



**HAL**  
open science

**Discours, objets et imaginaires autour des technologies  
big data : le cas des consultants en technologie, la  
construction d'une expertise à travers la raison  
algorithmique**

Alexandra Ducrot

► **To cite this version:**

Alexandra Ducrot. Discours, objets et imaginaires autour des technologies big data : le cas des consultants en technologie, la construction d'une expertise à travers la raison algorithmique. Sciences de l'information et de la communication. 2017. dumas-02414455

**HAL Id: dumas-02414455**

**<https://dumas.ccsd.cnrs.fr/dumas-02414455>**

Submitted on 16 Dec 2019

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

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

Copyright

## Master professionnel

Mention : Information et communication

Spécialité : Communication Médias

Option : Médias et numérique

### Discours, objets et imaginaires autour des technologies Big Data

*Le cas des consultants en technologie, la construction d'une  
expertise à travers la raison algorithmique*

Responsable de la mention information et communication  
Professeure Karine Berthelot-Guiet

Tuteur universitaire : Antoine Bonino

Nom, prénom : DUCROT Alexandra

Promotion : 2016-2017

Soutenu le : 18/10/2017

Mention du mémoire : Très bien



## REMERCIEMENTS

*Je remercie d'abord mon tuteur universitaire, Antoine Bonino, pour son suivi, ses conseils avisés qui ont su m'orienter dans ma réflexion et m'aider à construire celle-ci.*

*Je remercie aussi mon tuteur professionnel, Jérôme Denis, qui s'est rendu disponible lors de mes premiers tâtonnements réflexifs sur mon sujet et dont les cours en Master 2 m'ont été d'une grande aide pour bâtir mes pistes de réflexion autour du rapport entre technologies et société.*

*Je remercie enfin toutes les personnes avec qui j'ai eu l'occasion d'échanger à propos de ma problématique sur les technologies big data.*

*Mes remerciements s'adressent enfin au corps enseignant et pédagogique du CELSA.*

## RÉSUMÉ

Le terme « big data » étant souvent utilisé comme un terme étandard pour « faire valeur » dans les discours des acteurs qui le mobilisent, ce travail de recherche étudie les discours d'escorte des consultants en « technologie big data ». Autour de quels éléments de langage construisent-ils leur discours commercial pour promouvoir leur offre de services à leurs différents prospects (les entités dirigeantes des entreprises ou institutions) ? Il s'agit d'analyser comment ils légitiment leur discours et construisent une expertise à partir des « technologies big data ». L'analyse révèle aussi la construction d'un « ethos algorithmique », par lequel ces cabinets de conseil vont venir opérer différentes médiations au sein des entreprises – c'est-à-dire qu'ils donnent à voir des traductions d'une réalité à travers le filtre des technologies de traitement des données qu'ils proposent. L'étude souligne enfin comment circulent les discours de ces acteurs de la donnée dans les entreprises, comment les entités dirigeantes se les approprient pour légitimer eux-mêmes leurs propres décisions stratégiques et justifier un discours sur la performance auprès du public de l'entreprise.

MOTS-CLÉS : Technologies, Big data, Discours, Imaginaires, Consultants, Cabinets de conseil, Expertise, Raison algorithmique, Gouvernamentalité algorithmique, Entreprises, Ethos

# SOMMAIRE

REMERCIEMENTS .....	3
RÉSUMÉ.....	4
SOMMAIRE.....	5
INTRODUCTION .....	7

## **I. La construction de la figure de l'expert à travers les « technologies big data » : une stratégie de légitimation, au croisement de différents mythes et récits opérants.....18**

<u>1. Un processus de légitimation par des discours d'escorte nourris d'imaginaires et de mythes opérants : la construction d'un <i>ethos</i> algorithmique .....</u>	<u>19</u>
a. Une rhétorique ancrée dans des imaginaires techniques et scientifiques, liés une « rationalité occidentale », du nombre et du calcul.....	19
b. La fonction incantatoire de ces technologies : l'effet de croyance à l'œuvre dans les discours .....	22
c. La « puissance communicationnelle » du récit de l'innovation technologique : la réactivation du mythe du progrès et l'« impératif de croissance ».....	24
<u>2. La mise en scène et la monstration de l'expertise .....</u>	<u>28</u>
a. Une mise en scène de leur « ethos algorithmique » .....	28
b. Une mise en avant de l'univers de la « data-science » : le « Data Lab ».....	30
c. La mise en scène de compétences techniques, du savoir-faire de l'ingénieur, de la « prouesse technique » et de sa maîtrise .....	32
<u>3. Le double discours du consultant-ingénieur, à la fois scientifique et commercial : une instance énonciative hybride .....</u>	<u>34</u>
a. Une polyphonie énonciative : la posture ambivalente de ces consultants-ingénieurs .....	34
b. Le simulacre de la neutralité et de l'objectivité des « technologies big data » .....	36

## **II. Le renforcement de la légitimité du consultant-ingénieur par ces technologies performatives qui prétendent s'ériger en connaissance et en aide à la décision .....40**

<u>1. Les traductions à l'œuvre pour rendre les données intelligibles : du langage ingénieur au langage gestionnaire.....</u>	<u>41</u>
a. Performativité et effet d'« irréversibilité » des outils de représentation des « grosses données » : rendre indiscutable la connaissance issue des « technologies big data ».....	42
b. « Mise en nombre » et quantification du réel par les algorithmes et techniques de calcul .....	43
c. Synthétiser et simplifier le réel : la capacité de l'ingénieur à vulgariser l'objet technique ...	45
<u>2. Le rapport de pouvoir de ces consultants en technologie face aux décideurs d'entreprise : érigés au rang de devins ou de savants ? .....</u>	<u>47</u>
a. Le pouvoir du consultant ingénieur qui prétend interpréter le monde, analogue à l'exercice de l'art divinatoire .....	47

b. Un triple pouvoir heuristique, pragmatique et normatif ?.....	52
c. Des « scribes » de l'univers : une nouvelle élite savante ?.....	57
3. <u>L'ambition démiurgique de leur ethos algorithmique : prétendre ordonner le chaos</u>	
<u>informationnel par les technologies de traitement des données</u> .....	59
a. Ces consultants en technologie exerceraient plus qu'un pouvoir performatif : un acte démiurgique ?.....	59
b. L'influence des « signes » envoyés par la « raison algorithmique » dans le monde de l'entreprise .....	61
c. La prétention démiurgique à maîtriser et prévoir les évènements : des « simulacres » .....	63
<b>III. Les discours des décideurs sur la performance se réinventent par la circulation de la « raison algorithmique » au sein des entreprises</b> .....	66
1. <u>Une appropriation par les acteurs économiques : des outils de gouvernement au pouvoir ?</u> 66	
a. L'influence de la « raison algorithmique » sur les choix stratégiques des discours des décideurs : des technologies intégrées à un discours général sur la performance économique	67
b. Valoriser ses données pour survivre dans un environnement économique compétitif : le traitement de données au cœur des stratégies d'entreprise.....	69
c. L'utilisation politique : des outils de gouvernement qui assoient la légitimité des décideurs	71
2. <u>Une dimension messianique croissante : d'un ethos algorithmique à un ethos messianique ?</u> 73	
a. Un régime de croyance qui se manifeste par une symbolique messianique et religieuse : la « conversion numérique » .....	73
b. Une stratégie d'évangélisation <i>via</i> des discours idéologiques sur le « big data ».....	75
3. <u>Les limites des « technologies big data » : un manque de recul sur les pratiques</u> .....	76
a. Une « raison algorithmique » relativement peu mise à l'épreuve.....	77
b. Ne donner à voir qu'une traduction algorithmique du monde, une limite intrinsèque à la « raison algorithmique ».....	79
CONCLUSION.....	83
BIBLIOGRAPHIE .....	87
ANNEXES .....	91

## INTRODUCTION

« Les big data sont à la mode. Dans les médias comme dans le débat public, le terme est évoqué de manière récurrente. Des entreprises dédiées au traitement de ces masses d'informations sont créées chaque jour, d'autres réorganisent radicalement leurs activités autour de la gestion de ces données. Le domaine semble suffisamment porteur pour que, dans de nombreux pays, les pouvoirs publics investissent massivement dans ce domaine. À la suite d'entreprises privées, les administrations de nombreux pays sont désormais dotées d'un responsable de haut rang dédié à cette thématique. »

Étienne Ollion, Julien Boelaert, « Au delà des big data. Les sciences sociales et la multiplication des données numériques », *Sociologie*<sup>1</sup>.

Comme le souligne cet article de la revue *Sociologie*, le « big data » est une notion qui fait autorité dans les sphères économiques, managériales et entrepreneuriales et même gouvernementales. La multiplication et la circulation grandissante des discours sur l'innovation par le « big data », la fascination exercée par ces technologies qui traitent des volumes de données massifs et hétérogènes – cela peut prendre la forme de plateformes, de logiciels, d'experts en « big data », de « data-scientists », ou de création de postes dédiés (les « Chief Data Officer »<sup>2</sup>) – s'associent à une vision de ces technologies qui les pense souvent comme une rupture ou une « révolution ». Comme le souligne la revue en ligne *Maddyness* spécialisée dans les start-ups et l'innovation, le « big data » est un eldorado qui « est sur les lèvres de tous les managers et entrepreneurs, fait les gros titres des journaux et anime des conférences entières. Le *big data* s'est peu à peu imposé à tous, particuliers et entreprises, dans les dernières années. Tous ont le pressentiment que ce Big Data est vraiment un sujet central, qui préoccupe, tant ses impacts

---

<sup>1</sup> OLLION Étienne, BOELAERT Julien, « Au delà des big data. Les sciences sociales et la multiplication des données numériques », *Sociologie*, 2015, Vol. 6, pp. 295-310. URL : [hwww.cairn.info/revue-sociologie-2015-3-page-295.htm](http://www.cairn.info/revue-sociologie-2015-3-page-295.htm). Ils auteurs ajoutent : « Chief data officer aux États Unis, Chief digital officer responsable du Government Digital Service en Grande Bretagne, Administrateur général des données en France : ils sont en charge de s'assurer que le pays valorise ses données, présentées comme une source de croissance inexplorée »

<sup>2</sup> Ibid., À ce propos, le monde académique se réorganise autour du phénomène big data : « Des conférences sur le thème sont régulièrement organisées. Des postes destinés à des spécialistes de « big data » sont créés chaque année, et leurs titulaires sont amenés à enseigner dans des cursus de « data science », qui promettent à leurs étudiants de tirer profit des opportunités de ce phénomène »

peuvent être importants »<sup>3</sup>. Ce premier constat s'accompagne de l'observation d'un contexte économique et professionnel qui valorise les « technologies big data » et les fait se constituer progressivement en un secteur économique à part entière : dans les entreprises, les métiers de « data scientist », de « Chief Data Officer » (à Ipsos France par exemple<sup>4</sup>), de Responsable « Projets Big Data » (comme à Airbus) sont créés. Des entrepreneurs se lancent dans la vente de leur expertise en « big data », et créent leur cabinet de conseil spécialisé dans l'accompagnement des « projets big data » des entreprises. Des sociétés de conseil ainsi que des start-up se créent pour aider les entreprises à gérer leurs données, les cabinets de consulting en technologie se spécialisent dans ce « nouvel eldorado économique »<sup>5</sup> et émettent des discours sur la valeur que peut créer le « big data », comparé à une nouvelle ressource, un « nouveau pétrole »<sup>6</sup>, et la nécessité pour les entreprises de structurer ces données massives et les transformer en richesse. Les acteurs de la donnée pointent du doigt ces entreprises qui ne savent comment exploiter leurs « mégadonnées ». Aussi, pour faire connaître leur activité, ces acteurs travaillant sur la donnée produisent des discours sur les « big data » et sur leur caractère inouï, prescrivant aux entreprises de s'en saisir rapidement afin d'être compétitives économiquement et commercialement. Ainsi les discours de ces acteurs de l'industrie de la donnée décrivent-ils souvent une époque inédite, comme si nous étions entrés dans une ère du « big data » qui donnerait accès à une connaissance infinie et toujours renouvelée. Si l'on examine la cristallisation d'imaginaires et de discours autour des « technologies big data », un certain nombre de discours expriment ainsi de la fascination. Ainsi retrouve-t-on souvent des discours technocentrés souvent produits sur « le numérique » et « l'informatique communicante »<sup>7</sup> qui contiennent une « portée globalisante et déterministe sur les technologies et leurs effets »<sup>8</sup>. Nous avons remarqué que cette vision de l'objet technique tout-puissant qui a le potentiel de venir tout changer et améliorer, est souvent sous-jacente aux discours des cabinets de conseil en technologie et de leurs consultants-ingénieurs.

Pour tenter de cerner le phénomène « big data », il faut d'abord en comprendre le fonctionnement d'un point de vue technique. Ainsi peut-on proposer une définition basée sur la notion des « 3V » telle que l'explique Gilles Babinet<sup>9</sup>, c'est-à-dire volume, vitesse et variété.

---

<sup>3</sup> « Le big data est-il le nouvel eldorado des entreprises européennes ? », 5 juin 2015, Maddyne (www.maddyne.com/innovation/2015/06/05/big-data-eldorado)

<sup>4</sup> Cf. Création de poste dont nous avons témoigné lors de notre expérience de stage à Ipsos de février à juin 2017

<sup>5</sup> www.maddyne.com/innovation/2015/06/05/big-data-eldorado : « Le big data est-il le nouvel eldorado des entreprises européennes ? », 5 juin 2015

<sup>6</sup> www.maddyne.com/innovation/2015/07/23/big-data-vouchercloud : Infographie. Le big data et son influence sur le quotidien des consommateurs en un coup d'œil, 23 juillet 2015 (« ce pétrole qui profite pour le moment aux marques et aux Google, Facebook et autres acteurs qui en font commerce »)

<sup>7</sup> Cf. CANDEL Etienne, « Textualiser les interfaces. Opérativité et épistémologie d'une requalification », Volume 1, *Soutenance de HDR*, 7 décembre 2015

<sup>8</sup> Op. cit., p. 224

<sup>9</sup> Auteur de *Big Data, penser l'homme et le monde autrement*. C-Radar donne aussi une définition : « Le terme de *Big Data* est utilisé lorsque la quantité de données qu'une organisation doit gérer atteint une taille critique qui

Permis par le développement d'internet, l'informatisation des organisations, la diffusion des capteurs et objets connectés, et l'augmentation de la puissance des ordinateurs<sup>10</sup>, les technologies big data sont « un ensemble de méthodologies, processus, architectures et technologies, où des logiciels, algorithmes ou processus allant au-delà des techniques actuellement utilisées pour l'analyse de données sont nécessaires, afin de gérer des données volumineuses, analysées rapidement et variées, et finalement d'en tirer une valeur ajoutée »<sup>11</sup>. Comme l'explicitent Antoinette Rouvroy et Thomas Berns, le « big data » se nourrit à partir de données que « les gouvernements les récoltent à des fins de sécurité, de contrôle, de gestion des ressources, d'optimisation des dépenses... ; les entreprises privées recueillent quantité de données à des fins de marketing et de publicité, d'individualisation des offres, d'amélioration de leur gestion des stocks ou de leurs offres de service, bref, en vue d'accroître leur efficacité commerciale et donc leurs profits... ; les scientifiques collectionnent les données à des fins d'acquisition et d'amélioration de connaissances... ; les individus eux-mêmes partagent bénévolement leurs données sur les réseaux sociaux, les blogs, les « mailing lists »...<sup>12</sup> ». Nous étudierons dans ce travail de recherche les consultants des cabinets de consulting, et plus spécifiquement, les champs d'application des technologies « big data » concernant les entreprises. Ils peuvent être par exemple le fait de « révéler (segmentation stratégique, *clustering*), prédire (trafic, pollution) et réagir (marketing, sécurité) »<sup>13</sup>.

Il convient également de préciser que les « technologies big data » s'inscrivent dans un contexte technique, s'appuyant sur une série de changements technologiques tels que l'augmentation de la capacité de stockage des ordinateurs et des serveurs, la mise au point de logiciels capables de traiter de gros volumes de données hétérogènes<sup>14</sup>. Elles s'inscrivent aussi dans

---

nécessite de nouvelles approches technologiques pour leur stockage, leur traitement et leur utilisation. Volume, vitesse et variété sont souvent les trois critères qui permettent de qualifier une base de données de « *Big Data* » ([www.c-radar.com/lexique-de-lopen-data](http://www.c-radar.com/lexique-de-lopen-data))

<sup>10</sup> « Dans leur acception la plus générale, les big data désignent ces vastes ensembles de données, dont l'existence comme le traitement rapide ont été rendus possibles par une série de changements technologiques. Le développement d'internet, medium par lequel de nombreuses activités humaines peuvent être enregistrées, la multiplication des capteurs de toute sorte et l'informatisation croissante des organisations font que l'on peut disposer d'informations précises, récurrentes et massives sur d'innombrables pratiques. L'augmentation de la puissance des ordinateurs et l'apparition de nouvelles méthodes de traitement font qu'il est possible de les analyser rapidement. C'est l'existence et plus encore l'abondance de ces données qui a donné lieu aux déclarations enthousiastes sur les big data et l'avenir de la connaissance. » (OLLION Étienne, BOELAERT Julien, « Au delà des big data. Les sciences sociales et la multiplication des données numériques », *Sociologie*, 2015, Vol. 6, (pp. 295-310)

<sup>11</sup> Définition donnée par l'Etude Xerox menée par Forrester en mai 2015 auprès de 330 entreprises dans 5 pays européens (publiée sur Maddynews.com)

<sup>12</sup> ROUVROY Antoinette, BERNS Thomas, « Gouvernamentalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », *Réseaux*, 2013, n° 177, pp. 163-196 (p. 168)

<sup>13</sup> Maddynews, « Le Big Data est-il le nouvel eldorado des entreprises européennes ? »

<sup>14</sup> « La disponibilité de cette très grande quantité de données brutes, pouvant souvent rester anonymes, annonciatrice d'une « mémoire digitale totale » distribuée dans quantité de banques de données constamment enrichies et actualisées, est une condition préalable au fonctionnement du gouvernement algorithmique dans la mesure où ce n'est qu'à condition de disposer de ces très grandes quantités de données que les opérations de « data mining » sont concevables. » (ROUVROY Antoinette, BERNS Thomas, « Le nouveau pouvoir statistique. Ou

un contexte historique, comme l'exemplifie Jean-Gabriel Ganascia en faisant référence à Viktor Mayer-Schönberger : « la notion de données n'est pas neuve et leur recueil systématique est lui-même très ancien<sup>15</sup> ». En effet, « dès la plus haute antiquité, au temps des Assyriens, on procéda à des statistiques et la Bible évoque le recensement de la population du royaume qu'ordonna David en vue de lever l'impôt », diminuant l'aspect révolutionnaire des « big data » prôné par les acteurs économiques et marketing. Si le recueil et le traitement de données n'est pas nouveau, en revanche la méthode d'analyse de ces données semble avoir quant à elle changé. Jean-Gabriel Ganascia souligne une « révolution d'ordre épistémologique<sup>16</sup> », dans la mesure où la méthodologie d'analyse des « grosses données » rompt avec les méthodes statistiques traditionnelles. Soulignant ce propos même, dans *Big Data : La révolution des données est en marche*, Viktor Mayer-Schönberger et Kenneth Cukier expliquent que les big data rompent avec les méthodologies statistiques classiques d'analyse de données (le modèle d'échantillonnage ne s'appliquant plus<sup>17</sup>), travaillant à partir de bases de données « non structurées » et imprécises : « Il faut s'attendre à de grands changements dans notre interaction avec le monde, si l'on suit d'utiliser la totalité plutôt que des bouts d'informations et de raisonner en termes d'imprécision et de probabilités plutôt que d'exactitude<sup>18</sup> ». Ainsi « les big data, autrement dit des ensembles immenses de données déstructurées, nous rapprochent plus de la réalité que ne le faisaient des volumes plus petits de données aussi exactes que possible<sup>19</sup> ». Selon eux, avec le « big data » nous assisterions à changement de paradigme épistémologique et scientifique, avec la nécessité désormais d'accepter la notion de désordre qu'imposent ces volumes de données, ainsi que l'absence de modèle. En effet, « les « grosses données » révoquent ce schéma

