

---

Module 2 : Analyse des controverses

# Le Marketing à l'ère du Big Data

---

Groupe 2  
Clara Baillehache  
Laetitia Cammas  
Drifa Kesraoui

A l'attention de

Maryse Carmes  
Mathilde de Saint Léger  
Anne-Christine Pinchemel

# Table des matières

---

## 1. Cadre du projet

### 1.1 L'analyse sectorielle

### 1.2 L'action de veille

## 2 Construction d'une controverse et hypothèse

## 3. Méthodologie de constitution du corpus

### 3.1 Les sources

### 3.2 Justification du choix des sources

### 3.3 Méthodologie de traitement du corpus

## 4. Lieux de débat et acteurs

## 5 Mise en évidence des positions d'acteurs : qui parle de quoi ?

## 6 Dynamiques : coalitions et divergences

## 7. Cartographie du débat

### 7.1 Extraction terminologique de Calliope

### 7.2 Diagrammes stratégiques issus de Calliope

### 5.5 Les Clusters du diagramme stratégique

### 5.6 Croisement de champs de données

## Conclusion

### ! Résultats

### Retours méthodologiques

### ? Ouvertures et anticipations

---

# Introduction

Les données issues du Big data ont profondément changé le monde du marketing et constituent aujourd'hui une mine d'informations exploitables pour le marketing décisionnel et prédictif<sup>1</sup> ou "PredictiveMarketing" centré sur l'analyse du besoin des client.

Certains considèrent le traitement des données en masse comme une opportunité, un relais de croissance pour les études de marché. Pour d'autres c'est un challenge car les professionnels du secteur doivent redéfinir leurs stratégies, développer de nouveaux outils et s'adapter à de nouvelles techniques d'analyses.

Les données publiques (ex : données démographiques), les données personnelles non structurées du web (ex: forum) et les données clients sont croisées, analysées puis restituées, générant des modèles révélateurs sous forme de graphiques analytiques. Le traitement de ce grand volume de données permet de dresser un profil précis du consommateur, d'anticiper son comportement d'achat et de lui proposer des services personnalisés en temps réel.

L'exploitation des données personnelles est cependant soumise à un cadre législatif en évolution depuis la loi Informatique et Libertés en 1978, encadrée 1995 par la directive européenne 95/46/CE relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à leur libre circulation conçue avant l'essor d'internet, jusqu'à la révision de cette même directive en raison de son obsolescence technologique et de l'explosion des services en ligne<sup>2</sup>.

La nouvelle directive 2013/37/UE<sup>3</sup> du 26 juin 2013 dont les dispositions seront appliquées par les États membres à partir de juillet 2015 répond à un objectif économique de compétitivité afin de contribuer à la mise en œuvre le marché unique digital pour 2015 et à renforcer les obligations de sécurité de protection des données à caractère personnel. Elle donne un cadre plus contraignant pour les responsables de protection des données en entreprise (binding corporate rules) et permet pour toute personne physique à titre individuel ou en action collective d'exercer ses droits à la protection de ses données à caractère personnel, y compris le droit à l'oubli et le droit à la portabilité.

L'application de cette directive est donc susceptible dans les années à venir de modifier les pratiques des professionnels du secteur du marketing en particulier pour les traitements visant à analyser ou prévoir la situation économique d'une personne, sa localisation, son état de santé, la vidéosurveillance, les données relatives aux enfants ou de nature génétique ou biométrique.

---

<sup>1</sup> SPSS PredictiveMarketing : du marketing décisionnel au prévisionnel, Carmeni N., E-Marketing [en ligne] publié le 01/05/2003 [Consulté le 22/01/2014]. <<http://www.e-marketing.fr/Marketing-Direct/Article/SPSS-PredictiveMarketing-du-marketing-decisionnel-au-previsionnel-11498-1.htm>>

<sup>2</sup> Big data vs protection des données personnelles : ce que Bruxelles vous prépare, Ferraud-Clandet N. Les Échos, [en ligne] publié le 01/12/2013. [Consulté le 22/01/2014].

<<http://lecercle.lesechos.fr/entreprises-marches/high-tech-medias/informatiques/221185595/big-data-vs-protection-donnees-personne>>  
<sup>3</sup> Directive du Parlement Européen et du conseil 2013/37/UE du 26 juin 2013, EUR-Lex, Portail d'accès au droit de l'Union Européenne <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:FR:PDF>> [Consulté le 22/01/2014]

Nous nous sommes intéressés à la diversité des débats, la complexité des échanges créant un véritable tissu dont la densité constitue la controverse. Pour mener à bien cette analyse de controverses, et comprendre les impacts de cette profonde mutation liée au Big data, nous avons constitué un corpus, issu de différentes sources traitant toute l'actualité des acteurs du marketing sur le web, dans la presse spécialisée du marketing et des nouvelles technologies, les sites des principaux acteurs du marketing, sur les blogs spécialisés en marketing. Nous avons eu recours à des outils fouille des données textuelles (Calliope).

Afin de représenter l'ensemble des enjeux et acteurs du secteur marketing nous avons développé une cartographie basée sur les positions et les arguments échangés.

# 1. Cadre du projet

## 1.1 L'analyse sectorielle

Notre analyse porte sur le secteur des « Études de Marché et Sondages », un secteur producteur de services aux entreprises. Cette sous classe du code Naf 73.20Z concerne notre sujet « **Le Marketing à l'ère du Big Data** ».

L'analyse de ce secteur nous a permis de faire un état des lieux de la conjoncture et du dynamisme des entreprises productrices de ces services : un secteur qui se porte bien et croît régulièrement, un chiffre d'affaire qui atteint un record historique en avril 2013 et semble poursuivre son mouvement de progression.

## 1.2 L'action de veille

L'action veille recentrée autour des nouvelles physionomies du web impactant le marketing, nous a permis de :

- Repérer les acteurs principaux des études marketing : les entreprises du secteur, les organisations professionnelles, la presse spécialisée, les salons et événements.
- Découvrir les nouvelles pratiques et tendances des consommateurs : achat en ligne, connections internet, abonnements, Internet mobile, réseaux sociaux.
- Découvrir les grandes mutations et les nouvelles pratiques de ce secteur liées au traitement des données en masse : nouveaux outils de mesure du comportement des consommateurs et nouvelles méthodes d'analyse (entrepôts de données, plate-formes, logiciels, analyse en temps réel).
- Connaître les enjeux, et les défis scientifiques et technologiques auxquels les acteurs de ce secteur doivent faire face à l'heure du Big data.

## 2 Construction d'une controverse et hypothèse

Nous avons défini un ensemble de propositions (questions et sous questions) avancées, provisoirement, comme explication des rapports socio-sémantiques dans le cadre du développement du Big Data. Nous avons, ultérieurement, contrôlé cet ensemble de propositions en appliquant une méthodologie de veille et d'analyse textuelle statistique du corpus.

### Problématique / Question principale :

- Quels sont les actants des relations socio-sémantiques en jeu ?
- Quels sont les enjeux des nouvelles pratiques mobiles des consommateurs ?
- Les nouvelles méthodes marketing ne sont elles pas omniprésentes et intrusives ?
- Les consommateurs : entre consentement et sentiment d'exploitation : Digital labor <sup>4</sup>
- La confiance des consommateurs est un élément charnière pour les acteurs du marketing mais selon ces mêmes acteurs quels en sont les critères ?

