
Analyse exploratoire de réseaux d'inventeurs

L'impact de la collaboration sur la valeur des brevets canadiens en nanotechnologie

Ahmad Barirani, Catherine Beaudry, Bruno Agard

*École Polytechnique de Montréal
CP 6079, succ. Centre-ville, Montréal (Québec), H3C 3A7, Canada
bruno.agard@polymtl.ca*

RÉSUMÉ. L'analyse des réseaux d'innovation permet d'observer les interactions entre les innovateurs dans leur milieu de pratique et de voir comment différentes stratégies de collaboration ont un impact sur le rendement. Cet article propose une nouvelle méthode d'analyse exploratoire des réseaux d'inventeurs. Notre méthodologie consiste d'abord à extraire un ensemble de brevets à l'aide d'une liste de mots clés, puis de regrouper les inventeurs suivant la similitude dans leur stratégie de collaboration pour ensuite comparer le rendement de chaque groupe trouvé. La méthode sera appliquée au cas des inventeurs canadiens de brevets en nanotechnologie. Nous pourrions ainsi identifier les types de collaborations qui sont les plus productives au sein de la communauté d'inventeurs canadiens en nanotechnologie.

ABSTRACT. Innovation network analysis allows us to observe interactions between innovators within their environment of practice and see how different collaboration patterns have an impact on performance. This article proposes a new method for the exploratory analysis of inventor networks. Our methodology first consists in extracting a set of patents with a list of keywords, clustering inventors according to the similarity in their collaboration strategy and then comparing performance for each cluster. This method will be applied to the case of Canadian nanotechnology patent inventors. By doing so, we will be able to identify the most productive types of collaboration within the Canadian nanotechnology inventors' community.

MOTS-CLÉS : gestion de l'innovation, réseaux d'inventions, réseaux sociaux, exploration de texte, classification.

KEYWORDS: innovation management, invention networks, social network, text mining, clustering.

DOI:10.3166/JESA.46.855-875 © Lavoisier 2012

1. Introduction

À l'ère de la globalisation, un nouveau paysage économique marqué par l'hyper-compétitivité est en train de se former. Dans un premier temps, de grands pas dans le domaine de l'ingénierie industrielle et de la logistique ont offert la possibilité aux entreprises manufacturières de déplacer leurs activités de production dans des pays où la main d'œuvre était meilleur marché. Ainsi, après une longue période de rationalisation industrielle et de montée en compétence des fournisseurs, nous sommes face à une situation où ces pays jadis en voie de développement sont aujourd'hui des chefs de file dans plusieurs domaines technologiques de pointe. Cette nouvelle réalité a une incidence importante sur l'organisation industrielle des pays avancés qui sont forcés de développer des produits et services innovants à un rythme plus important pour assurer leur développement économique (Tomala et Sénéchal, 2002 ; Trentesaux, 2007).

Il devient ainsi important de revoir certaines pratiques en matière de gestion de l'innovation et de mettre en place des politiques qui favorisent l'innovation au sein des communautés scientifique et technologique des pays avancés. Les entreprises doivent développer des techniques qui leur permettent de comprendre leur environnement social et de développer leur ressources humaines (Addouche *et al.*, 2006 ; Gruat la Forme *et al.*, 2007). Dans cet article, nous proposons d'atteindre cet objectif en analysant les réseaux d'inventeurs canadiens en nanotechnologie afin de trouver les profils sociaux qui mènent à de meilleurs rendements en termes de production de brevets dans ce secteur.

Dans cet article, nous proposons une méthodologie qui permettra de caractériser un réseau social d'inventeurs, d'effectuer une classification des inventeurs sur la base de la similitude de leur stratégie de collaboration et de trouver les profils sociaux qui sont les plus productifs sur le nombre de brevets produits. Cette méthodologie sera appliquée au cas des réseaux d'inventeurs canadiens en nanotechnologie pour lesquels nous expliquerons les résultats.

Dans ce qui suit, nous présentons une revue de la littérature (section 2), nos questions et hypothèses de recherche (section 3), la méthodologie employée (section 4), nos résultats (section 5) et la conclusion de notre travail de recherche (section 6).

2. Revue de la littérature

En lien avec nos objectifs de recherche, nous effectuons un bref survol de la littérature qui, d'une part, soulève les questions théoriques se rapportant aux aspects collaboratifs de la gestion de l'innovation, et d'autre part, les techniques disponibles pour mesurer le progrès technologique et caractériser des réseaux de collaboration.

2.1. Gestion de l'innovation, collaboration et capital social

La littérature concernant la gestion de l'innovation peut être divisée en deux grandes écoles. La première voit la capacité d'innovation comme une caractéristique

intrinsèque de tout agent économique (Amin et Cohendet, 2004). Il s'agit là de l'approche de gestion stratégique où les gestionnaires sont les seules entités pensantes de la firme. Ceux-ci se basent sur les compétences clés de la firme pour prendre des décisions sur les directions stratégiques à prendre. Le rôle des gestionnaires revient alors à mettre en place des structures qui permettent à l'entreprise de protéger et de renforcer ses compétences. Par conséquent, les politiques nationales qui se rapportent à la favorisation de l'innovation sont limitées à l'identification et au support financier de quelques innovateurs potentiels ou importants. Par ailleurs, la deuxième école de pensée se préoccupe plutôt de ce que l'environnement a à offrir aux agents économiques (Amin et Cohendet, 2004). Ici, la capacité d'innovation d'un agent économique est fonction des connaissances que les autres agents ont à lui offrir. Il s'agit de reconnaître l'importance des externalités réseaux dans le processus de création et de diffusion de l'innovation. Ainsi, la principale qualité d'un innovateur est dans sa capacité d'absorption du savoir qui lui vient de l'extérieur (Cohen et Levinthal, 1990). Selon cette perspective, les réseaux d'innovations sont un lieu social où les membres développent, par le biais de la pratique, des habiletés qu'ils partagent avec les autres membres de la *communauté de pratique* (Brown et Duguid, 1991). Au fur et à mesure que les innovateurs accumulent un capital de compétences, ils accèdent à des positions de plus en plus centrales au sein du réseau. Ils passent ainsi d'un stade de nouveaux entrants à un stade de vétérans. Ce passage requiert toutefois que les membres innovent. S'ils n'apportent aucune innovation à la communauté, ils continuent d'occuper les positions périphériques du réseau, positions qu'ils peuvent même perdre au profit d'autres nouveaux entrants. Ils sont alors rejetés du réseau.

