

Étude de cas Big data

Gestion et analyse de données massives

Problématique

Reichert Technologies, est une filiale de AMETEK Inc, basé à Depew, dans l'État de New York au USA. La société est un leader mondial dans la conception et la fabrication d'instruments et d'équipements pour soigner la vue. La gamme de produits s'adressent principalement aux hôpitaux, aux ophtalmologistes et aux opticiens. En 2013, le chiffre d'affaires du groupe AMETEK était de 3,6 milliards de dollars, avec 15.000 collaborateurs répartis sur 120 sites de production dans le monde entier.

Dans le cadre de son développement, Jeff Bornheim, le nouveau directeur de Reichert Technologies souhaite s'implanter en France, notamment sur les marchés publics. Le but est de devenir un des principaux fournisseurs des hôpitaux. Pour atteindre cet objectif, le directeur envisage une alliance stratégique avec une entreprise déjà implanté depuis une dizaine d'années sur les marchés publics et très influente, afin de répondre aux appels d'offre en co-traitance.

La co-traitance est une situation dans laquelle des entreprises soumissionnaires peuvent s'allier dans un groupement d'entreprises, pour proposer une offre collective plus avantageuse que les offres individuelles. Pourtant, plusieurs études, montrent que 30 % à 70 % des alliances stratégiques échouent, c'est-à-dire qu'elles n'atteignent pas les objectifs fixés. Les conditions initiales de leur organisation et notamment le choix des partenaires serait déterminantes.

Mission

Vous venez d'être recruté(e) comme chargé d'étude au sein du service marketing. En collaboration avec vos nouveaux collègues, votre première mission consiste à élaborer une étude de marché. Cette étude doit aider le directeur de Reichert Technologies à sélectionner un partenaire pour former une alliance stratégique.

Pour identifier le bon partenaire, il vous est demandé de réaliser les tâches suivantes :

1. identifier des sources de données sur les marchés publics français ;
2. sélectionner une source de données ouvertes ;
3. collecter les données sur les transactions attribuées aux groupements d'entreprises ;
4. définir les critères de sélection du bon partenaire
5. analyser les données ;
6. recommander un partenaire ;
7. présenter vos résultats à travers une note de synthèse (3 pages minimum).

Les concepts mobilisés

Sommaire

A – Comment accéder un marché public ?	4
A1 - Les marchés publics	4
A2 – Les stratégies d'accès aux marchés publics	4
A3 - La formation des alliances stratégiques	6
B - Qu'apporte l'intelligence économique ?	7
B1 – Apports de l'intelligence économique	7
B2 - Conception de l'intelligence économique	7
B3 - Typologie des sources d'informations	8
B4 - Cycle du renseignement	8
C -Qu'est-ce que le Big data ?	9
C1 – L'open data (données ouvertes)	10
C2 – Définition du big data (données massives)	10
C3 – Conception du big data (données massives)	12
C4 - Les technologies du Big data	12
D– Comment prendre une décision ?	13
D1 - Qu'est-ce qu'une décision ?	13
D2 - Les difficultés de la prise de décision	13
D3 - Comment prendre la bonne décision ?	13
D4 - L'approche rationnelle du processus de décision	13
D5 - Les limites de l'approche rationnelle dans la prise de décision	14
D6 - L'approche émotionnelle de la prise de décision	15
E - Comment analyser les données ?	15
F - Quel est l'intérêt de l'analyse des réseaux ?	16
F1 – L'analyse des réseaux	16
F2- Visualisation des données avec Gephi	17
F3 – Propriétés des grands réseaux	17
F4 – Analyse prédictive à partir des réseaux	18
F5 - La théorie de la force des liens	19

A – Comment accéder un marché public ?

A1 - Les marchés publics

Les marchés publics représentent entre 8 % et 25 % du Produit Intérieur Brut (OECD, 2006) en fonction des pays, et 16 % pour l'Europe (CEC, 2008). Contrairement au secteur privé, les marchés publics sont tenus de respecter d'importantes contraintes légales (Arlbjørn & Freytag, 2012). En France, c'est le Code des marchés publics (2012), inspiré par les directives européennes (Gelderman et al., 2006), qui fixe le cadre légal. Le Code définit un marché public comme un contrat conclu, à titre onéreux, entre une personne publique et un opérateur économique privé (par exemple une PME, une multinationale, une association ou une alliance d'entreprises). Le but du contrat est de satisfaire des besoins en matière de travaux (construction de bâtiments ou génie civil, etc.), de fournitures (mobilier, matériels, consommables, etc.) ou de services (nettoyage de locaux, sécurité alarme, entretien de jardins, enlèvement des ordures ménagères, etc.).

Comme pour le secteur privé, la recherche des coûts d'approvisionnement aux meilleures conditions, reste un objectif fondamental de l'acheteur public (Loader, 2010). La mise en concurrence est le moyen utilisé depuis longtemps pour sélectionner des fournisseurs dans les marchés publics (Arlbjørn & Freytag, 2012). Cette mise en concurrence repose, généralement, sur des appels d'offres, c'est-à-dire des procédures d'enchères écrites, sous pli cacheté, avec remise d'une proposition unique (Mougeot et Cohen, 2001).

L'attribution d'un marché à l'un des candidats est donc une décision stratégique que prennent les pouvoirs publics, sur la base de l'offre considérée économiquement la plus avantageuse (Naegelen & Mougeot, 1998). Pour faire le meilleur choix, l'acheteur public se fonde sur les critères suivants : la qualité, le prix, la valeur technique, le caractère esthétique et fonctionnel, les performances en matière de protection de l'environnement, les performances en matière d'insertion professionnelle des publics en difficulté, le coût global d'utilisation, la rentabilité, le caractère innovant, le service après-vente et l'assistance technique, la date de livraison, le délai de livraison ou d'exécution (Code des marchés publics, 2012).

Tous ces critères n'ont pas la même importance. Ils peuvent être pondérés ou à défaut hiérarchisés. Pour cela, l'acheteur public définit l'équilibre entre les différents critères et leur valeur respective, au moyen par exemple, de l'application d'un pourcentage ou d'un coefficient.