---

quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques », *Multitudes*, 2010, n° 40, pp. 88-103. [www.cairn.info/revue-multitudes-2010-1-page-88.htm](http://www.cairn.info/revue-multitudes-2010-1-page-88.htm) (p. 91)

<sup>15</sup> GANASCIA Jean-Gabriel, « Les big data dans les humanités », *Critique*, 2015, n°819-820, pp. 627-636 (p. 628) : « Que signifie big data ? On l'a dit, la notion de données n'est pas neuve et leur recueil systématique est lui-même très ancien. Comme le rappelle Viktor Mayer-Schönberger, dès la plus haute antiquité, au temps des Assyriens, on procéda à des statistiques et la Bible évoque le recensement de la population du royaume qu'ordonna David en vue de lever l'impôt. Depuis l'avènement des premiers ordinateurs, les données représentées sous forme numérique binaire, comme des suites de 0 et de 1, se manipulent de façon automatique. Cette numérisation ne se limite pas à des données quantitatives : textes, images, sons, vidéos, voire sensations kinesthésiques se traduisent en flux d'information et se traitent comme tels. Là encore, rien de vraiment nouveau : il y a bientôt soixante-dix ans, Vannevar Bush, Alan Turing, Norbert Wiener et d'autres en eurent l'intuition. Bref, les notions de données, de recueil systématique, de numérisation et de traitement automatique apparaissent comme anciennes. L'innovation tient à l'automatisation et à la massification de la collecte par le truchement de machines. »

<sup>16</sup> Op. cit., p.634

<sup>17</sup> À la p. 43, les auteurs expliquent que l'échantillonnage statistique est « un concept vieux d'un siècle, élaboré pour résoudre un problème donné à un moment donné sous des contraintes technologiques données », en ajoutant à la p. 30 : « p. 30 « Un des buts de la statistique est, après tout, d'obtenir les résultats les plus pertinents en faisant appel à la plus petite quantité de données possible »

<sup>18</sup> MAYER-SCHÖNBERGER Viktor, CUKIER Kenneth, *Big data: la révolution des données est en marche*, R. Laffont, 2014, 296 p. (pp. 62-63, chapitre « Désordre »)

<sup>19</sup> Op. cit., (...) parce que nous ne pouvons collecter et étudier qu'une quantité limitée de données, nous n'étions pas confrontés à ce besoin irrésistible de tout voir sous tous les angles possibles et de tout avoir. » (p. 63). A la p. 53 : « Passer à un monde de big data nous demandera de changer de point de vue sur les mérites de l'exactitude. Effectuer des mesures dans ce monde numérique et connecté du XXIème siècle avec un état d'esprit conventionnel serait de passer à côté d'un point de vue essentiel ».

classique de la science expérimentale. La théorie ne prime plus, les dispositifs matériels de recueil d'observations ne sont pas conçus au regard d'une hypothèse que l'on cherche à prouver. La collecte des données se fait automatiquement, sans idée a priori. On exploite ensuite ces données à l'aide d'outils d'intelligence artificielle et d'apprentissage machine<sup>20</sup> ». C'est précisément ce que souligne Pierre Delort, ajoutant que « l'originalité de l'approche Big Data réside en ce qu'elle ne s'appuie pas sur les structurations préexistantes à la collecte de données, mais entend découvrir au sein de ces données des modèles<sup>21</sup> ».

Ce point de vue sur la fin des méthodes scientifiques traditionnelles rejoint celui de Chris Anderson dans le magazine *Wired* en 2008, « The End of Theory: the Data Deluge makes the Scientific method obsolete<sup>22</sup> ». La connaissance permise *via* les « technologies big data » serait tellement fine que nous n'aurions plus besoin de segments ni de modélisations. Cette vision technocentrée et utopiste, presque évangéliste, s'inscrit dans un contexte où nombre de discours de la révolution sont produits, comme à chaque nouvelle technologie qui apparaît à un moment donné de l'histoire : nous aurions à faire à une révolution des données avec des algorithmes de « machine-learning » capables de donner accès à un savoir jusque là inaccessible, des flux de données « s'érigeant comme l'instance primordiale de l'intelligibilité du réel (...) visant les plus hautes optimisation, fluidification et sécurisation des existences et des sociétés<sup>23</sup> ». Eric Sadin, dans *La vie algorithmique, critique de la raison numérique*, souligne ainsi l'espoir d' « une connaissance sans cesse approfondie [qui] s'instaure », avec des « technologies informationnelles » qui « imposent un mode de *rationalité* fondé sur la définition chiffrée de toute situation et sur une maîtrise indéfiniment accrue du cours des choses. Une raison numérique établie sur l'appréhension et l'évaluation en temps réel des faits ordonne désormais les pratiques du commerce ». Ces technologies computationnelles prétendent fournir une intelligence à partir de données déstructurées et contiennent des prétentions telles que le fait de « faciliter la prise de décision stratégique et gagner en compétitivité<sup>24</sup> », ou encore des promesses telles qu'optimiser l'activité et la production d'une société ou d'une industrie, améliorer la sécurité des produits et services pour les industries, accroître la fidélité des clients dans le secteur de la grande consommation et de la distribution... Les promesses de ces technologies sont donc multiples : aider à la prise de décision, améliorer la performance marketing et commerciale des entreprises, encourager les processus

---

<sup>20</sup> GANASCIA Jean-Gabriel, p. 634

<sup>21</sup> DELORT Pierre, *Le Big Data*. Presses Universitaires de France, « Que sais-je ? », 2015, 128 pages. URL : [www.cairn.info/le-big-data--9782130652113.htm](http://www.cairn.info/le-big-data--9782130652113.htm) (Chapitre I Informatique en entreprise [1] pp.7-9)

<sup>22</sup> ANDERSON Chris, « The End of Theory: the Data Deluge makes the Scientific method obsolete », *Wired*, 23 juin 2008: « This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear ». Plus loin, il ajoute : « faced with massive data, this approach to science — hypothesize, model, test — is becoming obsolete ».

<sup>23</sup> SADIN Éric, *La vie algorithmique. Critique de la raison numérique*, Paris : Éditions l'Échappée. DL 2015, 278 p. (quatrième de couverture)

<sup>24</sup> Maddyness, « Le Big Data est-il le nouvel eldorado des entreprises européennes ? »

d'innovation dans l'entreprise, prédire les comportements d'achat des consommateurs ou encore prédire des grandes tendances de société. Enfin, les créateurs de ces technologies font la promesse d'un progrès économique mais aussi social, à l'image de l'ONG française Bayes Impact qui a pour ambition de « tenter de résoudre les plus grands problèmes de société à l'aide de la data science »<sup>25</sup>.

Cette thématique de la révolution associée aux dispositifs numériques est explicitée dans l'HDR d'Etienne Candel : « aujourd'hui, le discours sur les dispositifs numériques est spécifique, en ce qu'il se mêle de loin en loin de thématiques de la révolution, de la transformation radicale, du retournement, de la fin d'un monde ou du début d'un autre. Quand ces idées cessent de fleurir sur les formats du blog ou du micro-blogging (...), c'est pour se déployer au sujet des données, de l'humain augmenté, de l'Internet des objets<sup>26</sup>. » Ainsi, constatant un manque de recul réflexif sur ces pratiques, nous avons voulu par ce travail de recherche nous intéresser à cette même notion, souvent convoquée par une multiplicité d'acteurs mais rarement analysée ou mise en perspective. Ce travail de recherche s'intéresse aux « technologies big data » en tant qu'elles sont une construction autour du « mythe de l'informatique comme communication aux implications radicales pour le vivre humain – mythe construit le plus souvent sous l'aspect de la révolution ou de la transformation, du basculement ou de la rupture<sup>27</sup> ». Ce mythe des « big data » s'inscrit dans une tendance à la production de discours qui font état d'un changement de paradigme, à la fois scientifique, économique et social. Qu'ils soient start-up ou cabinets conseil, les acteurs spécialisés dans le traitement des « big data » ou « mégadonnées », se présentent comme détenant des solutions pour aider les entreprises et acteurs économiques à développer leur activité, les accompagner pour s'acculturer au phénomène de « mégadonnées ». Ils se positionnent comme des experts qui détiennent un savoir fortement demandé et créateur de valeur, et c'est sur ce processus de construction de la figure d'autorité du savoir et de l'expert que cette recherche s'inscrit précisément. Il va s'agir à la fois de déconstruire les discours d'escorte des cabinets-conseil qui vendent ces « technologies big data », et de comprendre comment ces objets agissent sur les entreprises.

Les « big data » cristallisent les discours des sphères médiatiques, économiques et sociales : elles sont non seulement un objet de fascination, mais aussi d'inquiétude – avec la crainte que ces technologies échappent au contrôle de l'homme, rejoignant une question épineuse quant à la

---

<sup>25</sup> Sur leur site, on peut lire : « Bayes Impact est une entreprise technologique à but non lucratif qui conçoit des services sociaux en s'appuyant sur la big data. Nous utilisons la puissance de logiciels complexes et de la big data pour créer des solutions personnalisées à grande échelle et aider des millions de personnes à travers le monde. En tant qu'ONG, nous sommes convaincus que les nouvelles technologies peuvent faire davantage que de générer du profit. C'est pourquoi l'ensemble de notre travail est "open source" et dédié à l'intérêt public global. Les logiciels et la big data sont d'incroyables leviers pour le progrès social. Ils peuvent permettre à un petit groupe de personnes d'avoir un impact social positif à grande échelle. » ([www.bayesimpact.org/fr](http://www.bayesimpact.org/fr))

<sup>26</sup> CANDEL Etienne, « Textualiser les interfaces. Opérativité et épistémologie d'une requalification », Volume 1. *Soutenance de HDR*, 7 décembre 2015 (p. 22-23)

<sup>27</sup> Op. cit. p. 191

protection des données personnelles, ou l'exploitation de celles-ci à des fins commerciales. Le texte, les images, les vidéos que nous publions sur les forums d'entraide, sur Facebook, sur Instagram ou Twitter, notre navigation sur notre ordinateur, notre tablette ou notre mobile, les pages web que l'on consulte, notre acceptation des termes d'utilisation de ces plateformes ou des « cookies » sur notre navigateur... sont autant de gestes que nous effectuons derrière notre écran et qui participent à la création des « traces numériques<sup>28</sup> ». Dans *Traces numériques. De la production à l'interprétation*, Béatrice Galinon-Ménélec illustre ce phénomène de multiplication des « traces » numériques, d'abord caractérisé par des « enregistrements de données à des fins de gestion ou des fins de prévention de risques<sup>29</sup> ». Eric Sadin évoque une « puissance toujours plus totalisante détenue par les systèmes computationnels<sup>30</sup>, exemplifiant cette crainte d'une quantification et d'une marchandisation intégrales de la vie sous l'effet de l'industrie du traitement des données. De la même manière, prenant le contrepied de la vision utopiste de Chris Anderson, Dominique Cardon, dans son ouvrage *A quoi rêvent les algorithmes : Nos vies à l'heure des big data*, souligne l'avènement d'une « société du calcul », et exprime la crainte que les algorithmes viennent gouverner les comportements des individus, des acteurs économiques et sociaux : « Parmi les espoirs et les craintes que suscite la numérisation de nos sociétés, la constitution de grandes bases de données confère une place de plus en plus centrale aux algorithmes qui gouvernent les comportements de chacun<sup>31</sup> ». Cette notion de « gouvernementalité algorithmique » est explicitée dans les travaux d'Antoinette Rouvroy, qui critique cette vision technocentrée selon laquelle le monde, le réel, grâce aux « technologies de big data », seraient désormais réductibles à des bases de données. Une telle rationalité, qu'elle nomme « comportementalisme numérique », « se fonde sur la pure observation des données, complètement coupée du monde »<sup>32</sup>.

En théorisant une « critique de la gouvernementalité algorithmique »<sup>33</sup>, Antoinette Rouvroy met un nom sur ce champ de recherche encore peu étudié, puisqu'il « porte sur des phénomènes qui véritablement échappent au sens, dans la mesure où il s'agit pour nous d'un nouveau régime

---

<sup>28</sup> La notion de « trace » telle que l'entend Béatrice GALINON-MENELEC dans *Traces numériques. De la production à l'interprétation*, CNRS Éditions, 2013, 290 p. : « L'humain du XXI<sup>e</sup> siècle, même le moins tenté par les nouvelles technologies, laisse partout des traces de ses actions. Relevant de la sphère publique ou privée, les traces numériques de ce Nouveau Monde sont manipulables et susceptibles de ressurgir dans des contextes imprévisibles » (résumé)

<sup>29</sup> Op. cit., p. 7

<sup>30</sup> SADIN Eric, *La vie algorithmique, critique de la raison numérique*, Paris : Éditions l'Échappée. DL 2015, 278 p. (Résumé en quatrième de couverture) : « La quantification et la marchandisation intégrales de la vie qui s'instituent, soutenues par l'industrie du traitement des données, aujourd'hui dotée d'un pouvoir qui perturbe nombre d'acquis démocratiques fondamentaux »

<sup>31</sup> CARDON Dominique, *A quoi rêvent les algorithmes : Nos vies à l'heure des big data*, Paris : Seuil, La République des idées, 2015, 105 p. (Quatrième de couverture)

<sup>32</sup> Séminaire « Big data et anticipation : vers une gouvernementalité algorithmique ? », organisé le 16 avril 2015 à l'Espace éthique/IDF (disponible sur YouTube, vidéo publiée le 22 mai 2015 par Espace Éthique. URL : [www.youtube.com/watch?v=RUEkEsB6FJA](http://www.youtube.com/watch?v=RUEkEsB6FJA))

<sup>33</sup> Ibid.

d'intelligibilité et de gouvernement du monde, qui au départ est nourri par des données brutes principalement – qui sont souvent décontextualisées, bien souvent anonymisées, des données qui sont en quantité massive, qui sont complexes, c'est-à-dire des données qui prennent la forme de textes, d'images, de sons, de points de localisation, de trajectoires<sup>34</sup> ». Elle souligne bien l'idéologie « des gens qui font de la data » et qui souhaitent donc « que se produise cette substitution d'une réalité numérique au monde tel qu'il est, et une sorte d'indistinction croissante entre la représentation numérique et le monde tel qu'il est<sup>35</sup> ». Dans ce travail de recherche, nous nous concentrerons plus particulièrement sur les cabinets de consulting et consultants qui proposent des technologies et « logiciels big data » dans leur offre de conseil aux entreprises, en tentant d'analyser leurs discours à travers cet aspect idéologique et la grille d'analyse proposée par Antoinette Rouvroy.

Pour aller plus loin dans la réflexion, nous nous appuyerons aussi sur la notion de « technologies intellectuelles<sup>36</sup> » telle que développée par Pascal Robert, de manière à étudier les « big data » comme « outil de représentation<sup>37</sup> » - utilisé, entre autres pour cette étude, par les cabinets de consulting et les entreprises – à l'aune des logiques de médiations qu'elles exercent, puisqu'utiliser ces technologies implique un certain savoir-lire et un savoir-écrire. Ainsi les « big data » impliquent une certaine manière d'appréhender le monde<sup>38</sup>, et « s'interfacent à notre système de perception-cognition<sup>39</sup> ». De la sorte, nous nous attacherons à analyser simultanément technique et société, à considérer ces deux objets comme imbriqués l'un dans l'autre, à l'image de la sociologie des sciences et des techniques telle que proposée par Bruno Latour<sup>40</sup>. Il convient dès lors de penser les « technologies big data » à l'aune de leur contexte d'utilisation, de leurs pratiques, des différents acteurs qui les mobilisent ainsi que les rapports de pouvoirs qu'elles impliquent.

## PROBLÉMATIQUE

Dans un contexte de production de discours d'accompagnement qui attribuent des pouvoirs aux « technologies big data », portant la vision d'un l'objet technique tout-puissant qui

---

<sup>34</sup> Ibid.

<sup>35</sup> Ibid

<sup>36</sup> ROBERT Pascal, « Qu'est-ce qu'une technologie intellectuelle ? », *Communication & langages*, n°123, 2000, Dossier : Les médias en Russie (pp. 97-114). Il semble que « tout processus, dès lors qu'il relève d'une manière ou d'une autre du penser/classer, puisse être qualifié de technologie intellectuelle » (p. 102). Il précise : « l'intelligence a besoin d'outils pour appréhender le monde ; elle n'est pas en prise directe avec lui dès qu'il s'agit de la maîtriser » (p. 97).

<sup>37</sup> Op. cit., p. 105 : « Le travail d'une opération de traduction de l'événement en document »

<sup>38</sup> Op. cit., C'est en cela que les « big data » sont une « technologie intellectuelle » : il semble que tout processus, dès lors qu'il relève d'une manière ou d'une autre du penser/classer, puisse être qualifié de technologie intellectuelle » (p. 102)

<sup>39</sup> Op. cit., p. 108

<sup>40</sup> Il invite à « éviter la coupure entre deux sphères, celle de la société, réceptrice, passive, et celle de la technique, émettrice, active » (p. 111). LATOUR Bruno, « Les Vues de l'esprit », *Culture et technique, Réseaux*, volume 5, n°27, 1987, Questions de méthode. pp. 79-96

répondrait à la tentation d'avoir accès à une connaissance<sup>41</sup> toujours plus approfondie voire prédictive, il va s'agir d'étudier le discours des cabinets de conseil en technologie ainsi que leur positionnement en tant qu'instance énonciative sur le marché du traitement des « données ».

Alors que les discours d'accompagnement sur les « technologies big data » se multiplient, émanant d'entrepreneurs, de consultants et de parties prenantes de l'entreprise (managers, dirigeants, responsables du digital), comment les cabinets de conseil en technologie construisent-ils leur posture d'expertise sur le marché du traitement des données des entreprises, à partir d'une rationalité algorithmique pour se légitimer comme acteurs crédibles et incontournables pour l'aide à la décision ?

Autrement dit, à travers les interfaces et plateformes que ces acteurs proposent pour analyser les « grosses données » des entreprises, en quoi promettent-ils de fournir une vérité à partir de « traces numériques » et créent-ils des médiations algorithmiques qui se posent comme outils de prescription et d'intelligibilité du réel ?