Naturellement, les politiques qui découlent de ce mode de réflexion tentent de mettre en place des mécanismes qui favorisent la collaboration et le transfert de connaissances entre les agents innovants. Récemment, cette voie théorique a reçu beaucoup d'attention dans le milieu académique (Burt, 2004 ; Adler et Kwon, 2002 ; Inkpen et Tsang, 2005). En effet, les liens de collaboration ne sont pas distribués de manière uniforme au sein d'une communauté de pratique. L'ensemble des liens qu'un individu possède peut être perçu comme un *capital social* (Coleman, 1988 ; Nahapiet et Ghoshal, 1998). Alors que de nombreuses études montrent qu'il y a un lien entre le capital social d'un innovateur et sa performance (Adler et Kwon, 2001 ; Perry-Smith and Shalley, 2003 ; Cantner et Graf, 2006 ; Inkpen et Tsang, 2005), nous savons peu sur comment le capital social se développe et se manifeste en rendement. De plus, un débat a lieu autour de l'importance de la spécialisation et de l'intermédiation comme étant deux stratégies de collaboration distinctes. Le camp associé à la spécialisation stipule que la proximité sociale est associée à une meilleure compréhension mutuelle entre les collaborateurs, ce qui facilite l'échange de connaissances tacites (Coleman, 1988 ; Adler et Kwon, 2001 ; Rost, 2011). Il propose habituellement des stratégies en faveur de l'*exploitation* sociale, tel que le partenariat répété et entre acteurs ayant une similarité cognitive (Powell *et al.*, 1996 ; Uzzi, 1997 ; Hansen, 1999). L'autre camp stipule que la diversification et l'hétérogénéité des liens sociaux donne lieu à des possibilités de recombinaisons nouvelles de connaissances, ce qui génère des innovations radicales (Granovetter, 1973 ; Burt, 2004 ; Felin et Hesterly, 2007 ; Fleming *et al.*, 2007). Il prône plutôt

l'exploration sociale tels que les partenariats interdépartementaux ou interindustriels (Ahuja et Lampert, 2001 ; Nootboom *et al.*, 2007).

2.2. Indicateurs de valeur de brevets

Un bureau de brevet octroie un brevet à un cessionnaire lorsqu'une invention créée par ce dernier semble utile, non évidente et constitue une nouveauté. Ces caractéristiques font en sorte que les brevets peuvent être utilisés comme indicateurs d'activités innovantes (Griliches, 1990).

Toutefois, les brevets ne sont pas représentatifs de toute la production innovante d'un pays. En effet, il existe d'autres méthodes de protection de la propriété intellectuelle. De plus, le brevet n'est pas toujours synonyme d'innovation. De nombreux brevets ne se manifestent jamais sous forme de produits ou ne génèrent pas de bénéfices pour les firmes cessionnaires (Allison *et al.*, 2004 ; Lemley et Shapiro, 2005 ; Pénin, 2005). Afin de pouvoir discriminer entre les brevets de qualité et de moins bonne qualité, la littérature s'est penchée sur l'élaboration de nombreuses métriques indicatrices de l'importance économique des brevets (Lanjouw et Schankerman, 2004).

Pour qu'un brevet soit valide aux États-Unis, le cessionnaire se doit de payer des frais de renouvellement du brevet après la quatrième, huitième et douzième année. Le renouvellement d'un brevet peut être perçu comme un signal envoyé par le cessionnaire du brevet concernant l'importance qu'il accorde à l'invention (Lanjouw *et al.*, 1998). Cet intérêt peut être dû au fait que le cessionnaire prévoit de futures opportunités de commercialisation pour l'invention, le maintien des objectifs stratégiques tels que le blocage, mais aussi parce que celui-ci perçoit la technologie protégée par le brevet comme étant une voie de développement prometteuse. Le renouvellement d'un brevet peut donc être représentatif de la valeur privée d'un brevet (Baron et Delcamp, 2012).

Aux États-Unis, le nombre de revendications démarquent la portée légale d'un brevet et peuvent servir de mesure de qualité. Les déposants ont donc un incitatif à revendiquer le plus possible, alors que les examinateurs doivent s'assurer que toutes les revendications sont justifiées (Lanjouw et Schankerman, 2004). Tong et Frame (1994) montrent qu'un nombre plus grand de revendications est relié à une performance économique plus importante des brevets.

Une demande de brevet doit citer toutes les sources de connaissances qui sont reliées à l'invention. Le déposant n'est pas tenu d'effectuer un état de l'art antérieur, mais l'examineur du brevet doit s'assurer que toutes les sources pertinentes soient citées. De ce point de vue, un brevet qui reçoit des citations peut être vu comme étant une invention qui limite la portée des brevets qui le citent (Meyer, 2000 ; Shane, 2001). D'un autre point de vue, Trajtenberg (1990) voit la citation comme une indication de *retombées de savoir* générées par un brevet. De nombreuses études utilisent d'ailleurs le nombre de citations reçues par un brevet pour mesurer sa valeur économique (Harhoff *et al.*, 1999 ; Lanjouw et Schankerman, 2004 ; Shane, 2001). Hall *et al.* (2005) associent le nombre de citations reçues par un brevet à la

valeur marchande des actions de l'entreprise. Puisqu'un brevet peut être cité sans qu'il soit renouvelé, les citations peuvent représenter la valeur sociale d'un brevet. Dans ce cas, la valeur économique exprimée par les citations a une signification plutôt sociale que privée (Baron et Delcamp, 2012).

2.3. Analyse de réseaux sociaux

De nombreux projets de recherche utilisent les techniques d'analyse de réseaux sociaux pour étudier l'impact des interactions entre inventeurs sur le rendement de ceux-ci en termes d'innovation (Newman, 2003 ; Cowan et Jonard, 2003 ; Cantner et Graf, 2006 ; Fleming *et al.*, 2007 ; Rost, 2011).

Wasserman et Faust (1994) fournissent une description exhaustive des éléments requis lors de l'analyse de réseaux sociaux. Un réseau est défini par une paire d'ensembles $G = \{P, E\}$ où P est un ensemble de N nœuds P_1, P_2, \dots, P_n et E est un ensemble de liens qui connectent deux éléments de P .

Chaque nœud a un degré de distribution qui est défini par le nombre de liens qui le relient aux autres nœuds du réseau. Le nombre de liens réseaux qui séparent deux nœuds i et j s'appelle la distance géodésique. Le plus court chemin est la distance géodésique la plus courte qui sépare deux nœuds i et j . Par exemple, nous appelons proximité (*closeness centrality*) la distance géodésique moyenne qui sépare un nœud i de tous les autres nœuds du réseau. Le degré d'intermédiation (*betweenness centrality*) d'un nœud i est défini par :

$$C_B(i) = \sum_{j \neq k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \quad (1)$$

où σ_{jk} est le nombre de plus courts chemins entre les nœuds j et k et $\sigma_{jk}(i)$ est le nombre de plus courts chemins entre j et k qui passent par i . Le coefficient de groupement (*clustering coefficient*) C_i d'un nœud i est défini par :

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (2)$$

où E_i représente le nombre de liens entre les k_i nœuds qui sont liés au nœud i . Cette métrique montre à quel point les nœuds reliés à un nœud i sont reliés entre eux. Un ensemble de nœuds qui sont tous liés entre eux est appelé une clique. Un sous-ensemble de réseau pour lequel il existe au moins un chemin entre tous les nœuds est appelé *composante*. Pour un graph G donné, la composante qui comprend le plus grand nombre de nœuds est appelée *composante principale*.