Toutefois, l'attribution des marchés publics doit répondre à trois grands principes (Code des marchés publics, 2012). Premièrement, la liberté d'accès à la commande publique est garantie par l'acheteur public, afin de permettre à tous les candidats de pouvoir concourir. Deuxièmement, l'égalité de traitement des candidats doit être respectée pour ne pas fausser la concurrence entre les acteurs répondant à une demande. Participent de cette égalité de traitement, les procédures de consultation, l'analyse du contenu des réponses et la composition des commissions de sélection. Troisièmement, la transparence des procédures doit permettre de respecter l'équité de traitement et de favoriser une concurrence saine entre les entreprises. Cette transparence justifie la soumission des marchés publics à des procédures de publicité.

Source : Mamavi, O., et Morin, S. (2014). Quelle intelligence peut-on trouver dans les données massives ? La cas des marchés publics français. Revue Internationale d'Intelligence Économique, vol.6, pp. 131-42

A2 – Les stratégies d'accès aux marchés publics

L'accès aux marchés publics est un processus de décisions stratégiques comprenant : l'identification et la sélection de l'opportunité d'affaires, la décision d'engagement et, enfin, le choix d'une stratégie concurrentielle. Pour obtenir un marché public, les entreprises soumissionnaires recherchent la position concurrentielle qui leur permettra de proposer l'offre la plus avantageuse. Le choix d'une stratégie relationnelle peut se résumer au dilemme suivant : soit accepter la rivalité en stabilisant les règles du jeu concurrentiel, soit limiter l'intensité concurrentielle en perturbant les règles de ce jeu.

Koenig (1996) propose trois grands modes relationnels qu'une entreprise peut choisir par rapport à son

environnement concurrentiel : soit s'aligner sur les règles du jeu concurrentiel en essayant d'acquérir une meilleure position que ses rivaux (affrontement), soit contourner les règles du jeu de l'environnement en exploitant ses ressources pour éviter la concurrence (évitement), soit, enfin, s'adapter à l'environnement en collaborant avec certains acteurs (coopération). Le *tableau 1* récapitule les principales stratégies relationnelles mises en place par les entreprises pour accéder aux marchés publics.

Accepter la rivalité, c'est admettre que l'attribution du marché public se fera, uniquement, sur l'entreprise proposant un avantage concurrentiel. Cet avantage concurrentiel sera obtenu à partir de la différenciation sur les critères d'attribution préalablement définis lors de l'appel d'offres ; c'est-à-dire, dans la plupart des cas, le meilleur rapport qualité-prix. Par contre, limiter la rivalité permet d'éviter la diminution du profit et augmente la probabilité d'obtenir un marché public dans un processus d'enchères.

En effet, et comme le souligne Viau (2003), une entreprise diminue ses chances de remporter un marché public, vis-à-vis de ses concurrents, si elle le découvre seulement lors de l'annonce d'appel à concurrence. Plus que la réponse à l'appel d'offres, c'est la participation à l'élaboration du cahier des charges qui devient déterminant. Pour cela, l'entreprise soumissionnaire va chercher à se rapprocher des parties prenantes, afin de mieux intégrer la demande du donneur d'ordres, voire même l'aider dans l'étude de faisabilité, ou proposer le montage financier de l'opération. Viau (2003) précise les techniques d'influence utilisées par les entreprises pour réduire l'asymétrie d'information, anticiper la demande et ainsi éviter l'effet de la concurrence. Cette stratégie relationnelle s'appuie sur des relations directes avec le donneur d'ordres, ou par des relais en dehors des transactions économiques, par le biais d'événements sociaux (colloques, salons, clubs, voyages...), ou par la corruption.

L'entente entre acheteurs et fournisseurs étant prohibée dans les marchés publics, des formes plus subtiles de limitation de la rivalité concurrentielle peuvent exister à travers les relations interorganisationnelles. Ces relations correspondent, généralement, à de la coopération entre entreprises soumissionnaires à un appel d'offres, sous la forme de groupements d'entreprises. Ces groupements d'entreprises sont des alliances stratégiques. L'alliance stratégique peut être définie comme un lien entre plusieurs entreprises indépendantes, qui partagent des ressources afin de construire un avantage coopératif, dans le but d'atteindre des objectifs communs (Jolly, 2001). La notion d'alliance stratégique induit, d'une part, une coopération, c'est-à-dire l'action de participer à une œuvre commune ; et, d'autre part, un engagement mutuel, c'est-à-dire une union qui a été négociée.

Source : Mamavi, O. (2015). Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux. L'harmattan, collection Entreprise et Management, Paris

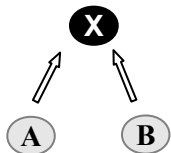
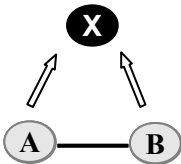
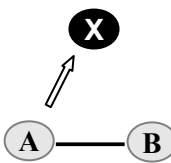
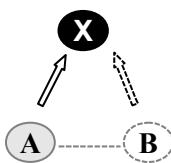
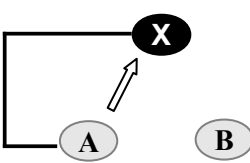




stratégies de différenciation	rivalité	
stratégies de coopération et de collaboration	cotraitance (alliances stratégiques)	
	sous-traitance	
stratégies d'évitement	retenue mutuelle	
	ententes/corruption	
 <i>donneur d'ordre</i>  <i>entreprises</i>  <i>soumission à un marché public</i>  <i>relations formelles ou informelles</i>		

Tableau 1 : Principales stratégies relationnelles pour accéder aux marchés publics

A3 - La formation des alliances stratégiques

Les alliances stratégiques se définissent ainsi comme des accords explicites, établis dans une perspective de plus ou moins longue durée. Au sein de ces formations, les entreprises partagent, combinent ou échangent des ressources, en vue de mener à bien un projet ou une activité spécifique. L'objectif de ces alliances est de bâtir et maintenir des avantages compétitifs plutôt que de réaliser ces activités de façon autonome (Philippart, 2001).