## HYPOTHÈSES

Nous avons dégagé nos hypothèses en étudiant les discours des cabinets de conseil à la lumière des textes structurants en SIC lors de notre recherche documentaire.

1 – Par leur discours d'escorte, ces cabinets de conseil mettent en avant la technologie et l'algorithme pour construire un *ethos* algorithmique, dans une stratégie de légitimation de leur offre de conseil.

2 – En traduisant, en modélisant et en représentant les « traces » numériques des entreprises, ces cabinets de conseil en technologie consolident leur posture d'expert, du fait du pouvoir heuristique et prescriptif exercé par ces « technologies intellectuelles » sur l'environnement de l'entreprise, ce qui viendrait renforcer la légitimité de ces consultants.

3 – Les consultants en « technologies big data », formant à ces technologies et les faisant ainsi circuler dans les entreprises, entraîneraient un renouvellement des discours de la part des entités dirigeantes.

## MÉTHODOLOGIE & CORPUS

---

<sup>41</sup> Notion qui rejoint celle d'Intelligence économique, qu'Harold Wilensky définit en 1967 comme « l'activité de production de connaissance servant les buts économiques et stratégiques d'une organisation »

Ce travail de recherche a d'abord été initié par l'observation d'une fascination exercée dès que le terme « big data » était convoqué, constatant à la fois un flou définitionnel et un manque de recul réflexif sur les pratiques liées à cette technologie. Remarquant que les outils « big data » se diffusent dans une diversité de structures, que ce soit dans les univers économiques ou gouvernementaux, c'est un objet qui reste assez peu interrogé, qu'on remet peu en question ou en perspective. Brandies comme un étendard et constatant que les « technologies big data » font autorité dans la sphère économique et médiatique, il nous est paru intéressant de tenter d'analyser sur quels ressorts repose cette légitimité, et pourquoi cet objet est tant mobilisé par une pluralité d'acteurs.

Nous avons d'abord effectué une recherche documentaire afin de cerner ce qu'étaient les « big data » et en comprendre les enjeux techniques, socio-économiques, et politiques. Nous avons compilé à la fois des articles des revues scientifiques issus des approches des SIC (*Communication & langages*) et des sciences sociales (revues telles que *Sociologie, Projet, Réseaux, Critique*), des articles de la presse économique, business et spécialisée dans les technologies (*Maddyness.com, Journal du Net.com, Alliancy.fr, Frenchweb.fr*). À ces premières lectures, s'ajoutent des ouvrages de chercheurs en SIC tels que *Traces numériques. De la production à l'interprétation* de Béatrice Galinon-Méléneq pour nous aider à analyser les « big data » dans une perspective de recherche liée au domaine des sciences de l'information et de la communication.

La démarche de ce travail s'est fondée sur la sélection et l'analyse d'un corpus constitué au départ de start-up spécialisées dans le traitement des données. Mais ayant recensé un trop grand nombre d'entre elles et remarquant que l'écosystème des start-ups pourrait constituer un sujet de recherche à part entière, nous avons décidé de réduire l'étude aux cabinets de conseil en technologie (en France). Aussi, la facilité d'accès à ce corpus a déterminé le choix de celui-ci. Nous avons en effet pu conduire plusieurs entretiens notamment l'un avec un cabinet de conseil, du fait de notre stage à Ipsos, qui ont permis de recueillir le point de vue de différentes parties prenantes sur le « big data » et la « data » en général, de comprendre comment elles voient leurs métiers et d'identifier leurs logiques de discours propres : un entretien qualitatif avec Novagen ainsi qu'un document de présentation papier<sup>42</sup>, un entretien avec Agnès Gilbert, « Chief Digital Officier » et Charlotte Zaepfel, « data-scientist » à Ipsos. Ces trois entretiens ont permis de recueillir différents types de discours : l'un davantage commercial et technocentré, le second davantage teinté marketing, et le troisième plus technique. Enfin, le témoignage de Vincent Cottin, développeur

---

<sup>42</sup> C'est un cabinet qui souhaite collaborer avec Ipsos et a fait une présentation dans leurs locaux (en avril 2017). Ils se définissent comme « cabinet de conseil IT et Big Data ».

web PHP à Atsukè<sup>43</sup>, a aussi participé à ce projet de recherche et à mieux cerner les contours techniques des « technologies big data ».

En complément du recueil de discours *via* une enquête de terrain qualitative<sup>44</sup>, nous avons recueilli des captures d'écran des sites web et documents produits par différents cabinets de conseil en technologie (en France principalement) afin de pouvoir étudier le matériau que ces acteurs donnent à voir, et d'analyser les sémiotisations et les logiques de médiations à l'œuvre dans chacun d'entre eux (comment ils présentent leurs « solutions big data », sous quels atours, quelles représentations ils donnent à voir des « big data »). Les documents recueillis peuvent parfois prendre la forme de tableaux de bord, de graphiques, de vidéos de démonstrations, de livres blancs, ou plus simplement de fiches de présentation – que certains acteurs nomment « datasheets ». Nous avons ainsi constitué un corpus composé de textes, d'images, et de contenus graphiques pour pouvoir démontrer quelles sont les sémiotisations à l'œuvre chez les cabinets de conseil en technologie, comment ils mettent en discours leur expertise, et quels types de prétentions contiennent leurs discours. Pour analyser ce corpus, il va s'agir d'étudier les éléments de langage et les thématiques récurrentes dans la communication de ces cabinets de conseil, et de compléter cela par une analyse sémiologique des supports de communication sur lesquels ils s'appuient afin de se faire connaître et de diffuser les pratiques qu'ils promeuvent.

## PLAN

Ainsi peut-on dégager trois axes de réflexion sur les offres de « big data » qui sont élaborées et construites par les cabinets de consulting.

Il va d'abord s'agir d'étudier comment se construit la figure de l'expert à travers les « technologies big data », *via* une stratégie de légitimation qui s'appuie sur différents mythes et récits opérants, dont des imaginaires techniques et scientifiques.

Puis, nous analyserons comment se renforce la légitimité du consultant-ingénieur *via* ces technologies performatives, grâce auxquelles il prétend faire valeur de connaissance et d'aide à la décision auprès des décideurs stratégiques.

Enfin, nous étudierons comment les décideurs s'approprient ces discours sur les « technologies big data » et comment la « raison algorithmique » circule dans l'entreprise, avec un discours sur la performance qui semble se réinventer.

---

<sup>43</sup> Start-up qui crée des solutions de commerce web et mobile, qui « exploite également la richesse du Big Data et devient l'interface avec les objets connectés » (cf. Présentation de la société : [www.atsuke.com/agency/groupe.html](http://www.atsuke.com/agency/groupe.html))

<sup>44</sup> Mon stage à Ipsos m'a par ailleurs permis de connaître la méthodologie liée au métier de qualitatifiste

# I. LA CONSTRUCTION DE LA FIGURE DE L'EXPERT À TRAVERS LES « TECHNOLOGIES BIG DATA » : UNE STRATÉGIE DE LÉGITIMATION, AU CROISEMENT DE DIFFÉRENTS MYTHES ET RÉCITS OPÉRANTS

En premier lieu, il va s'agir de déconstruire les discours d'escorte des acteurs qui vendent ces technologies, ici les cabinets de conseil en technologie. Quelles sont les prétentions et médiations qui se cachent derrière les « technologies big data » et les acteurs qui les commercialisent ? Dans le but de s'auto-légitimer et d'apparaître comme experts sur le marché des « big data », les discours d'accompagnement de ces consultants en technologie se fondent sur différents leviers, dont la mise en avant d'un « ethos<sup>45</sup> » algorithmique, nourri d'imaginaires empruntant aux univers mathématiques et scientifiques. Il semble aussi qu'ils s'appuient sur une histoire de la rationalité occidentale et de la base de données, pour construire un récit de l'innovation et du progrès, participant à une stratégie de légitimation de leur discours, les faisant se positionner comme instances énonciatives expertes dans le secteur du conseil en technologie. La notion d'*ethos*, empruntée à la rhétorique antique grecque, permet ici de déconstruire les discours des cabinets de conseil de technologie, en analysant leur manière de mettre en discours les technologies qu'ils vendent, ainsi que leur manière de se constituer en une instance énonciative caractérisée par une « rationalité algorithmique ». En examinant leurs leviers discursifs et « procédés expressifs<sup>46</sup> », il s'agit de comprendre comment ils tentent de persuader et de susciter l'adhésion à leurs discours commerciaux – notamment en incarnant des valeurs rationnelles, mathématiques et algorithmiques.

---

<sup>45</sup> « Pour la rhétorique antique, l'*ethos* en grec « personnage » désigne l'image de soi que l'orateur construit dans son discours pour contribuer à l'efficacité de son dire, autrement dit, pour exercer une influence sur son auditoire ». Aussi, Dominique Maingueneau ajoute que l'*ethos* est « l'attitude d'un énonciateur donné dans une situation de communication donnée et face à des coénonciateurs donnés ». (ALSAFAR Ali. *Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012. Linguistique*. Université Paul Valéry - Montpellier III, 2014) (p. 23)

<sup>46</sup> Nous appliquons ici la définition de l'*ethos* politique évoquée par Patrick Charaudeau dans la revue *Corpus* est éclairante (DE CHANAY Hugues, « Patrick Charaudeau. — Le discours politique. Les masques du pouvoir. Paris : Vuibert, 256 p. », *Corpus*, 4, 2005. URL : <http://corpus.revues.org/320>) (pp. 207-213), le définissant comme les « images des acteurs politiques construites dans et par leur discours », marqués par des signifiants (« procédés expressifs »).

## 1. Un processus de légitimation par des discours d'escorte nourris d'imaginaires et de mythes opérants : la construction d'un *ethos* algorithmique

En visitant les sites Internet des cabinets de conseil proposant d'accompagner les entreprises dans leurs « projets big data » (Novagen<sup>47</sup> et Affini-Tech) leur point commun est qu'ils mettent en avant leur connaissance et leur expertise concernant les « technologies big data ». Par exemple, la société de conseil Novagen souligne « un savoir-faire éprouvé sur les métiers de l'IT et du Web », « une expertise de l'écosystème *Big Data* au service des besoins Métiers de nos clients » ou encore « un savoir-faire opérationnel<sup>48</sup> ». Le cabinet Affini-Tech met en avant sur la page d'accueil de son site web leur solution « combinant Data-Science et expertise technologique<sup>49</sup> ». Ce champ lexical de l'expertise et de la connaissance participe à construire une figure d'expert, s'auto-légitimant dans ses discours d'escorte. Dans cette stratégie de légitimation de leur posture d'expert, quels sont les principaux leviers rhétoriques sur lesquels ces consultants en technologie s'appuient ?

### a. Une rhétorique ancrée dans des imaginaires techniques et scientifiques, liés une « rationalité occidentale », du nombre et du calcul

En analysant les discours d'escorte de ces cabinets de conseil, nous remarquons qu'ils construisent leurs discours à partir d'imaginaires liés à la technologie qui font autorité, et que la légitimité de leurs discours se trouve peu questionnée ou remise en perspective. Les consultants-ingénieurs de ces sociétés de conseil ont souvent recours à une rhétorique qui prend racine dans des imaginaires techniques et scientifiques qui sont autant de mythes opérants faisant valeur dans le discours, et font écho à notre « rationalité occidentale<sup>50</sup> ». À ce titre, le cabinet Affini-Tech met en avant les « modèles de Machine Learning » appliqué à un projet de prévision des ventes<sup>51</sup> ou encore l'« Application d'algorithmes prédictifs sur des flux d'évènements en temps réel<sup>52</sup> ». En convoquant dans leurs discours d'accompagnement les algorithmes de « machine-learning », ces cabinets de

<sup>47</sup> Cf. Description de leur activité dans les résultats de page Google indique : « Novagen est un cabinet de conseil IT et Big Data. La Data Factory de Novagen développe des solutions Data pour PME-PMI, ETI et Grands Groupes » & « Affini-Tech, Consulting et Integration de services Data, Bigdata, Datasience »

<sup>48</sup> Cf. Annexe 1. Corpus d'étude : Sites internet des cabinets de conseil en technologie. a. Site web Novagen – Page d'accueil ([www.novagen.tech](http://www.novagen.tech))

<sup>49</sup> Cf. Annexe 1. Corpus d'étude. e. Site web Affini-Tech – Page d'accueil ([www.affini-tech.com](http://www.affini-tech.com))

<sup>50</sup> Cette « rationalité occidentale » peut se définir comme à la lumière de la conception de la raison hégélienne et le positivisme de Comte. Ioanna KUÇURADI dans la revue Diogène explicite cette notion : on peut la comprendre comme une « comme une forme particulière d'expliquer et justifier les choses », lorsque l'Homme explique les phénomènes naturels non pas « théologiquement » ou « métaphysiquement », mais par l'observation et l'expérimentation, ou par la « science positive » (« Rationalité' » et « rationalités » dans le cadre du débat modernisme – post-modernisme », 2003, n° 202, pp. 12-20)

<sup>51</sup> Cf. Annexe 1. Corpus d'étude. g. Site web Affini-Tech – Projets - Prédiction des ventes

<sup>52</sup> Cf. Annexe 1. Corpus d'étude. g. Site web Affini-Tech – Projets – Maintenance prédictive

consulting jouent sur l'effet de fascination exercé par les « technologies big data » pour promouvoir leur offre de services en technologie à l'égard des organisations. En plus d'exploiter les imaginaires sur le numérique faisant des « big data » une « prouesse technique<sup>53</sup> », cette rhétorique fait aussi écho à notre raison occidentale en empruntant des termes liés à l'univers mathématico-scientifique. En effet, notre « rationalité occidentale » serait marquée par la conception de la raison hégélienne et le positivisme de Comte comme le souligne Ioanna Kuçuradi dans « *Rationalité' » et « rationalités » dans le cadre du débat modernisme – post-modernisme* », nous serions tentés d'expliquer le monde qui nous entoure non plus « théologiquement » ou « métaphysiquement », « mais par l'observation et l'expérimentation, ou par la « science positive »<sup>54</sup>. Lorsque le cabinet Affini-Tech mobilise cette rhétorique liée à la science qui convoque la « data science » et les « modèles de Machine Learning<sup>55</sup> », il entre en résonance avec notre « rationalité occidentale » qui s'est construite à partir de l'universalisme scientifique, la philosophie kantienne, la raison hégélienne et le positivisme de Comte, reconnaissant la valeur des mathématiques et de la physique, et de la démonstration rationnelle (théorie selon laquelle « l'homme doit se borner à ce qu'il peut savoir de manière certaine, grâce à la science<sup>56</sup> »). Dans la même logique, lorsque le cabinet Affini-Tech convoque la notion de « data », celui-ci résonne de manière symbolique dans nos imaginaires occidentaux : « Simplifions vos Data. De l'idée au service, nous accélérons votre business avec vos données ! ». Nos imaginaires, nourris par une image de la base de données infinie fantasmant une construction infinie de la connaissance, font partie d'un système de représentations qui valorise cette notion même de « data ». Comme le souligne Stéphan-Eloïse Gras dans son cours de *Prospective digitale*<sup>57</sup>, l'explosion des « big data » traduit un certain modèle de la confiance dans nos sociétés, c'est-à-dire le fait de céder notre confiance à un objet extérieur – ici, le numérique et l'informatique – et d'adhérer à ce système de représentations. Cette confiance est liée à notre raison occidentale et à une histoire du calcul rationnel. Dans un article intitulé « Data, l'onction du nombre », Erik Bertin dans le magazine *Influencia* évoque : « la place croissante de la *data* dans l'imaginaire social interroge de manière centrale la valeur du nombre dans notre modèle culturel », lequel semble agir comme une « nouvelle forme de pouvoir fondé sur la rationalité absolue du nombre et du calcul »<sup>58</sup>. Les

---

<sup>53</sup> CANDEL Etienne, dans son HDR (« « Textualiser les interfaces. Opérativité et épistémologie d'une requalification »), souligne la vision utopique et technocentrée à l'égard des interfaces techniques : « se déploie, de façon à la fois très explicite et discrète, l'idée générale que recouvrent ces interfaces dans cet usage spécifique : une prouesse technique, un exploit, placés au profit et au service des utilisateurs, et porteurs de promesses inattendues et inespérées pour lui » (p. 23)

<sup>54</sup> Op. Cit. Ioanna KUÇURADI

<sup>55</sup> Cf. Annexe 1. e. Site web Affini-Tech – Page d'accueil

<sup>56</sup> Cf. Le positivisme d'Auguste Comte ici : [la-philosophie.com/positivisme-auguste-comte](http://la-philosophie.com/positivisme-auguste-comte)

<sup>57</sup> Cf. Prises de notes de son cours dans le Master Communication, Médias et Numérique

<sup>58</sup> Cf. BERTIN Erik, « Data, l'onction du nombre », *Influencia*, 16 octobre 2016. URL : <http://www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html>

« technologies big data » inspirent confiance du fait de leur caractère « infaillible »<sup>59</sup>. En effet, il semble qu'on puisse s'en remettre aux chiffres sans hésiter et sans risque, comme l'explique Erik Bertin, le Directeur Général Adjoint de l'agence MRM Worldwide, sur *Influencia* : « il se pourrait que l'infaillibilité fantasmée de la data prenne le relais dans l'imaginaire collectif »<sup>60</sup>. Dans quelle mesure la « data » serait-elle au cœur des représentations contemporaines ? Il convient aussi de préciser que l'idéologie du « système web » et son essence même « reposent sur un paradigme mathématique et statistique : le calcul, la mesure et le comptage, qui permettent la modélisation et la reproductibilité. [...] Le système numérique, c'est d'abord le domaine du nombre, un monde régi par le chiffre plus que par le texte »<sup>61</sup>. Nous pouvons peut-être comprendre l'importance accordée aux « dispositifs big data » en comprenant d'abord sur quels piliers repose notre société numérique, en tentant d'examiner pourquoi ce « système web », qui produit des millions de téraoctets de traces qu'on nomme « données », est plébiscité et occupe une place récurrente dans les discours des acteurs économiques tels que les sociétés de conseil.

Lorsque les cabinets de conseil en technologie convoquent la notion de « data » et d'algorithmes, il s'opère une autolégitimation de leurs discours, puisque le système de représentations de notre raison occidentale est ancré dans une histoire du calcul rationnel et une confiance en la base de données. À ce titre, l'histoire de la base de données telle qu'expliquée par Stéphan-Eloïse Gras est éclairante. C'est une forme culturelle ancienne qui a été créée pour structurer l'information, qui est une structure informationnelle nécessaire au système de confiance lié à Internet. D'abord inspirée d'un modèle humain (avec la création de la Memex en 1945 qui automatise la capacité neuronale du cerveau à créer des liens entre différentes informations), elle glisse vers un imaginaire d'une immense base de données<sup>62</sup>, à l'origine de la conception du World Wide Web. Les liens entre les informations ne sont plus limités à une table centrale mais conçus comme infinis. Ainsi se construit le principe d'une base de données liquide, déstructurée, dont la forme n'est plus celle d'un tableau centralisé – c'est par ailleurs l'image qui fonde la notion de « big data ». Ces bases de données qui font autorité dans notre raison occidentale sont ainsi des « technostructures » qui héritent de représentations culturelles construites historiquement, et sont au fondement de nos systèmes de confiance occidentaux. Stéphan-Eloïse Gras évoque une « expansion d'une culture « logicielle » et d'une « raison computationnelle »<sup>63</sup>, invitant à considérer les « artefacts numériques » à la lumière de leur « rôles dans la constitution des mondes sociaux et

---

<sup>59</sup> Ibid.