### Sous questions :

- Quelles sont les forces en présence qui déterminent le positionnement des acteurs de la profession ?
- L'évolution des pratiques du marketing liées au Big Data sont-elles compatibles avec le respect de la vie privée ?
- Les réseaux sociaux, qui représentent une opportunité pour le marketing, font l'objet de nouvelles méthodes d'analyses par le secteur du marketing. Cependant, les méthodes et les résultats ne font pas l'objet de validation scientifique.
- Les données personnelles ont-elles d'avantage vocation à être exploitées ou protégées ?
- Quels impacts l'évolution de la législation sur la protection des données personnelles aura sur les pratiques marketing?

### Hypothèse :

- **Les Internautes consommateurs** sont devenus le pivot de la nouvelle physionomie des études marketing. Cela marque la transition d'une dominance B to C vers le C to B reconnue sous la notion de "insight consommateur" (*consumer insight*).
- **La confiance des consommateurs est un enjeu majeur** pour les professionnels du secteur du marketing qui exploitent leurs données personnelles.

---

<sup>4</sup> Qu'est-ce que le digital labor ? , Antonio A. Casilli, BodySpaceSociety, [en ligne] Publié le 01/04/2013, [Consulté le 22/01/2014] <http://www.bodyspacesociety.eu/2013/03/26/slides-quest-ce-que-le-digital-labor/>

## 3. Méthodologie de constitution du corpus

### 3.1 Les sources

Le plan de veille nous a permis d'établir un Tableau de sourcing pour collecter et structurer les sources. Cela nous a permis de lister les sources selon le type de source auquel elles appartiennent, le type d'information accessibles, en donner une succincte description et noter l'accès à la source.

Un re-ciblage des sources sur la presse spécialisée et les acteurs eux-mêmes les entreprises ou organisations professionnelles (communication institutionnelle) a été utile pour élaborer le corpus qui servira dans le cadre de l'analyse des débats (traitement par *text mining* via Calliope).

Les sources sont constituées de :

- La presse spécialisée du marketing et des nouvelles technologies
- Les sites des principaux acteurs du marketing
- Les blogs spécialisés en marketing

Le Tableau de sourcing est accessible depuis la page de publication (Wiki) du projet :

<[http://ticri.univ-lorraine.fr/tp-intd.fr/index.php/Le\\_Marketing\\_à\\_l'ère\\_du\\_Big\\_Data#3.6\\_Identification\\_des\\_sources](http://ticri.univ-lorraine.fr/tp-intd.fr/index.php/Le_Marketing_à_l'ère_du_Big_Data#3.6_Identification_des_sources)>

### 3.2 Justification du choix des sources

Afin de comprendre les impacts de la mutation liée au traitement des données en masse (*Big data*), nous nous sommes attachés à connaître l'actualité des acteurs du marketing selon deux axes complémentaires :

#### Sources issues des professionnels du secteur du marketing

- Que disent les principaux acteurs du marketing eux-mêmes ?
- Sur quels points précis communiquent-ils (nouvelles méthodes, nouvelles structures, nouveaux acteurs) ?

#### Sources issues de la presse spécialisée

- Qu'est-ce que la presse spécialisée met en exergue ?
- Quelles sont les nouvelles pratiques couvertes par la presse spécialisée ?
- Quelle est la place des nouvelles technologies, du consommateur et des professionnels ?
- Y a-t-il des éléments de discours dans la presse spécialisée qui permettent de comprendre le sens de l'évolution de la relation triangulaire liant les professionnels, les entreprises client et le consommateur ?

### 3.3 Méthodologie de traitement du corpus

#### Méthode des mots associés et fouille de données textuelles :

##### Démarche :

La démarche consiste à rendre les données textuelles calculables pour faire des opérations de tri, de classification et de regroupement des informations utiles extraites du corpus.

##### But :

Il est question de réaliser une analyse thématique du domaine, vérifier les hypothèses de travail, détecter des thèmes émergents et des controverses (acteurs-réseaux, rapports de forces entre les actants), c'est à dire transformer les données textuelles en connaissances utiles, mises en perspectives selon une problématique.

##### Outil :

L'approche statistique via le logiciel Calliope permet de comptabiliser les mots associés dans le corpus et de comptabiliser leurs associations.

##### Méthode des mots associés :

Extraction d'une terminologie significative permet de mettre en évidence des co-occurrences entre les mots clefs, répartition des co-occurrences dans un diagramme stratégique (densité et centralité des clusters).

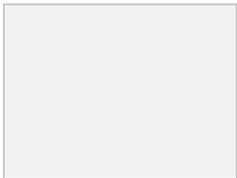
## 4. Lieux de débat et acteurs

<b>Lieux de débat</b>	<b>Actants:</b> - Acteur - intermédiaires - Acteurs-traducteurs (portes parole)
<b>WEB :</b> <b>1- Lieux d'expression publique des individus</b> (Forum, sites, blog, réseaux sociaux )  <b>2- Lieux de circulation des données</b> (Dépôt et récoltes des données sur les comportements des consommateurs, internautes, mobinaute, etc.)	<b>WEB:</b> <b>intermédiaires 1</b> , expression, avis : - Consommateurs, notes, avis et commentaires - Professionnels du marketing (ex. blog de pro)  <b>intermédiaires 2</b> , traces : - Données de géolocalisation, données sur les comportements d'achat, de paiement, de circulation,...

<p><b>Secteur économique du marketing :</b> Communication institutionnelle</p>	<p><b>Acteurs :</b></p> <ul style="list-style-type: none"> <li>- Cabinets d'études de marketing</li> <li>- Entreprises de Business Intelligence</li> <li>- Éditeurs de logiciels</li> <li>- Entrepôts de données</li> </ul> <p><b>intermédiaires :</b></p> <ul style="list-style-type: none"> <li>- Organisations professionnelles</li> <li>- Associations de consommateurs</li> <li>- Codes de déontologie</li> <li>- Codes de bonnes pratiques</li> </ul>
<p><b>Secteur de la recherche scientifique</b> Technologies et Sciences Humaines et Sociales (SHS)</p>	<p><b>Acteurs-traducteurs</b>, portes parole scientifique :</p> <ul style="list-style-type: none"> <li>- Laboratoires de recherche</li> </ul> <p><b>intermédiaires :</b></p> <ul style="list-style-type: none"> <li>- Publications académiques et scientifiques</li> </ul>
<p><b>Presse :</b></p> <ul style="list-style-type: none"> <li>- Presse spécialisée dans le domaine marketing</li> <li>- Presse spécialisée dans le domaine information et documentaion : pages technologie et web des grands journaux (qui vulgarisent les grands sujets et permettent aux citoyens de prendre part aux débats)</li> </ul>	<p><b>Acteurs-traducteurs</b>, portes paroles des observateurs de la sphère publique (médias) :</p> <ul style="list-style-type: none"> <li>- Publications d'actualités</li> <li>- Articles de presse</li> </ul>
<p><b>Instances officielles</b></p>	<p><b>intermédiaires :</b></p> <ul style="list-style-type: none"> <li>- Commission Nationale Informatique et Liberté (CNIL)</li> <li>- Ministères délégué chargé de l'économie sociale et solidaire et de la consommation</li> <li>- Ministères concernés (de la Recherche, de l'Économie numérique, du Redressement productif)</li> <li>- Insatances législatives nationales, européennes et internationales : législation sur la protection des données et le respect de la vie privée..</li> </ul>

## 5 Mise en évidence des positions d'acteurs : qui parle de quoi ?

Nous nous sommes attachés à mettre en évidence les positions des acteurs autour des méthodes et des impacts du traitements des données en masses dans le secteur du marketing. Il s'agit, pour rendre compte de la problématisation, de traduire un phénomène en éléments discursifs. Nous avons transcrit dans le tableau ci-dessous les éléments positifs (en vert) ou négatifs (en rouge) du discours classés selon la nature des argument (colonne centrale) et les sphères émettrices de ces discours (colonnes à l'extrême gauche et à l'extrême droite du tableau ci-dessous).