De manière générale, et malgré les controverses, un consensus se dégage dans la littérature sur les motivations des partenaires à former des alliances (Meier, 2009) : la minimisation des coûts de transactions, l'accès aux ressources et compétences, le transfert et l'apprentissage organisationnel, l'adaptation à l'environnement et à la structure compétitive d'un secteur. De plus, la littérature identifie de nombreux critères pour sélectionner un partenaire (Cherni, 2010). Néanmoins, Geringer (1988) distingue deux grandes catégories de critères de sélection : des critères liés aux ressources à partager (similitude ou complémentarité), et des critères liés aux comportements du partenaire (niveau d'engagement, compatibilité, contrôle de l'alliance). Pour une entreprise qui ne dispose pas des capacités ou des spécialités nécessaires à l'exécution d'un marché, la coopération constitue donc, un moyen privilégié pour obtenir des ressources. Au-delà des ressources qui contribuent directement à construire l'avantage concurrentiel, Geringer (1988) souligne que les actifs relationnels peuvent, également, être pris en compte. Par exemple, la connaissance du marché local, du contexte institutionnel, de même que les liens établis avec les acteurs locaux peuvent influencer le choix du partenaire. De plus, le niveau d'engagement dans l'accord d'alliances est, selon la

littérature, un critère fondamental dans le choix du partenaire. L'engagement correspond à l'implication des partenaires pour fournir un effort, afin d'atteindre les objectifs communs. La confiance permettra d'appréhender le niveau d'engagement mutuel des partenaires dans la coopération et d'éviter les comportements opportunistes. Enfin, Geringer (1988) souligne l'importance de la compatibilité sur les plans stratégique et organisationnel dans la sélection des partenaires. La compatibilité des dirigeants, la culture d'entreprise et les procédures organisationnelles sont déterminantes dans la survie des alliances stratégiques.

Source : Mamavi, O. et Meier, O. (2014). *Le principe d'attachement préférentiel dans la formation des alliances stratégiques*, *Revue Française de Gestion*, vol. 4 n°241, p. 79-91

B - Qu'apporte l'intelligence économique ?

B1 – Apports de l'intelligence économique

Le choix d'une stratégie concurrentielle, et plus particulièrement le déploiement d'une alliance, est complexe. Thiétart et coll. (2009) distinguent trois causes principales de cette complexité. La première est due à l'incertitude liée à l'impossibilité de raisonner simplement en termes de causalité linéaire. La deuxième est liée à la multitude de facteurs pour comprendre l'efficacité d'une stratégie. La troisième est la difficulté à saisir les interactions et les interdépendances entre les parties prenantes.

L'organisation d'une alliance stratégique suppose, de la part du dirigeant d'entreprise, une aptitude permanente à réagir à l'environnement, à s'adapter aux évolutions et à anticiper les changements. Face à un environnement dynamique et instable, l'apport de *l'intelligence économique* peut être considéré comme la solution qui permettra la réduction de la part d'incertitude dans la prise de toute décision stratégique (Moinet, 2011). Comme le précise Morin (cité par Massé et al., 2006): être intelligent, c'est pouvoir trouver une solution dans un environnement complexe. Dans la pratique, cela correspond à la *capacité d'absorption* (Zahra and George, 2002) de l'information à des fins stratégiques, c'est-à-dire à son acquisition, son assimilation, sa transformation et son exploitation.

En effet, les entreprises qui voudront former des alliances stratégiques devront développer leur capacité à créer, et/ou activer, des relations de coopération dans le cadre d'un projet défini. Ces combinaisons de liens ne constituent qu'une matière première relationnelle. L'intelligence économique naît à partir du moment où l'un des acteurs, au moins, oriente et active ses liens dans le cadre d'un projet, et est à la recherche d'un avantage compétitif (Massé et al., 2006).

Source : Mamavi, O. (2015). *Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux*. L'harmattan, collection *Entreprise et Management*, Paris

B2 - Conception de l'intelligence économique

L'environnement des entreprises est complexe. Cette complexité est due, notamment, à l'intensité croissante de la concurrence, aux interactions entre acteurs, à l'incertitude des comportements, à l'instabilité des positions, aux raccourcissements des cycles de vie des produits... L'intelligence économique naît de la prise de conscience que la maîtrise de l'information devient vitale, pour permettre à une entreprise d'identifier les opportunités et les menaces de son environnement. Cette maîtrise de l'information, qui passe par une démarche systématique et globale du traitement de l'information stratégique, est de l'intelligence économique.

L'intelligence économique est d'abord un concept empirique qui s'appuie sur des processus, des actions et des comportements en entreprises, dont le rapport Martre viendra donner, en France, une première formalisation (Lebrument, 2011). Ensuite, un ensemble de travaux de recherche tentera de compléter cette conception pour lui donner un cadre théorique. D'ailleurs, Larivet (2002) montre l'existence des différentes approches et pratiques à travers la polysémie de la définition d'intelligence économique. Ce qui fait dire à Bruté de Rémur

(2009) que l'intelligence économique est encore à la recherche de son fondement épistémologique en tant que science sociale.

Néanmoins, même si le concept n'est pas encore stabilisé, nous pouvons considérer, comme Marcon & Moinet (2011), que l'intelligence économique répond principalement à deux finalités : (1) comprendre l'environnement externe ou interne et , (2) faciliter l'aide à la décision stratégique en influençant l'environnement (*tableau 2*).

Concept	Dimension	Composantes
Intelligence économique	renseignements	Veilles (stratégiques, commerciales, concurrentielles, technologiques,...)
		Gestion des connaissances (KM, cartographie, aide à la décision,...)
	influences	Protection de l'information (sécurité, protection de l'innovation,)
		Communications stratégiques (lobbying, relations publiques,...)

Tableau 2: conception de l'intelligence économique

L'intelligence économique est un processus et un produit (Vedder and Guynes, 2001). Un processus, parce qu'il permet de traiter les informations concernant les environnements à partir de méthodes légales et éthiques. Un produit, parce que c'est une démarche et un outil de gestion de connaissances qui permet d'améliorer la compétitivité (Bournois and Romani, 2000).

Source : Mamavi, O. (2015). Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux. L'harmattan, collection Entreprise et Management, Paris

B3 - Typologie des sources d'informations

L'intelligence économique se nourrit d'informations. Cela a conduit les praticiens à développer une typologie de couleur pour caractériser les différentes sources d'informations (blanche, grise ou noire). Cette typologie permet une catégorisation en fonction de différents critères (*tableau 3*).