<sup>60</sup> Ibid.

<sup>61</sup> Ibid.

<sup>62</sup> Théorisé par Berbers-Lee en 1997

<sup>63</sup> GRAS Stéphan-Eloïse, « Éthique computationnelle et matérialisme numérique : l'apport des Software Studies », *Critique* 2015, n° 819-820, pp. 667-679 (p. 671)

culturels<sup>64</sup> ». En considérant ici les « big data » à travers leurs représentations et leur effectivité dans la société, cela permet d'analyser les discours d'accompagnement des cabinets de conseil sur leur offre de « services big data » à l'aune des imaginaires qu'ils mobilisent, et de comprendre en quoi ils construisent leur discours sur des mythes qui opèrent dans la société. Ainsi, lorsque le cabinet Affini-Tech affiche dans sa présentation sur la page d'accueil de son site web : « Nous mettons en œuvre des expertises technologiques sur des moyens modernes de traitement de la donnée<sup>65</sup> » pour « Explorer les données, trouver de nouveaux gisements de valeur » - on remarque qu'il prend appui sur le système de confiance ancré dans notre raison occidentale, étroitement liée à l'histoire du calcul rationnel et à cet imaginaire de la base de données infinie. Dans leurs discours d'accompagnement, ils valorisent ainsi la technologie en action basée sur des algorithmes de « machine-learning », et qui fait la prouesse de créer des « gisements de valeur » à partir d'une méthodologie exploratoire de « données ». Ces « gisements de valeur » semblent toujours en construction, puisque basés sur une base de donnée liquide et déstructurée, à l'image de la vision du *World Wide Web* qui fait la promesse de donner accès à une connaissance infinie.

Ainsi, à partir d'une rhétorique qui s'ancre dans une histoire de la « rationalité occidentale » et emprunte à l'univers de la science – consistant notamment à valoriser la technologie et les algorithmes qui tournent – les cabinets de conseil et leurs consultants en technologie tentent de construire des discours d'escorte qui font autorité, dans un processus de légitimation de ceux-ci.

#### b. La fonction incantatoire de ces technologies : l'effet de croyance à l'œuvre dans les discours

Dans cette logique qui vise une stratégie de légitimation, nous remarquons que les discours d'accompagnement des cabinets de conseil en technologie prennent non seulement appui sur un modèle de confiance lié à la raison occidentale, mais aussi sur un phénomène de croyance. À l'image de la thèse développée par Stéphan-Eloïse gras à propos de notre système de confiance occidental, raison et croyance ne peuvent pas être pensés indépendamment l'un de l'autre. En cela, les notions de « big data » et de « data science » conviennent d'être analysées à travers l'effet de croyance qu'elles suscitent. Lorsque l'objet « big data » est mobilisé dans les discours, il semble revêtir une valeur incantatoire, et semble faire autorité par le fait même de l'invoquer.

Lorsque Novagen Conseil ou Affini-Tech convoquent les images liées aux bases de données, ils invitent à intégrer leur offre de conseil « big data » dans notre système de confiance – et ici, celui des acteurs économiques et des entreprises. Ce modèle de confiance, construit à partir d'un système de représentations qui adhère au principe de la base de données et aux structures qui

---

<sup>64</sup> Ibid.

<sup>65</sup> Cf. Annexe 1. Corpus d'étude. f. Site web Affini-Tech – *Nos Services*

produisent de la valeur à partir des données, sous-tend précisément un effet de croyance. Ainsi une véritable isotopie autour de la base de données se retrouve dans leurs discours, à l'image du cabinet Novagen :

« Maîtriser sa complexité par l'Analyse de Données textuelles de données massives

Extraire de l'Intelligence de vos bases de données internes, de votre GED, Historique Documentaire... »

« Maîtriser son Marché par le traitement Analytique des Données du WEB

Vous voulez découvrir de nouveaux usages en vous ouvrant à des données externes et prêtes à l'emploi ?<sup>66</sup> »

Le cabinet cite un nombre de « données » de différentes natures, comme pour montrer que leur offre de conseil couvre tout le champ de la « data » (données textuelles, massives, internes, externes, issues du web). Il cite même une autre manière de travailler les données qu'avec le « big data » lorsqu'il cite les « technologies Smart Data<sup>67</sup> » dans son champ de compétences. Ainsi Novagen crée-t-il une cohérence isotopique dans son discours, montrant que la donnée est au cœur de son modèle économique. Plus encore, leurs discours, relevant du registre de la croyance, peuvent se comparer à une formule incantatoire, dans la mesure où le terme « données » est répété plus de six fois dans leur rubrique « Data Factory ». Les consultants en technologie structureraient ainsi leurs discours selon une logique qui convoque la data telle une incantation qui serait récitée, à l'image de ce qu'on trouve sur le site web de Novagen : « Des équipes dédiées au Nettoyage et à la fouille des Données. Epurer, nettoyer, reformater, agréger les données ». Répété deux fois de suite (*via* « Data Analysis » et « Data Science<sup>68</sup> ») à la manière d'une formule psalmodiée, elle est accompagnée de gestes rituels (épurier, nettoyer, agréger) – qui seraient ici la méthodologie employée dans le Data Factory de Novagen Conseil. Celle-ci suit en effet un enchaînement précis : source > collecte de « datas » > traitement des données > analyse<sup>69</sup>. L'incantation de cette formule ainsi que la ritualisation de leurs procédures méthodologiques viseraient ainsi à faire advenir, une valeur contenue dans ces données mêmes.

En outre, le flou définitionnel de la notion même de « big data » permet d'invoquer le terme sans le définir précisément. La notion se trouve ainsi convoquée comme si elle allait de soi, et comme si le fait de la convoquer provoquait un certain effet, celui de créer de la valeur, à l'image de la formule du cabinet Affini-Tech qui revêt en cela une valeur incantatoire: « Explorer les données, trouver de nouveaux gisements de valeur<sup>70</sup> ». On voit ainsi que les « technologies big

---

<sup>66</sup> Pour illustrer, voir Annexe 1. b. Site web Novagen – Data Factory – *Ce qu'apporte le data lab* ([www.novagen.tech/data-factory](http://www.novagen.tech/data-factory))

<sup>67</sup> Ibid.

<sup>68</sup> Voir Annexe 1. d. Site web Novagen – *Conseil IT* ([www.novagen.tech/conseil-it](http://www.novagen.tech/conseil-it))

<sup>69</sup> Cf. Annexe 1. b. Site web Novagen – *Data Factory – Notre Méthodologie* ([www.novagen.tech/data-factory](http://www.novagen.tech/data-factory))

<sup>70</sup> Cf. Annexe 1. f. Site web Affini-Tech - *Nos Services « Imaginer, Expérimenter, Construire & Exécuter »*

data » sont un objet presque magique mobilisé dans les discours de ces cabinets. Invoquées à la manière d'une formule incantatoire, on se retrouve face à un effet magique du discours qui convoque les « big data ». Ces acteurs de la « data » entretiennent ainsi un système de croyances et entrent en cohérence avec celui-ci. L'article d'*Influencia* à propos de « L'onction du nombre » souligne cette valeur incantatoire de la « data », caractérisée par une « invocation permanente sans la mettre à l'épreuve<sup>71</sup> ». En cela les « technologies big data » s'inscrivent dans le registre de la croyance, une croyance qui fait autorité aujourd'hui. L'article explique cet effet de croyance lié à la fascination pour la « data », en évoquant que nous aurions l'« espoir d'une transformation magique d'un savoir en pouvoir, à caractère certain, puisque fondée sur la rationalité du nombre et du calcul ». C'est cette « transformation magique d'un savoir en pouvoir » qui semble à l'œuvre dans les discours des cabinets de conseil lorsqu'ils convoquent les « technologies big data ». Cette valeur incantatoire serait aussi liée au fait que ces technologies sont issues de l'univers du calcul rationnel, ce qui rassure et leur confère un caractère quasi « infaillible », fantasmé dans l'imaginaire collectif. Ainsi, derrière la croyance à l'égard de la « data », des « grosses données » et des bases de données, il existerait ainsi « une nouvelle forme de pouvoir fondé sur la rationalité absolue du nombre et du calcul ». En promettant une rationalité « infaillible », le système technologique des « big data » paraît se muer en un système de croyance, registre dans lequel s'inscrivent les consultants-ingénieurs en technologie.

Les discours d'accompagnement des cabinets de conseil sur leur offre de service « big data » se construisent à partir des mythes opérants dans la société. En convoquant régulièrement la notion de « données » ou de « data » dans leurs discours, ils entrent en cohérence avec un imaginaire social qui fantasme la « data », avec un modèle culturel qui prône une « rationalité du nombre et du calcul<sup>72</sup> », et avec les imaginaires du « numérique ». Ainsi les discours d'escorte de ces cabinets de conseil entrent-ils en résonance avec une utopie technologique de nos sociétés, où valeur du nombre, modélisation et calcul se trouvent valorisés et légitimés.

### c. La « puissance communicationnelle » du récit de l'innovation technologique : la réactivation du mythe du progrès et l'« impératif de croissance<sup>73</sup> »

Pour ces cabinets qui vendent leur offre de conseil en « big data », il s'agit pour eux de construire un « ethos algorithmique<sup>74</sup> » afin de susciter l'adhésion du public auquel ils s'adressent

---

<sup>71</sup> Cf. BERTIN Erik, « Data, l'onction du nombre », *Influencia*, 16 octobre 2016. URL : <http://www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html>

<sup>72</sup> Cf. Ibid.

<sup>73</sup> Expression tirée du philosophe Hermut Rosa

<sup>74</sup> Cf. ALSAFAR Ali (URL : [tel.archives-ouvertes.fr/tel-01146590/document](http://tel.archives-ouvertes.fr/tel-01146590/document), dans *Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012* », explique

(ici les entreprises) à leurs idées (qui peuvent se résumer à la puissance des « technologies big data »). Pour cela, on a vu qu'ils s'appuient sur une certaine stratégie discursive fondée sur la rationalité du nombre et du calcul. À cela, s'ajoute la construction d'un récit sur l'innovation technologique, accompagné d'une rhétorique de l'impact et de la révolution, qui vient réactiver le mythe du progrès apparu pendant la révolution scientifique mécaniste du dix-septième siècle<sup>75</sup> et qui a marqué la philosophie de la modernité. Dans notre entretien avec les co-fondateurs de Novagen, ce discours de la « révolution » est présent, ils répètent des formules telles que c'est « extraordinaire<sup>76</sup> » (cinq fois : « une valeur ajoutée qui est assez extraordinaire », « Il y a aujourd'hui un foisonnement extraordinaire de technologies ») et parlent de disruption quand ils évoquent le « big data », avec la phrase suivante : « il y a des approches vraiment disruptives, en rupture avec ce qui se faisait auparavant ». Aussi, quand ils affirment « qu'aujourd'hui on peut vraiment cibler des comportements très complexes, ou avoir la victoire de Google au jeu de Go, ou encore l'histoire de la machine sur l'homme, ça laisse à penser qu'on change quand même de monde », nous voyons qu'ils ont une certaine vision technocentrée de la société, avec un changement de paradigme qui se ferait par la technologie.

Ces éléments de langage produits par ces cabinets de conseil en technologie autour de la révolution liée à ces « nouvelles technologies » telles que Hadoop, Spark<sup>77</sup>, autour de « ces nouvelles plateformes technologiques qui permettent de faire des choses qu'on ne pouvait pas faire avant » revêt une « puissance communicationnelle<sup>78</sup> » en tant qu'ils rencontrent un écho dans l'imaginaire collectif contemporain lorsqu'ils évoquent l'innovation par le « big data ». À la lumière de la notion de « prédiscours » telle qu'évoquée par Alice Krieg-Planque<sup>79</sup>, les discours sur les « technologies big data » s'ancreraient dans « des croyances et savoirs partagés » et « dans un

---

la notion d'ethos dans l'art rhétorique grec : « désigne l'image de soi que l'orateur construit dans son discours pour contribuer à l'efficacité de son dire », et dans les sciences du langage, Maingueneau définit l'ethos comme « ce que l'orateur prétend être, il le donne à entendre et à voir » (p. 7)

<sup>75</sup> L' « idée de Progrès », apparu lors de la révolution scientifique mécaniste du XVIIe siècle et conçue comme « mouvement ascendant et nécessaire » de la société, marque la philosophie de l'histoire de la modernité selon l'historien des idées Frédéric Rouvillois. URL : [www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php](http://www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php)

<sup>76</sup> Voir Annexe 2. Entretien avec Novagen, cabinet de conseil IT et Big Data

<sup>77</sup> Ibid., « Derrière le terme de big data, il y a un aspect qui est technologique, qui est en fait représenté par des nouvelles technologies qui sont accessibles, des nouvelles plateformes. Alors ça s'appelle Hadoop, Spark, Scala, avec des langages associés. Ce sont tous ces nouveaux outils, toutes ces nouvelles plateformes technologiques qui permettent de faire des choses qu'on ne pouvait pas faire avant »

<sup>78</sup> Dans *Communication et organisation*, Patrick Chaskiel évoque la « puissance communicationnelle » liée à l'emploi du terme risque dans les discours (n°45, p. 9). On peut appliquer ici cette notion de puissance liée à l'emploi du terme « innovation » dans les discours des cabinets de consulting. (Voir GARDÈRE Élisabeth et D'ALMEIDA Nicole, « Risques mineurs, changements majeurs », *Communication et organisation*, 45, 2014, mis en ligne le 27 octobre 2014. URL : [communicationorganisation.revues.org/4456](http://communicationorganisation.revues.org/4456))

<sup>79</sup> KRIEG-PLANQUE Alice, *Analyser les discours institutionnels*, Paris, Armand Colin « ICOM », 2012 (chapitre 6. Polyphonie, dialogisme, interdiscours : l'ouverture du discours, p.186-210) : La notion de « prédiscours » « permet de penser la question des données antérieures à la mise en discours ». L'auteur ajoute : « La notion aide à interroger les questions liées au sens commun, à la doxa, aux préjugés, ou encore à la mobilisation d'une mémoire du sens ou des discours. »

ensemble de cadres sémantiques collectifs ». Ce qui expliquerait la puissance sémantique des discours sur les « technologies big data ». La notion de « prédiscours » met en lumière le fait que « chaque mot, en tant que sédimentation d'usages et résultante instable de ses trajectoires antérieures, constitue une « arène en réduction », un lieu de conflictualité sociale<sup>80</sup> ». C'est en ce sens que les discours de ces cabinets sur les « technologies big data » revêtent une « puissance communicationnelle », où se cristallisent des systèmes de valeur et imaginaires associés à un imaginaire de la technique et de l'innovation, à une certaine idée de la modernité caractérisée par la disruption par les technologies – qui permettrait de rendre une entreprise plus efficace, plus performante, plus productive.

Ce récit de l'innovation permise par les nouvelles technologies se voit non seulement dans l'entretien avec le cabinet mais aussi sur leur site web, lorsqu'ils présentent leur « Data Factory » sous l'angle là encore de l'innovation : « Agir comme une fabrique de solutions Data innovantes en usages et en technologies Smart Data – Big Data<sup>81</sup> » ; « Vous faire profiter des dernières innovations à travers des « Formations axées mise en œuvre et le partage de nos savoir-faire<sup>82</sup> » ; « Le croisement des innovations et des explorations est créateur de valeur ». Cette thématique de l'innovation est centrale dans le discours de Novagen, qui associe l'innovation à de la création de valeur pour les entreprises. Aussi, en définissant sa baseline autour de la notion d'innovation (« Novagen, Générateur d'innovation »), le cabinet-conseil Novagen met cette thématique au cœur de son identité, faisant de l'innovation l'un des fondements de son « ethos algorithmique », et se positionne comme instance énonciative qui propose à ses clients potentiels de « faire profiter des dernières innovations ». On voit bien que ce type de discours est associé à l'imaginaire autour des nouvelles technologies qui supposent que la société connaît un phénomène d'innovation croissant, une représentation d'un sens de l'histoire qui serait linéaire et ascendant. Ici ces cabinets véhiculent l'idée que les « technologies big data » vont dans ce sens ascendant du progrès, s'ancrent dans cette modernité qui fantasme une efficacité économique des entreprises qui peut toujours être amélioré, optimisé. Cette valorisation de l'innovation, notion centrale dans le discours de Novagen et faisant partie intégrante de la construction de son « ethos algorithmique », se retrouve aussi dans d'autres cabinets, tel que sur le site de la société C-Radar : « Intégrez à haute dose l'innovation technologique et digitale dans votre processus de ciblage<sup>83</sup> ». Ce type de discours sonne presque comme une injonction à l'innovation, un impératif. Comme si pour atteindre cet idéal d'innovation, il fallait passer par un processus de ciblage marketing, qui permettrait d'atteindre un « impératif de croissance », tel que théorisé par le philosophe et sociologue Harmut Rosa. En effet,

---

<sup>80</sup> Ibid.

<sup>81</sup> Cf. Annexe 1. Corpus d'étude. b. Site web Novagen – Data Factory ([www.novagen.tech/data-factory](http://www.novagen.tech/data-factory))

<sup>82</sup> Cf. Annexe 1. Corpus d'étude. c. Site web Novagen – Conseil Data ([www.novagen.tech/conseil-data](http://www.novagen.tech/conseil-data))

<sup>83</sup> Société spécialisée en marketing prédictif B2B. [www.c-radar.com/systeme-predictif-marketing-b2b](http://www.c-radar.com/systeme-predictif-marketing-b2b)

ce fantasme d'une société qui se dirigerait vers une innovation constante, peut être mis en regard des travaux de H. Rosa lorsqu'il parle de la « pathologie » de l'accélération propre à nos sociétés modernes : « comme la société moderne et capitaliste ne reproduit sa structure qu'à travers l'innovation et l'accélération – ce que je qualifie de « dynamisation du temps » –, on a besoin de croître et d'innover pour préserver les structures économiques, sociales, politiques<sup>84</sup> ». Comme l'évoque Julien Tassel qui a travaillé sur l'histoire d'entreprise et la sociologie de l'innovation<sup>85</sup>, l'innovation peut s'analyser comme une mythologie. Ce sont des récits qui se donnent comme allant de soi, comme ayant une certaine « naturalité », mais qui configurent la réalité de ce que peut être l'organisation<sup>86</sup>. Les éléments de langage des consultants en technologie, *via* leur récit de l'innovation, se veulent être un discours pour l'action, diffuser une vision organisationnelle pour l'avenir de leurs entreprises clientes. Nobert Alter évoque aussi la tendance à adopter une vision de l'entreprise comme étant en perpétuel mouvement, en changement et en adaptation – elle doit être flexible et créatrice, et cela correspondrait à un récit de l'entreprise contemporaine, selon lequel l'entreprise doit être de plus en plus un lieu créatif où s'invente la nouveauté en permanence. Dans cette perspective, le cabinet de consulting Affini-Tech qui, par l'exploration des données propose de « trouver de nouveaux gisements de valeur », contient cet « impératif de croissance » et se positionne dans le mythe opérant de la croissance propre à notre société moderne. Il met en avant des technologies et plateformes capables de produire cette croissance, d'atteindre cet idéal d'innovation, et cristallise « les espoirs d'une nouvelle forme de puissance pour la société » comme l'évoque l'article d'*Influencia* « Data, l'onction du nombre ». C'est ainsi sur ces leviers que semble reposer « la puissance communicationnelle » des discours de ce cabinet sur les « technologies big data ».

