-Société*, 5(27), 79-96.

Latour, B. (1996). Sur la pratique des théoriciens. *Savoirs théoriques et savoirs d'action*, 55, 116-116.

Latour, B. (1999). *Pandora's hope: essays on the reality of science studies*. Harvard university press.

Laudel, G. (2001). Collaboration, creativity and rewards: why and how scientists collaborate. *International Journal of Technology Management*, 22(7-8), 762-781.

Laudel, G., & Bielick, J. (2019). How do field-specific research practices affect mobility decisions of early career researchers?. *Research Policy*, 48(9), 103800.

Laudel, G., & Gläser, J. (2008). From apprentice to colleague: The metamorphosis of early career researchers. *Higher education*, 55(3), 387-406.

Laville F. et Maddi A., La position scientifique et technologique de la France dans la recherche en chimie, L'état de l'Enseignement supérieur, de la Recherche et de l'Innovation en France (n°12 - Juillet 2019).

Lawson, C., & Shibayama, S. (2015). Appointment, Promotion, and Mobility of Bioscience Researchers in Japan. In *Global mobility of research scientists* (pp. 239-269). Academic Press.

Lawson, C., & Shibayama, S. (2015). International research visits and careers: An analysis of bioscience academics in Japan. *Science and Public Policy*, 42(5), 690-710.

LCC CNRS Laboratoire de Chimie de Coordination UPR 8241 (<https://www.lcc-toulouse.fr/>)

Lebart, L., & Salem, A. (1988). Analyse statistique des données textuelles: questions ouvertes et lexicométrie. Paris: Dunod.

- Lebart, L., Salem, A., & Berry, L. (1998). Visualization of Textual Data. In *Exploring Textual Data* (pp. 101-128). Springer, Dordrecht.
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of vocational behavior, 45*(1), 79-122.
- Lent, R. W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. *Career choice and development, 4*, 255-311.
- Lent, R. W., Sheu, H. B., Singley, D., Schmidt, J. A., Schmidt, L. C., & Gloster, C. S. (2008). Longitudinal relations of self-efficacy to outcome expectations, interests, and major choice goals in engineering students. *Journal of Vocational Behavior, 73*(2), 328-335.
- Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management science, 50*(11), 1477-1490.
- Levin, S. G., & Stephan, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review, 114*-132.
- Lévi-Strauss, C. (1967). Les Structures élémentaires de la parenté. 2e éd. *Paris-La Haye, Mouton, 108*.
- Levy, R. (2015). *L'Université et ses territoires : Dynamismes des villes moyennes et particularités de site*. PUG.
- Li, E. Y., Liao, C. H., & Yen, H. R. (2013). Co-authorship networks and research impact: A social capital perspective. *Research Policy, 42*(9), 1515-1530.
- Lin, N. (1995). Les ressources sociales: une théorie du capital social. *Revue française de sociologie, 685*-704.
- Lincoln, J. R., Gerlach, M. L., & Ahmadjian, C. L. (1996). Keiretsu networks and corporate performance in Japan. *American sociological review, 67*-88.
- Lissoni, F., Mairesse, J., Montobbio, F., & Pezzoni, M. (2011). Scientific productivity and academic promotion: a study on French and Italian physicists. *Industrial and corporate change, 20*(1), 253-294.
- Livingstone, D. N. (1995). The spaces of knowledge: contributions towards a historical geography of science. *Environment and Planning D: Society and Space, 13*(1), 5-34
- Long, R., Crawford, A., White, M., & Davis, K. (2009). Determinants of faculty research productivity in information systems: An empirical analysis of the impact of academic origin and academic affiliation. *Scientometrics, 78*(2), 231-260.
- Loubère, L. (2016). L'analyse de similitude pour modéliser les CHD. In *Journées internationales d'Analyse statistique des Données Textuelles*.
- Loubère, L., & Ratinaud, P. (2014). Documentation IRaMuTeQ 0.6 alpha 3 version 0.1. *Toulouse, França*.
- Louvel, S., & Valette, A. (2014). Les carrières à l'université. *Revue d'anthropologie des connaissances, 8*(3), 523-546.
- Mahroum, S. (2000). Highly skilled globetrotters: mapping the international migration of human capital. *R&D Management, 30*(1), 23-32.

- Mahroum, S. (2000). Scientific mobility: An agent of scientific expansion and institutional empowerment. *Science Communication*, 21(4), 367-378.
- Mahroum, S. (2000). Scientists and global spaces. *Technology in society*, 22(4), 513-523.
- Maisonobe, M. (2015). *Étudier la géographie des activités et des collectifs scientifiques dans le monde: de la croissance du système de production contemporain aux dynamiques d'une spécialité, la réparation de l'ADN* (Doctoral dissertation).
- Maisonobe, M., Grossetti, M., Milard, B., Eckert, D., & Jégou, L. (2016). L'évolution mondiale des réseaux de collaborations scientifiques entre villes: des échelles multiples. *Revue française de sociologie*, 57(3), 417-441.
- Mali, F., Kronegger, L., Doreian, P., & Ferligoj, A. (2012). Dynamic scientific co-authorship networks. In *Models of science dynamics* (pp. 195-232). Springer, Berlin, Heidelberg.
- Mansfield, E. (1991). Academic research and industrial innovation. *Research policy*, 20(1), 1-12.
- Marchand, P., & Ratinaud, P. (2012). L'analyse de similitude appliquée aux corpus textuels: les primaires socialistes pour l'élection présidentielle française (septembre-octobre 2011). *Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles. JADT, 2012*, 687-699.
- Markowitz, D. G. (2004). Evaluation of the long-term impact of a university high school summer science program on students' interest and perceived abilities in science. *Journal of science education and technology*, 13(3), 395-407.
- Marmolejo-Leyva, R., Perez-Angon, M. A., & Russell, J. M. (2015). Mobility and international collaboration: case of the Mexican scientific diaspora. *PloS one*, 10(6), e0126720.
- Mayer, E. N., Lenherr, S. M., Hanson, H. A., Jessop, T. C., & Lowrance, W. T. (2017). Gender differences in publication productivity among academic urologists in the United States. *Urology*, 103, 39-46.
- Meijer, I., Mejlgaard, N., Woolley, R., Rafols, I., & Wroblewski, A. (2016). Monitoring the evolution and benefits of responsible research and innovation (MoRRI)—a preliminary framework for RRI dimensions & indicators.
- Melin, G. (2004). Postdoc abroad: inherited scientific contacts or establishment of new networks? *Research Evaluation*, 13(2), 95-102.
- Melin, G. (2004). Postdoc abroad: inherited scientific contacts or establishment of new networks?. *Research Evaluation*, 13(2), 95-102.
- Melin, G. (2005). The dark side of mobility: negative experiences of doing a postdoc period abroad. *Research Evaluation*, 14 (3), 229-237.
- Melin, G., & Persson, O. (1996). Studying research collaboration using co authorships. *Scientometrics*, 36(3), 363-377.
- Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, 141-183
- Merton, R. K. (1942). A note on science and democracy. *J. Legal & Pol. Soc.*, 1, 115.
- Merton, R. K. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American sociological review*, 22(6), 635-659.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810), 56-63.