INFORMATION	BLANCHE	GRISE	NOIRE
Accès	Public	Restreint	Strictement limité
disponibilité	80 %	15 %	5 %
Classification	Non protégé	Diffusion restreinte	Confidentiel - secret
Acquisition	Légale sous réserve de respecter les droits de propriété	Domaine juridique non clairement défini.	Illégal. Acquisition relevant de l'espionnage
Sources	Ouvertes	Autorisées – fermées	Clandestines

Tableau 3: typologie des sources d'informations - source: Lebrument (2012) adaptée de Bulinge (2002)

Source : Mamavi, O. (2015). Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux. L'harmattan, collection Entreprise et Management, Paris

B4 - Cycle du renseignement

En entreprise, l'une des principales pratiques de l'intelligence économique consiste à mettre en place une

démarche de transformation de l'information, en connaissance utile pour prendre des décisions stratégiques (Larivet, 2002). Issu des méthodes du renseignement militaire, le cycle du renseignement est un modèle qui doit permettre aux entreprises la surveillance de leur environnement, notamment concurrentiel. Ce modèle constitue le socle des pratiques de la veille stratégique. Il doit être alimenté par des sources d'informations légales.

Le cycle du renseignement est un processus continu qui peut être décomposé en quatre étapes (*figure 12*):

1. l'expression des besoins,
2. l'acquisition de l'information,
3. l'exploitation de l'information,
4. la diffusion de l'information.

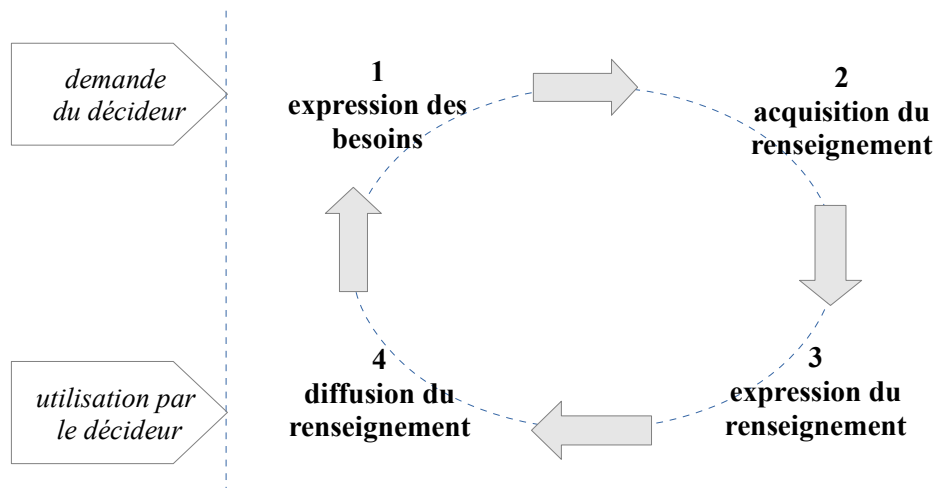


Figure 1: cycle du renseignement (d'après Lebrument, 2011)

Pour une entreprise, un des apports du cycle du renseignement, et plus généralement de l'intelligence économique, est de pouvoir analyser les jeux des concurrents et détecter les mouvements des acteurs du marché (Baumard, 2000). Mais, comme le soulignent Salvetat & Le Roy (2007), au-delà du renseignement, l'intelligence économique permet également à l'entreprise de développer une capacité d'influence sur son environnement, notamment en contribuant à mettre en œuvre des manœuvres coopératives.

Source : Mamavi, O. (2015). Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux. L'harmattan, collection Entreprise et Management, Paris

C -Qu'est-ce que le Big data ?

C1 – L'open data (données ouvertes)

Initié aux USA et en Angleterre en 2009, le mouvement de l'open data consiste pour une organisation à rendre accessible certaines de ses données (transports, météo, santé, marchés publics,...) pour qu'elles soient réutilisables par des entreprises ou des citoyens (Simón et al., 2014). Au delà des enjeux sur la transparence de l'action publique, l'ouverture des données peut être une source d'innovation importante en favorisant, par exemple, l'interconnexion (Streeter et al., 1996) ou l'utilisation des compétences d'une foule - crowdsourcing (Lebraty et Lobre, 2010).

A la différence de l'open source (Gosain, 2003) ou du big data (Wamba et al., 2015), le mouvement de l'open data se distinguent très nettement des autres concepts issues des nouvelles façons de produire et de partager de l'information. En effet, l'open source visent à fournir des applications tandis que l'open data cherche à exploiter des données (Lobre & Lebraty, 2012). Cependant, pour rendre les données, souvent massives, accessibles à tous, elles doivent être transcrites dans un format connu et facilement manipulable.

Plusieurs mesures ont été prises pour améliorer la qualité, l'accès, la documentation et l'échange de données (Davies, 2010 et Yu et Robinson 2012). Toutefois, il reste encore de nombreux défis à relever: adoption et l'utilisation des normes, outils et infrastructures pertinentes, formation du personnel, gestion du changement...

Selon McKinsey, la valeur annuelle de l'open data se situe entre 3 220 et 5 290 milliards de dollars. A titre de comparaison, le PIB de l'Allemagne, 4e puissance mondiale est de 3 747 milliards. Tous les secteurs économiques sont concernés et les opportunités en termes d'amélioration de performance concernent à la fois la productivité et les revenus : trouver de nouveaux clients, améliorer le marketing mix, réduire les coûts marketing, concevoir de nouvelles offres .

En France, l'utilisation de l'open data se développe très rapidement depuis quelques années. L'État met à disposition ses données (www.data.gouv.fr), des start-up construisent des services sur la base de données ouvertes (Data-publica) et des acteurs historiques ouvrent leurs données à leur écosystème (SNCF, Orange).

Source : www.opendata-commercial.strikingly.com

C2 – Définition du big data (données massives)

Auteur	Définition
(IBM, 2012)	Big Data: data captured from sensors, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals, etc.
(B. D. Johnson, 2012)	Big Data: extremely large sets of data related to consumer behaviour, social network posts, geotagging, sensor outputs (p. 21).
(Thomas H. Davenport, et al., 2012)	Big Data: data from everything including click stream data from the Web to genomic and proteomic data from biological research and medicine
(Manyika, et al., 2011)	Big Data: datasets with a size that is beyond the ability of typical database software tools to capture, store, manage, and analyse
(Rouse, 2011)	Big Data: description of the voluminous amount of unstructured and semi-structured data a company creates or data that would take too much time and cost too much money to load into a relational database for analysis
(IDC, 2013)	Big Data has three main characteristics of Big Data: the data itself, the analytics of the data, and the presentation of the results of the analytics. Then there are the products and services that can be wrapped around one or all of these Big Data elements (p. 1)

(Boyd & Crawford, 2012)	<p>Big Data: a cultural, technological, and scholarly phenomenon that rests on the interplay of:</p> <ul style="list-style-type: none"> (1) Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets. (2) Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims. (3) Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy.” (p. 663).
-------------------------	---

Source : Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. (2015). "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study," International Journal of Production Economics.