Puisant dans ces imaginaires sociaux collectifs, les cabinets de consulting en technologie tentent de se positionner comme une instance énonciative légitime sur le marché des « grosses données », participant à construire un ethos propre à l'univers des « technologies big data », ce que nous avons appelé un « ethos algorithmique ». La construction de cet ethos propre aux cabinets de consulting en technologie s'appuie sur un processus de légitimation par le discours qui s'ancre dans ces différents imaginaires et mythes liés aux nouvelles technologies et à l'innovation, qui se trouvent valorisés dans la sphère socio-économique. Ce processus de légitimation *via* les

---

<sup>84</sup> Le Monde, « Hartmut Rosa : Plus on économise le temps, plus on a la sensation d'en manquer », 1<sup>er</sup> Avril 2016. URL : [www.lemonde.fr/m-perso/article/2016/04/01/hartmut-rosa-plus-on-economise-le-temps-plus-on-a-la-sensation-d-en-manquer\\_4893818\\_4497916.html](http://www.lemonde.fr/m-perso/article/2016/04/01/hartmut-rosa-plus-on-economise-le-temps-plus-on-a-la-sensation-d-en-manquer_4893818_4497916.html). Voir aussi : WAHL Iodie, « Hartmut Rosa, Accélération. Une critique sociale du temps », Lectures, Les comptes rendus, 2010, mis en ligne le 16 avril 2010, consulté le 23 août 2017. URL : <http://lectures.revues.org/990>

<sup>85</sup> Cf. Cours de Master 1 Communication des entreprises et des institutions

<sup>86</sup>Op. cit., il évoque notamment les travaux de SCHUMPETER (1912)

imaginaires est aussi complété par une mise en spectacle de leur expertise, une monstration de leurs compétences techniques et en tant qu'ingénieurs spécialistes de la donnée.

## 2. La mise en scène et la monstration de l'expertise

L'enjeu pour ces cabinets de consulting et leurs consultants est d'être reconnu comme experts, et l'expert est celui qu'on reconnaît, dont on ne questionne pas la légitimité du savoir, lequel fait autorité. Dans cet objectif de se positionner comme experts dans le domaine du conseil en technologie, ces acteurs mettent en spectacle leurs compétences – techniques telles que la maîtrise des « technologies big data ». Nous observons que leur « ethos algorithmique » est ainsi mis en scène et fait partie d'une scénographie plus générale, qui met en avant leurs compétences et leur savoir-faire liés aux technologies et aux « data ».

### a. Une mise en scène de leur « ethos algorithmique »

Dans leur stratégie de légitimation de leur offre de conseil et dans l'objectif de se positionner comme instance énonciative reconnue comme experte sur le marché des données, ces consultants-ingénieurs mettent en scène leur « ethos algorithmique » dans une « scénographie ». Celle-ci, selon Maingueneau, est « la scène de parole » qui conditionne l'énonciation du discours et qui permet que le discours soit énoncé, c'est-à-dire « le processus par lequel une scénographie légitime un énoncé qui, en retour doit le légitimer<sup>87</sup> ».

Pour les cabinets de conseil, la finalité du discours est à la fois didactique et persuasive, ces cabinets s'adressant à leurs clients potentiels pour les informer sur leurs services, et les convaincre d'avoir recours à leurs services pour leurs « projets big data ». Ainsi la « scénographie » d'Affini-Tech repose d'abord sur un énoncé didactique qui vise à informer les interlocuteurs ciblés (les entreprises susceptibles de faire appel à eux) : son site web présente d'abord ses services, les exemples de projets que les consultants proposent, puis les technologies qu'ils utilisent. Il présente enfin ses clients actuels qui leur ont fait confiance, visant à persuader leurs prospects *via* une accumulation visuelle sur la page des logos des entreprises qui lui ont déjà fait confiance (des grandes entreprises qui font autorité dans la sphère économique : SNCF, Leroy Merlin, Essilor). Nous avons ici une mise en scène de la crédibilité du cabinet de conseil, visant à bâtir une crédibilité, à inspirer confiance. Il s'agit d'un processus de légitimation à travers des instances

---

<sup>87</sup> Cf. ALSAFAR Ali. *Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012. Linguistique*. Université Paul Valéry - Montpellier III, 2014 : voir p. 18 sur « ethos et scène d'énonciation » selon D. Maingueneau

externes. Ainsi l'« ethos algorithmique » d'Affini-Tech est mis en scène autour de ces leviers didactiques et *via* des figures actoriales externes.

Le site Internet du cabinet Novagen est construit sur un schéma d'énonciation plus nuancé. Il commence lui aussi par présenter leurs services, afin de situer leur offre, de cadrer le discours et démontrer qu'ils sont en accord avec les besoins des entreprises qui visitent leur site web. Novagen s'appuie sur une démonstration au caractère davantage didactique et démonstratif, puisque leur site comporte trois pages différentes (à l'inverse d'Affini-Tech qui est construit sur une seule page fixe) qui correspondent à trois offres différentes : *Conseil Data*, *Conseil IT*, *Data Factory*. Dans leur offre *Conseil Data*, ils se focalisent sur leur approche de « consulting in situ » et leur approche « Datalab », les étapes du projet, comme pour rassurer sur la mise en œuvre concrète lorsqu'on travaille avec eux, pour montrer une facilité à travailler avec eux, une fluidité dans les relations de collaboration. Ils mettent en visibilité la thématique de la collaboration, de l'accompagnement, et leur capacité à identifier les besoins de leurs clients, qui sont autant de sujets qui structurent leur discours. Pour illustration, la rubrique « Services » affirme : Nous identifions avec vous les étapes et jalons de votre stratégie Data et Big Data. Nous accompagnons nos clients PME-PMI / ETI et Grands Groupes dans l'élaboration de cette feuille de route ». Novagen met aussi en scène ses compétences en tant que cabinet capable d'identifier et de comprendre les besoins de ses clients. Dans son offre Conseil IT, il met en avant ses consultants, avec une expertise détaillée des consultants définis par spécialité : consultants ingénieurs études, développeurs tech, AMOA, PMO, « data analyst », « data scientist » ; formation, « Connaissance approfondie des besoins clients »... Est ainsi mis en scène le capital humain, ses différentes expertises et domaines d'interventions, et des ingénieurs formés à avoir l'« engagement de résultats ». L'« ethos algorithmique » de Novagen est ainsi mis en scène autour des thématiques des compétences, de l'expertise, de la maîtrise de son métier (capacité à comprendre les besoins clients, maîtrise de différents sujets, variété d'expertises). Enfin, dans son offre *Data Factory*, il explique avant tout la nature de son offre : sous la rubrique « Pourquoi une Data Factory ? » il démontre « Ce qu'apporte le Data Lab » ainsi que sa méthodologie appuyée par un schéma qui permet de visualiser les différentes étapes. La rubrique « Pourquoi nous ? », appuyée par une démonstration chiffrée en pourcentages, et « Les Cas d'usage », les « Exemples de prestations » dans son document de présentation<sup>88</sup>, son « positionnement »<sup>89</sup> et son « approche »<sup>90</sup> où il explique point par point sa manière de travailler avec leurs clients<sup>91</sup> : « On réfléchit avec vous », « On vous accompagne », « On développe un Partenariat »), et rassure sur l'accompagnement de ses clients sur leurs projets data (cf. « les

---

<sup>88</sup> Cf. Annexe 8. Support de présentation de Novagen (slides Powerpoint en PDF) – slides de la p. 12 à la p. 14

<sup>89</sup> Op. cit., slide p. 4

<sup>90</sup> Op. cit., pp. 8-9

<sup>91</sup> Op. cit., slide p. 8

premiers pas avec un client »<sup>92</sup>). Ce sont autant d'éléments de langage de la démonstration et de l'explication, qui visent à aider les clients potentiels à se projeter dans son offre de services. Nous avons donc ici un discours avec une tonalité à la fois pédagogique et persuasive, fondé sur des explications, des justifications, des chiffres, un schéma. Novagen met en avant ces éléments explicatifs pour aider à la compréhension de son offre de services, et pour contribuer à convaincre son auditoire d'entreprises qui visitent son site Internet – auditoire auprès duquel Novagen tente de faire valoir les compétences de ses ingénieurs « data-scientists » ainsi que leur posture d'expertise.

#### b. Une mise en avant de l'univers de la « data-science » : le « Data Lab »

Toujours dans cette même logique de mise en scène et « scénographie » de leur « ethos algorithmique » pour qu'ils soient reconnus comme experts dans leur domaine (les problèmes technologiques et liés aux données), ces cabinets de consulting mettent en avant l'univers de la « data-science » sur leurs sites Internet. Serait-ce pour provoquer un effet de fascination pour mieux faire adhérer à leur discours commercial ? Il convient de rappeler que ces cabinets de consulting, fondés par des ingénieurs, ont une vision très « data » de l'action, ils conçoivent l'action à partir de la donnée, et s'appuient sur la donnée au quotidien. À ce titre, on observe l'utilisation d'un lexique dédié à l'univers de la « data » : « Conseil Data », « Stratégie Data », « Projets Big Data », « Smart Data », « Data analyst », « data-science », « Data Lake », « Data Lab », « Data Factory », « architectures Big Data », « collecte data ». Des mots-valises constitués à partir d'un préfixe base qui ne bouge pas (« data »), utilisation linguistique qu'on peut analyser comme étant à l'épicentre de leur stratégie discursive, duquel découlent ensuite les autres éléments de langage. Outre ce lexique dédié, le cabinet Novagen met en avant son « Data Factory » qu'il nomme aussi « Data Lab », le substantif « Lab » étant utilisé comme pour évoquer un laboratoire où se fabrique le savoir, c'est l'imaginaire du chercheur faisant des découvertes, ici un chercheur qui travaille la donnée pour en faire émerger un savoir nouveau. Ce « Data Lab » ou « Data Factory » est mis en avant sur leur site Internet comme un lieu spécifique où le travail autour de la donnée se fait, les problèmes se résolvent, des découvertes de nouvelles solutions sont faites. Dans ce lieu, Novagen donne à voir « Des équipes dédiées au Nettoyage et à la fouille des Données », son discours nous fait nous figurer des équipes d'ingénieurs en action, qui travaillent activement à « épurer, nettoyer, reformater, agréger les données », à « évaluer et valider les modèles statistiques<sup>93</sup> ». La « Data

---

<sup>92</sup> Op. cit, slide p. 9

<sup>93</sup> Annexe 1. Corpus d'étude. d. Site web Novagen – Conseil IT ([www.novagen.tech/conseil-it](http://www.novagen.tech/conseil-it))

Factory » est visuellement représentée par un schéma<sup>94</sup> qui donne à voir la méthodologie particulière qu'ils tentent de construire (sélection des sources de données, collecte, traitement des données, analyse) : il montre un savoir-faire en action, qui fabrique (« prototypage », « développement des projets ») – le propre d'une « Factory », d'une usine donc, étant de produire, de fabriquer des objets. Le choix des icônes vient compléter cette vision d'une fabrique d'un livrable à partir de « data » : une icône wifi, qui signifie le réseau, le nuage qui signifie le « cloud », soit l'accès aux données ; une boîte à tiroirs pour signifier la collecte de données et la classification ; des pièces d'engrenage cylindriques signifiant le mouvement mécanique, l'action propre à l'univers de la machine, évoquant la transmission mécanique par des courroies intégrée à un dispositif plus général, celui de l'usine, de la production automatisée. Enfin, l'icône du tableau avec un diagramme dessiné pourrait signifier les compétences d'analyse du cabinet. Novagen met ainsi en scène la « data-science » en action, en empruntant à l'univers de la méthodologie scientifique (le laboratoire), de la mécanique (engrenages), de l'ingénierie (prototypage). Le « Data Lab » est ainsi mis en scène comme lieu dans lequel les « data-scientists » exercent leurs compétences en « data science ». Ce « data lab » constituant une rubrique à part entière sur le site web de Novagen<sup>95</sup>, met aussi en avant « les bénéfices des Architectures Big Data ». Nous avons donc à la fois une mise en scène du savoir qui se fabrique dans ce lieu dédié à la « data », et une valorisation de la technologie qui permet de tirer des « bénéfices » - sans pour autant préciser la nature de ceux-ci. Aussi, Novagen se proclame d'une méthodologie scientifique, comme l'illustre l'expression « Sciences et Big Data » de leur Datalab interne<sup>96</sup>, l'« utilisation avancée de la théorie des Graphes » (cf. application de Social Media Intelligence<sup>97</sup>). Le cabinet montre qu'il sait utiliser une théorie mathématique. Les cabinets de conseil en technologie se dotent de signes qui appartiennent à l'univers de la science (le laboratoire, la fiole, le tube, la découverte avec le bulbe chez Affini-Tech par exemple), et mettent avant le métier de « data-scientist » (sur le site de Novagen).

La notion d'« opérativité » de ces représentations symboliques expliquée par Etienne Candel est ici éclairante : l'ingénieur « data-scientist » remplit son « rôle » d'expert, c'est-à-dire qu'il fait exactement les postures qu'on attendrait de lui. Ces postures sont la capacité à travailler la donnée, l'utilisation d'un lexique technique, le lieu de travail de ces consultants-ingénieurs (ici le « Data Lab »). Aussi, le concept de « prétention » de Yves Jeanneret est éclairant pour expliquer ce phénomène d'expertise, il écrit : « un métier, une compétence ne peuvent s'imposer sans communication ; leur valeur et leur image sont construites par un récit<sup>98</sup> ». En effet, ces cabinets ne

---

<sup>94</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

<sup>95</sup> Ibid.

<sup>96</sup> Cf. Annexe 8. Support de présentation de Novagen (slides Powerpoint en PDF) – slide p. 3

<sup>97</sup> Cf. Annexe 8. Support de présentation de Novagen (slides Powerpoint en PDF) - slide p. 12

<sup>98</sup> Cf. TËTU Jean-François, « Yves Jeanneret, *Penser la trivialité*. Volume 1 : *La vie triviale des êtres culturels* », *Questions de communication*, 2009, mis en ligne le 17 janvier 2012, consulté le 19 novembre 2016. URL :

sont experts qui si on les reconnaît comme tels, qu'à travers le regard de ceux qui les identifie experts. Les cabinets de consulting ont ainsi tout intérêt à construire un récit qui mettent en scène leurs compétences et leur savoir-faire, afin d'être en cohérence avec une posture d'expert en « data-science », rationnelle et technique.

### c. La mise en scène de compétences techniques, du savoir-faire de l'ingénieur, de la « prouesse technique » et de sa maîtrise

Pour renforcer cette posture d'expert en « data-science », il semble que ces cabinets mettent en scène deux éléments qui vont créer un effet de fascination à l'égard de ces cabinets de conseil : démontrer la « prouesse technique » que constituent ces « technologies big data », et démontrer la maîtrise de ces technologies (et leur connaissance technique). C'est aussi sur ces leviers que les cabinets de consulting construisent leur légitimité et renforcent leur posture d'expert. Cela va passer par la mise en avant de leur savoir-faire dans leur présentation auprès des entreprises, comme le cabinet Novagen pour tenter de convaincre Ipsos de recourir à leurs services « big data »<sup>99</sup>.

Avec des éléments de langage comme : « Maîtrise des technologies Big Data. Croisement des compétences Data Science / Machine Learning. Veiller pour rester au fait des dernières avancées », nous avons une mise en scène de leur savoir-faire, mis en exergue dans deux slides à part entière<sup>100</sup> : à la p. 3, le savoir-faire est donné à voir *via* quatre sphères qui représentent chacune une compétence technique (*big data, data science, machine learning, deep learning*). Ce schéma fait penser à un ensemble de planètes, formant un système interdépendant dont les compétences forment une orbite autour d'un noyau central, celui des « projets et solutions innovantes ». Ce serait alors cette combinaison de ces compétences particulières qui crée le savoir-faire « innovant » de Novagen. La slide en p. 7 est moins visuelle mais donne à voir sous forme de texte un savoir-faire qui serait basé sur la « maîtrise », le « croisement de compétences Data Science / Machine Learning », sur la veille sur les « dernières avancées ». On voit ici l'effort de mise en discours et en mise en visibilité d'un savoir-faire technique (du « machine-learning », du « big data », de la « data science »), avec le souci de valoriser celui-ci et le rendre lisible auprès de leurs clients potentiels. Enfin, ces cabinets proposent des formations à ces « technologies big data », ce qui vient renforcer leur parole d'expert. Novagen propose une offre de formation qui s'adresse aux « décideurs », une formation

---

<http://questionsdecommunication.revues.org/176> (cf. Effeillage, « La trivialité pensée par Yves Jeanneret », <http://effeuillage-la-revue.fr/portfolio-item/la-trivialite-pensee-par-yves-jeanneret/>)

<sup>99</sup> Cf. Annexe 8. Support de présentation de Novagen – slide n°3 et n°7

<sup>100</sup> Op. cit., slides aux p. 3 et p. 7

« Spark et Scala », Python, Hadoop, Elastic Search<sup>101</sup>. Cette démonstration du savoir-faire passe aussi par un vocabulaire technique spécifique : « entrepôt Hadoop »<sup>102</sup>, « Définition d'une architecture technique et applicative (passage en mode SaaS) »<sup>103</sup>, « les briques infra up-to-date »<sup>104</sup>, les « briques applicatives state of the art ». On a là un lexique technique qui confère un certain pouvoir rhétorique à leurs locuteurs, et participent à crédibiliser le discours de ces consultants en technologie.

Ils mettent en scène leur expertise technique, en montrant qu'ils font fonctionner ces plateformes technologiques, qu'ils tournent des algorithmes, en représentant une certaine mécanique de la machine, une vision entre automatisation, délégation à la machine, et intervention humaine. Soit avec des verbes d'action à l'infinitif (*Collecter, Traiter, Stocker, Segmenter, Activer, Visualiser* dans les exemples de projets sur le site web d'Affini-Tech), on a des éléments de langage qui nous font nous figurer ces technologies en fonctionnement, qui performant. Les « technologies big data » sont représentées comme un système en mouvement permanent *via* des flèches circulaires sur leur page *Conseil Data*<sup>105</sup>, qui montrent à la fois un processus automatisé (les flèches en mouvement circulaire sur elles-mêmes) et une intervention humaine (les mains). L'icône de l'engrenage mécanique de la « Data Factory »<sup>106</sup> rappelle ce processus automatisé propre à la machine qu'ils appliquent à leur « méthodologie ». Ils présentent une méthodologie basée sur la mise en spectacle de la technologie elle-même, en train de faire, de performer, d'agir : ils donnent successivement à voir une technologie qui « source », qui « collecte », qui « traite », qui « analyse ». En représentant visuellement ces technologies qui performant, ils mettent en scène une certaine « prouesse technique » - et donnent à voir des techniques de collecte, de traitement, d'extraction et de valorisation des données, au caractère circulaire et mécanique, en mouvement. Cette performance de la technologie vient appuyer leur stratégie de légitimation et le renforcement de leur posture d'expert.