Afin d'illustrer les diverses métriques, nous utilisons l'exemple de réseau social présenté à la figure 1. Les nœuds 1, 2 et 3 font partie d'une clique. Les nœuds 1 et 3 ont un coefficient de groupement de 1,0, alors que pour le nœud 2 il est de 1/3. Les degrés d'intermédiation pour les nœuds 2 et 4 sont de 8 et de 9,5 respectivement puisque les nœuds 1 et 3 doivent 'passer' par ceux-ci pour atteindre les nœuds 5, 6

ou 7. Ces derniers sont des nœuds plutôt périphériques, ils se retrouvent moins souvent sur un plus court chemin et ont donc un degré d'intermédiation plus petit.

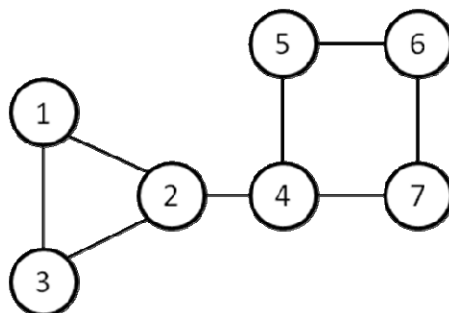


Figure 1. Réseau social comprenant 7 nœuds et 8 liens

Réseaux aléatoires et petits-mondes

Dans un réseau composé de n nœuds, il y a une possibilité totale de $n(n-1)/2$ liens entre les nœuds. Toutefois, les différents types de réseaux que nous trouvons dans la nature ne possèdent pas autant de liens entre les nœuds. Un réseau aléatoire est un réseau pour lequel les liens sont choisis aléatoirement parmi tous les nœuds possibles entre les n nœuds. En changeant la probabilité p avec laquelle deux nœuds peuvent être connectés, on peut obtenir une gamme de réseaux aléatoires qui se situe entre le réseau régulier ($p = 0$) et le réseau complet ($p = 1$).

La plupart des réseaux trouvés dans la nature sont de type petit-monde (*small-world*) (Watts et Strogatz, 1998), *i.e.* le degré de proximité moyen des nœuds du réseau est petit et le coefficient de groupement est plus grand que dans un réseau aléatoire.

Les petits-mondes sont reconnus pour favoriser le transfert de ce qui peut être transmis d'un nœud à l'autre dans le réseau. Ainsi, un réseau d'inventeur qui aurait les caractéristiques d'un petit-monde favoriserait la diffusion plus rapide des connaissances entre des inventeurs du réseau (Cowan et Jonard, 2003). De plus, dans un réseau d'inventeurs de type petit-monde, on observe que le niveau de connaissances agrégées des inventeurs est plus grand que dans les réseaux aléatoires.

2.4. Classification

Les techniques de classification consistent à regrouper un ensemble d'observations de façon à ce que les éléments similaires se retrouvent dans le même groupe (Berry et Linoff, 2004 ; Denimal et Boussu, 2002 ; Sbihi *et al.*, 2005 ; Pronzato, 2006). Il s'agit là d'une technique d'apprentissage non supervisée. En partant d'un réseau social, les techniques de classification peuvent donc être utilisées

pour regrouper les individus qui ont des caractéristiques semblables (Girvan et Newman, 2002).

Il existe plusieurs types de méthodes de classification. Toutes les méthodes basées sur la similarité ont en commun le fait de requérir une mesure de distance entre deux éléments. La distance euclidienne, la distance Manhattan et le coefficient de corrélation de Pearson en sont des exemples. La classification vise à ce que la distance intergroupe soit maximisée et que la distance intragroupe soit minimisée (Manning *et al.*, 2008). Les sections suivantes font état de l'état de l'art pour les trois méthodes de classification les plus répandues.

2.4.1. Classification hiérarchique

La classification hiérarchique regroupe les éléments d'observations sous une structure en arbre après un certain nombre d'itérations (Berry et Linoff, 2004). La classification peut se faire soit par agglomération (algorithmes CHA, CURE) ou par division (algorithmes DIANE, BIRCH). Dans les méthodes par agglomération, chaque observation constitue initialement un groupe. À chaque itération, les segments les plus proches sont fusionnés pour former de nouveaux segments élargis. Le processus se termine lorsqu'il n'y a plus qu'un seul groupe. Les méthodes par division, à l'inverse, partent avec un seul groupe qui contient tous les éléments. À chaque itération, les segments sont divisés de manière à ce que les éléments d'un groupe soient les plus distants possibles des éléments des autres segments. Le processus se termine lorsque tous les segments ne contiennent qu'un seul élément.

2.4.2. Classification par partitionnement

La classification par partitionnement (ex : *k*-means) regroupe les données selon un nombre fixe (*k*) de segments. Ce nombre de segments peut être prédéfini par l'utilisateur ou trouvé après quelques tentatives avec différents essais. Le processus de partitionnement commence en assignant un élément à un groupe. Cet élément deviendra le noyau du groupe. Ensuite, chaque élément restant est assigné à un groupe selon la distance de l'élément avec le noyau du groupe. Il y a alors une itération où un nouveau noyau est choisi pour chaque groupe. Les éléments sont réassignés à des segments toujours selon leur distance avec le noyau (Berry et Linoff, 2004). Le processus s'arrête après un nombre d'itérations prédéfini ou lorsque le processus converge vers un optimum local.

2.4.3. Classification par densité

La classification par densité diffère de la classification par partitionnement du fait qu'elle n'impose pas un nombre prédéfini de segments. Les segments sont plutôt définis par des régions qui comprennent un nombre suffisant d'éléments proches (Ester *et al.*, 1996 ; Ankerst *et al.*, 1999). Ces méthodes prennent comme paramètres d'entrée le seuil de similarité ϵ (la distance maximum au-delà de laquelle deux éléments ne peuvent être considérés comme étant proches) et le nombre d'éléments minimum pour former un groupe. Le processus de classification débute en choisissant un élément qui fera partie d'un groupe. Tous les éléments qui seront à

l'intérieur de ε seront affectés au groupe. De même, tous les éléments qui sont à une distance ε de tout élément du groupe y sont affectés. Lorsqu'il n'y a plus d'éléments à considérer pour ce groupe, un élément qui n'est pas assigné à un groupe est choisi pour faire partie d'un autre groupe. De la même façon, tous les éléments qui sont à une distance ε d'un des éléments de ce groupe en feront partie. Le processus se termine lorsque tous les éléments sont assignés à un groupe ou lorsqu'aucun des éléments non-assignés n'est en deçà d'une distance ε d'un groupe.

3. Objectif de recherche

L'objectif de cet article est de proposer une nouvelle méthode pour l'analyse exploratoire des réseaux d'inventeurs qui permettra d'explorer certains aspects du travail collaboratif dans les communautés technologiques. À la lumière de notre revue de la littérature, l'analyse exploratoire des réseaux d'invention peut permettre d'identifier différentes stratégies de collaboration et de voir comment les inventeurs qui adoptent une certaine stratégie performant comparativement aux autres.