Milard, B. (2010). Les citations scientifiques: des réseaux de références dans des univers de références. L'exemple d'articles de chimie. *Redes. Revista hispana para el análisis de redes sociales*, 19, 69-93.

Milard, B. (2013). Les écrits scientifiques: des ressorts relationnels pour la recherche. *Sciences de la société*, (89), 18-37.

Milard, B. (2013). Quelles sociabilités derrière les références bibliographiques? Citations et relations sociales. *Socio-logos. Revue de l'association française de sociologie*, (8).

Milard, B. (2014). Réseaux et entourages citationnels en chimie: quatre types d'implications dans la recherche. *L'Année sociologique*, 64(1), 15-46.

Ministère de l'enseignement supérieur et de la recherche (16 Octobre 2013). ESPACE EUROPÉEN DE LA RECHERCHE (E.E.R.) Présentation de l'E.E.R..

Ministère de l'enseignement supérieur et de la recherche (17 Octobre 2013). ESPACE EUROPÉEN DE LA RECHERCHE (E.E.R.), Le cadre législatif et réglementation européenne.

Ministère de l'enseignement supérieur et de la recherche (2020). La mobilité des chercheurs et des enseignants-chercheurs.

Ministère de l'enseignement supérieur et de la recherche (23 Octobre 2019). Espace Européen de l'Enseignement Supérieur (E.E.E.S.) Présentation de l'E.E.E.S.

Ministère de l'enseignement supérieur, de la recherche et de l'innovation (Avril 2006), Etude de la promotion 2005 des qualifiés aux fonctions de professeur des universités et de maître de conférences.

Ministère de l'enseignement supérieur, de la recherche et de l'innovation (19 Juillet 2019), L'état de l'Enseignement supérieur, de la Recherche et de l'Innovation en France (n°12 - Juillet 2019).

Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. (2011, 15 mai). ÉNERGIE, DÉVELOPPEMENT DURABLE, CHIMIE ET PROCÉDÉS.

Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. (2018, 24 Avril). Fiches démographiques des sections de sciences - Année universitaire 2016/2017.

Mohammadamin, E., Ali, R. V., & Abrizah, A. (2017). Co-authorship network of scientometrics research collaboration. *Malaysian Journal of Library & Information Science*, 17(3), 73-93.

Mongeon, P., Brodeur, C., Beaudry, C., & Larivière, V. (2016). Concentration of research funding leads to decreasing marginal returns. *Research Evaluation*, 25(4), 396-404.

Morano-Foadi, S. (2005). Scientific mobility, career progression, and excellence in the European research area. *International migration*, 43(5), 133-162.

Morrison, L. J., Lorens, E., Bandiera, G., Liles, W. C., Lee, L., Hyland, R., ... & Levinson, W. (2014). Impact of a formal mentoring program on academic promotion of Department of Medicine faculty: a comparative study. *Medical teacher*, 36(7), 608-614.

Mullins, N. C. (1972). The development of a scientific specialty: The phage group and the origins of molecular biology. *Minerva*, 10(1), 51-82.

- Musselin, C. (2008). Vers un marché international de l'enseignement supérieur?. *Critique internationale*, (2), 13-24.
- Musselin, C. (2009). Les réformes des universités en Europe: des orientations comparables, mais des déclinaisons nationales. *Revue du MAUSS*, (1), 69-91.
- Musselin, C. (2010). The impact of changing recruitment practices on academic profiles. *Academic and professional identities in higher education: The challenges of a diversifying workforce*, 125-137.
- Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the national academy of sciences*, 101(suppl 1), 5200-5205.
- Nguyen, V., Marmor, R. A., Ramamoorthy, S. L., Blair, S. L., Clary, B. M., & Sicklick, J. K. (2018). Academic surgical oncologists' productivity correlates with gender, grant funding, and institutional NCI comprehensive cancer center affiliation. *Annals of surgical oncology*, 25(7), 1852-1859.
- Nowak, J. K., Lubarski, K., Kowalik, L. M., & Walkowiak, J. (2018). H-index in medicine is driven by original research. *Croatian medical journal*, 59(1), 25-32.
- Odell, T., Toor, H., Takayanagi, A., Zampella, B., Siddiqi, J., Jalal, S., ... & Khosa, F. (2019). Gender disparity in academic neurosurgery. *Cureus*, 11(5).
- OECD. (2010), *Measuring Innovation: A New Perspective*, OECD Publishing, Paris
- OST (2019), *Dynamics of scientific production in the world, in Europe and in France, 2000-2016*, Hcéres, Paris
- Pagel, P. S., & Hudetz, J. A. (2011). An analysis of scholarly productivity in United States academic anaesthesiologists by citation bibliometrics. *Anaesthesia*, 66(10), 873-878.
- Parker, J. N., Allesina, S., & Lortie, C. J. (2013). Characterizing a scientific elite (B): publication and citation patterns of the most highly cited scientists in environmental science and ecology. *Scientometrics*, 94(2), 469-480.
- Pashkova, A. A., Svider, P. F., Chang, C. Y., Diaz, L., Eloy, J. A., & Eloy, J. D. (2013). Gender disparity among US anaesthesiologists: are women underrepresented in academic ranks and scholarly productivity?. *Acta Anaesthesiologica Scandinavica*, 57(8), 1058-1064.
- Patinkin, D. 1968 "A 'Nationalist' model", in W. Adams (Ed.), *The Brain Drain*, Macmillan, New York: Chapter 6, 92-108.
- Petersen, Alexander Michael, et al. "Reputation and impact in academic careers." *Proceedings of the National Academy of Sciences* 111.43 (2014): 15316-15321.
- Pettigrew, A. M. (1990). Longitudinal field research on change: Theory and practice. *Organization science*, 1(3), 267-292.
- Pezzoni, M., Sterzi, V., & Lissoni, F. (2012). Career progress in centralized academic systems: Social capital and institutions in France and Italy. *Research Policy*, 41(4), 704-719.
- Pfeiffer, M., Fischer, M. R., & Bauer, D. (2016). Publication activities of German junior researchers in academic medicine: which factors impact impact factors?. *BMC medical education*, 16(1), 1-10.
- Piore, M. J., & Sabel, C. F. (1986). *The second industrial divide: possibilities for prosperity*.
- Polanyi, M. (1958). *Personal Knowledge: "Towards a Post-Critical Philosophy*. University of Chicago Press, Chicago