En donnant à voir ces opérations techniques (algorithmes qui tournent, modèles de calculs appliqués, processus d'automatisation, « machine-learning »), ils viennent légitimer leur posture d'ingénieur « data-scientist », ainsi cette mise en avant de compétences techniques viennent renforcer cette posture de l'ingénieur qui maîtrise les techniques de collecte et de traitement des données. On voit alors que ces cabinets construisent leur « ethos algorithmique » en donnant à voir des processus mécaniques et automatisés à l'œuvre, une certaine délégation à la machine, une expertise technique à la fois automatisée et avec une intervention humaine. La figure d'expert de

---

<sup>101</sup> Op. cit., slide p. 10

<sup>102</sup> Op. cit., slide p. 13

<sup>103</sup> Op. cit., slide p. 14

<sup>104</sup> Op. cit., slide p. 5

<sup>105</sup> Cf. Annexe 1. Corpus d'étude. c. Site web Novagen – Conseil Data ([www.novagen.tech/conseil-data](http://www.novagen.tech/conseil-data))

<sup>106</sup> Voir Annexe 1. Corpus d'étude. b. Site web Novagen – Data Factory ([www.novagen.tech/data-factory](http://www.novagen.tech/data-factory))

ces consultants en technologie se construit donc à travers ces différents leviers discursifs qui se réclament d'un univers scientifique, d'une méthodologie mathématique, d'un savoir-faire technique.

### 3. Le double discours du consultant-ingénieur, à la fois scientifique et commercial : une instance énonciative hybride

Nous avons vu que l'expertise de ces consultants se met en scène de différentes manières : par des discours d'accompagnement qui font partie d'une « scène d'énonciation » qui met en avant l'univers de la science, la mise en scène d'un savoir-faire propre à la « data-science », et d'une certaine « prouesse technique » des « technologies big data ». Or nous remarquons aussi qu'ils adoptent une posture énonciative ambivalente, puisque empruntant à deux univers étroitement liés entre eux, le technique et le commercial, car les « technologies big data » doivent en effet être promues et vendues à travers un discours commercial, et le discours commercial légitimé par la technologie. En analysant notre corpus, nous constatons une énonciation hybride qui mélange à la fois un *ethos* commercial, qui vend son offre de conseil et sa technologie, et un *ethos* ingénieur, qui s'appuie sur un savoir-faire technique, un vocabulaire spécifique.

#### a. Une polyphonie énonciative : la posture ambivalente de ces consultants-ingénieurs

Nous postulons que les discours des cabinets de conseil en « technologie big data » se caractérisent par la présence de plusieurs acteurs *incarnés* (un acteur ingénieur et un acteur commercial), par une « polyphonie énonciative », en référence à la définition qu'en donne Bakhtine : « Dans le domaine littéraire, l'hybridité se produit lorsque deux ou plusieurs discours se disputent l'autorité de l'énonciation<sup>107</sup> ». Ce qu'il définit comme l'existence d'une pluralité d'acteurs *incarnés* dans le discours, nous le retrouvons lors de notre entretien avec le cabinet de conseil Novagen, au cours duquel nous avons identifié deux postures énonciatives différentes, correspondant à deux acteurs *incarnés*, l'ingénieur et le commercial. On pourrait postuler que cette imbrication d'instances énonciatives *incarnées* dans le discours est propre à l'ethos algorithmique tel qu'on l'a étudié précédemment, avec des éléments de langage liés à l'*ethos* ingénieur : la précision du lexique utilisé, l'imaginaire de la technique (« processus de transformation et les algorithmes de

---

<sup>107</sup> Voir sur le site du comptoir des presses d'universités, Résumé. URL : [www.lcdpu.fr/livre/?GCOI=27000100318130](http://www.lcdpu.fr/livre/?GCOI=27000100318130) sur Le texte hybride, Presses Sorbonne Nouvelle, 160 p.

valorisation »<sup>108</sup>, « stratégies data lake, stratégies de réconciliation des données »). À ceux-ci vient s'ajouter une seconde strate discursive qui *incarne l'ethos* commercial, et qui utilise des éléments de spécifiques pour vendre son offre de services, avec l'emploi de superlatifs tels que « une valeur ajoutée extraordinaire », « c'est un champ formidable, c'est-à-dire se confronter à de la donnée peu structurée et rendre la communication interne plus efficace<sup>109</sup> », « des usages et des apprentissages qui sont assez extraordinaires à faire », « ça oblige une réflexion beaucoup plus transversale, et avec une valeur ajoutée qui est assez extraordinaire », « des histoires incroyables à raconter autour de ce qu'on fait dans les données », « aujourd'hui un foisonnement extraordinaire de technologies »<sup>110</sup>. L'*ethos* commercial va mettre en avant l'offre et les avantages pour les clients et prospects (« les *big data* peuvent vous offrir l'avantage très compétitif en les appliquant à vos propres processus internes »). On sent dans leur discours que les cabinets de conseil essaient de nous convaincre d'utiliser ces technologies, qui permettent de « mieux connaître ses clients, on va être plus efficace, faire des précisions sur les ventes beaucoup plus fines, on peut optimiser un métier industriel<sup>111</sup> » - en se référant à l'entretien que nous avons mené avec Novagen.

Cette dualité énonciative propre à la posture de ces consultants transparait dans le discours de Hubert Stefani, le « Chief Innovation Officier » du cabinet, qui nous confie qu'il a un profil à la fois ingénieur en informatique et technico-commercial<sup>112</sup> : « J'aime bien la dualité, être à la fois pertinent techniquement, avoir une large culture technique, et le mettre au service de produits. Depuis 6-7 ans, je suis entre les deux, à un moment les mains dans le code et quatre minutes après je discute avec une Direction stratégique. Il y a une vraie place pour les acteurs techniques au profil ingénieur qui savent vulgariser les enjeux du moment. » C'est cette ambivalence de la figure de l'expert qui a « les mains dans le code » et qui sait négocier avec des décideurs stratégique qu'on retrouve dans les discours d'accompagnement de ces cabinets de conseil, et qui fait la particularité de leur posture de consultant, à mi-chemin entre un discours ingénieur et un discours commercial. On voit bien que les compétences intrinsèques à cet « ethos algorithmique » est de posséder une culture technique (sur les « technologies big data », les algorithmes, le langage du code, les architectures des systèmes d'information) et savoir vulgariser celle-ci dans un discours technico-commercial. L'entretien avec la « data-scientist » d'Ipsos rejoint l'importance de cette capacité à vulgariser les connaissances techniques : « Il faut vulgariser pour le client, et en même temps ça donne plus de légitimité si j'utilise un langage technique<sup>113</sup> ». On retrouve donc cette posture

---

<sup>108</sup> Voir Annexe 2. Entretien avec Novagen, cabinet de conseil IT et Big Data

<sup>109</sup> Ibid.

<sup>110</sup> Ibid.

<sup>111</sup> Ibid.

<sup>112</sup> Cf. Annexe 2. Entretien avec Novagen : « Pour ma part j'ai un profil ingénieur en informatique, j'ai travaillé sur les bases de projets informatiques et de développement web. Puis j'ai évolué vers des fonctions technico-commerciales. »

<sup>113</sup> Cf. Annexe 3. Entretien avec Charlotte Zaepfel

énonciative ambivalente qui jongle sur ces deux types de discours, qui entrent en polyphonie l'un avec l'autre.

Il semble que c'est ainsi que s'auto-légitiment les discours d'accompagnements de ces acteurs de la donnée, en faisant entendre un discours technique, propre aux sciences de l'ingénieur qui apporte une caution technico-scientifique, et en faisant entendre un autre discours plus commercial, vulgarisé, afin de faire comprendre les « technologies big data » qu'ils utilisent dans leur offre de service. Ces deux types de discours se trouvent en interrelation et semblent indissociables l'un de l'autre, faisant de la particularité de la posture d'énonciation – hybride – de ces cabinets de conseil en technologie.

#### b. Le simulacre de la neutralité et de l'objectivité des « technologies big data »

Sous couvert de légitimité technique, mathématique, rationnelle ou scientifique, ces cabinets semblent dissimuler leur intention commerciale. Comment ces cabinets construisent-ils un tel simulacre de la neutralité et font croire à un objet technologique neutre dans leurs discours ?

L'entretien avec la « data-scientist » Charlotte Zaepfel nous a révélé sa posture critique à l'égard des consultants en « data » qui se réclament d'une méthodologie de « data-science » qui utilise les algorithmes de « machine-learning », nous confiant qu'elle se méfiait de l'utilisation trop fréquente des termes « machine-learning » ou « algorithmes » par certains acteurs de la donnée qui veulent impressionner leurs prospects. Selon elle, ces termes sont imprécis et ne désignent par la nature de l'algorithme utilisé (si c'est un algorithme génératif ou analytique par exemple), néanmoins elle reconnaît qu'ils qui font valeur par leur dimension technique et mathématique. Elle souligne le caractère auctorial de ces termes qu'on questionne assez peu : « Après le « big data », ça fait vendre... » et « big data ce n'est qu'un big word<sup>114</sup> ». Ainsi les cabinets de conseil les utilisent comme éléments de langage qui font autorité, et font passer leurs méthodes de travail basées sur la « data » comme objectives et ayant une certaine naturalité, comme si les « technologies big data » allaient de soi, étaient évidentes car issues d'une méthodologie ingénieure et mathématique (que ce soit leurs « architectures Big Data », le « machine-learning » ou le théorème des graphes mentionné dans le support de présentation de Novagen<sup>115</sup>). Dans leurs discours d'accompagnements ils leur confèrent un caractère rationnel et certain, ce qui a pour effet de produire un simulacre d'objectivité, de naturalité. Leurs discours induisent qu'on reconnaisse ces technologies comme objectives, neutres, puisqu'ils leur attribuent une certaine valeur de vérité. Or il convient de signaler

---

<sup>114</sup> Cf. Annexe Entretien avec Charlotte Zaepfel

<sup>115</sup> Cf. Annexe 8. Support de présentation de Novagen

la dimension commerciale des technologies et « plateformes big data » que vendent ces cabinets de consulting, reconnaître qu'elles sont un construit de discours, qui les travaillent tour à tour et leur donnent de la valeur. Il convient aussi de reconnaître que derrière le dispositif technique, il y a une intention humaine qui a conçu ce dispositif et y le travaille avec une certaine subjectivité.

Nous pouvons ainsi dire que les cabinets de conseil construisent leurs services *via* les « technologies big data » qu'ils commercialisent auprès des entreprises comme un simulacre de méthode scientifique, c'est-à-dire qu'ils présentent leurs technologies sous l'angle de la science et d'une méthode mathématique, prétendant en cela détenir une vérité à partir des bases de données. À ce titre, Affini-Tech les promeut de cette manière : « Le savoir-faire de nos Data-Engineer et de nos Data-scientists vous éclairera sur les leviers business, et l'intégration des plateformes nécessaires<sup>116</sup> ». Ils font la promesse d'éclairer les « leviers business » à partir de ces technologies. Novagen, avec son « Data Lab », propose d' « Extraire de l'Intelligence de vos bases de données internes, de votre GED, Historique Documentaire... », de « Maîtriser son Marché par le traitement Analytique des Données du WEB » - émet la promesse qu'à partir des bases de données (travaillées dans le « Data Lab »), on puisse extraire une « intelligence », « maîtriser son marché », disposer d' « outils de Business Intelligence, temps réel<sup>117</sup> ». Cette mise en avant d'une promesse de résultats à partir de l'action de la technologie vient occulter l'action humaine, de l'expert. Cette absence de la figure de l'humaine au profit de la technologie est intéressante : en réalité, l'action et l'interprétation humaine y ont une place importante, mais sont occultées, ne sont pas ici mises en avant sur leur site web – mais évoquées dans les entretiens avec ces acteurs. Ces consultants en technologie créent ainsi, sans le savoir, un simulacre de vérité : ces technologies prétendent détenir une vérité à partir de ces « grosses données », ils supposent une naturalisation de la donnée, or celle-ci est toujours un construit : une donnée est issue d'un processus de sélection des sources de ces données, le consultant « data-scientist » y applique des opérations d'algorithmes, par exemple la « data-scientist » d'Ipsos y applique des « opérations vectorielles » qu'elle a choisies en amont, implique de recommencer, elle nous confie que c'est un processus itératif<sup>118</sup> : « C'est en donnant des exemples d'apprentissage à la machine que l'algo génératif se crée. On crée et on utilise des algorithmes pour que la machine construise des modèles en lien avec l'achèvement d'une tâche. Au niveau des résultats, idéalement on voudrait que l'algo soit aussi bien que le serait un humain. » Elle mentionne que l'humain est à l'origine de la création de l'algorithme, que c'est lui qui décide des modèles d'apprentissage à appliquer à l'ordinateur. Cette tendance à l'effacement de la figure de l'humain derrière l'algorithme et l'automatisation est soulignée par Agnès Gilbert, « Chief Data Officier » d'Ipsos : « On a l'impression qu'il y a avant tout de l'automatisation, des processus

---

<sup>116</sup> Cf. Annexe 1. Corpus d'étude. e. Site web Affini-Tech – Page d'accueil

<sup>117</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

<sup>118</sup> Cf. Annexe 3. Entretien avec Charlotte Zaepfel

automatisés, or en réalité l'humain intervient beaucoup derrière<sup>119</sup> ». Bien que ces consultants nous fassent croire que la machine agit en quasi autonomie, ce sont ces consultants qui décident des modèles de calcul à appliquer. Il s'agit en effet d'un procédé « itératif » qui est chronophage pour eux, or on fait croire que la machine travaille rapidement. Comme le souligne Novagen, le travail de la donnée « est vraiment conséquent », il faut la structurer, la « normaliser ». Puis vient le choix de l'algorithme le plus approprié pour traiter une base de donnée selon un objectif stratégique défini en amont<sup>120</sup>. Ils affirment à ce propos : « il y a une expertise humaine qui est liée à la compréhension du modèle mathématique qui se cache derrière les algorithmes. Et ensuite il faut pouvoir itérer, confronter les résultats qu'on obtient en confrontant les données à des algorithmes, avec les phénomènes observés : est-ce que la traduction est bonne ou mauvaise, et selon l'état, on va entrer encore dans des cycles, on va ajuster les paramètres d'entrée, ajuster notre stratégie de prévision, et à la fin aboutir à un résultat, celui qui est attendu<sup>121</sup> ». En creusant les discours, on comprend alors que l'intervention humaine est centrale et se fait tout au long du processus de traitement des données – le consultant en technologie décide de l'algorithme, l'analyse, l'ajuste, confronte les résultats qu'il observe. Aussi, ces cabinets promettent une lecture par les algorithmes qui s'appliquerait à tous les domaines, ce qui entraîne à considérer la donnée comme une juste description de la réalité. Or les modèles de calculs ne s'appliquent pas à tous les sujets, tels que la lecture des émotions. Alors que Novagen évoque cette capacité de la machine à détecter les émotions<sup>122</sup>, Charlotte Zaepfel souligne au contraire que cette compétence est propre à l'humain, car « il y a des tâches où l'homme ou la machine sont l'un et l'autre meilleur. L'homme pour détecter les émotions, les intuitions de langage et de grammaire ».

Enfin, ces cabinets construisent un simulacre de prédiction lorsqu'ils promettent de prédire avec justesse un chiffre d'affaires par exemple<sup>123</sup>. Il convient de rappeler la marge d'erreur, et que la prédiction algorithmique n'est rien sans la capacité de jugement, l'algorithme ne fait que mettre en œuvre une série de calculs, à grande échelle. Agnès Gilbert le rappelle : « La machine ne peut pas aller sur le chemin de l'arbitrage, qui relève du domaine de l'expert. L'algorithme ne fait que répliquer ce que l'humain lui dit de faire. Même s'il y a une forme d'intelligence artificielle, avec un apprentissage de la machine, elle n'aura jamais de libre-arbitre ». Elle rappelle le rôle central de

---

<sup>119</sup> Cf. Annexe 4. Entretien avec Agnès Gilbert, « Chief Digital Officer » à Ipsos

<sup>120</sup> Cf. Entretien avec Novagen : « une fois qu'on a fait tout ce travail qui est vraiment conséquent et qui nous mobilise beaucoup, là on peut le pousser dans un algorithme. Là aussi l'intervention humaine, c'est de bien choisir l'algorithme, selon qu'on veut reconnaître des objets dans une image, ou faire de la traduction, ou analyser un sentiment dans un texte... on ne va pas être sur les mêmes algorithmes. On ne va précisément pas être sur la même chaîne de traitement de l'information »

<sup>121</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>122</sup> Ibid.

<sup>123</sup> Ibid., « avec cet algorithme, vous pouvez faire de la prédiction sur les années 11, 12, 13, 14, 15 qui arrivent demain ». Et « Quand on tombe sur une combinaison d'algorithmes qui donnent des informations exactes sur le passé, on peut se projeter sur l'avenir. Donc à ce moment là, on va donner une vraie information à la direction des ventes et à la direction générale d'une entreprise »

l'expert, du geste humain qui prime sur celui de la machine, qui décide et dirige celle-ci, et la machine répliquant la subjectivité du consultant. Ce brouillage des postures énonciatives témoigne que les technologies que ces consultants vendent sont en partie construites dans les discours, qui ont tendance à occulter la part de la subjectivité de l'expert au profit d'une objectivité de l'algorithme. Leurs discours d'escorte prétendent de fournir à partir de bases de données multiples un résultat présenté comme scientifique et comme vérité objective, légitimée par le biais d'opérations techniques (c'est pour cela qu'ils donnent à voir des algorithmes qui tournent, des processus d'automatisation, qu'ils emploient des termes évocateurs comme « machine-learning »). Tout cela participe à venir légitimer la posture du consultant-ingénieur, qui se trouve en position de se prévaloir d'une vérité.

On a ainsi analysé comment les cabinets de conseil en technologie produisent des discours d'escorte qui font autorité dans notre système de représentations, qu'ils font partie d'un processus de légitimation de leur offre de conseil. Leurs discours se construisent selon plusieurs logiques de légitimation : puiser dans les mythes opérants dans nos imaginaires sociaux, liés à une rationalité du nombre et du calcul, à la place de la science dans l'histoire, à notre modèle de confiance en la base de données. Mais aussi entretenir le régime de croyance qui ne met pas à l'épreuve ces technologies, qui croit en une « infailibilité des big data, capables de créer des « nouveaux gisements de valeur », qui « cristallisant tous les espoirs d'une nouvelle forme de « puissance » pour la société<sup>124</sup> ».