C3 – Conception du big data (données massives)

Dimension	Caractéristique	Exemple
Volume	Large volume of data that either consume huge storage or consist of large number of records	<ul style="list-style-type: none"> - Tesco generates more than 1.5 billion new items of data every month (Manyika, et al., 2011). - Wal-Mart's data warehouse includes some 2.5 petabytes of information (Manyika, et al., 2011). - Dell initiated to develop a database that includes 1.5 million records related with sales and advertisements (Thomas H Davenport, 2006)
Variety	Data generated from greater variety of sources and formats, and contain multidimensional data fields	<ul style="list-style-type: none"> - Procter & Gamble, created a group consisting of more than 100 analysts from such functions as operations, supply chain, sales, consumer research, and marketing to improve total business performance by analysing interrelationships among functional areas(Thomas H Davenport, 2006). - Tata Motors analyzes 4 million text messages every month, spanning everything from product complaints to reminders about service appointments to announcements about new models and also connected with customer satisfaction polling (Agarwal & Weill, 2012).
Velocity	Frequency of generation and/or frequency of data delivery	- Retailers can now track individual customer's data including clickstream data from the Web and can leverage from their behavioral analysis. Moreover, retailers are now capable of updating such increasingly granular data in near real time to track changes in customer behavior
Veracity	Inherent unpredictability of some data requires analysis of big data to gain reliable prediction	- eBay Inc. faced an enormous data replication problem, with between 20 to 50 fold versions of the same data scattered throughout its various data marts. Later, eBay developed an internal website (data hub) which enables managers to filter data replication (Thomas H. Davenport, et al., 2012).
Value	The extent to which big data generates economically worthy insights and or benefits through extraction and transformation.	<ul style="list-style-type: none"> - Premier Healthcare Alliance used enhanced data sharing and analytics to improve patient outcomes while reducing spending by US\$2.85 billion (IBM, 2012a). - Match.com reported more than 50% increase in revenue in the last two years' time, with more than 1.8 million paid subscribers in its core business, most of which driven through data analytics (Kiron & Shockley, 2011).

Source : Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. (2015). "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study," *International Journal of Production Economics*.

C4 - Les technologies du Big data

Pour optimiser les temps de traitement sur des bases de données géantes, plusieurs solutions peuvent entrer en jeu :

- Des bases de données NoSQL (comme MongoDB, Cassandra ou Redis) qui implémentent des systèmes de stockage considérés comme plus performants que le traditionnel SQL pour l'analyse de données en masse (orienté clé/valeur, document, colonne ou graphe).
- Des infrastructures de serveurs pour distribuer les traitements sur des dizaines, centaines, voire milliers de nœuds. C'est ce qu'on appelle le traitement massivement parallèle. Le framework Hadoop est sans doute le plus connu d'entre eux. Il combine le système de fichiers distribué HDFS, la base NoSQL HBase et l'algorithme MapReduce.
- Le stockage des données en mémoire (Memtables) permet d'accélérer les temps de traitement des requêtes.

Source : Delort, P. (2015) *Le big data. Collection Que sais-je ? P.U.F.*

D– Comment prendre une décision ?

D1 - Qu'est-ce qu'une décision ?

La décision est le fait d'un acteur (ou d'un ensemble plus ou moins cohérent d'acteurs) qui effectue un choix entre plusieurs solutions susceptibles de résoudre le problème ou la situation auxquels il est confronté (Wikipédia).

Un exemple de décision stratégique : les choix stratégiques

- cible et positionnement sur le marché (segments et niches)
- manœuvres stratégiques (domination par le prix, la différenciation, la focalisation, l'innovation, les alliances, l'internationalisation, le changement des règles du jeu, etc)
- formulation stratégique : le Business Model est une **représentation partagée** relative à la **Génération** de la valeur, à la **Rémunération** de la valeur, et au **Partage** de la valeur

Un exemple de décision stratégique : l'évaluation d'un projet

- les critères objectifs de l'évaluation (grille)
- l'influence de l'environnement (concurrence, dynamique, etc)
- l'influence du comportement stratégique (proximité, affinité, peurs, etc)

D2 - Les difficultés de la prise de décision

- La complexité de l'environnement rend la prise de décision difficile. Cette complexité est due à la multiplicité de ses éléments et de leurs interactions, mais aussi à la diversité de ses comportements dynamiques (Sternan, 2000).
- Dans les systèmes complexes, les actions menées créent souvent des effets qui diffèrent des résultats attendus et désirés, même lorsque les décideurs tentent d'agir au mieux en fonction des objectifs à atteindre.

D3 - Comment prendre la bonne décision ?

D'une manière générale, les processus de décision stratégique peuvent s'inscrire dans deux démarches processuelles distinctes : les démarches synoptiques versus les démarches incrémentales.

D4 - L'approche rationnelle du processus de décision

La démarche synoptique est considérée comme une extension du modèle traditionnel rationnel (Simon, 1945 – modèle IMC), qui se centre sur l'analyse. Elle requiert ainsi un fort degré de complétude tant en termes d'information que d'analyse.

Cette démarche se caractérise par les étapes suivantes :

1. Définir le problème

La première étape vers un processus de prise de décision consiste à définir le problème. Évidemment, il ne serait pas nécessaire de prendre une décision sans avoir un problème. Ainsi, la première chose qu'on a à faire est de poser le problème sous-jacent qui doit être résolu. Vous devez aussi indiquer clairement le résultat ou le but que vous désirez après avoir pris la décision. C'est une bonne façon de commencer, car en précisant vos objectifs vous aider à clarifier vos pensées.

2. Rechercher des alternatives

La situation de prise de décision se pose parce qu'il y a beaucoup d'alternatives disponibles. Par conséquent, la prochaine étape après avoir défini le problème principal serait d'indiquer les alternatives disponibles pour cette situation particulière. Ici, vous n'avez pas à vous limiter à réfléchir aux options évidentes, plutôt, vous pouvez utiliser vos talents créatifs et avoir des solutions qui peuvent sembler un peu hors-jeu mais importantes.