Qui parle?	+	Nature de l'argument	-	Qui parle?
		<i>Méthodologique</i> <i>Scientifique</i>	<p>Le manque de scientificité des méthodes de traitement des données.</p> <p>La rapidité de l'évolution des techniques et le manque de recul pour observer et extraire une information scientifique autour de ces nouvelles pratiques des cabinets d'études.</p>	<b>Sphère scientifique, Sciences Humaines et Sociales (SHS)</b>
<b>Sphère scientifique et technologique</b>	<p>Le contexte est favorable au développement de la recherche.</p> <p>Innover scientifiquement et techniquement est une nécessité pour la société</p>	<i>Technologique</i> <i>Scientifique</i> <i>Économique</i>		
<b>Cabinets d'études Marketing</b>	<p>Le contexte offre une meilleure connaissance des comportements des consommateurs (notamment en temps réel).</p> <p>Le contexte permet une meilleure connaissance des mouvements d'opinion (selon de nouvelles granularités).</p> <p>L'analyses des données en temps réel est un enjeu très fort.</p>	<i>Méthodologique</i> 	Le contexte contraint les professionnels du marketing à une révision de leurs méthodes de mesure (ils ne peuvent plus se baser sur les panels de population).	<b>Cabinets d'études et de marketing</b>
<b>Cabinets d'études Marketing</b>	<p>Le recueil passif des données représente un avantage pour l'acteur.</p> <p>La masse de données</p>	<i>Économique</i> <i>Politique</i>	L'internaute est exploité par les agents de marketing à travers ses données ( <i>Digital Labor</i> ). En effet, si	<b>Sphère publique :</b> - Association de consommateurs

	sur les comportements des consommateurs se développe de façon exponentielle.		les données personnelles ont une valeur économique, le consommateur ne devrait-il pas être rémunéré pour leur exploitation?	
<b>Amont du secteur:</b> Entrepôts de données	Le développement d'une industrie de stockage des données en masse est une opportunité économique,	<i>Économique</i> 	Le coût du stockage de données est un coût croissant .	<b>Secteur économique du marketing :</b> - Cabinets d'études de marketing - Entreprises utilisant les services des études de marketing
<b>Secteur économique du marketing :</b> - Acteurs du marketing - Service marketing des Entreprises	Le développement de nouveaux outils (solutions) est une opportunité de développement et de renouvellement.	<i>Économique</i> 	Le challenge est difficile à relever, il n'est pas facile pour les acteurs du marketing de rester compétitif.	<b>Média spécialisé dans ce secteur</b>
		<i>Économique</i>	Le coût des restructuration des services pèse sur les petites structures d'études de marketing.	<b>Observateurs des entreprises :</b> - Insee - Xerfi
<b>Sphère publique :</b> - Consommateurs - Associations de consommateurs	Le conseil personnalisé et les recommandations ciblées sont des services très appréciés et utilisés régulièrement. Les attentes et les besoins des consommateurs sont mieux ciblés.	<i>Relation entre le consommateur et les professionnels du marketing</i> 	L'omniprésence des capteurs des comportements et des opinions est vécue comme anxiogène et développe chez les usagers méfiance et sentiment d'être "tracé" .  La technologie va plus vite que la loi.	<b>Sphère publique :</b> - Consommateurs - Associations de consommateurs
<b>Secteur économique du marketing :</b> - Les cabinets d'études de marché	il existe une législation qui protège les données personnelles et la vie privée.  Codes de déontologie et bonnes pratiques au sein de la profession.	<i>Relation entre les législateurs et les professionnels du marketing</i> 	Difficulté de coordination des législateurs au niveau international  Les frontières du web ne sont pas les frontières territoriale.	<b>Instances officielles de contrôle :</b> - CNIL - Législation / Protection des données : Directive européenne 2013/37/UE du 26 juin 2013 modifiant la directive 2003/98/CE concernant la

				réutilisation d'informations détenues par le secteur public <sup>5</sup> .
<b>Instances officielles :</b> Ministères : Économie numérique, recherche, Redressement productif	Volonté de développer d'innovation  Nécessité de rester compétitif	<i>Technologique et économique</i>  		

Légende :  = éléments conflictuels, polémiques ou en contradiction.

<sup>5</sup> Open data : une nouvelle directive européenne modifie le cadre juridique, Cousin A., JDN, [en ligne] publié le 22/08/2013, [consulté le 22/01/2014]  
<<http://www.journaldunet.com/ebusiness/expert/55029/open-data---une-nouvelle-directive-europeenne-modifie-le-cadre-juridique.shtml>>  
Big data vs protection des données personnelles : ce que Bruxelles vous prépare, Ferraud-Clandet N. Les Échos, [en ligne] publié le 01/12/2013 [consulté le 22/01/2014].  
<<http://lecercle.lesechos.fr/entreprises-marches/high-tech-medias/informatiques/221185595/big-data-vs-protection-donnees-personne>>

## 6 Dynamiques : coalitions et divergences



### Quelles divergences?

- Tension idéologique :
  - Technophobie **VS** Technophilie
  - Sphère publique **VS** Sphère privée
- Tension économique :
  - Opportunité de développement **VS** Coût des restructurations
- Tension scientifiques :
  - Remise en cause de la scientificité des méthodes liées au traitement des données en masse **VS** La volonté et la nécessité de développer l'économie à travers l'innovation l'expérimentation
- Tension juridique :
  - Respect de la vie privée, protection des données et droit à l'oubli **VS** Mouvement naturellement expansif de la multiplication des traces, recoupement et regroupement de données de sources hétérogènes (géolocalisation, transactions d'achats, données administrative, données médicales...)

### Quelles coalitions ?



- Coalition cabinets d'études marketing + Entreprises clients :
  - La nécessité de s'emparer de ce marché
  - La prépondérance de l'analyse en temps réel sur le web 2.0 comme mode d'analyse des mouvements d'opinion
- Coalition citoyen + législateurs :
  - Le droit à l'oubli
  - Le respect de la vie privée
  - La protection des données personnelles
- Coalition sphère scientifiques + sphère institutionnelle + cabinets d'études marketing :
  - L'intérêt de l'exploitation des données en masse
  - L'explosion des données non structurées avec le web 2.0
  - La prépondérance des réseaux sociaux et des données non structurée pour l'analyse d'opinion