Ces consultants en « data » vont aussi mettre en scène un ethos de l'expert rationnel et le « scénographe ». L'enjeu est d'être reconnu comme instance énonciative experte, comme des experts en « data-science », dont la parole est peu remise en question. Il s'agit de faire autorité dans leur domaine en construisant un « ethos algorithmique » qui joue sur des imaginaires sociaux et culturels forts (l'imaginaire du scientifique qui fait des découvertes dans son laboratoire, qui fabrique l'innovation). Ils vont ainsi jouer sur cette légitimité de l'univers scientifique et mathématique, pour se prévaloir d'une vérité, d'une objectivité à partir de bases de données et de traitements d'algorithmes. Leurs discours d'escorte font croire à un nouveau savoir produit par les données, légitime, or nous avons démontré que ces données sont précisément liées à une subjectivité, à un libre-arbitre humain, aux décisions que prend le l'ingénieur-consultant.

---

<sup>124</sup> Cf. BERTIN Erik, « Data, l'onction du nombre », *Influencia*, 16 octobre 2016 (URL : [www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html](http://www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html))

## II. LE RENFORCEMENT DE LA LÉGITIMITÉ DU CONSULTANT-INGÉNIEUR PAR CES TECHNOLOGIES PERFORMATIVES QUI PRÉTENDENT S'ÉRIGER EN CONNAISSANCE ET EN AIDE À LA DÉCISION

C'est en parvenant à construire un ethos « algorithmique » que ces consultants en technologie peuvent faire autorité auprès des entreprises qu'ils prospectent, se prévaloir d'une vérité (économique ou organisationnelle par exemple). Ils peuvent ainsi exercer leur rationalité algorithmique et opérer des médiations au sein des entreprises avec lesquelles ils travaillent. Leur ethos d'expert se construit non seulement *via* ces outils et des pratiques algorithmiques qu'ils utilisent comme pour montrer les signes d'une technicité, d'une méthode, d'une scientificité, mais se renforce aussi par le fait qu'ils produisent une « intelligence<sup>125</sup> » à partir de ces mêmes outils algorithmiques. Nous allons ainsi analyser les médiations algorithmiques qui sont à l'œuvre lorsque ces consultants en « data-science » proposent de traiter les « grosses données » des entreprises. En effet *via* ces technologies qu'ils présentent comme outils de compréhension, d'appréhension et d'intelligibilité du réel, ils donnent à voir une traduction du monde particulière. Nous tenterons d'examiner comment ces experts imposent leurs savoir-faire et savoir-lire aux entreprises. Nous analyserons aussi ces technologies de traitement des « grosses données » comme des dispositifs performatifs qui font exister la réalité qu'ils représentent, en examinant comment ils semblent asseoir leur légitimité *via* ce pouvoir performatif, dans la mesure où ils donnent à voir des réalités complexes et abstraites. Ils apparaissent ainsi comme les seuls médiateurs d'un monde difficilement intelligible et accessible autrement qu'en passant par ces outils de traitement des données (ou « traces numériques » pour reprendre la notion telle qu'évoquée dans *Traces numériques. De la production à l'interprétation*<sup>126</sup>). En cela il apparaît qu'ils détiennent un pouvoir performatif, celui de rendre présent sous nos yeux – et ici, sous les yeux des décideurs stratégiques – une réalité qui n'était jusque-là pas portée à leur connaissance. Un autre enjeu pour ces consultants experts en technologie semble d'être capable de traduire cette « intelligence » des données pour être compris par les décideurs auxquels ils s'adressent. S'adresser aux acteurs économiques en passant d'un langage ingénieur à un langage plus gestionnaire constitue précisément une traduction, résultant d'une activité interprétative et cognitive particulière (cette activité d'interprétation d'un corpus de

---

<sup>125</sup> Novagen revendique la capacité de son Data Lab à « Extraire de l'Intelligence de vos bases de données internes, de votre GED, Historique Documentaire » (cf. Annexe 1. b. Site web Novagen – Data Factory)

<sup>126</sup> GALINON-MÉLÉNEC Béatrice, ZLITNI Sami, *Traces numériques. De la production à l'interprétation*, CNRS Éditions, 2013, 290 p.

« grosses données » étant le fruit de procédés de médiations qui consistent à mettre en nombre, quantifier, simplifier et synthétiser) et qu'ils retranscrivent par le moyen de supports de représentation, tels que les tableaux de bord qui offrent une vue synthétique du réel, parle à la rationalité gestionnaire des décideurs qui ont une logique de lecture stratégique (donc synthétique et transversale). Par cela même ils effectuent des médiations algorithmiques, passant d'une rationalité ingénieure à une rationalité gestionnaire, et dans lesquelles on peut lire l'expression d'un certain pouvoir, celui de produire de la connaissance pour les acteurs économiques et les aider à l'action.

## 1. Les traductions à l'œuvre pour rendre les données intelligibles : du langage ingénieur au langage gestionnaire

Il semble que l'enjeu pour les cabinets de conseil en technologie soit de rendre les données intelligibles auprès des entreprises, pour qu'elles puissent faire valeur de connaissance, être reconnues comme un savoir à part entière et puissent servir d'aide à la décision. Lorsque ces consultants experts en traitement des données proposent de traiter les « grosses données » de leurs entreprises clientes, plusieurs traductions semblent s'opérer, c'est-à-dire qu'ils proposent de traduire les « données brutes » des entreprises en une information intéressante pour l'entreprise, qui peut porter sur les tendances du marché, les comportements des consommateurs, la réputation d'un produit ou d'une marque sur le web. Ils proposent de « scripturer<sup>127</sup> » les données de l'entreprise d'une certaine manière, de les organiser et de donner à voir l'information qui en résulte, et qui s'inscrit dans le cadre interprétatif de leur rationalité algorithmique. Ils proposent ainsi aux entreprises des modèles de pensée et des schémas interprétatifs particuliers, souvent rationnels, impliquant une vision du monde quantitative, simplifiée, réduite dans des classeurs de données numériques, des tableurs, des graphiques, des tableaux de bord, qui fournissent des « cadres de classement pertinents qui améliorent la perception que l'homme a du monde<sup>128</sup> ». Ainsi les consultants technologues améliorent la perception qu'ont les décideurs stratégiques de l'environnement économique et social de leur entreprise. Nous voyons que s'opèrent des médiations par la traduction des données brutes en informations exploitables et opérationnelles

---

<sup>127</sup> Nous reprenons le terme de « scribe » employé par Emmanuël Souchier à propos des ingénieurs-cartographes, qu'il compare à des « scribes » dans *Traces numériques, de la production à l'interprétation*. Il revient sur la naissance de l'écriture idéographique qui est le « résultat d'une activité consciente, rationnelle, pratique et réflexive, qui a notamment consisté à organiser l'univers dans lequel l'homme vivait ». Il souligne que les scribes ont élaboré « un processus de « rationalisation » des observations initiales et l'élaboration d'une véritable « science » (Bottéro, 1974 : 168 sq.). Les scribes ont procédé à une comparaison des données et, parallèlement, élaboré des traités théoriques rendant compte de cette activité spéculative » (p. 220)

<sup>128</sup> Ibid.

pour les décideurs économiques. Il s'opère aussi des procédés de traduction du fait d'une traduction d'un langage ingénieur utilisant sa rationalité ingénieure et des termes techniques, à un langage gestionnaire aux termes davantage ancrés dans une réalité économique. Ce procédé de traduction d'un langage à un autre semble essentiel pour rendre compréhensible l'interprétation des données auprès des décideurs d'entreprise, afin qu'ils puissent s'approprier l'interprétation qui est faite de leurs données, et concevoir un plan d'actions stratégique à partir des résultats donnés à voir par ces dispositifs technologiques.

a. Performativité et effet d' « irréversibilité » des outils de représentation des « grosses données » : rendre indiscutable la connaissance issue des « technologies big data »

Nous pouvons analyser les « technologies big data » comme des dispositifs performatifs qui font exister des réalités abstraites complexes en représentant celles-ci, au sens où, en les rendant présentes sous nos yeux et en s'appuyant sur des outils de représentation graphiques – qui peuvent être un tableur numérique de bases de données, un diagramme, un histogramme, ou encore un tableau de bord. En cela on peut dire qu'ils ont un « effet d'irréversibilité<sup>129</sup> », c'est-à-dire qu'une fois qu'ils donnent à voir la réalité sous une certaine forme, ils organisent l'information selon un certain modèle, fruit de plusieurs procédés de traduction, de modification, d'ajustement. Cette mise en forme graphique des « grosses données » que le cabinet de consulting donne à voir, s'ancre dans l'esprit de celui qui voit (ici le décideur de l'entreprise). Une fois que l'information est représentée, cela génère une représentation mentale qui rend difficile de se représenter cette information autrement.

Les consultants-ingénieurs en traitement des données promettent de rendre tangible l'insaisissable : par exemple pour Affini-Tech, comprendre le lectorat par la « collecte de données d'audience », « Identifier des clusters d'intentionnistes » ou « Construire une vue 360° cross-canal » avec une « donnée segmentée et activable en quelques minutes après la collecte<sup>130</sup> ». La notion de « technologies intellectuelles » de Jack Goody et Pascal Robert résolvent ce problème de ne pas saisir une réalité qui échappe aux acteurs économiques, puisqu'elles font exister des réalités complexes, les représentent ; elles sont en cela des dispositifs performatifs. Par un tableau de bord, un histogramme ou un graphique, ces consultants en technologie promettent de faire émerger des choses qu'on ne pourrait rencontrer en dehors de sa représentation graphique, à l'image des « clusters d'intentionnistes » chez Affini-Tech, la « profondeur d'une communauté web », ou la

---

<sup>129</sup> Notion de Pascal Robert, qui a été évoquée par Thomas Grignon lors de son cours de « Cartographie du web », en Master 2 Médias et Numérique

<sup>130</sup> Cf. Annexe 1. Corpus d'étude. g. Site web Affini-Tech – Projets

« propagation d'une thématique sur le web<sup>131</sup> » chez Novagen, qui ne sont pas des objets concrets que l'on rencontre dans notre quotidien. Les clusters d'intentionnistes ou la communauté web n'existent pas en soi, mais seulement dans le graphique ou le graphe qui les représentent. Grâce à ces dispositifs technologiques qui les représentent, ils acquièrent une réalité concrète et tangible. Pascal Robert parle même d'un « effet d'irréversibilité » : plus on voit ces objets, plus on a du mal à les voir autrement.

Bruno Latour, dans *Les vues de l'esprit*, ajoute que les supports de visualisation permettent de faire advenir le consensus autour d'une vérité scientifique, de la rendre indiscutable. La représentation graphique permet d'asseoir une affirmation scientifique – ici la représentation graphique des « big data » permet d'asseoir les affirmations des consultants-experts en technologie, de rendre indiscutables les résultats qu'ils présentent à leurs entreprises clientes. En cela, ce sont des outils qui renforcent leur position d'expert, leur confère une posture de détenteurs d'un savoir que leurs interlocuteurs n'ont pas. Ce sont des outils qui semblent renforcer la posture de l'expert, car ils appellent une parole du consultant qui explicite les résultats donnés à voir, n'étant pas directement intelligibles par les décideurs. Ils permettent ainsi à ces consultants experts d'inclure un rapport de pouvoir, en faisant sentir cet écart entre l'expert (le consultant en technologie) et le profane (le décideur). Cela leur donne un pouvoir rhétorique qui re-légitime le consultant en technologie, par l'utilisation de ce langage technique (« machine-learning », « algorithmes », « traitement batch »...) qui le rend crédible auprès des acteurs de l'entreprise. C'est dans ce moment de parole qui commente les « technologies big data » que se met en spectacle l'expertise.

Afin d'être en cohérence avec l'« ethos algorithmique » qu'ils promeuvent auprès des entités dirigeantes, ces consultants en technologie donnent ainsi à voir des supports de représentation des données propres à leur « rationalité algorithmique » : classeurs numériques de données, tableurs, tableaux de bord, diagrammes, graphiques, histogrammes. Nous allons analysons comment ces supports conditionnent l'information de manière quantitative et synthétique.

#### b. « Mise en nombre » et quantification du réel par les algorithmes et techniques de calcul

En donnant à voir les résultats issus des « technologies big data » qu'ils ont appliqués aux entreprises, les consultants-ingénieurs imposent un certain savoir-lire fondé sur une vision quantifiée du monde (la « raison algorithmique » évoquée par Eric Sadin<sup>132</sup>), façonnant l'information qu'ils représentent selon des techniques de calcul. En examinant les sites Internet des

---

<sup>131</sup> Voir leur solution de Social media intelligence (cf. Annexe 8. Support de présentation de Novagen - slide p. 12)

<sup>132</sup> Cf. SADIN Eric, *La vie algorithmique, critique de la raison numérique*, Paris : Éditions l'Échappée. DL 2015, 278 p. (quatrième de couverture)

cabinets de conseil en technologie, Novagen et Affini-Tech, nous remarquons la prétention de leurs outils de traitement des données à calculer le réel, à le soumettre à des algorithmes (les « algorithmes de classification » ou « prédictifs » évoqués sur le site d’Affini-Tech<sup>133</sup>) ou à des modèles de calculs (« modèles de machine-learning », « modélisation des équipements industriels pour de la simulation<sup>134</sup> ») et à appréhender les événements en temps réel (« traitement batch et temps réel<sup>135</sup> » d’Affini-Tech, qui traite les données « par lots » de manière automatique sans intervention de l’opérateur). Dans « L’onction du nombre<sup>136</sup> », Erik Bertin met en lumière la prétention des « technologies big data » à tout calculer et tout englober dans leur modèle algorithmique : « Cette possibilité de tout calculer éclaire l’intention sous-jacente du système de la « data », qui est celle de l’optimisation des conduites, des comportements. Toute conduite, toute action est optimisable, du point de vue du calcul et de la modélisation. L’algorithme est la clé d’accès à un savoir cumulatif sur soi et sur la société, dans un but d’amélioration ». Quand Affini-Tech mentionne la « modélisation des équipements industriels pour de la simulation » sur son site, l’expression sous-tend cette prétention soulignée par l’article *Influencia*, mais aussi par Eric Sadin et Dominique Cardon. Ils imposent ainsi leur « raison algorithmique » à l’univers de l’entreprise, dont les événements, les métiers et les procédures internes se trouvent modélisés, calculés en vue d’être améliorés et optimisés. Dominique Cardon, dans *À quoi rêvent les algorithmes : Nos vies à l’heure des big data*, dénonce ces mêmes techniques de calcul qui façonnent nos sociétés, jusqu’à parler de « société du calcul » qui caractériserait notre société contemporaine : « Classement, personnalisation publicitaire, recommandation de produits, mesures corporelles, orientation et déplacements... Les « calculateurs » s’immiscent partout<sup>137</sup> ». Eric Sadin dénonce de la même manière ces « technologies informationnelles », qui « imposent un mode de rationalité fondé sur la définition chiffrée de toute situation et sur une maîtrise indéfiniment accrue du cours des choses<sup>138</sup> ». À la lecture de ces concepts, il semble que les consultants-ingénieurs en « data » exercent leur « raison algorithmique » selon ces logiques mêmes.

Pour aller plus loin dans l’analyse, les travaux d’Antoinette Rouvroy mettent en lumière ce phénomène de « gouvernementalité algorithmique » qui semble caractériser notre époque contemporaine : « Ce phénomène de traduction du monde physique et de ses habitants en données métabolisables par les systèmes informatiques n’est désormais plus limité, ni même freiné » à cause de raisons techniques (capacités de stockage notamment). Il existerait une tentation des acteurs

---

<sup>133</sup> Cf. Annexe 1. g. Site web Affini-Tech – Projets

<sup>134</sup> Ibid.

<sup>135</sup> Ibid.

<sup>136</sup> BERTIN Erik, « Data, l’onction du nombre », *Influencia*, 16 octobre 2016 (URL : [www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html](http://www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html))

<sup>137</sup> CARDON Dominique, *À quoi rêvent les algorithmes ? : Nos vies à l’heure des big data*, Paris, Seuil, La République des idées, 2015, 105 p. (Quatrième de couverture)

<sup>138</sup> SADIN Eric, *La vie algorithmique, critique de la raison numérique* (Quatrième de couverture)

économiques et sociaux à traduire le monde en « données métabolisables » (autrement dit qu'on peut transformer en informations opérationnelles, activer en actions concrètes), dans un contexte de « banalisation d'une surveillance démocratique et égalitaire qui ne prétend plus cibler personne a priori, mais s'applique à tout le monde par défaut, ont tôt fait d'éroder les réticences et résistances<sup>139</sup> ». Elle poursuit en révélant que « l'usage de plus en plus systématique du *data mining* et du profilage répond à un glissement stratégique ou tactique dans la gestion de l'incertitude, que ces nouveaux dispositifs sont symptomatiques d'une nouvelle forme de gouvernementalité<sup>140</sup> ». Le fait de mettre en nombre, de quantifier le monde à travers les « technologies big data » et de les traduire en informations pour les entités dirigeantes traduit une volonté de gérer l'incertitude, phénomène qui caractériserait la « gouvernementalité algorithmique » de notre société. Cette tendance à opérer des traductions du monde en chiffres comme manière de gérer l'incertitude, s'accompagne aussi d'un procédé de traduction propre à la « raison algorithmique » des consultants-ingénieurs : la simplification et la synthèse du réel.

### c. Synthétiser et simplifier le réel : la capacité de l'ingénieur à vulgariser l'objet technique

Il est intéressant d'analyser les outils de traitement des « big data » des consultants ingénieurs comme des outils de simplification. Un tableau de bord ou un graphique ont en effet pour objectif d'offrir une vue synthétique du réel et d'être facilement identifiables par leur interlocuteur. Ils permettent d'organiser le monde dans un certain cadre, qui doit être intelligible par le plus grand nombre, qui doit lui permettre de s'y retrouver. Ainsi le schéma, la liste, le diagramme, le tableau de bord sont des « instruments de simplification efficaces dans le social<sup>141</sup> ». À ce titre, Affini-Tech promet d'offrir cette vue synthétique afin d'être en conformité avec la rationalité gestionnaire des entités dirigeantes, qui visent à collecter des informations stratégiques d'ensemble : « Disposer d'une vue globale sur chaque client et prospect », « Construire une vue 360° cross-canal : Rapprocher les données de magasin et de site web pour identifier des parcours clients<sup>142</sup> ». Ainsi les tableaux de bord et diagrammes produits par les consultants en « data » donnent à voir le web et les données internes des entreprises sous une forme synthétique, afin de garantir l'efficacité de leur communication avec leurs interlocuteurs économiques.