3. Identifier la meilleure alternative

Cela peut être considéré comme l'une des étapes les plus importantes du processus décisionnel. C'est l'étape où il faut analyser chaque option que vous avez trouvée. Vous devez connaître les avantages et les inconvénients de chaque option. Cela peut se faire que par la recherche que vous avez faite sur cette alternative particulière. A ce stade, vous pouvez également filtrer les options que vous pensez sont impossibles ou ne répondent pas à vos besoins. L'évaluation de chaque option avec un chiffre numérique vous aiderait également dans le processus de filtration.

4. Prenez la décision

C'est le stade où le travail acharné que vous aviez commencé prend fin. Le processus d'évaluation pourrait vous aider lorsque vous observez les options disponibles de façon claire et que vous avez à choisir ce qui vous semble les plus pertinentes.

5. Mettre en œuvre la solution

La prochaine étape évidente après avoir choisi une option serait la mise en œuvre de la solution. Juste prendre la décision ne donnerait pas le résultat qu'on veut. Au contraire, vous devez travailler sur la décision que vous aviez choisie. Il s'agit d'une étape très cruciale, car toutes les personnes impliquées dans la mise en œuvre d'une solution doivent avoir connaissance des implications de la prise de décision.

6. Surveiller votre solution

Juste prendre la décision et la mise en œuvre n'est pas la fin de la prise de décision, il est très important de surveiller régulièrement votre décision. A ce stade, vous devez garder un œil attentif sur l'évolution de la solution adoptée et si elle a abouti aux résultats que vous attendiez.

D5 - Les limites de l'approche rationnelle dans la prise de décision

- Les décisions dans les systèmes complexes sont généralement caractérisées par un fort degré d'incertitude, en ce sens que « les résultats des décisions ne peuvent être calculés avec certitude, étant donné que les états futurs du système sont difficilement prévisibles et que la complexité du système est trop élevée pour permettre le traitement de toutes les données » (Größler, 2004, p. 319).
- Les décideurs disposent rarement de toute l'information requise (Jones et Gross, 1996; Simon, 1945). Ils restreignent le nombre d'options considérées et l'information utilisée pour les évaluer (Krabuanrat et Phelps, 1998; Simon, 1945) et ne peuvent anticiper tous les champs de comportements possibles (Simon, 1945).
- Par conséquent, l'idéal synoptique ne semble que peu adapté aux difficultés que pose la résolution de problèmes complexes, étant donné l'imperfection de l'information, le coût de l'analyse (en termes de temps et de ressources) et la capacité limitée des individus à résoudre tous les enjeux interconnectés par une analyse globale et complète.
- Dans les systèmes complexes, l'action doit donc se focaliser sur des changements graduels et mineurs et les décideurs tendent ainsi à suivre une démarche incrémentale (intérêt de la recherche en sciences de gestion).

D6 - L'approche émotionnelle de la prise de décision

- Contrairement à Descartes qui considère que la raison et les émotions ne fonctionnaient pas ensemble mais en opposition (discours de la méthode), il existe l'hypothèse que les émotions jouent un rôle dans la prise de décision.
- Les émotions sont profondément associées à nos décisions. Elles sont un mouvement (*e movere*, en latin) qui aide le passage à l'acte. En fait, nous ne pourrions décider sans elles. Pourquoi? Parce que, contrairement à la vision cartésienne longtemps adoptée, notre centre de décision implique directement l'émotion comme source d'informations complémentaires à l'analyse, et aide au tri parmi les options possibles.
- Par exemple, vous l'avez sans doute noté, trancher est plus facile quand nous sommes mus par une émotion simple et forte (*Faut-il engager ce candidat ? Comment être sûr ? Ai-je raison de choisir cette stratégie commerciale ? Et si je me trompais ? Ce partenariat présente beaucoup d'avantages sur le papier, mais j'ai un doute; comment confirmer ma décision ?*).
- Le mécanisme de prise de décision inclue, de manière unique ou combinée, deux cheminements :
 - a) la raison avec l'analyse des faits, des options possibles, de leurs conséquences
 - b) l'émotion qui complète l'information via l'activation de la mémoire émotionnelle, avec les marqueurs somatiques
- L'émotion participe donc à la décision en réduisant les possibilités entrevues par la raison ; en d'autres termes elle lui simplifie le travail... ou elle limite ses choix, selon le cas !
- Les émotions permettent la cognition rapide, raccourci puissant face aux méandres du raisonnement parfois lent. C'est ce qu'on appelle **l'intuition**.

E - Comment analyser les données ?

Le traitement des données massives est un processus qui comprend plusieurs étapes : la collecte, la normalisation, la construction d'un entrepôt de données et l'analyse. Les caractéristiques des différentes étapes sont récapitulées dans le *tableau 4*.

étapes	Collecter	Normaliser	Stocker	Analyser
objectifs	extraire des données d'une source primaire	préparer les données pour être exploitables	mémoriser les données pour les rendre accessibles	traitement structural et statistique pour fournir des indicateurs
tâches	- acquisition de données auprès du BOAMP	- sélectionner - nettoyer - authentifier - dédoublonner	- construire un entrepôt de données	- indexer - classer - tester
output	données sources	données formatées	données validées	informations

Tableau 4: processus de traitement des données issues des marchés publics

Les marchés publics constituent un environnement complexe, lié notamment aux nombres d'acteurs, à la dynamique des transactions et aux interactions entre les entreprises. L'observation des marchés publics a donc nécessité un protocole particulier. Ce protocole (*figure 2*) s'appuie sur une observation documentaire, et il est composé de quatre étapes principales concernant les données : la collecte, la normalisation, le stockage

et l'analyse.

Avec plus d'une centaine de milliers de transactions entre des donneurs d'ordre publics (état, collectivités, établissements publics) et leurs fournisseurs, les données ouvertes sur les marchés publics fournissent chaque année une masse considérable d'informations. Ces données sont non structurées, incomplètes, dispersées dans différentes sources et ne respectent aucune règle de normalisation. L'enjeu de notre travail a consisté à transformer ces données, considérées comme des « big data », pour en extraire de l'intelligence.