- Polanyi, M. (1962). Tacit knowing. *Philosophy Today*, 6(4), 239-262.
- Powell, A., Dainty, A., & Bagilhole, B. (2012). Gender stereotypes among women engineering and technology students in the UK: lessons from career choice narratives. *European Journal of Engineering Education*, 37(6), 541-556.
- Powell, W. W., & Brantley, P. (1992). Networks and organizations: structure, form, and action. *Competitive Cooperation in Biotechnology: learning through networks*.
- Powell, W. W., & Smith-Doerr, L. (1994). Networks and economic life. *The handbook of economic sociology*, 368, 380.
- Puuska, H. M. (2010). Effects of scholar's gender and professional position on publishing productivity in different publication types. Analysis of a Finnish university. *Scientometrics*, 82(2), 419-437.
- Quermonne, J. L. (1981). *Etude générale des problèmes posés par la situation des personnels enseignants universitaires: rapport*. Centre national de documentation pédagogique.
- Rallet, A., & Torre, A. (2001). Proximité géographique ou proximité organisationnelle? Une analyse spatiale des coopérations technologiques dans les réseaux localisés d'innovation.
- Reagans, R., & Zuckerman, E. W. (2001). Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization science*, 12(4), 502-517.
- Reiner, C., & Meyer, S. (2015). The return on scientific mobility: An econometric analysis based on CV data.
- Reinert, M. (1993). Les «mondes lexicaux» et leur «logique» à travers l'analyse statistique d'un corpus de récits de cauchemars. *Langage et société (Maison des Sciences de l'Homme)*, 66, 5-39.
- Reyes Gonzalez, L., González Brambila, C. N., & Veloso, F. (2018). Birth of prominent scientists. *PloS one*, 13(3), e0193374.
- Ricoeur, P. (1983). *Temps et récit: la configuration dans le récit de fiction* (Vol. 2). Seuil.
- Robinson-Garcia, N., Cañibano, C., Woolley, R., & Costas, R. (2016). Tracing scientific mobility of Early Career Researchers in Spain and The Netherlands through their publications. *arXiv*, arXiv-1606.
- Robinson-Garcia, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, 13(1), 50-63.
- Rouse, J. (1987). Knowledge and power: Toward a political philosophy of science.
- Rowley, T., Behrens, D., & Krackhardt, D. (2000). Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic management journal*, 21(3), 369-386.
- Sabatier, M. (2010). Do female researchers face a glass ceiling in France? A hazard model of promotions. *Applied Economics*, 42(16), 2053–2062.
- Saint-Blancat, C. (2018). Making sense of scientific mobility: How Italian scientists look back on their trajectories of mobility in the EU. *Higher Education Policy*, 31(1), 37-54.

- Salter, A., D'Este, P., Pavitt, K., Scott, A., Martin, B., Geuna, A., ... & Patel, P. (2000). Talent, not technology: the impact of publicly funded research on innovation in the UK. *SPRU, University of Sussex*.
- Sambunjak, D., & Marušić, A. (2009). Mentoring: what's in a name?. *JaMa*, 302(23), 2591-2592.
- Sambunjak, D., Straus, S. E., & Marušić, A. (2006). Mentoring in academic medicine: a systematic review. *Jama*, 296(9), 1103-1115.
- Samuelson, P. A. (1954). The pure theory of public expenditure. *The review of economics and statistics*, 387-389.
- Sandnes, F. E. (2018). Do Norwegian academics who publish more earn higher salaries?. *Scientometrics*, 115(1), 263-281.
- Santos, G., & Dang Van Phu, S. (2019). Gender and Academic Rank in the UK. *Sustainability*, 11(11), 3171.
- Sarigöl, E., Pfitzner, R., Scholtes, I., Garas, A., & Schweitzer, F. (2014). Predicting scientific success based on coauthorship networks. *EPJ Data Science*, 3(1), 9.
- Sauermann, H., & Roach, M. (2012). Science PhD career preferences: levels, changes, and advisor encouragement. *PLoS one*, 7(5), e36307.
- Sauermann, H., & Stephan, P. (2013). Conflicting logics? A multidimensional view of industrial and academic science. *Organization science*, 24(3), 889-909.
- Saxenian, A. (1994). Regional networks: industrial adaptation in Silicon Valley and route 128.
- Saxenian, A. (2002). Brain circulation. How high-skill immigration makes everyone better off. *Brookings Review*, 20(1), 28-31.
- Scellato, G., Franzoni, C., & Stephan, P. (2015). Migrant scientists and international networks. *Research Policy*, 44(1), 108-120.
- Schwartz, R. (2008). Commission de réflexion sur l'avenir des personnels de l'Enseignement supérieur. *Rapport à Madame la ministre de l'Enseignement supérieur et de la Recherche*.
- Sent, E. M. (1999). Economics of science: survey and suggestions. *Journal of Economic Methodology*, 6(1), 95-124.
- Shen, C. (2010). *The patterns, effects and evolution of player social networks in online gaming communities*. University of Southern California.
- Simmel, G. (1955). Conflict and the Web Affiliations (KH Wolff and R. Bendix, trans.).
- Soulié, C. (1996). Précarité dans l'enseignement supérieur. *Actes de la recherche en sciences sociales*, 115(1), 58-64.
- Stark, O. (2004). Rethinking the brain drain. *World Development*, 32(1), 15-22.
- Stark, O., Helmenstein, C., & Prskawetz, A. (1998). Human capital depletion, human capital formation, and migration: a blessing or a "curse"?. *Economics Letters*, 60(3), 363-367.
- Stephan, P. (2013, March). The Economics of the Postdoctoral Position. In *Keynote address to the National Postdoctoral Association 11th Annual Meeting, Charleston March* (pp. 15-17).