## 7. Cartographie du débat

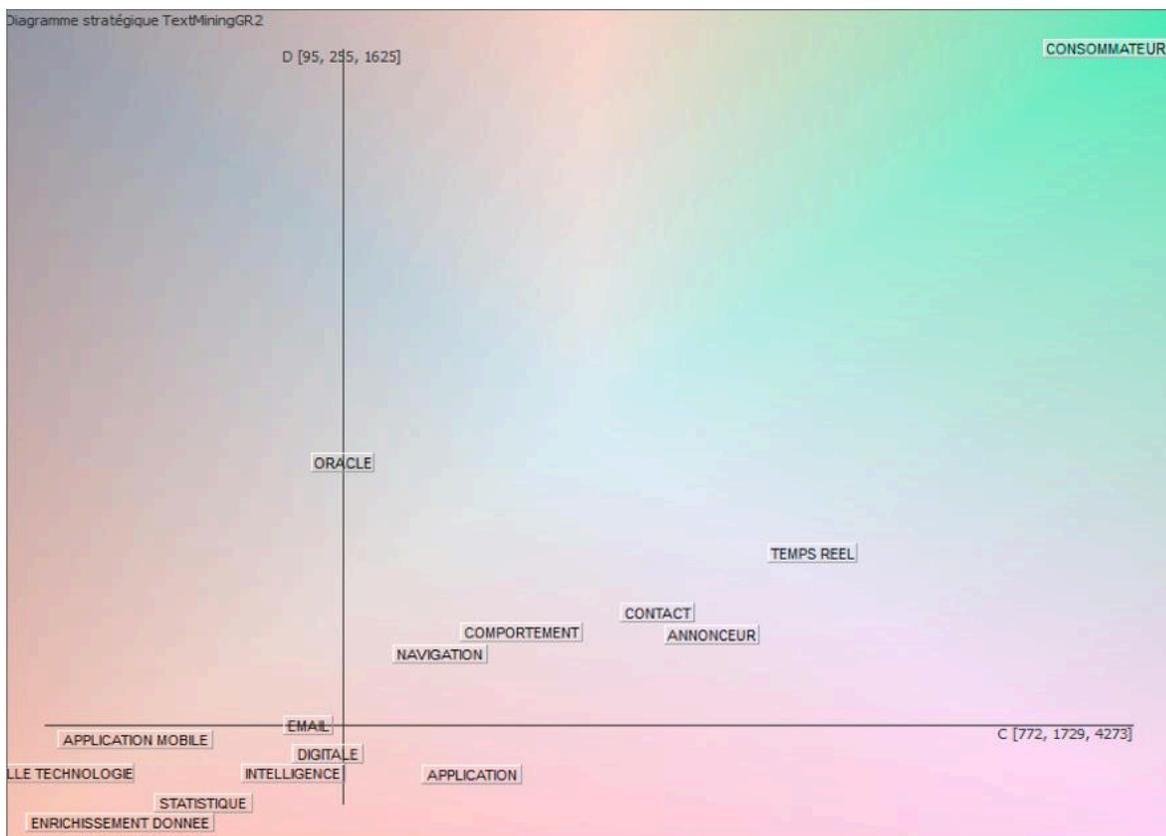
### 7.1 Extraction terminologique de Calliope

La démarche induite par ce choix d'outil, permet de construire un "dictionnaire de terminologies" (issu d'une extraction terminologique) en partant du texte intégral des articles du corpus. Des groupes de mots pertinents sont extraits du corpus (par la méthode des segments répétés) et rendus homogènes (par lemmatisation et construction d'un dictionnaire de report de mots) afin de constituer le lexique de référence.<sup>6</sup>

Les mots clefs sont regroupés en clusters selon le principe de co occurrences de ces mots dans les articles du corpus. Les clusters sont ensuite distribués sur un diagramme stratégique.

### 7.2 Diagrammes stratégiques issus de Calliope

Un diagramme stratégique est construit selon l'axe des abscisses (médiane de la centralité) et l'axe des ordonnées (médiane de la densité). La distribution des clusters dans les différents quadrants donne alors une synthèse de la dynamique du domaine qu'ils représentent.<sup>7</sup>



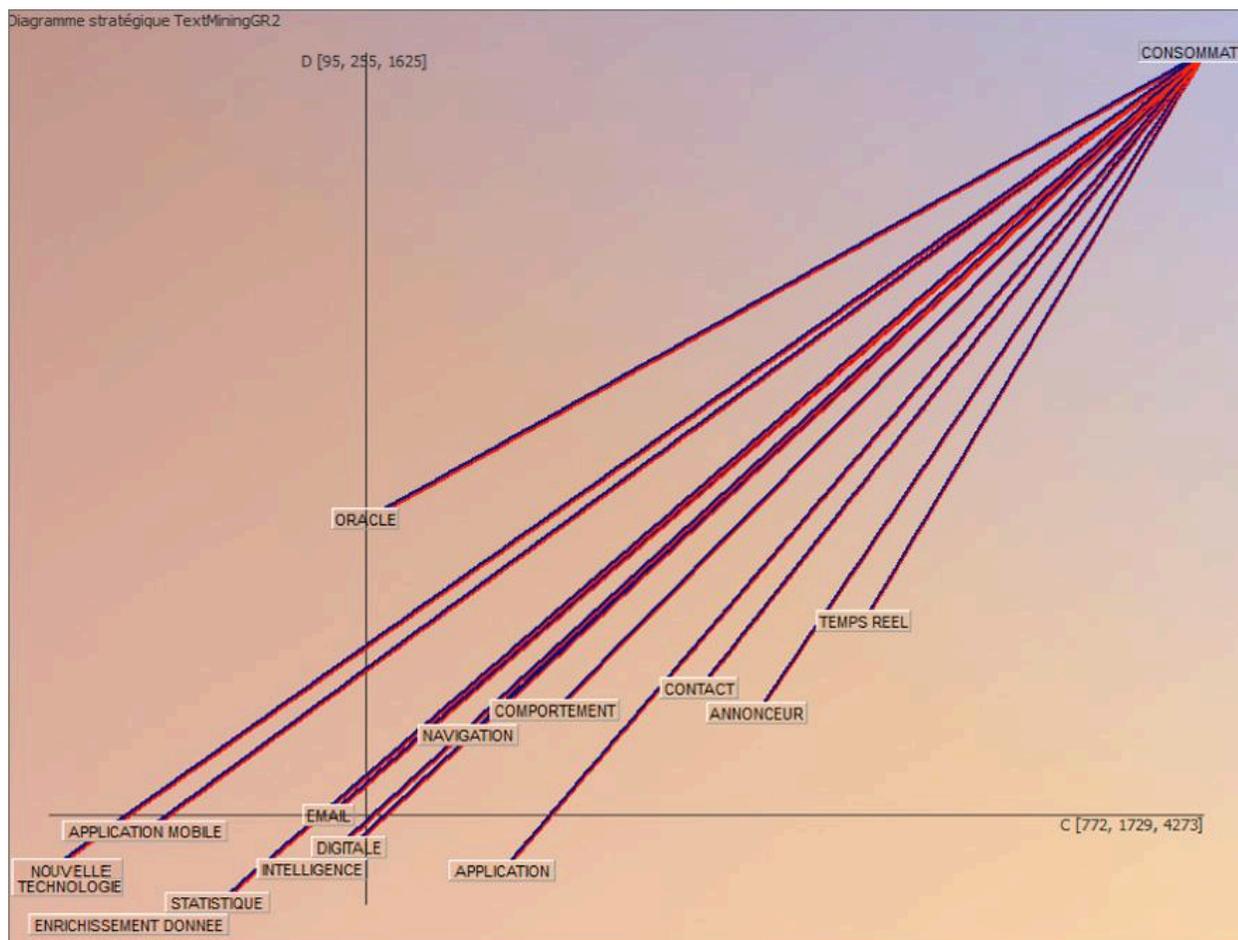
<sup>6</sup> Mathilde de Saint Léger et Karl M. van Meter , *Cartographie du premier congrès de l'ASF avec la méthode des mots associés*, Bulletin de méthodologie sociologique, 85 | 2005, [en ligne], Mis en ligne le 08 juillet 2008. Consulté le 13 janvier 2014, <<http://bms.revues.org/1050>>.