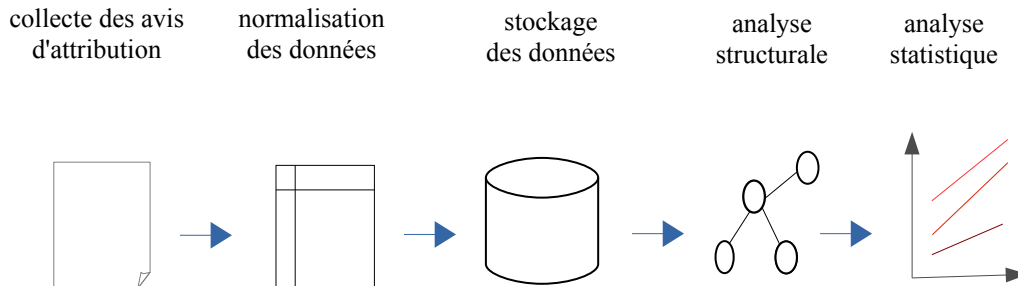


Figure 2: apport méthodologique

Source : Mamavi, O. (2015). *Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux*. L'harmattan, collection Entreprise et Management, Paris

F - Quel est l'intérêt de l'analyse des réseaux ?

F1 – L'analyse des réseaux

L'essor de l'Internet et du Web a popularisé l'observation de grands réseaux sociaux. Les nouvelles technologies de l'information permettent d'étudier et de surveiller tous les systèmes de relations, notamment : les réseaux de connaissance (deux individus sont reliés s'ils se connaissent), les réseaux de contact physique (deux individus sont reliés s'ils ont été physiquement en contact), les réseaux de collaboration (deux individus sont reliés s'ils ont travaillé ensemble), les réseaux d'échanges (deux entités sont reliées si elles ont échangé un courrier électronique par exemple).

L'analyse des réseaux sociaux considère la société comme un système d'individus reliés par des relations que l'on peut représenter sous la forme d'un graphe. Un graphe est un schéma constitué par un ensemble (supposé fini) de points et par un ensemble de flèches reliant chacune deux points. Les points, qui peuvent correspondre à des individus ou des structures, sont appelés « nœuds » et les flèches, « arcs » du graphe. Le graphe est donc un outil de visualisation des réseaux de relations, mais il est surtout un instrument pour calculer des propriétés relationnelles. On peut ainsi classer les nœuds et explorer les liens.

L'analyse des réseaux est donc une méthode adaptée pour comprendre et formaliser les phénomènes complexes faisant intervenir un système de relations interactif. En effet, cette méthode permet de décrire, et de reconstituer, un réseau, de façon simplifiée à travers un graphe. Le graphe représente les interactions entre des objets reliés par des liens.

Pour analyser les réseaux, trois grandes dimensions (*tableau 1*) peuvent être retenues .

1. La première consiste à identifier les réseaux et à décrire la manière dont les structures de ces réseaux font peser des contraintes sur leurs membres. La connexité est un indicateur qui permet de définir la frontière du réseau au sein d'un graphe. Il y a un réseau, s'il existe toujours un chemin pour relier deux sommets de l'ensemble. Un réseau est donc une composante connexe d'un graphe.
2. La seconde dimension permet d'identifier la position, plus ou moins dominante, d'un acteur dans le réseau. Elle peut s'évaluer à partir des approches de centralité définies par Freeman (Freeman 1979). La centralité de degré exprime la popularité d'un acteur dans le réseau, c'est-à-dire le nombre de

connexions directes d'une entreprise aux autres. La centralité de proximité identifie les entreprises les plus proches des sources de pouvoir et d'influence, c'est-à-dire rapidement joignables par l'ensemble des autres membres du réseau.

3. La troisième vise à dégager la cohésion des groupes homogènes au sein du réseau. Il s'agit d'analyser les similarités structurales dans le réseau. 2 acteurs sont donc structuralement équivalents s'ils ont des relations identiques avec les autres acteurs du réseau (White, 1981). Des techniques de partitionnement du réseau permettent de détecter des classes d'acteurs ayant une équivalence structurale.

Tableau 5 cadre d'analyse des réseaux

analyse des réseaux	propriété des liens	Clinques, n-cliques
		Intensité relationnelle
		équivalence structurale
	position des nœuds	centralité de degré
		Centralité d'intermédiarité
		centralité de proximité
	structure du réseau	Densité, distance
		connexité

Source : Mamavi, O. (2015). *Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux*. L'harmattan, collection Entreprise et Management, Paris

F2- Visualisation des données avec Gephi

Gephi (Bastian et al., 2009) est un logiciel open source pour une analyse et une représentation graphique des réseaux. Il utilise un moteur de rendu 3D pour afficher les réseaux en temps réel et accélérer leur exploration. Une architecture flexible et multitâche permet de travailler avec des ensembles de données complexes et de produire des résultats visuels intéressants, grâce à des fonctions de spatialisation et de navigation.

Source : Bastian M., Heymann S., Jacomy M. (2009). *Gephi: an open source software for exploring and manipulating networks*. International AAAI Conference on Weblogs and Social Media.

Télécharger le logiciel : www.gephi.org

F3 – Propriétés des grands réseaux

Il existe un ensemble de propriétés mathématiques qui ont été relevé lors de l'étude des grands graphes terrain (Strogartz, 2001). En effet, tous les grands réseaux sociaux possèdent tous :

- Le principe de l'attachement préférentiel (preferential attachment – Barabasi, 1999) constate que « les riches deviennent plus riches », c'est-à-dire que dans beaucoup de cas les nœuds les plus connectés gagnent aussi plus facilement des nouvelles connexions. L'attachement préférentiel se caractérise par une hétérogénéité des degrés et un faible degré moyen.
- L'effet du « petit monde » (Milgram, 1967) montre que dans une société de masse, pratiquement tous les individus sont reliés les uns aux autres dans un vaste réseau, et que la distance moyenne entre deux individus quelconques devrait être d'environ 5 intermédiaires. L'effet du « petit monde » se caractérise donc par des chemins courts entre les sommets.