- Stephan, P. E. (1996). The economics of science. *Journal of Economic literature*, 34(3), 1199-1235.
- Stephan, P. E. (2008). Science and the university: Challenges for future research. *CESifo Economic Studies*, 54(2), 313-324.
- Stephan, P. E., & Levin, S. G. (1997). The critical importance of careers in collaborative scientific research. *Revue d'économie industrielle*, 79(1), 45-61.
- Stephan, P. E., & Levin, S. G. (2001). Exceptional contributions to US science by the foreign-born and foreign-educated. *Population research and Policy review*, 20(1-2), 59-79.
- Stephan, P. E., Gurnu, S., Sumell, A. J., & Black, G. (2007). Who's patenting in the university? Evidence from the survey of doctorate recipients. *Econ. Innov. New Techn.*, 16(2), 71-99.
- Stephan, P., Scellato, G., & Franzoni, C. (2015). International competition for PhDs and postdoctoral scholars: What does (and does not) matter. *Innovation Policy and the Economy*, 15(1), 73-113.
- Suitor, J. J., Wellman, B., & Morgan, D. L. (1997). It's about time: How, why, and when networks change. *Social networks*, 19(1), 1-7.
- Sullivan, S. E., & Baruch, Y. (2009). Advances in career theory and research: A critical review and agenda for future exploration. *Journal of management*, 35(6), 1542-1571.
- Susarla, S. M., Rada, E. M., Lopez, J., Swanson, E. W., Miller, D., Redett, R. J., & Kumar, A. R. (2017). Does the H index correlate with academic rank among full-time academic craniofacial surgeons?. *Journal of surgical education*, 74(2), 222-227.
- Swihart, R. K., Sundaram, M., Höök, T. O., DeWoody, J. A., & Kellner, K. F. (2016). Performance benchmarks for scholarly metrics associated with fisheries and wildlife faculty. *PloS one*, 11(5), e0155097.
- Tani, M. (2008). Short-Term Skilled Labour Movements and Economic Growth. *International Migration*, 46(3), 161-187.
- Thorelli, H. B. (1986). Networks: between markets and hierarchies. *Strategic management journal*, 7(1), 37-51.
- Thorn, K., & Holm-Nielsen, L. B. (2008). International mobility of researchers and scientists: Policy options for turning a drain into a gain. *The international mobility of talent: types, causes, and development impact*, 145-167.
- Tourbeaux J., Rossi P., Thirion J. (2012), Origine des enseignants chercheurs recrutés lors de la campagne 2012 Session synchronisée. Récupéré du Ministère de l'enseignement supérieur et de la recherche.
- Uddin, S., Hossain, L., & Rasmussen, K. (2013). Network effects on scientific collaborations. *PloS one*, 8(2), e57546.
- Valsangkar, N., Fecher, A. M., Rozycki, G. S., Blanton, C., Bell, T. M., Freischlag, J., ... & Koniaris, L. G. (2016). Understanding the barriers to hiring and promoting women in surgical subspecialties. *Journal of the American College of Surgeons*, 223(2), 387-398.
- Van Den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: a longitudinal case study. *Scientometrics*, 106(1), 143-162.

van den Besselaar, P., & Sandström, U. (2017). Vicious circles of gender bias, lower positions, and lower performance: Gender differences in scholarly productivity and impact. *PloS one*, *12*(8), e0183301.

Van Heeringen, A., & Dijkwel, P. (1987). The relationships between age, mobility and scientific productivity. Part I: Effect of mobility on productivity. *Scientometrics*, *11*(5-6), 267-280.

Van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research policy*, *40*(3), 463-472.

VanMeter-Adams, A., Frankenfeld, C. L., Bases, J., Espina, V., & Liotta, L. A. (2014). Students who demonstrate strong talent and interest in STEM are initially attracted to STEM through extracurricular experiences. *CBE—Life Sciences Education*, *13*(4), 687-697.

Veugelers, R., & Van Bouwel, L. (2015). The effects of international mobility on European researchers: Comparing intra-EU and US mobility. *Research in Higher Education*, *56*(4), 360-377.

Vieira, E. S., & Gomes, J. A. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, *4*(1), 1-13.

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research policy*, *34*(10), 1608-1618.

Wang, J., Hooi, R., Li, A. X., & Chou, M. H. (2019). Collaboration patterns of mobile academics: The impact of international mobility. *Science and Public Policy*, *46*(3), 450-462.

Whittington, K. B. (2018). A tie is a tie? Gender and network positioning in life science inventor collaboration. *Research Policy*, *47*(2), 511-526.

Wong, K. Y., & Yip, C. K. (1999). Education, economic growth, and brain drain. *Journal of Economic Dynamics and Control*, *23*(5-6), 699-726.

Woolley, R., & Turpin, T. (2009). CV analysis as a complementary methodological approach: Investigating the mobility of Australian scientists. *Research Evaluation*, *18*(2), 143-151.

Woolley, R., Turpin, T., Marceau, J., & Hill, S. (2008). Mobility matters: Research training and network building in science. *Comparative Technology Transfer and Society*, *6*(3), 159-184.