<sup>7</sup> Mathilde de Saint Léger , « Comment ont évolué les thématiques des 99 premiers numéros de *BMS* ? », *Bulletin de méthodologie sociologique*, 100 | 2008, [en ligne]. Mis en ligne le 01 octobre 2011. Consulté le 13 janvier 2014, <<http://bms.revues.org/3153>>.

Les clusters issus du corpus “Le Marketing à l’ère du Big Data” sont majoritairement distribués selon une diagonale S-O / N-E, des termes périphériques ou émergeant vers des termes fédérateurs.

Cependant deux termes se distinguent :

- ORACLE est plus isolé mais très structuré. Il a quitté la sphère des termes dominant pour intégrer la sphère des termes spécialistes.
- APPLICATION est dans la sphère des termes connectés mais peu structurés.



Le cluster CONSOMMATEUR se détache des autres sur le diagramme stratégique; c’est un terme attracteur.<sup>8</sup>

<sup>8</sup> Mathilde de Saint Léger , « Comment ont évolué les thématiques des 99 premiers numéros de BMS ? », *Bulletin de méthodologie sociologique*, 100 | 2008, [en ligne]. Mis en ligne le 01/10/2011. Consulté le 14/01/2013. <<http://bms.revues.org/3153>>.

## 5.5 Les Clusters du diagramme stratégique

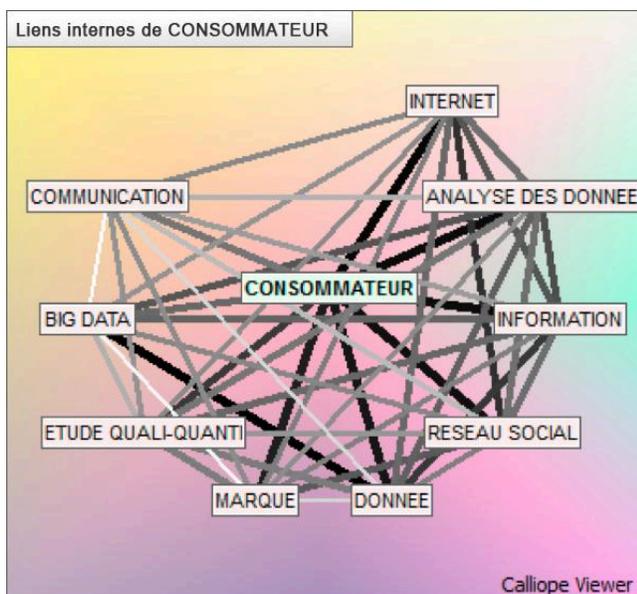
Au-delà d'un diagramme stratégique, la cartographie donne une vision globale du corpus analysé, il est nécessaire de *zoomer* sur le contenu de chaque cluster pour avoir une vision détaillée et donc une analyse plus fine des thématiques du corpus.

Quatre clusters du diagramme stratégique de cette période arrivent en tête sur le diagramme stratégique, il s'agit des clusters CONSOMMATEUR, COMPORTEMENT, ANNONCEUR et NAVIGATION. Ils sont les plus centraux et les plus denses.

E-MAIL se démarque dans le cadran sud-est, c'est à dire un terme périphérique émergent, central et structuré, ORACLE est à la frontière du cadran nord-est c'est à dire plus spécialiste.

Les clusters sont présentés ci-dessous dans l'ordre croissant de leur densité.

### 5.5.1 Cluster CONSOMMATEUR



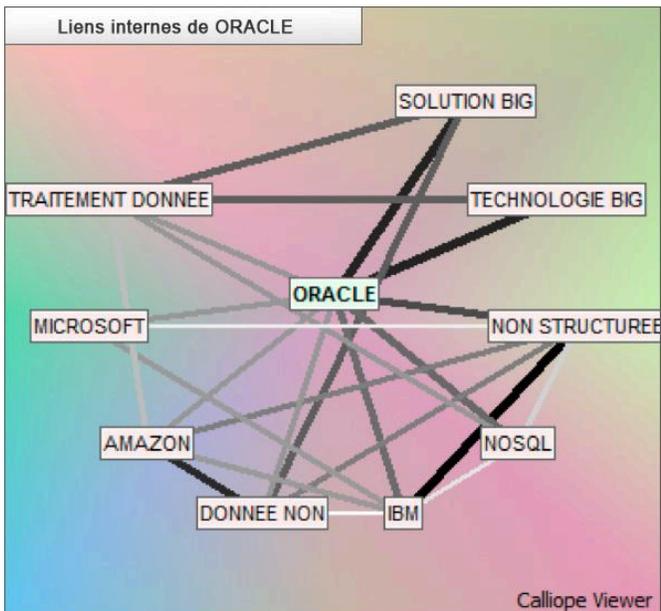
Le terme le plus fort (plus forte densité et plus forte concentration) laisse apparaître une grande complexité des relations intra-cluster (grand nombre de co occurrences).

Les liens les plus denses vont vers :

- INTERNET
- ANALYSE DE DONNÉES
- INFORMATION
- RÉSEAU SOCIAL
- DONNÉES
- MARQUE

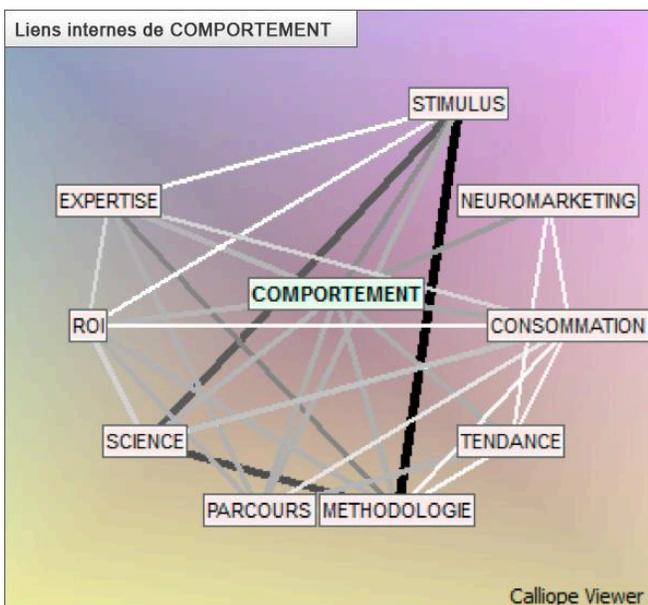
Notons également la présence de la valeur forte "Big Data".

### 5.5.2 Cluster ORACLE



ORACLE(10) à une place important dans le 1er graphique (noeud de cluster en nord ouest) . Alors que IBM (48) ou MICROSOFT (8) n'apparaissent pas. ORACLE a des liens de co occurrences plus forts avec les mots clefs du sujet que IBM ou MICROSOFT qui renvoient à des prestataires plus généralistes que l'éditeur de solution Oracle. Oracle est un acteur spécialisé qui se démarque dans le secteur, et il présente de nombreuses co occurrences avec les autres termes.