- L'effet d'agglomération (clustering) correspond au principe suivant : mes amis ont une forte probabilité d'être eux-mêmes amis. L'effet d'agglomération (clustering) se caractérise par un réseaux avec une densité globale faible et une densité locale forte
- Le principe de l'équivalence structurale (White, 2008) s'illustre lorsque deux entités sont structurellement équivalentes au sein d'un réseau si elles y occupent la même place, ou la même position, c'est-à-dire si elles ont les mêmes relations (ou des relations relativement similaires) avec un ensemble d'acteurs données . La notion d'équivalence structurale permet donc de retrouver la notion classique de rôle (ou de position) mais d'un point de vue strictement structurel, par une analyse de réseau, sans faire d'hypothèse sur les contenus de ces rôles.
- La force des liens faibles est un concept popularisé par Granovetter (1973). Pour lui, ce sont leurs liens faibles qui procurent aux individus des informations qui ne sont pas disponibles dans leur cercle restreint : « Les individus avec qui ont est faiblement lié ont plus de chances d'évoluer dans des cercles différents et ont donc accès à des informations différentes de celles que l'on reçoit. »
- Les trous structuraux (Burt, 1992) correspond à une propriété d'intermédiaire au sein d'un réseau. Par exemple, une entreprise A, en relation avec deux autres B et C, sans que ces dernières ne le soient entre elles, dispose d'un avantage qu'elle est tentée d'exploiter à son profit. Cette position d'intermédiaire procure une autonomie à l'entreprise A. Les marchés sont donc organisés par le fait que des acteurs cherchent à accroître leur «autonomie structurale».

Source : www.publicprocurement.fr

F4 – Analyse prédictive à partir des réseaux

Au delà de la visualisation des critères d'analyse des réseaux, la démarche que nous envisageons pour construire un système de recommandation de liens sera la suivante (*figure 42*). A partir d'une étude longitudinale sur plusieurs années, nous pourrions prédire les nouveaux liens qui vont apparaître et qui relieront des entreprises déjà présentes dans les réseaux mais qui n'ont jamais été liés directement auparavant. A partir d'un échantillon d'apprentissage, des techniques d'*apprentissage supervisés* (Benchettara et al., 2010) pourraient, alors, être appliquées pour construire un modèle de prédiction de nouveaux liens dans les réseaux afin d'aider les entreprises à former des coalitions gagnantes.

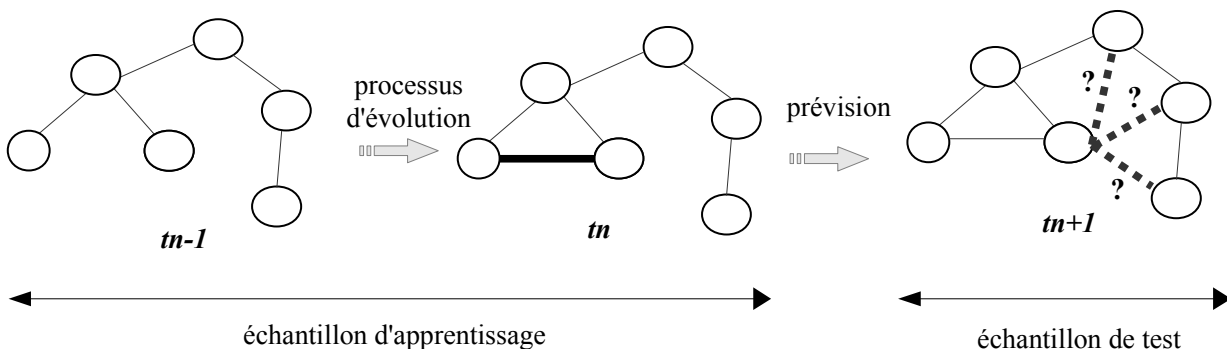


Figure 3: apprentissage automatique pour la prédiction de liens - source : adapté de Benchettara et al. (2010)

Source : Mamavi, O. (2015). *Les coulisses des marchés publics : manœuvres d'alliances et influence des réseaux*. L'harmattan, collection Entreprise et Management, Paris

F5 - La théorie de la force des liens

Selon Granovetter (1973), « *La force d'un lien est une combinaison (vraisemblablement non linéaire) du temps accumulé, de l'intensité émotionnelle, de l'intimité (confiance mutuelle) et des services réciproques qui caractérisent ce lien* ». La thèse principale de Granovetter (2000) est de démontrer que, dans les sociétés modernes, l'économie demeure encadrée dans des relations sociales qui en assurent le fonctionnement. Il développe, notamment, des analyses tout à fait originales du marché du travail, dans lesquelles il identifie les rapports sociaux et les institutions qui permettent à ce marché de fonctionner. Il montre, ainsi, que l'obtention d'un emploi est déterminée, en grande partie, par les relations personnelles. Il souligne l'importance de l'identité, non seulement des personnes que l'individu connaît et des relations qu'il a avec elles, mais aussi de l'ensemble des personnes connues par ses relations et ainsi de suite. La structure et la dynamique d'un tel réseau, aussi difficile que soit leur analyse, déterminent largement quelle information est à la disposition d'un individu et dans quelle mesure une opportunité s'offre à lui. Granovetter (2000) précise que la force d'un lien dépend de la quantité de temps, de l'intensité émotionnelle et des services réciproques qui caractérisent ce lien. Pour lui, les acteurs qui sont faiblement liés aux autres, sont plus susceptibles d'évoluer dans des cercles sociaux différents et sont, par conséquent, exposés à des informations différentes de celles dont ils peuvent disposer. C'est donc la force des liens faibles qui favorise l'accès à de nouvelles connaissances, la propagation rapide des informations, et permet ainsi la diffusion des innovations par exemple. L'argument de Granovetter est principalement structural. Il considère que les liens faibles constituent des ponts, donnant plus facilement accès à d'autres réseaux que son seul réseau de relations directes. Le concept de pont constitue le seul chemin possible reliant deux points.

A contrario des liens faibles, Granovetter (1973) considère que si un individu A a un lien fort, simultanément avec un individu B et un individu C, il est très probable que B et C se connaissent et aient, eux aussi, un lien. Dans ce cas, l'argument de Granovetter (1973) est également structural et repose sur un principe de transitivité. Autrement dit : plus il y a un lien fort entre deux individus, plus la proportion d'individus avec qui ils sont tous les deux liés est grande. Car, selon lui, plus le lien entre deux individus est fort, plus ces derniers se ressemblent de diverses manières (homophilie). En conséquence, des acteurs reliés par des liens forts supposent des contacts fréquents, une fourniture de services réciproques, et cela justifie le partage des ressources. En effet, comme l'expliquent Gulati et Singh (1998), les acteurs économiques ont tendance à échanger, en priorité, avec des partenaires qu'ils connaissent, et à accumuler de l'expérience, afin de limiter les risques inhérents à toutes transactions. Le principal intérêt de liens forts avec un partenaire est la confiance qui favorise un apprentissage organisationnel.

Source : Mamavi, O., Meier, O. et Zerbib, R. (2016). Alliance management capability: the roles of alliance control and strength of ties. Management Decision (à paraître)