Xie, Z., Ouyang, Z., Li, J., Dong, E., & Yi, D. (2018). Modelling transition phenomena of scientific coauthorship networks. *Journal of the Association for Information Science and Technology*, *69*(2), 305-317.

Yan, E., Ding, Y., & Zhu, Q. (2010). Mapping library and information science in China: A coauthorship network analysis. *Scientometrics*, *83*(1), 115-131.

Zimmer, A., Lemerrier, C., & Cénac, P. (2020). Enseignement et recherche sont inséparables. *La Vie des Idées*, en-ligne.

Zucker, L. G., Darby, M. R., & Armstrong, J. (1998). Geographically localized knowledge: spillovers or markets?. *Economic Inquiry*, *36*(1), 65-86.

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1994). *Intellectual capital and the birth of US biotechnology enterprises* (No. w4653). National Bureau of Economic Research.

Table des matières

Remerciements	3
Sommaire	7
Introduction générale	9
Chapitre 1 : Revue de la littérature la carrière scientifique et les comportements collaboratifs	24
I- Littérature et carrière scientifique	25
A) Carrière scientifique : état des lieux de la littérature	25
1- Méthodologie de l'analyse lexicométrique	25
2- Analyse des similitudes : détection des communautés lexicales relatives à la carrière scientifique.....	27
3- Classification des enjeux de la carrière scientifique soulevés dans la littérature.....	28
B) Différentes approches de la carrière scientifique	33
1- Approche économique : progression de carrière d'un agent autonome.....	33
2- Approches psychologiques et sociologiques de la carrière.....	36
C) Impact de l'évolution de l'environnement universitaire sur le métier d'enseignant-chercheur .	39
II- Politiques de mobilité géographique et comportement collaboratif du chercheur : interaction entre les enjeux politiques et individuels	41
A) Les différentes mesures politiques à la mobilité scientifique	41
1- La création d'une communauté de chercheurs européens	41
2- La France et ses programmes internationaux non européens	44
3- Autonomie du programme de mobilité universitaire	46
B) Incitations politiques de la mobilité : effets attendus à l'échelle universitaire et territoriale.....	48
1- Approche néoclassique : la mobilité en termes de réallocation des connaissances	49
3- Synthèse : deux conceptions politiques de la mobilité	55
C) La mobilité scientifique : quels effets sur l'activité scientifique et la géographie des collaborations ?	57
1- Débats autour du recours de la mobilité géographique comme indicateur de performance ...	58
2- Débats sur l'articulation entre la mobilité géographique et la géographie des collaborations	63
III- Les dynamiques relationnelles au sein des réseaux de co-authorship	65
A) Réseau scientifique : de l'outil managérial à l'analyse du réseau de co-authorship	65
1- Le réseau est un outil managérial en sciences sociales	65
2- Taille du réseau de co-authorship et capital social du chercheur.....	66

3- Force des liens collaboratifs et positionnement dans le réseau.....	68
4- Réseau de co-authorship et cohésion inter-individuelle	70
5- Synthèse : Conception du capital social et structure du réseau de co-authorship.....	72
B) Analyses empiriques de la structure du réseau de co-authorship.....	72
1- Réseaux complets et performance scientifique.....	72
2- Méthodologie commune	77
C- Dynamiques de collaboration au fil de la carrière.....	78
IV- Synthèse de la revue de littérature.....	80

Chapitre 2 : Présentation de la base de données.....82

I- Création de la base de données et choix méthodologiques82

A) Choix de l'échantillon.....	83
1- Etude de cas de la recherche en chimie	83
2- Etude de cas de deux laboratoires.....	92
3- Choix de l'échantillon.....	99
B) Construction de la base de données	100
1- Extraction des publications et la création des paires de co-authorship.....	100
2- Codage et informations supplémentaires	102
3- Regroupement en cohortes.....	105

II- Statistiques descriptives de la base de données..... 107

A- Description de la population étudiée.....	107
B) Description d'alter et tendances collaboratives.....	109
C) Description de la géographie des réseaux	110

III- Réseaux personnels et temporalité de carrière 116

A) Réseaux de co-authorship par année d'activité : Approche de panel.....	116
B) Réseaux de co-authorship par période d'activité : Approche de l'analyse de réseaux	120
1- Segmentation des carrières	120
2- Lecture graphique des réseaux ego-centrés	124
3- Statistiques descriptives.....	126

IV- Conclusion..... 132

Chapitre 3 : Carrières des chercheurs et dynamique des réseaux scientifiques : une mise en intrigue à partir des cas de cinq chimistes émérites 135

I- Contribution méthodologique : articulation des échelles d'analyse et des matériaux empiriques..... 135

A) Combiner les échelles d'action et de temps : pour une analyse processuelle des carrières scientifiques	136
B) Combiner données bibliométriques et matériaux qualitatifs : une analyse relationnelle de l'activité de publication.....	137

1- Les paires de co-authorship issues de la base de données bibliométrique	137
2- Les Curriculum Vitae.....	138
3- Les entretiens semi-directifs	138
II - Le cas de cinq chimistes émérites de l'Université de Poitiers	139
A- Constitution de l'échantillon.....	139
B) Statistiques descriptives de l'activité de publication et de collaboration	141
1- Statistiques descriptives statiques	141
2- Statistiques descriptives dynamiques.....	142
III- Résultats : contextualisation des relations collaboratives	145
A- Les années 1960-1970 : L'« âge d'or » de la recherche et l'insertion dans les réseaux scientifiques des « patrons»	145
1- L'émergence de la chimie à Poitiers.....	145
2- Une carte française de la chimie très lisible : Poitiers bien encadrée dans les réseaux.....	146
B) Les années 1980-1990 : Autonomisation scientifique et multiplication des collaborations	147
1- La « course à la publication »	147
2- Construction de l'équipe de recherche en tant qu'animateur d'un collectif de recherche	147
3- La mobilité scientifique et à l'internationalisation des réseaux	148
C) A partir des années 2000 : rationalisation et fusion des laboratoires	150
IV-Conclusion.....	150