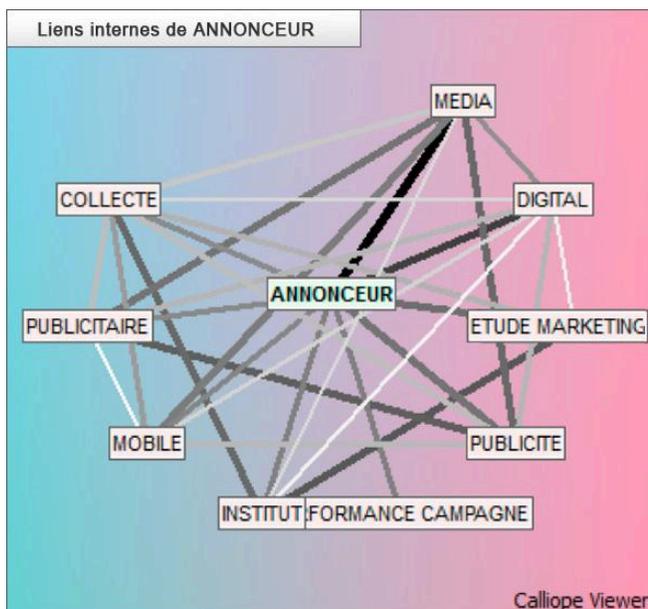
### 5.5.3 Cluster COMPORTEMENT



Les relations intra-cluster autour du terme COMPORTEMENT sont faiblement charpentés (prédominances de liens faibles en étoile vers STIMULUS, NEUROMARKETING, CONSOMMATION, PARCOURS , ROI, EXPERTISE) à l'exception des trois plus fortes dynamiques en trinagle et périphériques :

- STIMULUS / METHODOLOGIE
- METHODOLOGIE / SCIENCE
- SCIENCE / STIMULUS

#### 5.5.4 Cluster ANNONCEUR



Le cluster ANNONCEUR présentent des liens forts et en étoile.

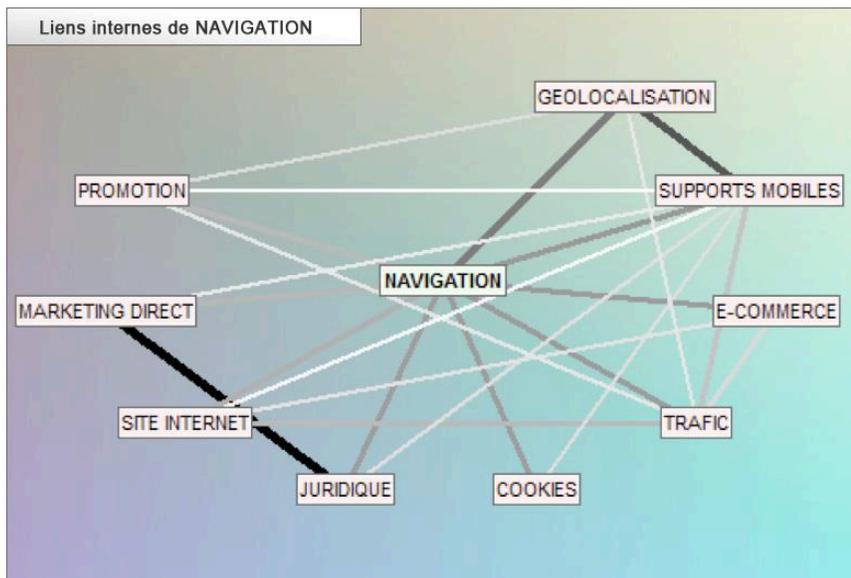
Les liens les plus fort mettent en relief deux axes :

- ANNONCEUR / MEDIA
- ANNONCEUR / DIGITAL.

Deux liens assez forts marquent des axes périphériques :

- ETUDE MARKETING / INSTITUT( de sondage, *ndlr*)
- INSTITUT( de sondage, *ndlr*) / COLLECTE

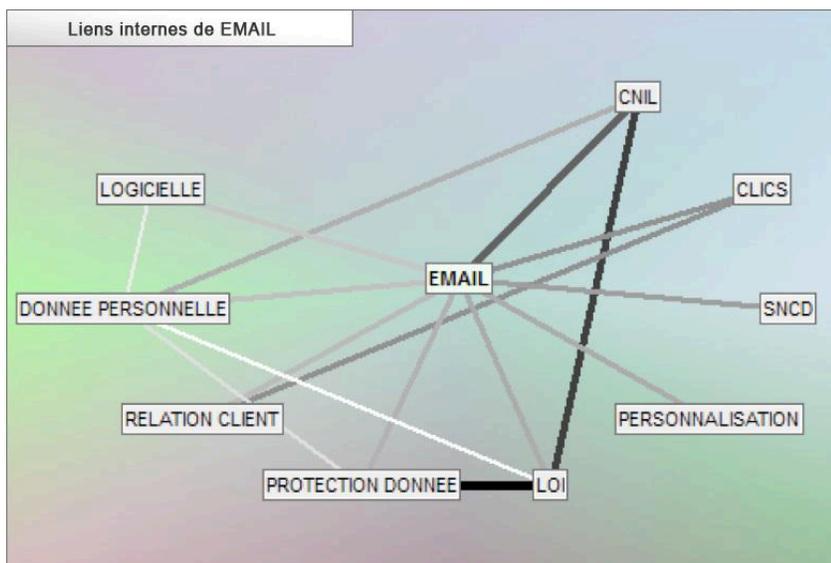
### 5.5.5 Cluster NAVIGATION



Le cluster NAVIGATION est peu charpentés (liens intra-cluster faibles) mais cependant polarisés par deux axes parallèles forts :

- · GEOLOCALISATION / SUPPORTS MOBILES
- · JURIDIQUE / MARKETING DIRECT

### 5.5.6 Cluster E-MAIL



Si dans son positionnement dans le diagramme stratégique, le cluster EMAIL est neutre (à la conjonction des axes), les dynamiques intra-cluster les plus fortes révèlent l'importance des liens de ce

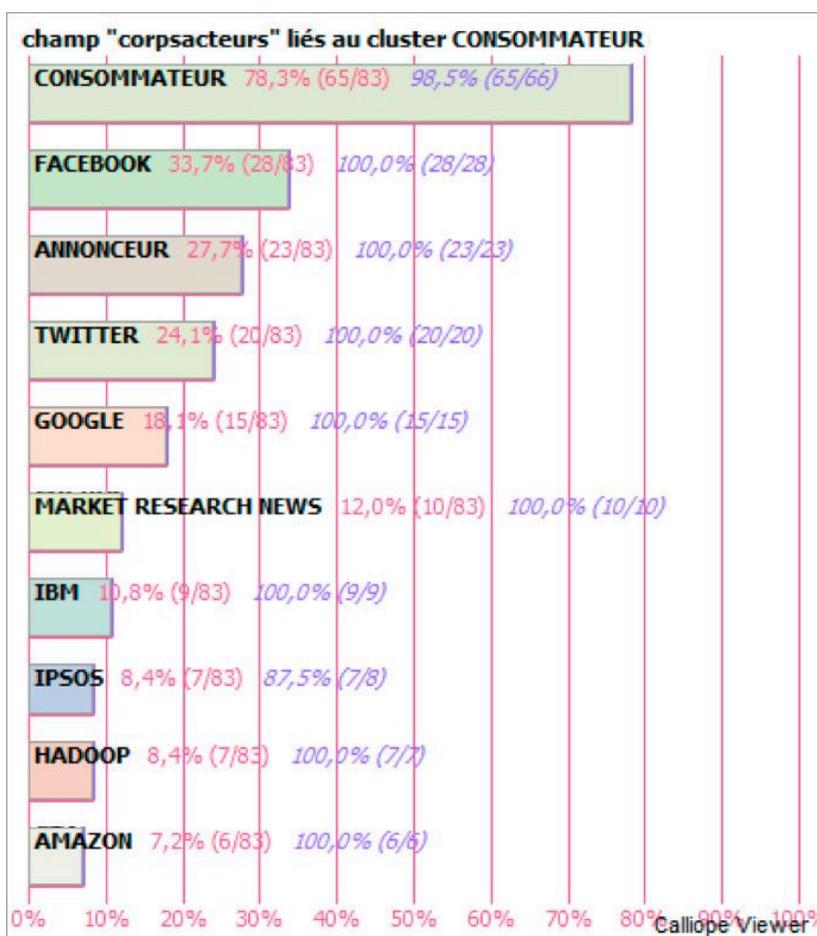
terme avec PROTECTION DES DONNEES, LOI et CNIL. Le terme EMAIL , véhicule de données personnelles, est porteur de tensions socio-sémantiques au coeur notamment de la sphère législative.