4. Méthodologie

Notre méthode consiste à extraire les brevets pertinents à un domaine technologique, d'en monter les réseaux d'inventeurs et de trouver des stratégies de collaboration qui pourraient donner lieu à de meilleurs rendements d'innovation pour les inventeurs. Nous assumons ainsi que les liens de co-inventions sont le résultat de l'activité collaborative des inventeurs et façonnent ainsi le capital social de ceux-ci (Fleming *et al.* 2007). La méthodologie proposée compte quatre étapes telles qu'illustrées à la figure 2.

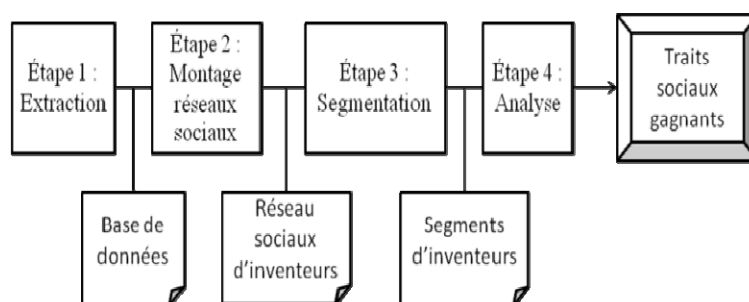


Figure 2. Méthodologie

4.1. Étapes successives

Étape 1. Dans un premier temps, nous procédons à l'extraction et au nettoyage des données provenant de bases de données de brevets de l'United States Patent and

Trademark Office¹. Les États-Unis représentent le marché technologique le plus important au monde et sont le principal partenaire économique du Canada. Ainsi, Li *et al.* (2007) observent que les entreprises canadiennes déposent plus souvent de brevets aux États-Unis qu'ils en déposent en Europe. Nous extrayons de cette base de données : les titre, résumé, dates d'application et d'obtention, nombre de réclamations, références, citations, nom, ville et pays des inventeurs ainsi que nom, ville et pays des cessionnaires pour chaque brevet d'intérêt. Les brevets sont sélectionnés en se basant sur une liste de mots clés, pour les brevets émis par un cessionnaire canadien. Nous utilisons la méthode d'extraction booléenne, ce qui signifie que les brevets extraits contiendront au moins l'un des mots clés apparaissant dans la liste.

Étape 2. En un second temps, nous montons et étudions la structure des réseaux sociaux d'inventeurs canadiens en nanotechnologie. À cet effet, le réseau social des inventeurs sera un réseau dans lequel les nœuds sont des inventeurs provenant de notre échantillon d'inventeurs extrait à l'étape 1 et dont les liens sont définis par la relation de co-invention identifiée par la liste des inventeurs de chaque brevet. Dans le cadre de la présente étude, nous définissons le capital social de chaque inventeur par son degré de distribution, d'intermédiation, de proximité et par le coefficient de groupement. L'objectif est de caractériser la stratégie de collaboration de chaque inventeur, de voir à quel point le réseau social d'inventeurs est fragmenté et à quel point il présente les caractéristiques d'un petit-monde.

Étape 3. La troisième étape de notre méthodologie consiste à effectuer une classification des inventeurs d'après la similitude de leur stratégie de collaboration en utilisant la méthode de classification par modèle de Fraley et Raftery (2002). Pour ce faire, nous sélectionnons la composante principale du réseau et nous utilisons les métriques de capital social de chaque inventeur comme attributs des éléments à segmenter. La plus importante composante est celle qui contient le plus grand nombre d'inventeurs. Travailler sur la plus importante composante permet dans un premier temps d'améliorer la précision de notre échantillon initial de brevets. En effet, notre méthode d'extraction basée sur les mots clés aura forcément un certain nombre de faux positifs (c'est-à-dire des brevets qui ne sont pas *réellement* des brevets de nanotechnologie mais qui possédaient tout de même un des mots clés utilisés lors de l'extraction). La probabilité que des *faux* inventeurs soient co-inventeurs avec des *vrais* inventeurs étant faible, la composante principale du réseau aura forcément une concentration plus grande de *vrais* inventeurs que l'échantillon initial. Dans un second temps, la taille d'un réseau ayant un impact sur sa complexité, il est judicieux de ne pas comparer les métriques réseaux pour des inventeurs provenant de deux composantes distinctes. Par exemple, un coefficient de groupement de 1 n'a pas la même signification dans une composante comprenant une centaine d'inventeurs que dans une autre comprenant trois inventeurs.

Étape 4. Finalement, une fois les groupes obtenus, nous procédons à la caractérisation de ceux-ci en effectuant le test en U (sommes des rangs) de Mann-

1. <http://uspto.gov>.

Whitney-Wilcoxon (Conover, 1980) afin de voir si les variables d'indicateurs de stratégie de collaboration proviennent des mêmes distributions d'un groupe à l'autre. Cette comparaison sera faite d'après la *cote-z* du test en U lors de la comparaison de deux groupes. Par exemple, un groupe qui aura un rang du coefficient de groupement significativement plus élevé que les autres contiendra principalement des spécialistes qui œuvrent au sein d'une clique. En effet, l'existence d'une clique indique qu'il y a une affinité suffisamment grande entre les membres de la clique. C'est cette affinité qui fait que les membres de la clique soient liés les uns aux autres (Granovetter, 1973). À l'inverse, un groupe qui aura un rang du degré d'intermédiation plus élevé que les autres sera composé d'inventeurs intermédiaires (Fleming *et al.* 2007). Par la suite, nous effectuons des tests en U sur les indicateurs de valeur de brevets. Les métriques utilisées seront le nombre de brevets par inventeur, le nombre minimum et maximum de citations obtenues pour les brevets de chaque inventeur au cours des 4 années suivant leur date d'octroi ainsi que le pourcentage de renouvellement, à la 4^e année, des brevets produits pour un inventeur.

4.2. Outils et techniques utilisés

Les brevets proviennent des bases de données de l'USPTO. Ce choix est motivé par l'importance que représente le marché américain sur le plan technologique et commercial. Pour effectuer notre recherche par mot clés et extraire les données, nous utilisons le logiciel PatentBot développé à l'École polytechnique de Montréal. Ce logiciel exécute l'interface web de l'USPTO pour exécuter les requêtes et extraire les données relatives aux brevets ciblés. De plus, nous avons eu accès à la base de données de l'Observatoire des sciences et des technologies (OST) pour obtenir l'information relative aux citations des brevets extraits par le biais de PatentBot.

Afin de caractériser ce réseau, nous allons utiliser le module R *igraph* (Csardi et Nepusz, 2006) en fournissant l'ensemble des liens entre les inventeurs obtenus à partir des données USPTO. Ce logiciel permet de tracer le graphe social d'inventeurs et de calculer les métriques que nous allons utiliser durant l'étape de classification. Ensuite, nous utilisons le module R *mclust* (Fraley *et al.*, 2012) pour employer la méthode de classification par modèle (Fraley et Raferty, 2002). Ce choix est motivé par le fait que nous ne voulons pas, a priori, imposer un nombre strict de profils sociaux dans lequel chaque inventeur devra se retrouver. Cette méthode permet de laisser libre court à toute la diversité qu'il peut y avoir dans la structure sociale des inventeurs.