Chapitre 4 : Est-ce que la mobilité géographique façonne le réseau de co-authorship ? Une approche de panel des carrières scientifiques..... 153

I- Intégration de la mobilité géographique dans la trajectoire du chercheur	153
A) Enjeux scientifiques de la mobilité et de la diversification du réseau	154
B) Evaluation de la mobilité dans notre base de données	155
C) Premiers résultats : la mobilité ne semble pas nécessaire à la diversification du réseau	156
II- Présentation du modèle économétrique.....	159
III- Résultats	163
IV- Conclusion et implications politiques.....	166

Chapitre 5 : « Existe-t-il plusieurs façons de collaborer au fil de la carrière ? Une analyse de réseaux sociaux de co-authorship des chimistes »..... 168

I- Structures des réseaux et types de collaboration	169
A) Indicateurs des réseaux ego-centrés.....	169
B) Analyse de composantes principales.....	172
1- Description des variables actives	172
2 – Application de l'Analyse de composantes principales.....	174
3- Représentation graphique des réseaux les plus contributifs.....	177

C) Description de la méthodologie de la classification hiérarchique ascendante	178
II- Caractérisation des quatre structures de réseau.....	180
A) Caractérisation des classes.....	180
1- Caractérisation de la classe 1 : Réseaux à très forte densité	180
2- Caractérisation de la classe 2 : Réseaux éclatés.....	182
3- Caractérisation de la classe 3 : Densité et diversité des relations	185
4- Caractérisation de la classe 4 : Réseaux gravitaires et dualité des relations	187
B) Le processus de classification	189
1- Contextualisation de la création des classes 3 et 4	191
2- Contextualisation de la création de la classe 2.....	192
3- Contextualisation de la création de la classe 1.....	192
III- Styles de collaboration et carrière scientifique.....	195
A) Statistiques descriptives des quatre classes.....	195
1- Structure du réseau et caractéristiques d'ego.....	196
2- Structure du réseau et diversité des alter.....	197
B) Estimation du style collaboratif en fonction de la temporalité de carrière et de la description du co-authorship.....	199
1- Présentation des modèles	200
2- Les résultats	205
C) Estimation du style collaboratif en fonction de la dynamique de carrière	211
1- Statistiques descriptives des dynamiques relationnelles des réseaux	211
2- Présentation des modèles	217
3- Les résultats	222
IV-Conclusion.....	229
Conclusion générale	233
I- Caractériser la dynamique des collaborations scientifiques grâce à des données individuelles et longitudinales.....	233
A) Contextualiser l'analyse des collaborations scientifiques.....	233
B) La création d'une base de données bibliométrique originale.....	234
C) Des méthodes d'analyse variées et complémentaires	235
II- Compréhension du comportement collaboratif des chercheurs au fil de la carrière	236
A) Appréhension de la carrière par année de publication : Contextualisation des dynamiques sociales et géographiques collaboratives.....	236
1- Explication de la dynamique collaborative en fonction la composition de l'équipe de publication.....	236
2- Explication de la dynamique collaborative via la contextualisation des collaborations	236
3- Explication de la dynamique collaborative via la mobilité géographique du chercheur	237
B) Création d'une typologie de styles collaboratifs	238

III- Limites et perspectives de recherche possible	239
A) Elargir l'échantillon	239
B) Multiplier la mise en intrigue de carrière sur différents groupes de l'échantillon	240
C) Intégrer des données micro complémentaires	240
D) Qualifier les thématiques scientifiques de l'activité de publication.....	241
Annexes.....	242
Annexe 1 : Sections CNRS de la chimie.....	242
Annexe 2 : Indicateurs des réseaux ego-centrés.....	243
A) La position des alter dans le réseau : Quelles sont les différentes relations de collaboration qu'ego entretient avec ses co-auteurs ?.....	243
1- Relations « duos » entre <i>ego</i> et <i>alter</i> : Quelles sont les différentes relations d' <i>ego</i> avec chacun de ses <i>alter</i> ?.....	244
2- Relations « trios » entre <i>ego</i> et chaque couple d' <i>alter</i> : Dans quelle mesure certains <i>alter</i> sont-ils plus structurants que d'autres dans les relations collaboratives qu' <i>ego</i> a avec son réseau ?	246
B) L'organisation des liens collaboratifs	248
1- Comment <i>ego</i> organise-t-il ses liens collaboratifs à l'échelle globale et locale ?.....	248
2- Comment <i>ego</i> organise-t-il ses liens collaboratifs autour de ses sous-espaces collaboratifs ?	251
Annexe 3 : Statistiques descriptives des 13 variables actives centrées et réduites	256
Annexe 4 : Choix du nombre de valeurs propres.....	256
Annexe 5 : Illustration graphique des deux axes de l'ACP	259
Annexe 6 : Processus de classification	261
Annexe 7 : Critères de classification	261
A) 1 ^{er} critère : Favoriser la proximité des individus au sein des classes	261
B) 2 ^{ème} critère : Favoriser une distance entre les réseaux les différentes classes différentes	264
Annexe 8 : Variables les plus contributrices de la CAH	267
Annexe 9 : Exemple de calculs des indicateurs de trajectoire structurelle du réseau	268
Bibliographie.....	271
Table des matières	292
Table des tableaux.....	297
Table des figures.....	300