## 5.6 Croisement de champs de données

Une nouvelle extraction terminologique représentant les acteurs a été effectuée ainsi qu'une nouvelle indexation du corpus dans le champs "Corpsacteur".

Cette démarche permet de faire apparaître de façon statistique les Acteurs-Traducteurs (sociologie de la traduction) dotés d'une forte visibilité.

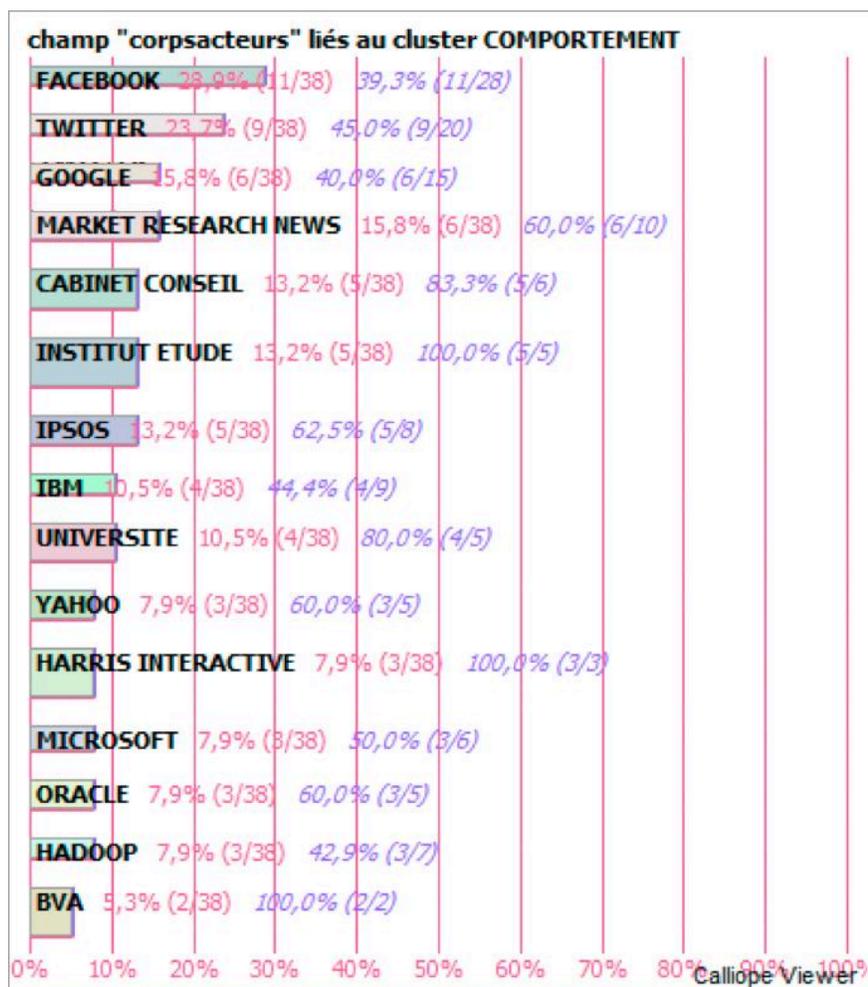
### 5.6.1 Acteurs-Traducteurs liés au cluster CONSOMMATEUR



Dans le cluster CONSOMMATEUR, les principaux acteurs cités sont :

- CONSOMMATEUR
- FACEBOOK
- ANNONCEUR
- TWITER

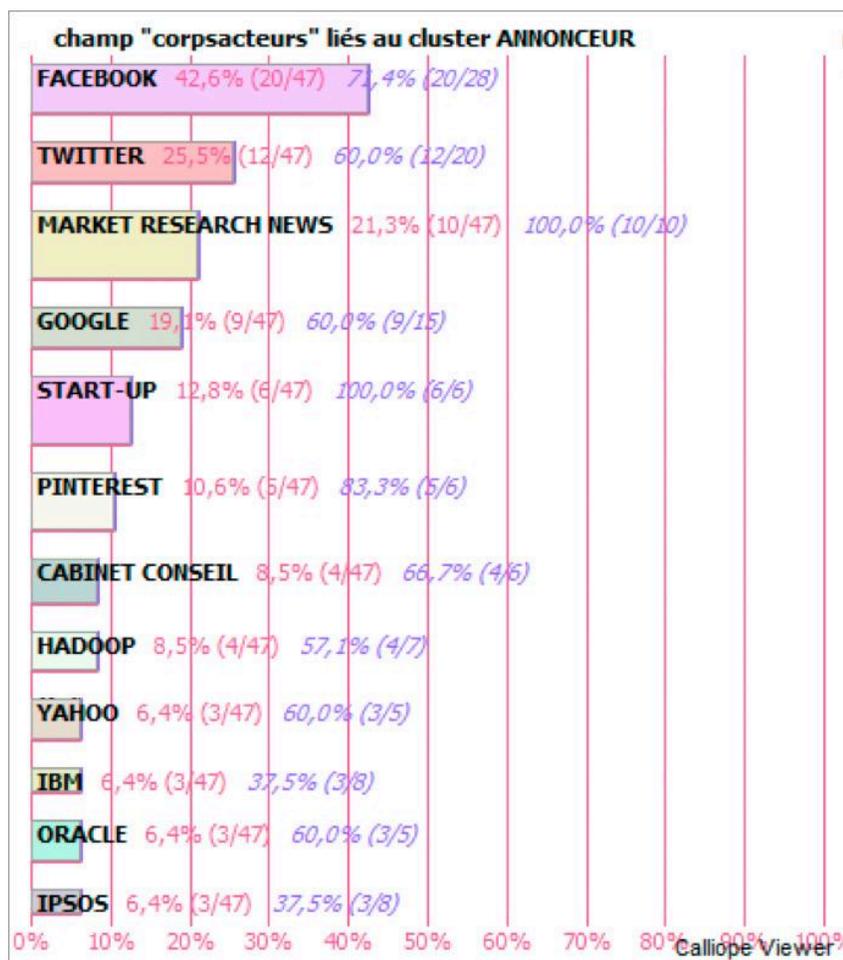
## 5.6.2 Acteurs-Traducteurs liés au cluster COMPORTEMENT



Dans le cluster COMPORTEMENT, les principaux acteurs cités sont :