5. Cas d'application

5.1. Contexte

Nous avons choisi le domaine des nanotechnologies pour effectuer notre analyse puisque c'est un domaine technologique multidisciplinaire en émergence dans lequel les différents sous-domaines ne sont pas encore tous en interaction. Par

nanotechnologie, nous entendons toute technologie basée sur la manipulation de la matière à l'échelle du μm ou du nm . Les industries en émergence sont reconnues pour connaître initialement des phases turbulentes où beaucoup d'innovations sont introduites mais où le niveau de réussite de ces innovations est très variable. L'industrie de la nanotechnologie est donc riche en données qui permettent de faire une bonne discrimination entre les inventeurs qui y évoluent, ce qui en fait un exemple approprié pour y appliquer les techniques de classification.

Finalement, les nanotechnologies sont généralement reconnues comme étant l'un des domaines de pointe les plus importants sur le plan technologique pour les années à venir. Comme il a été discuté précédemment dans cet article, la capacité des pays avancés à assurer leur compétitivité dans ce domaine est cruciale pour leur avenir technologique. Le Canada ne fait pas exception, ce qui nous motive à regarder de plus près la structure du réseau technologique canadien dans ce domaine.

5.2. Extraction des données

Nous avons utilisé la liste de mots clés apparaissant au tableau 1. Ces mots clés ont été obtenus à partir des mots clés utilisés dans au moins deux études bibliographiques antérieures concernant les nanotechnologies (Alexcar *et al.*, 2007 ; Fitzgibbons et McNiven, 2006 ; Mogoutov et Kahane, 2007 ; Porter *et al.*, 2008 ; Schmoch *et al.*, 2003 ; Zitt et Bassecouard, 2006). Ces mots clés ont été utilisés pour extraire les brevets déposés de 1976 à 2009 en nanotechnologies dont le cessionnaire ou l'un des inventeurs est canadien (a une adresse au Canada). Ainsi, les brevets octroyés à des firmes non canadiennes mais dans lesquels un inventeur canadien est nommé sont aussi considérés dans notre analyse.

Tableau 1. Mots clés nanotechnologies

Mots clés	Brevets	Mots clés	Brevets
atom* force microscop*	88	quantum	1 287
biosensor	231	scanning prob*	41
mesoporous material*	31	scanning probe microscop*	16
molecular beam epitaxy	95	self assem*	219
molecular switch	25	single electron*	85
nano*	4 568	supramolecular chemistry	18
nems	9	tunnel* microscop*	2
photonic*	969	Photonic*	969
polymer composite*	379	scanning prob*	41
polymer dna	10	single electron*	85
polymer rna	3		

* = signe générique représentant toute combinaison de caractères alphanumériques

La base de données de brevets extraits est ensuite nettoyée afin de corriger les duplicatas, les entrées incomplètes ou mal formatées. Après le nettoyage, sur les 8 076 brevets du tableau 1, 4 196 ont été retenus pour l'étape suivante. Ces brevets ont été obtenus entre 1990 à 2005 et ont été inventés par plus de 5 902 inventeurs, ce qui fait une moyenne de 1,41 inventeurs par brevet. Dans notre échantillon d'inventeurs, plus de 2 809 résident au Canada.

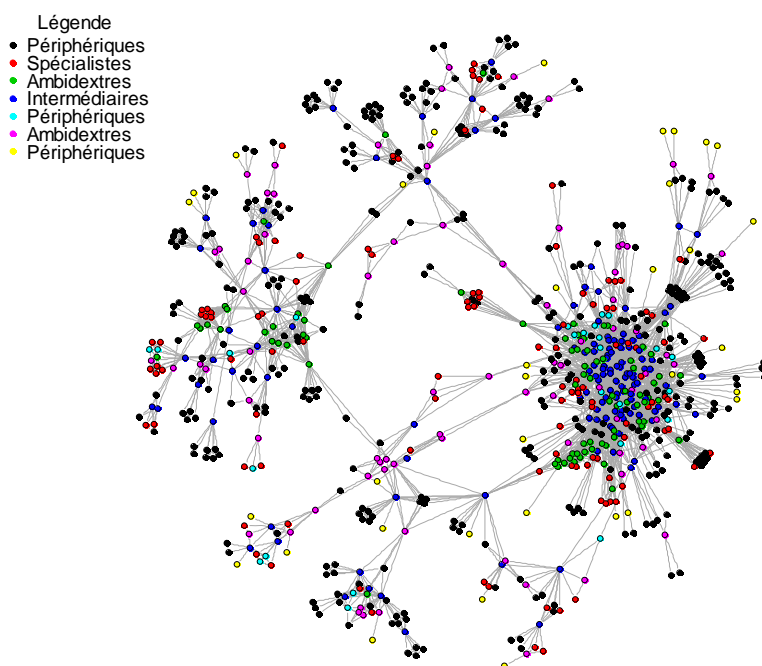


Figure 3. Graphe du réseau social des inventeurs canadiens en nanotechnologie. Les couleurs représentent le groupe auquel l'inventeur est assigné. La zone encerclée représente une communauté de spécialistes avec un nombre élevé d'intercollaboration entre les inventeurs

5.3. Réseaux sociaux

Les liens entre inventeurs se font par l'intermédiaire de brevets, peu importe qu'ils soient octroyés à des entreprises canadiennes ou non. Les inventeurs qui résident au Canada et qui participent à des inventions octroyées à des entreprises non canadiennes peuvent quand même être reliés à des inventeurs qui résident au Canada mais qui brevettent pour des entreprises canadiennes. Aussi, dans le cas des multinationales, il peut arriver que des inventeurs canadiens travaillent conjointement avec des inventeurs qui sont à l'étranger sur quelques brevets. Notons

aussi que notre échantillon contient des brevets sur une période de 15 ans (de 1990 à 2005) et le réseau social qui en résulte (figure 3) tient compte de tous les liens cumulés sur ces années. Un lien de co-invention implique un échange de connaissances entre les co-inventeurs. L'accumulation du capital social signifie alors une accumulation de connaissances. En tenant compte de l'ensemble des liens sur une longue période, nous mesurons aussi l'impact de l'accumulation des liens sociaux passés sur les performances futures (Agrawal *et al.*, 2006).

Ces inventeurs sont liés par 45 966 liens entre eux, ce qui fait une moyenne de 8,28 liens par inventeur. Ainsi, sur les 5 902 inventeurs de notre échantillon, 353 (6%) sont des singletons, *i.e.* qu'ils n'ont aucun lien avec d'autres inventeurs. Nous avons sélectionné la composante réseau principale qui comprend 811 inventeurs soit 13,7% de tous les inventeurs du réseau. On peut ainsi supposer que le réseau social des inventeurs canadiens en nanotechnologie est très fragmenté et que beaucoup d'inventeurs n'ont pas de liens de co-invention avec d'autres inventeurs. Nous mesurons le coefficient de groupement, le degré de distribution, le degré de proximité et le degré d'intermédiation pour chaque inventeur de la composante pour l'étape suivant de notre analyse.