Table des tableaux

Tableau 1 : Organisation de la thèse.....	22
Tableau 2 : Applications empiriques des théories économiques relatifs à la carrière scientifique	35
Tableau 3 : Programmes nationaux de mobilité de personnel d'enseignement supérieur.....	45
Tableau 4 : Synthèse des conceptions néoclassique et évolutionniste de la mobilité géographique.....	55
Tableau 5 : Articles traitant de la corrélation entre la mobilité et la performance scientifique individuelle	61
Tableau 6 : Formes de capital selon Bourdieu	67
Tableau 7 : Conception du capital social et structure du réseau de co-authorship	72
Tableau 8 : Articles portant sur la performance scientifique et la structure de réseaux de co-authorship complets	74
Tableau 9 : Méthodologie de 10 articles qui mobilisent des réseaux complets <i>dans l'évaluation de la performance scientifique</i>	77
Tableau 10 : Positionnement de la France dans la recherche en chimie	86
Tableau 11 : Nombre moyen de co-auteurs par publication par discipline entre 2000 et 2004 et 2012 et 2016.....	90
Tableau 12 : Equipes de l'IC2MP	93
Tableau 13 : Equipes du LCC	94
Tableau 14 : Description de l'échantillon	99
Tableau 15 : Données supplémentaires à l'échelle individuelle.....	102
Tableau 16 : Composition des cohortes de génération des chercheurs	106
Tableau 17 : <i>Activité de publication et tendances collaboratives selon le genre et la position</i>	108
Tableau 18 : <i>Activité de publication et tendances collaboratives selon la génération</i>	109
Tableau 19 : Tendances collaboratives en fonction du grade et du rôle d'alter dans la carrière d'ego	111
Tableau 21 : Description des stades de carrière	121
Tableau 22 : Distribution des réseaux selon la génération et le stade de carrière du chercheur.....	123
Tableau 23 : Statistiques descriptives des réseaux	129
Tableau 24 : Dispersion des tailles de réseaux (selon le nombre d' <i>alter</i>).....	131
Tableau 25 : Description des trois études empiriques réalisées à partir de la base de données	134
Tableau 26 : Statistiques descriptives sur la carrière et le co-authorship	141
Tableau 27 : Géographie des relations <i>ego-alter</i>	158

Tableau 28 : Variables dépendantes	160
Tableau 29 : Variables explicatives relatives à la mobilité	161
Tableau 30 : Synthèse des modèles estimés	161
Tableau 31 : Statistiques descriptives des variables dépendantes, explicatives et de contrôle	162
Tableau 32 : Estimations de panel (impact de la mobilité sur l'entrée et la sortie d'alter dans le réseau d'ego)	164
Tableau 33 : Synthèse des treize indicateurs de réseaux <i>ego</i> -centrés : position des <i>alter</i> dans le réseau	170
Tableau 34 : Synthèse des treize indicateurs de réseaux <i>ego</i> -centrés : structure des liens collaboratifs	171
Tableau 35 : Statistiques descriptives des 13 variables actives.....	172
Tableau 36 : Variables contributrices à la caractérisation de la dimension 1.....	174
Tableau 37 : Variables contributrices à la caractérisation de la dimension 2.....	176
Tableau 38 : Caractérisation de la classe 1.....	180
Tableau 39 : Caractérisation de la classe 2.....	182
Tableau 40 : Caractérisation de la classe 3.....	185
Tableau 41 : Caractérisation de la classe 4.....	187
Tableau 42 : Répartition des 259 réseaux en fonction de leur structure et des caractéristiques d' <i>ego</i>	196
Tableau 43 : Diversification sociale et géographique des collaborations au sein de chaque réseau ...	198
Tableau 44 : Variables dépendantes : Classej	200
Tableau 45 : Rapports de corrélation élevés entre la description du co-authorship et le stade de carrière	202
Tableau 46 : Synthèse des modèles estimés	203
Tableau 47 : Statistiques descriptives des variables.....	204
Tableau 48 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 1 ou 2 en %	209
Tableau 49 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 3 ou 4 en %	210
Tableau 50 : Distribution des 179 réseaux au sein des 16 dynamiques structurelles possibles	211
Tableau 51 : Structure du réseau au stade de chercheur confirmé	213
Tableau 52 : Structure du réseau au stade de senior.....	214
Tableau 53 : Structure du réseau au stade de senior avancé.....	216
Tableau 54 : Variables dépendantes Classe-dynamiquej	217
Tableau 55 : Synthèse des modèles estimés	219
Tableau 56 : Statistiques descriptives des variables.....	221
Tableau 57 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 1 en %.....	225
Tableau 58 : Effets marginaux associés aux probabilités d'avoir un réseau de la classe 2 en %.....	226

Tableau 59 : Effets marginaux associés aux probabilités d’avoir un réseau de la classe 3 en %.....	227
Tableau 60 : Effets marginaux associés aux probabilités d’avoir un réseau de la classe 4 en %.....	228
Tableau 61 : Classe 1 : Réseaux « très denses ».....	230
Tableau 62 : Classe 2 : Réseaux « éclatés ».....	231
Tableau 63 : Classe 3 : Réseaux « denses aux relations individuelles ».....	231
Tableau 64 : Classe 4 : Réseaux « gravitaires ».....	232
Tableau 65 : Sections CNRS de la chimie.....	242
Tableau 66 : Statistiques descriptives des sous-espaces collaboratifs.....	253
Tableau 67 : Statistiques descriptives des 13 variables actives centrées et réduites.....	256
Tableau 68 : Part de l’inertie expliquée par chaque axe.....	256
Tableau 69 : Contribution des variables dans la classification hiérarchique ascendante.....	267
Tableau 70 : Dynamique structurelle du réseau de Paul au fil de sa carrière.....	269

Table des figures

Figure 1 : Distribution rang-fréquences des formes lexicales	26
Figure 2 : Arbre maximum de l'analyse de similitude (indice de co-occurrence) des 50 formes les plus récurrentes dans le corpus	28
Figure 3 : Dendrogramme issu de la classification hiérarchique descendante de Reinter.....	29
Figure 4 : Illustration des liens faibles et forts	69
Figure 5 : Illustration de trous structuraux	70
Figure 6 : Synthèse du chapitre 1 : revue de la littérature sur la carrière scientifique et le comportement collaboratif	81
Figure 7 : Indices de spécialisation et d'activité de la chimie.....	88
Figure 8 : Impact de citation normalisé des publications et revues.....	89
Figure 9 : Distribution des publications de France selon la géographie de l'équipe et la discipline, 2016	91
Figure 10 : Impact des co-publications françaises par type de collaboration et discipline, 2014	91
Figure 11 : Carte de la France et laboratoires IC2MP et LCC	92
Figure 12 : Composition de l'effectif des deux laboratoires (Annuaire de Juin 2020).....	98
Figure 13 : Composition moyenne de chaque équipe de publication par laboratoire	99
Figure 15 : Réseau complet des publications réalisées par les 42 chercheurs qui étaient affiliés au LCC en 2018.....	101
Figure 14 : Réseau complet des publications réalisées par les 38 chercheurs qui étaient affiliés à l'IC2MP en 2018.....	101
Figure 16 : Allocation des localisations géographiques à l'échelle micro	104
Figure 17 : Regroupement en cohortes.....	106
Figure 18 : Distribution des collaborations fractionnées parmi les villes de la France métropolitaine (Hormis Poitiers et Toulouse)	114
Figure 19 : Distribution des collaborations internationales (en compte fractionné)	115
Figure 20 : Nombre moyen de publications annuelles par génération au fil des années (a) et des stades de carrière (b)	117
Figure 21 : Nombre moyen de co-auteurs par article par génération au fil des années (a) et des stades de carrière (b).....	118
Figure 22 : Décomposition sociale (a) et géographique (b) de l'équipe de publication au fil de la carrière	119