- FACEBOOK
- TWITTER
- GOOGLE
- MARKET RESEARCH NEWS

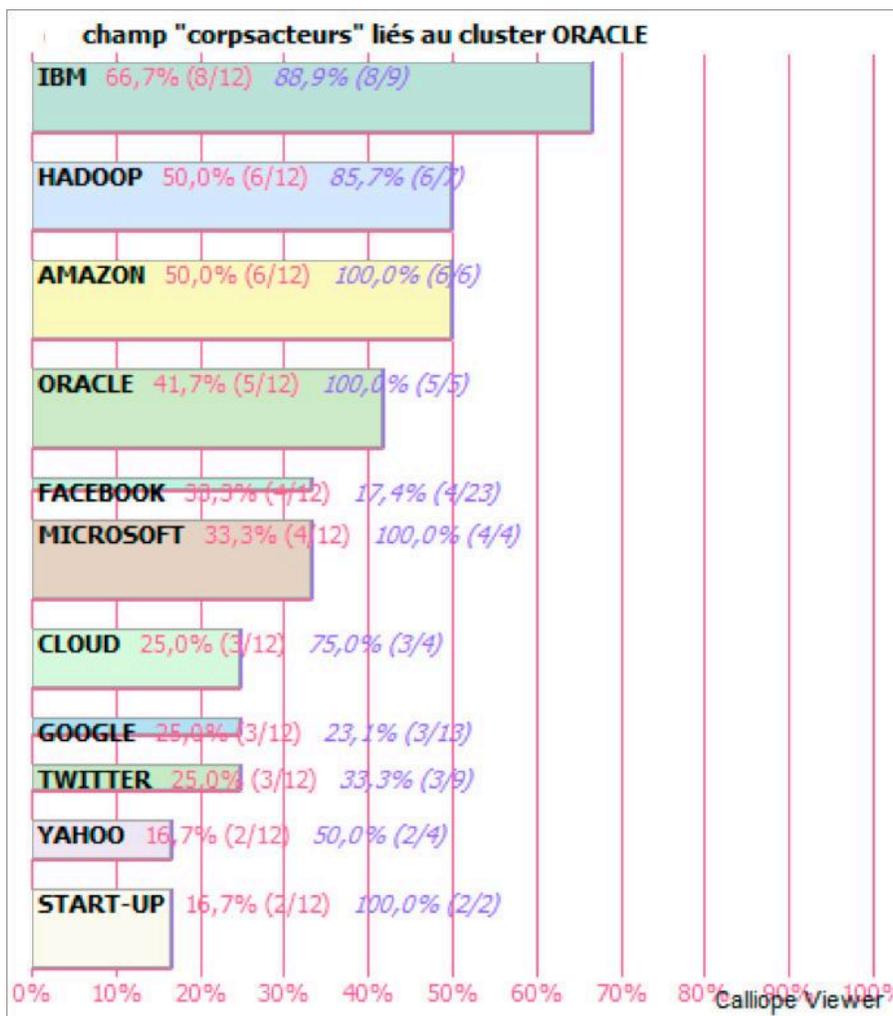
### 5.6.3 Acteurs-Traducteurs liés au cluster ANNONCEUR



Dans le cluster ANNONCEUR, les principaux acteurs cités sont :

- FACEBOOK
- TWITTER
- MARKET RESEARCH NEWS
- GOOGLE

## 5.6.4 Acteurs-Traducteurs liés au cluster ORACLE



Dans le cluster ORACLE, les principaux acteurs cités sont :

- IBM
- HADOOP
- AMAZON
- ORACLE

# Conclusion

## ! Résultats

- **Les cabinets d'études de marketing** utilisent les nouvelles méthodes de *data mining* et développent l'analyse des données en temps réel (*real time marketing*), ce qui représente une réelle opportunité de développement et également un enjeu majeur.
- **Les entreprises client** apprécient ces méthodes d'analyses de comportements et de flux d'opinions en temps réel qui deviennent des supports pour de nouvelles stratégies commerciales.
- **Les internautes consommateurs** sont devenus le pivot de la nouvelle physionomie des études marketing. Cette transition d'une dominance B to C vers le C to B souligne le rôle central de l'expérience du consommateur (*consumer insight*).

Les relations triangulaires entre les professionnels du marketing, les entreprises client et l'internaute consommateur sont en pleine mutation du fait du développement du traitement des données en masse.

## Retours méthodologiques

Les différentes méthodes et approches mobilisées, choisies par l'équipe dans le cadre du projet de veille:

- **Apports**
  - la méthode des mots associés permet de visualiser les dynamiques entre des mots-clés thématiques et des mots-clés désignant les acteurs. L'extraction terminologique est une démarche à la fois automatique et intellectuelle.
- **Limites**
  - l'approche statistique nécessite un corpus de taille conséquente.
- **Remise en cause des hypothèses**

Le corpus pourrait être d'avantage élargi aux codes de déontologie des professionnels du secteur afin d'identifier les réseaux sociaux-sémantiques liés aux critères de confiance, de protection et sécurisation des données personnelles.

## ? Ouvertures et anticipations

Comment cette controverse, replaçant le consommateur au centre des intérêts de la sphère économique et marketing, peut-elle vraisemblablement évoluer, compte tenu des équilibres mouvants en jeu ?

- D'avantage de fiabilité des méthodes marketing et de scientificité des techniques et des analyses de fouilles de données.
  - ex. : Quel est l'indice de fiabilité des notations et commentaires des usagers sur le web ?
  - ex. : Une nouvelle granularité des populations (tweet, commentaires...) en défaveur des panels de consommateurs classiques (CSP).
  - ex. : La prolifération des solutions marketing.
- Stabilité de l'opacité des nouvelles méthodes d'analyses de donnée<sup>9</sup>.
  - ex. : Multiplication des offres : IBM, Médiamétrie, Nielsen ...
  - ex. : La complexité des algorithmes.
- Le Big Data est-il vraiment le nouvel *Eldorado* des cabinets d'études de marché ?
  - ex. : La rentabilité économique des nouvelles offres regroupant des plate-formes, des logiciels et des services d'analyses de données.
  - ex.: Confusion dans la nature des offres pour les clients.
- L'utilisation généralisée des traces numériques des internautes (et leur traitement par recoupement et regroupement) peut-elle se faire dans respect la vie privée ?
- L'évolution et l'usage exponentielle des analyses en temps réel.
  - ex. : Croisement des données d'affiliation des consommateurs avec les données de géolocalisation, développement de stratégie multi-canal et cross-canal.

**La relation Marketeur - Entreprise client - Internaute consommateur est en pleine évolution et n'a pas encore trouvé un équilibre.**

Malgré l'évolution de la législation autour de la protection des données et du respect de la vie privée, ces nouvelles méthodes marketing laissent apparaître une omniprésence autour des traces numériques des internautes.

Comment les Cabinets d'études marketing vont-ils faire face à la méfiance croissante des consommateurs?

Les données personnelles auront-elles un nouveau statut lié à leur valeur marchande sur le net ?

Les législateurs pourront-ils toujours contraindre les acteurs économique en faveur du respect de la vie privée ?

La courbe de développement de la technologie va-t elle atteindre un palier de stabilisation ?

---

<sup>9</sup> Boullier D., Lohard A., 2012, *Opinion mining et sentiment analysis - Méthodes et outils*. OpenEdition Press, , Sciences Po | médialab, Chap 2. Les controverses autour du sentiment analysis , [en ligne]. Consulté le 14/01/2014, <<http://www.jstor.org/stable/1879431>>.