5.4. Classification

Étant donnée la grande variation dans la magnitude des indicateurs de capital social, nous avons standardisé les attributs de capital social de chaque inventeur de la composante principale. Les degrés de distribution et d'intermédiation, n'étant pas gaussiens, ils ont d'abord été normalisés en appliquant le logarithme naturel.

Chaque inventeur peut se distinguer des autres par la diversité dans son capital social. Par exemple, un inventeur peut faire partie d'une clique tout en ayant un degré de distribution et d'intermédiation élevé. D'autres joueront un rôle plutôt périphérique dans le réseau et auront un degré de distribution plus petit que les autres. Nous allons utiliser 3 attributs pour représenter le capital social de chaque inventeur, les possibilités de trouver plusieurs types de profils sociaux sont importantes. Puisque l'un des objectifs de cet article est de découvrir des profils sociaux, nous ne voulions pas imposer un nombre strict de segments dans lesquels les inventeurs seraient regroupés. Notre objectif est plutôt de faire un regroupement naturel des inventeurs qui ont un profil social suffisamment proche. Afin de pouvoir atteindre cet objectif, nous avons choisi d'adopter la méthode de classification par modèle de Fraley et Raferty (2002) avec le nombre maximum de groupes égal à 9. Cette méthode propose une division optimale comportant 7 groupes.

5.5. Analyse des segments

Le tableau 2 montre la matrice contenant le signe des *cotes-z* du test en U de Mann-Whitney-Wilcoxon indiquant si deux groupes proviennent d'une même distribution. Chaque case où le test en U s'applique contient 3 valeurs séparées par des virgules : ces valeurs représentent le signe des *z-scores* du test en U pour le

degré d'intermédiation, le coefficient de groupement et le degré de distribution respectivement. La valeur « n.s. » signifie que l'hypothèse nulle d'égalité de distribution entre les deux groupes ne peut être rejetée avec une probabilité inférieure à 0,1. Par exemple, les valeurs de la case montrant le résultat des tests en U entre les groupes 1 et 2 (« -, n.s., -«) signifient que : le rang du degré d'intermédiation du groupe 1 est significativement plus bas que celui du groupe 2 ; le rang du coefficient de groupement du groupe 1 et 2 sont de même niveau ; et le rang du degré de distribution du groupe 1 est plus bas que celui du groupe 2.

Tableau 2. Caractérisation des groupes trouvés

	Différence entre groupes (Degré d'intermédiation, coefficient de groupement, degré de distribution)						Type de profil social
	1	2	3	4	5	6	
1							Périphériques
2	-, n.s., -						Spécialistes
3	-, +, -	-, +, -					Ambidextres
4	-, +, -	-, +, -	-, +, -				Intermédiaires
5	-, +, -	+, +, -	+, -, +	+, -, +			Périphériques
6	-, +, -	-, +, -	-, n.s., +	+, -, +	-, +, +		Ambidextres
7	+, +, +	+, +, +	+, +, +	+, +, +	+, +, +	+, -, +	Périphériques

En analysant les attributs sociaux de chaque groupe, on observe des différences dans la stratégie de collaboration des inventeurs qui figurent dans chaque groupe. Cette distinction permet de distinguer plus de quatre types de groupes : les périphériques, les spécialistes, les intermédiaires et les ambidextres. L'analyse qui suit décrit l'assignation de chaque groupe à ces stratégies.

On remarque tout d'abord que les groupes 1, 5 et 7 ont des rangs des degrés d'intermédiation et de distribution plus bas et des rangs de coefficients de groupements plus élevés que les autres groupes. Ces groupes représentent des inventeurs périphériques : ils ont peu de liens, ce qui les place dans les périphéries de la composante principale telle que montrée à la figure 3.

Le groupe 4 regroupe les intermédiaires : ces inventeurs ont des rangs de degré d'intermédiation plus élevé et des rangs de coefficient de groupement plus bas que les autres. À l'opposé, les inventeurs du groupe 2 ont des rangs de degré d'intermédiation plus bas et des rangs de coefficient de groupement plus élevé que les autres. Ils sont donc des spécialistes qui ont la majorité de leur travail exécuté au sein d'un groupe hermétique d'inventeurs fortement interreliés. Entre les groupes 2 et 4 se situent les inventeurs des groupes 3 et 6. Ils ont des coefficients de groupement plus bas que les spécialistes, mais des degrés d'intermédiation plus petit que les intermédiaires. Ils sont donc classés en tant qu'inventeurs ambidextres.

Le résultat de notre classification est conforme avec la littérature sur les communautés de pratiques dans le sens où nous retrouvons quatre types de profils sociaux espérés : les inventeurs périphériques (segments 1, 5 et 7), les inventeurs intermédiaires (groupe4), les spécialistes (segment 2) et les ambidextres (segments 3 et 6).

5.6. Analyse de la valeur des brevets

Les tableaux 3 et 4 montrent les résultats des tests en U pour les indicateurs de valeur sociale et privée des brevets, respectivement. Ces tableaux sont organisés de la même manière que le tableau 2. Toutefois, il y a deux valeurs par case : le nombre minimum et maximum de citation par brevets pour le tableau 3, et le pourcentage de renouvellement et le nombre de revendications par brevets pour le tableau 4.

Tableau 3. Comparaison des indicateurs de valeur sociale des brevets pour les groupes trouvés

	Différence entre groupes (nombre minimum, maximum de citations)						Type de profil social
	1	2	3	4	5	6	
1							Périphériques
2	n.s., n.s.						Spécialistes
3	n.s., -	+, -					Ambidextres
4	n.s., -	+, -	n.s., -				Intermédiaires
5	n.s., -	n.s., -	-, n.s.	-, +			Périphériques
6	n.s., -	n.s., -	n.s., +	-, +	n.s., n.s.		Ambidextres
7	n.s., n.s.	n.s., n.s.	-, +	-, +	n.s., +	n.s., +	Périphériques

Ces résultats montrent que les stratégies de collaboration ont des impacts différents dépendamment de si l'on tient compte des gains sociaux ou privés. Par exemple, on voit dans le tableau 3, que les inventeurs intermédiaires produisent des brevets qui ont des rangs du nombre maximum de citations plus élevés que les autres segments. Par contre, lorsqu'on tient compte du nombre de citations minimum, les inventeurs spécialistes ont un rang plus élevé que les intermédiaires. En d'autres termes, les inventions qui ont des retombées sociales plus importantes sont plus souvent produites par des inventeurs intermédiaires. Par contre, ces mêmes inventeurs intermédiaires produisent aussi les inventions qui ont le moins de répercussion sociale. Pour ce qui est des gains privés, les inventeurs intermédiaires semblent moins bien performer que les spécialistes. En effet, leurs brevets semblent être moins souvent renouvelés et ont en moyenne moins de revendications que les brevets produits par les spécialistes.