Figure 23 : Activité de publication au fil des stades de carrière.....	121
Figure 24 : Segmentation des carrières	122
Figure 25 : Réseau personnel dans lequel <i>ego</i> est représenté.....	124
Figure 26 : Réseau personnel ego-centré.....	124
Figure 27 : Lecture du réseau ego-centré	126
Figure 28 : Trajectoire professionnelle des chercheurs	140
Figure 29 : Renouvellement du co-auteursip	143
Figure 30 : Evolution de la géographie du co-authorship	144
Figure 31 : Dans quelle mesure le contexte historique de recherche et le stade de carrière impactent les dynamiques collaboratives ?	152
Figure 32 : Volume des évènements de mobilité par chercheur et durant la carrière	156
Figure 33 : Structure des réseaux et axes de l'ACP	178
Figure 34 : Perte relative d'inertie selon le nombre de classes	179
Figure 35 : Distribution graphique de la classe 1	181
Figure 36 : Réseaux caractéristiques de la classe 1 (les plus éloignés du centre de gravité des autres classes)	182
Figure 37 : Distribution graphique des classes 1 et 2.....	183
Figure 38 : Réseaux caractéristiques de la classe 2 (les plus éloignés du centre de gravité des autres classes)	184
Figure 39 : Distribution graphique des classes 1,2 et 3	185
Figure 40 : Réseaux caractéristiques de la classe 3 (les plus éloignés du centre de gravité des autres classes)	186
Figure 41 : Distribution graphique des classes 1,2,3 et 4	187
Figure 42 : Réseaux caractéristiques de la classe 4 (les plus éloignés du centre de gravité des autres classes)	188
Figure 43 : Matrice synthétique de caractérisation des classes	189
Figure 44 : Recontextualisation des créations de classe dans le processus de classification	190
Figure 45 : Reconstruction du dendrogramme et du processus de classification	194
Figure 46 : Dynamiques structurelles des réseaux au stade de chercheur confirmé	213
Figure 47 : Dynamiques structurelles des réseaux au stade de senior.....	215
Figure 48 : Dynamiques structurelles des réseaux au stade de senior avancé.....	216
Figure 49 : Typologie des styles de collaboration.....	239
Figure 50 : Représentation graphique de l'équivalence structurelle	244
Figure 51 : Représentation graphique de la transitivité locale	250
Figure 52 : Taille moyenne des sous-espaces collaboratifs.....	254

Figure 53 : Décomposition de l'inertie totale	257
Figure 54 : Individus de l'ACP	257
Figure 55 : Variables de l'ACP.....	258
Figure 56 : Illustration graphique des 5 réseaux qui contribuent le plus à la dimension 1.....	259
Figure 57 : Illustration graphique des 5 réseaux qui contribuent le plus à la dimension 2.....	260
Figure 58 : Dendrogrammes et nuages de points : choix de la classification.....	261
Figure 59 : Représentation graphique de cinq réseaux distincts avant la classification.....	262
Figure 60 : Représentation graphique de cinq réseaux distincts après la classification	263
Figure 61 : Exemple : dendrogramme de la classification hiérarchique des cinq réseaux	264
Figure 62 : Exemple de la trajectoire structurelle du réseau de co-authorship de l'individu "Paul" ...	268
Figure 63 : Dynamiques structurelles des réseaux de Paul	270

Dynamiques sociales et spatiales des collaborations scientifiques : une étude de cas de chercheurs en chimie

Résumé : Ce travail vise à contribuer à la compréhension de la dynamique de collaboration des chimistes. Nous appréhendons le comportement collaboratif du chercheur comme non singulier et nous nous attachons à décrire sa multiplicité et son évolution au fil de la carrière scientifique. A partir d'une base de données bibliométrique originale, nous identifions et expliquons les dynamiques collaboratives via l'utilisation de données contextuelles, géographiques et sociales. Alors que beaucoup de travaux se positionnent sur l'étude de la compétitivité du chercheur au travers d'études macro, nous affinons notre perception de la carrière à partir de données micro et longitudinales dans une démarche descriptive et compréhensive. A partir de trois chapitres empiriques issus de diverses méthodologies (méthodologie mixte, économétrie des données de panel, analyse de réseaux sociaux, analyse de données), nous illustrons différentes dynamiques collaboratives reposant sur des relations sociales et des géographies particulières.

Mots clefs : co-authorship – réseaux - trajectoire – carrière scientifique – mobilité géographique

Social and geographical patterns of scientific collaborations: a case study of chemists

Abstract: This work aims to better understand the collaborative patterns of chemists. We look at the collaborative behavior of the researcher as not singular and we attach to describe its diversity and its evolution along the scientific career. From an original bibliometric database, we identify and explain collaborative dynamics using contextual, geographical and social data. While many articles focus on the researcher's competitiveness across macro studies, we improve our perception of career from individual and longitudinal data in a descriptive and comprehensive approach. Along three empirical chapters, using different methodologies (mixed methodology, econometrics of panel data, social network analysis, data analysis), we illustrate a variety of collaborative dynamics based on social and geographic relationships.

Key words: co-authorship – network - trajectory – scientific career – Geographical mobility