Tableau 4. Comparaison des indicateurs de valeur privée des brevets pour les groupes trouvés

	Différence entre groupes (% renouvellement, nombre de revendications par brevet)						Type de profil social
	1	2	3	4	5	6	
1							Périphériques
2	-, -						Spécialistes
3	+, +	+, +					Ambidextres
4	+, +	+, +	+, +				Intermédiaires
5	+, +	+, +	n.s., n.s.	-, -			Périphériques
6	+, +	+, +	n.s., -	-, -	n.s., -		Ambidextres
7	n.s., n.s.	n.s., n.s.	-, -	-, -	-, -	-, -	Périphériques

Il importe de mettre ces résultats en perspective par rapport au nombre de brevet produits. En effet, il semble y avoir une corrélation entre le nombre maximum de citations et le nombre de brevets produits par un inventeur. Ceci est tout à fait naturel puisque produire plus de brevets augmente les probabilités qu'un des brevets perçoive. Il semble donc y avoir un lien endogène entre intermédiation et longévité pour un inventeur : plus un inventeur est actif, plus il tissera des liens de collaborations, ce qui augmentera les chances de produire des innovations qui auront une portée sociale plus importante. Bien entendu, d'un point de vue des communautés de pratiques, pour qu'un inventeur se positionne au centre d'un réseau, il importe que celui-ci innove. Un inventeur qui n'innoverait pas se retrouverait dans les périphéries du réseau, produirait peu d'innovations et ses performances dépendraient de ces quelques innovations produites.

Il faut noter que la corrélation de *Pearson* entre le nombre de brevets produits et les autres indicateurs de valeur de brevets (nombre minimum de citations, pourcentage de renouvellement et nombre de revendications) sont faibles. Pour ce qui est de la valeur privée des brevets, l'avantage relié à la spécialisation ne semble pas être relié au nombre de brevets produits.

6. Conclusion

Nous avons utilisé une technique de classification par modèle pour regrouper les inventeurs suivant la similitude dans leur stratégie de collaboration. Cette méthode nous a permis d'identifier les inventeurs qui sont périphériques, ceux qui sont des spécialistes, ceux qui sont ambidextres et ceux qui sont plutôt des intermédiaires. Une analyse exploratoire des groupes montre que le capital social a un impact différent suivant le fait qu'on considère les gains sociaux ou privés d'un brevet. Les inventeurs intermédiaires produisent plus de brevets et des brevets qui ont d'importantes répercussions au niveau social, mais ils n'arrivent pas toujours à

s'approprier les retours sur ces brevets. À l'inverse, les inventeurs stratégiques produisent des brevets qui sont moins variés dans le bénéfice social qu'ils génèrent mais qui engendrent plus de bénéfices privés.

Ainsi, la notion de qualité d'une innovation dépend du point de vue de l'innovateur. Un innovateur issu du milieu industriel aura tout à fait intérêt à promouvoir la stratégie de spécialisation dans les collaborations de l'organisation puisque cette voie offre la meilleure opportunité d'appropriation des bénéfices sur l'activité innovante. Par contre, pour un organisme public, l'intermédiation semble la stratégie la plus prometteuse puisqu'elle génère le plus de retombées au niveau social : cela produit des brevets qui reçoivent plus de citations, mais qui revendiquent moins et qui sont moins souvent renouvelés.

Notre méthodologie comprend un certain nombre de limitations. Tout d'abord, notre liste de mots clés n'est pas exhaustive et contient seulement un sous-ensemble des mots clés qui représentent la nanotechnologie. De plus, nous n'avons pas utilisé de mots clés d'exclusion sur le terme « nano* », ce qui peut introduire des brevets non pertinents dans notre échantillon initial de brevets. La deuxième limitation de notre étude provient de l'ambiguïté des noms et initiales d'inventeurs liée à leur identification. Alors que cette limitation n'a pas de conséquences sérieuses sur un petit échantillon ciblant une discipline particulière, cela peut faire augmenter le niveau de connectivité dans un réseau de co-invention dans un échantillon très large. Des techniques de résolutions de l'ambiguïté de noms d'inventeurs (Trajtenberg, 2005 ; Fleming *et al.*, 2007) peuvent être adoptées pour pallier ce problème. Finalement, la mise en place d'une méthode automatisée qui pourra ordonner les groupes de manière automatique suivant leur rang au niveau des tests en U permettra d'analyser des échantillons plus grands qui contiennent un nombre plus grand de groupes.

Remerciements

Nous remercions le Conseil de recherches en sciences humaines, les Instituts de recherche en santé du Canada et le Conseil de recherches en sciences naturelles et en génie du Canada pour son support financier. Nous remercions l'Observatoire des sciences et des technologies et particulièrement Vincent Larivière pour l'accès et le support aux données concernant les citations de brevets USPTO. Nous remercions Ricard-Olivier Moreau et Terence-Shawn Ramine Kananian pour leur longue et laborieuse contribution dans le nettoyage et l'administration de notre base de données. Finalement, nous remercions Ilse-Carolina Molina Gutiérrez pour la préparation de la liste de mots clés en nanotechnologie.

Bibliographie

- Addouche S., Dafaoui E., El Mhamedi A. (2006). Démarche analytique d'amélioration continue des processus d'entreprise. Identification de la relation de causalité entre indicateur et inducteur de performance. *Journal Européen des Systèmes Automatisés*, vol. 40, n° 1, p. 33-50.

- Adler P. S., Kwon S.-W. (2002). Social capital - Prospects for a new concept. *The Academy of Management Review*, vol. 27, n° 1, p. 17-40.
- Agrawal A., Cockburn I., McHale J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, vol. 6, p. 571-591.
- Ahuja G., Lampert C.M. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, vol. 22, p. 521-543.
- Alencar M.S.M., Porter A.L., Antunes A.M.S. (2007). Nanopatenting patterns in relation to product life cycle. *Technological Forecasting & Social Change*, vol. 74, p. 1661-1680.
- Allison J.R., Lemley M.A., Moore K.A., Trunkey R.D. (2004). Valuable Patents. *Georgetown Law Journal*, vol. 92, n° 3, p. 435-479.
- Amin A., Cohendet P. (2004). *Architecture of Knowledge: Firms, Capabilities and Communities*. NY, Oxford University Press.
- Ankerst M., Breunig M.M., Kriegel H.-P., Sander J. (1999). OPTICS: Ordering Points To Identify the Clustering Structure. *ACM SIGMOD international conference on Management of data*, p. 49-60.
- Berry M.J., Linoff G.S. (2004). *Data Mining Techniques For Marketing Sales and Customer Relationship Management*. Wiley.
- Brown J.S., Duguid P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science - Special Issue Organizational Learning: Papers in Honor of (and by) James G. March*, vol. 2, n° 1, 40-57.
- Burt R.S. (2004). Structural Holes and Good Ideas. *The American Journal of Sociology*, vol. 110, n° 2, 349-399.
- Cantner U., Graf H. (2006). The network of innovators in Jena: An application of social network analysis. *Research Policy*, vol. 35, p. 463-480.
- Cohen W., Levinthal D. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administration Science Quarterly*, vol. 35, p. 128-152.
- Coleman J.S. (1988). Social Capital in the Creation of Human Capital. *The American Journal of Sociology - Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure*, vol. 94, p. S95-S120.
- Conover W.J. (1980). *Practical Nonparametric Statistics*. John Wiley & Sons.
- Cowan R., Jonard N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics & Control*, vol. 28, p. 1557-1575.
- Csardi G., Nepusz T. (2006) The igraph software package for complex network research. *InterJournal*, vol. Complex Systems 1695.
- Denimal J., Boussu F. (2002). Estimation dynamique des comportements de vente d'articles textiles. *Journal Européen des Systèmes Automatisés*, vol. 36, n° 2, p. 223-244.
- Ester M., Kriegel H.-P., Sander J., Xu X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Simoudis E., Han J., Fayyad U.M.

Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), p. 226-231.

- Felin T., Hesterly W. (2007). The knowledge-based view, nested heterogeneity, and new value creation: Philosophical considerations on the locus of knowledge. *The Academy of Management Review*, vol. 32, n° 1, p. 195-218.
- Fitzgibbons K., McNiven C. (2006). Towards a Nanotechnology Statistical Framework. *Blue Sky Indicators Conference II*.
- Fleming L., Mingo S., Chen D. (2007). Collaborative brokerage generative creativity and creative success. *Administrative Science Quarterly*, vol. 52, p. 443-475.
- Fraley C. Raftery A.E., Murphy T.B., Scrucca L. (2012). *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*. Technical Report No. 597, Department of Statistics, University of Washington.
- Fraley C., Raftery A.E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, vol. 97, p. 611-631.
- Girvan M., Newman M.E.J. (2002). Community structure in social and biological networks. *PNAS*, vol. 99, n° 2, p. 7821-7826.
- Granovetter M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, vol. 78, n° 6, p. 1360-1380.
- Griliches Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, vol. 28, n° 4, p. 1661-1707.
- Gruat la Forme F., Botta-Genoulaz V., Campagne J. (2007). Problème d'ordonnement avec prise en compte des compétences. Résolution monocritère pour indicateurs de performance industriels et humains. *Journal Européen des Systèmes Automatisés*, vol. 41, n° 5, p. 617-642.
- Hall B.H., Jaffe A., Trajtenberg M. (2005). Market value and patent citations. *The RAND Journal of Economics*, vol. 36, n° 1, p. 16-38.
- Hansen M. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, vol. 44, n° 1, p. 82-111.
- Harhoff D., Narin F., Scherer F.M., Vopel K. (1999). Citation Frequency and the Value of Patented Inventions. *The Review of Economics and Statistics*, vol. 81, n° 3, p. 511-515.
- Inkpen A.C., Tsang E.W.K. (2005). Social capital networks, and knowledge transfer. *The Academy of Management Review*, vol. 30, n° 1, p. 146-165.
- Lanjouw J.O., Pakes A., Putnam J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. *The Journal of Industrial Economics*, vol. 46, n° 4, p. 405-432.
- Lanjouw J.O., Schankerman M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, vol. 114, n° 495, p. 441-465.
- Lemley M.A., Shapiro C. (2005). Probabilistic Patents. *The Journal of Economic Perspectives*, vol. 19, n° 2, p. 75-98.

- Li X., Lin Y., Chen H., Roco M.C. (2007). Worldwide nanotechnology development: a comparative study of USPTO, EPO and JPO (1976–2004). *Journal of Nanoparticle Research*, vol. 9, n° 6, 977-1002.
- Manning C.D., Raghavan P., Schütze H. (2008). *An Introduction to Information Retrieval*. Cambridge University Press.
- Meyer M. (2000). What is Special about Patent Citations? Differences between Scientific and Patent Citations. *Scientometrics*, vol. 49, n° 1, p. 93-123.
- Mogoutov A., Kahane B. (2007). Data search strategy for science and technology emergence: A scalable and evolutionary query for nanotechnology tracking. *Research Policy*, vol. 36, p. 893-903.
- Nahapiet J., Ghoshal S. (1998). Social Capital, Intellectual Capital, and the Organizational Advantage. *The Academy of Management Review*, vol. 23, n° 2, p. 242-266.
- Newman M. (2003). The structure and function of complex networks. *SIAM Review*, 45, n° 2, 167-256.
- Noteboom B., Haverbeke W.V., Duysters G., Gilsing V., van den Oord A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, vol. 36, p. 1016-1034.
- Penin J. (2005). Patents versus ex post rewards: A new look. *Research Policy*, vol. 34, p. 641-656.
- Perry-Smith and Shalley (2003). The Social Side of Creativity: A Static and Dynamic Social Network Perspective. *The Academy of Management Review*, vol. 28, n° 1, p. 89-106.
- Porter A.L., Youtie J., Shapira P., Schoeneck D.J. (2008). Refining search terms for nanotechnology. *Journal of Nanoparticle Research*, vol. 10, p. 715-728.
- Powell W.W., Koput K.K., Smith-Doerr L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, 41, 116-145.
- Pronzato L. (2006). Sélection séquentielle de conditions expérimentales non contrôlées, *Journal Européen des Systèmes Automatisés*, vol. 40, n° 2, p. 197-210.
- Rost K. (2011). The strength of strong ties in the creation of innovation. *Research Policy*, vol. 40, p. 588-604.
- Sbihi M., Moussa A., Postaire J., Sbihi A. (2005). Approche markovienne pour la classification automatique non supervisée de données multidimensionnelles. *Journal Européen des Systèmes Automatisés*, vol. 39, n° 9-10, p. 1133-1154.
- Schmoch U., Heinze T., Hinze S., Rangnow R. (2003). *Mapping Excellence in Science and Technology across Europe: Nanoscience and Nanotechnology*. Centre for Science and Technology Studies.
- Shane S. (2001). Technological Opportunities and New Firm Creation. *Management Science*, vol. 47, n° 2, p. 205-220.
- Tomala F., Sénéchal O. (2002). Modèles de produits et de processus pour l'analyse de l'impact de l'introduction d'une innovation dans un produit existant. *Journal Européen des Systèmes Automatisés*, vol. 36, n° 2, p. 273-298.
- Tong X., Frame J. (1994). Measuring national technological performance with patent claims data. *Research Policy*, vol. 23, n° 2, p. 133-141.

- Trajtenberg M. (2005). The mobility of inventors and the productivity of research. *Working paper, Tel Aviv University*; NBER and CEPR, NBER SI.
- Trajtenberg (1990). A Penny for Your Quotes: Patent Citations and the Value of Innovations. *The RAND Journal of Economics*, vol. 21, n° 1, p. 172-187.
- Trentesaux D. (2007). Les systèmes de pilotage hétérarchiques. *Journal Européen des Systèmes Automatisés*, vol. 41, n° 9-10, p. 1165-1202.
- Uzzi B. (1997). Social structure and competition in interfirm networks - The paradox of embeddedness. *Administrative Science Quarterly*, vol. 42, n° 1, p. 35-67.
- Wasserman F., Fraust K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Watts D., Strogatz S. (1998). Collective Dynamics of 'small-world' networks. *Nature*, vol. 393, 4 June, p. 440-442.
- Zitt M., Bassecoulard E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing and Management*, vol. 42, p. 1513-1531.
- Zucker L., Darby M., Armstrong J. (1998). Geographically localized knowledge: Spillovers or markets? *Economic Inquiry*, vol. 36, n° 1, p. 65-86.