Les consultants-ingénieurs font donc des opérations de traduction de l'environnement des entreprises, traduisant leurs données « brutes », « massives » ou « internes » pour les rendre

---

<sup>139</sup> ROUVROY Antoinette, BERNS Thomas, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques »... », *Multitudes*, 2010, n° 40 (pp. 88-103) (p. 90)

<sup>140</sup> Ibid., p. 92

<sup>141</sup> Cf. Cours dispensés par Thomas Grignon en Master 2

<sup>142</sup> Cf. Annexe 1. g. Site web Affini-Tech – Projets

intelligibles et les transformer en connaissance activable, à l'image d'Affini-Tech qui promet d'éclairer les acteurs de l'entreprise « sur les leviers business, et l'intégration des plateformes nécessaires », et leur « permettre de les concrétiser rapidement<sup>143</sup> ». Ils vont ainsi proposer de visualiser ces données de manière simplifiée, sous une forme synthétique qui aide leurs clients à appréhender les informations que leurs données brutes contiennent, à comprendre leur sens de lecture de celles-ci, de manière rapide, et de sorte à ce qu'ils puissent les appliquer dans leurs décisions stratégiques. L'un des enjeux stratégiques pour ces consultants en technologie est en effet d'élaborer des outils d'aide à la décision qui s'adressent à des entités dirigeantes ou des équipes « métiers », pour les aider à repérer les tendances, apprécier un état au fil du temps, ou encore à prédire le chiffre d'affaires. Ils doivent adopter un langage gestionnaire afin d'être compris par ces interlocuteurs, et être capables de « vulgariser » leur culture technique en des résultats opérationnelles pour les entreprises. Comme l'affirment les fondateurs de Novagen dans notre entretien, « la finalité va être de rendre un service opérationnel. Cela passera vraisemblablement sur la capacité à faire des tableaux de bord, ou en tout cas des outils qui, connectés à la donnée, leur apportent une vraie valeur ajoutée. C'est dans ce cas là que souvent, la fin de notre présentation se termine par une démonstration live des tableaux de bord, de notre capacité à matérialiser tout ce qui est acquisition et valorisation des données. On montre que pour arriver à ce produit qui réponde à leurs attentes stratégiques, il y a une démarche d'accompagnement et de formation qu'il faut bien avoir à l'esprit<sup>144</sup> ». On voit bien ce procédé de traduction à l'œuvre, qui opère des va-et-vient entre le langage ingénieur et le langage gestionnaire, et l'importance de maîtriser ces deux langages pour les consultants, afin de permettre aux décideurs de comprendre cette « valeur ajoutée » qui est portée à leur connaissance *via* les « technologies big data ». Comme l'affirme Hubert, le « Chief Innovation Officer » de Novagen qui « à un moment les mains dans le code et quatre minutes après je discute avec une Direction stratégique. Il y a une vraie place pour les acteurs techniques au profil ingénieur qui savent vulgariser les enjeux du moment<sup>145</sup> », maîtriser ces deux langages, à la fois technique et gestionnaire, leur permet d'être compris et fait partie du processus qui les reconnaît comme experts. La maîtrise du langage gestionnaire permet aux consultants d'être compris et assure que les résultats qu'ils donnent à voir soient reconnus comme connaissance, fassent valeur de savoir auprès des décideurs d'entreprise - que ce soit par des classeurs numériques de données, tableaux de bord, graphiques... le schéma, la liste, le diagramme, l'infographie, la carte – ces « instruments de simplification efficaces dans le social » qui offrent une vue synthétique du réel et leur permet d'être immédiatement compréhensibles. Emmanuël Souchier évoque ces « technologies abrégatives qui, à l'instar de la page ou du tableau, ont pour

---

<sup>143</sup> Cf. Annexe 1. e. Site web Affini-Tech – Page d'accueil

<sup>144</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>145</sup> Ibid.

vertu d'exprimer le monde dans un espace préhensible, à dimension humaine (Souchier et Zali, 2002) »<sup>146</sup>.

## 2. Le rapport de pouvoir de ces consultants en technologie face aux décideurs d'entreprise : érigés au rang de devins ou de savants ?

Du fait de cette connaissance rendue presque indiscutable, traduite du langage ingénieur au langage gestionnaire, et reconnue comme savoir légitime par les entités dirigeantes économiques, on peut dire que ces consultants-ingénieurs entretiennent un rapport de pouvoir particulier face à ce savoir qu'ils font advenir *via* les « technologies big data ». Les entreprises leur attribuent une valeur car ils répondent à leur besoin de connaissance sur le marché, sur les comportements des consommateurs, sur un diagnostic de l'existant. Leur « raison algorithmique » semble conférer un pouvoir à ces acteurs de l'industrie du traitement de la donnée. Ils permettraient de faire advenir un savoir inédit, inépuisable, toujours renouvelé – Eric Sadin évoque une « anthropologie de l'exponentiel » dans le dernier chapitre de son ouvrage. Cela fait écho à la comparaison des « big data » à un « nouveau pétrole » évoqué dans l'introduction – Affini-Tech propose d'ailleurs de « trouver de nouveaux gisements de valeur » par l'exploration des données. Lorsque Eric Sadin souligne que « dorénavant, les flux de *data* témoignent de la quasi-intégralité des phénomènes, s'érigeant comme l'instance primordiale de l'intelligibilité du réel<sup>147</sup> », peut-on pour autant dire que les « technologies big data » s'imposent comme instance principale d'intelligibilité du réel ? Nous tenterons d'analyser les rapports de pouvoir que les consultants-ingénieurs entretiennent avec les entités dirigeantes et les acteurs économiques, du fait de l'effet de fascination à l'égard des technologies, de la figure auctoriale de l'expert en technologie et de la maîtrise des technologies. La posture du consultant-ingénieur pourrait ainsi se comparer à celle du devin de l'Antiquité ou d'un savant qui prodiguerait un savoir à partir de « signes » que lui seul est à même d'interpréter, en soulignant aussi l'impensé de la médiation qu'ils opèrent.

### a. Le pouvoir du consultant ingénieur qui prétend interpréter le monde, analogue à l'exercice de l'art divinatoire

---

<sup>146</sup> GALINON-MENELEC Béatrice, *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013, 290 p. (p. 221)

<sup>147</sup> SADIN Eric, *La vie algorithmique, critique de la raison numérique* (Quatrième de couverture)

Ces consultants en technologie, se positionnant comme médiateurs qui permettent de comprendre les « données brutes » et massives, prétendent produire un savoir du fait de leur maîtrise des « technologies big data ». Ils exerceraient une activité de lecture des « grosses données » à la manière des devins de l'Antiquité qui interprètent à partir d'une base (le vol des oiseaux, les entrailles des animaux – ici, les « données brutes ») et fabriquent leur interprétation à partir des signes qu'ils observent, à la manière de l'art divinatoire chinois. Emmanuël Souchier, en analysant la posture des ingénieurs-cartographes dans *Traces numériques, de la production à l'interprétation*, analyse cette prétention à faire valeur de savoir et à faire advenir celui-ci à la manière des devins de l'Antiquité<sup>148</sup> : « Notre « cartographe » est dans une posture analogue à celle du devin chinois qui crée les signes sur les carapaces de tortue, signes qu'il va ensuite trier, comparer, hiérarchiser, interpréter. Tous deux cherchent à donner du sens au monde qui nous entoure<sup>149</sup> ». Il poursuit : « Tous deux élaborent à leur façon une science et une écriture, une technologie intellectuelle qui a pour objectif de comprendre le monde – au deux sens du terme : le cerner autant que le rendre intelligible<sup>150</sup> ».

Emmanuël Souchier, dans « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », met en lumière cette pratique divinatoire comme un « processus intellectuel » qui ne relève pas de la voyance au sens habituel du terme : « Il ne s'agit pas d'une divination inspirée ou d'une divination prédictive, mais d'une pratique déductive reposant sur un examen précis, minutieux, de faits et de signes observés sur les objets analysés – qu'il s'agisse de carapaces de tortue pour la Chine ou de foies d'animaux pour Sumer. [...] L'activité consiste donc à observer, comparer, interroger et non pas à deviner au sens où nous l'entendons communément<sup>151</sup> ». Ainsi les consultants-ingénieurs en technologie exerceraient la divination comme pratique déductive qui observe, compare et analyse les données, relevant donc d'une activité technique particulière, un savoir-faire propre aux consultants-ingénieurs. Il compare cette activité d'interprétation des signes à la divination « artificielle » (et non la divination « naturelle » qui se fait à partir des oracles divins), qui est « exercée par les haruspices, les devins, les observateurs de la foudre et les astrologues<sup>152</sup> ». Nous comprenons ainsi l'activité de ces consultants en technologie comme une activité divinatoire, et en la comparant à l'activité des ingénieurs-cartographes comme analysée dans l'article<sup>153</sup>, nous constatons que ces consultants en technologie produisent, à partir de « traces<sup>154</sup> » leurs propres objets visuels (tableaux de bord, histogrammes) en

---

<sup>148</sup> SOUCHIER Emmanuël, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture, *Traces numériques, de la production à l'interprétation*, pp. 213 à 229 (p. 226)

<sup>149</sup> Op. cit., p. 229

<sup>150</sup> Op. cit., p. 230

<sup>151</sup> Op. cit. p. 216

<sup>152</sup> Op. cit., p. 217

<sup>153</sup> Op. cit. p. 222

<sup>154</sup> Ibid.

interprétant ces traces comme autant de manifestations du « social » (cela peut être le comportement des consommateurs, les acteurs économiques, le marché, les ventes...). Remarquons aussi le phénomène de croyance pensant que c'est le « social » qui parle à travers ces traces numériques. Cette prétention à faire valeur de connaissance se retrouve chez Novagen : « dans le data lake, vous allez partir d'une matière brute et vous allez en créer une nouvelle qui va être portée à la connaissance des décideurs<sup>155</sup> ». Ils évoquent une « matière brute » qui n'a pas de sens en soi et qui attend d'être interprétée par le consultant-ingénieur, qui va lui donner un sens, faire advenir une signification à la manière des devins étrusques dans l'Antiquité romaine qui, chargés d'observer le ciel pour y lire les signes, interprétaient les auspices à travers le vol des oiseaux (l'ornithomancie). Ainsi peut-on rapprocher métaphoriquement les consultants ingénieurs en « big data » avec les « augures » romaines, dans la mesure où une augure est une « personne qui se prétend en mesure de prédire l'avenir et d'avancer des prévisions dans un domaine particulier<sup>156</sup> » ou encore un « « prêtre dont la charge était d'observer le vol des oiseaux » et de faire des « prédictions » d'après certains signes.

Si l'on considère la divination comme une forme de connaissance « provenant de facultés puissantes d'observation et de déduction<sup>157</sup> », on voit qu'elle est étroitement liée à des compétences particulières, à une expertise – celle de savoir lire les signes. Ici, on peut comparer les consultants ingénieurs à des devins, dans la mesure où ils exercent une compétence spécifique et reconnue comme rare, celle de savoir lire et interpréter les signes que donnent à voir les « technologies big data », à extraire une « intelligence<sup>158</sup> » à partir de données brutes. Ainsi les consultants ingénieurs peuvent se rapprocher de l'analyse qu'en fait l'article dans *Traces numériques*, à propos des « ingénieurs-cartographes » : « Les bârû entretenaient vis-à-vis du savoir un rapport de pouvoir comparable à celui de nos « ingénieurs-cartographes » quand ils exercent leur compétences pour interpréter la cartographie du web. Franck Ghitalla souligne que la pratique « confère ainsi aux cartographes de l'atelier une forme d'expertise stratégique inédite, qu'il s'agisse d'observer le réseau (activités de veille), ou de le transformer », constatant qu'un certain pouvoir d'action se révèle » à travers cette pratique de cartographie du web.

Où réside le pouvoir qu'on attribue à ces consultants en technologie qui font advenir un savoir *via* la technologie ? Comment leur savoir fait-il valeur et est-il reconnu comme légitime ? Il apparaît que la maîtrise de la technologie leur confère une posture de savant, on leur reconnaît le pouvoir de révéler un savoir sur le monde, comme on reconnaissait aux devins antiques, les « bârû sumériens », qui exerçaient l'art divinatoire chinois. Emmanuël Souchier dans *Traces numériques, de la*

---

<sup>155</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>156</sup> Définition sur le site du CNRTL ([www.cnrtl.fr/definition/Augure](http://www.cnrtl.fr/definition/Augure))

<sup>157</sup> Selon le CNRTL : [www.cnrtl.fr/definition/divination](http://www.cnrtl.fr/definition/divination)

<sup>158</sup> Cette « intelligence » est mise en avant par Novagen à propos de son « Data Lab » (cf. Annexe 1. b. Site web Novagen – Data Factory)

*production à l'interprétation*, fait remarquer que « les technologues agissent comme les devins de l'Antiquité, ils créent un système de cohérence entre des signes ». Producteurs d'un sens entre les signes (ici les « traces numériques, les « données brutes », les différents résultats issus des algorithmes génératifs), ils exercent une médiation, c'est-à-dire qu'ils partent d'une base (le dos de la tortue pour les barû sumériens, et les « données brutes » pour les consultants-ingénieurs) et à partir de celle-ci ils fabriquent leur interprétation. Il convient de souligner un phénomène propre à l'activité interprétative de ces acteurs qui prétendent interpréter les signes du monde, et qui soutient un certain pouvoir du consultant-ingénieur face aux entités dirigeantes, qui est « l'impensé de la médiation<sup>159</sup> » : « Mais la cohérence de cette activité n'est possible qu'à partir du point aveugle qui la fonde. Comment en effet cette pratique d'interprétation peut-elle être validée dès lors que les devins sont eux-mêmes les producteurs des signes qu'ils interprètent ? Le phénomène n'est viable que s'il repose sur un socle de croyances permettant l'adhésion du collectif et la légitimité des acteurs. Cette pratique qui fait du scribe le lecteur et l'interprète des signes divins trouve sa cohérence dans l'impensé radical de la médiation qui la fonde. » Par « l'impensé de la médiation », Emmanuël Souchier souligne l'effacement des cadres de l'énonciation par les ingénieurs-cartographes qui font oublier leur activité d'interprétation des données, le fait qu'ils soient à la fois producteurs et interprètes des signes qu'ils analysent. Pour autant, ils font passer leur parole pour neutre, objective, comme si leur parole résultait d'une simple transmission du sens que leur ont donné les technologies de traitement de données. Ils se positionnent comme « porte-parole » de la parole des dieux ou de l'univers, et se dédient de toute responsabilité de production du savoir – « puisque je ne produis pas le sens, mais le subis, me contentant de le « transmettre<sup>160</sup> ». Aussi, « le médiateur ne pense pas plus son action que sa présence, ce sont les signes ou les dieux qui parlent à travers lui. Mais cet aveuglement n'est pas uniquement d'ordre idéologique, il est avant tout opératoire et relève de la « mémoire de l'oubli », qui est une caractéristique nécessaire au fonctionnement de nos pratiques de communication (Souchier, 2012a)<sup>161</sup> ». Cette « mémoire de l'oubli » quand il s'agit de penser l'activité de construction de sens des consultants-ingénieurs fait ainsi croire à une sorte de naturalité des données valorisés par ces acteurs, comme si les informations qu'ils produisent pour les acteurs économiques allaient de soi, et étaient autant de manifestations d'une « rationalité algorithmique » neutre issue des technologies.

---

<sup>159</sup> SOUCHIER Emmanuël Souchier, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture, *Traces numériques, de la production à l'interprétation* (pp. 213 à 229) (p. 222)

<sup>160</sup> Ibid., p. 222. Emmanuël Souchier précise : « Si je ne suis pas l'énonciateur, si les cadres d'énonciation me sont « naturels », alors la parole que je profère est une parole des dieux ou de l'univers. Je ne suis que leur « porte-parole », dédié de toute responsabilité – puisque je ne produis pas le sens, mais le subis, me contentant de le « transmettre ». Je maîtrise néanmoins les cadres de l'énonciation, car c'est moi qui active les éléments matériels ou symboliques qui permettent la divination – ou la « cartographie ». J'énonce uniquement sous couvert de mon art ou de ma science »

<sup>161</sup> Ibid., p. 222 (« Exhibition de la technique, impensé de la médiation »)

On pourrait même dire qu'une posture de savant les investit, puisqu'on constate que ces consultants en technologie sont investis d'une mission de révéler le monde, d'en comprendre la complexité. Novagen affiche ainsi sur la page de son site web qu'il est possible de « Maîtriser sa complexité par l'Analyse de Données textuelles de données massives » et de « Maîtriser son Marché par le traitement Analytique des Données du WEB<sup>162</sup> ». Ils sont considérés comme une figure auctoriale savante à la manière du « barû » sumérien, de l'« ingénieur-cartographe », du devin de l'Antiquité, qui fait accoucher un certain savoir. Quand ils proposent d'« extraire de l'Intelligence de vos bases de données internes », ils proposent de faire advenir une « intelligence », une nouvelle connaissance, ce qui les dote d'une posture de savant, à la manière du barû ingénieur-cartographe qui donne un éclairage nouveau sur une problématique, une situation donnée : « Selon Jean Bottéro le barû sumérien « n'était en vérité pas autre chose qu'un savant », sa pratique de la divination déductive ayant un caractère proprement « scientifique » (p. 127, 125). En Mésopotamie, « la divination déductive, sous toutes ses formes, apparaît, par son mécanisme, une activité disons "rationnelle" et, par ses spécialistes, une technique, et une "science" qui trouve en elle-même sa raison d'être (p. 139) ». À l'instar de nos « élites » contemporaines, les barû devaient suivre « d'interminables études pour posséder à fond toute la panoplie des traités de la science et des manuels de la pratique divinatoire (Bottero, 1974 : 130)<sup>163</sup> ». Nous pouvons considérer les consultants-ingénieurs comme ces « barû », ces élites sociales qui s'auto-légitimaient par leur savoir fondé sur de la techno-science. L'article d'Emmanuel Souchier nous éclaire aussi sur le fait que les pratiques divinatoires sont une série d'opérations techniques, qu'on peut considérer comme des « techniques intellectuelles » pour comprendre le monde : « ces pratiques divinatoires débouchent sur l'élaboration d'une pensée scientifique qui repose sur des techniques intellectuelles, des techniques rationnelles de comparaison, de classification, de hiérarchisation ou d'ordonnement, lesquelles s'apparentent aux pratiques documentaires contemporaines<sup>164</sup> ». Cela met à jour l'activité interprétative des consultants-ingénieurs, qui classifient, organisent et hiérarchisent l'information qu'ils donnent à voir. Ils opèrent des choix de construction de sens que les acteurs de l'entreprise ne voient pas, en cela ils détiennent un certain pouvoir sur la « donnée ». Emmanuel Souchier précise que « ce processus intellectuel consistant, originellement, à produire et comparer les signes visuels », induit « l'élaboration d'une pensée scientifique rationnelle qui s'organise elle-même à travers la mise en place d'une ingénierie documentaire », en ajoutant que « cette geste intellectuelle s'est toujours configurée dans un rapport étroit au pouvoir en ce qu'elle prétend interpréter le monde ». Il mentionne ici la prétention des ingénieurs-cartographes (ici les consultants-ingénieurs des données) à comprendre le monde et à en fournir une version

---

<sup>162</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

<sup>163</sup> Op. cit., p. 218

<sup>164</sup> Op. cit., p. 225

intelligible, avérée, fondée sur un processus intellectuel rationnel et scientifique, qui témoignerait d'une fiabilité et qu'il serait inutile de remettre en question, puisque émanant par essence des technologies et d'une « rationalité algorithmique » qui produisent un sens « naturellement » à partir des données.

Ainsi, comme Emmanuël Souchier, on pourrait dire que « les cartographes du Web cherchent un pouvoir, le pouvoir de dire un monde qu'ils élaborent. Ce pouvoir de signification est également un pouvoir économique. Il leur donnera prise sur leurs congénères, car au travers de la « maîtrise » technique ils sauront interpréter ce que les autres ignoreront. Les cartographes du Web sont nos nouveaux devins. Ils inventent une écriture et les prérogatives qui l'accompagnent<sup>165</sup> ». Ce passage éclaire notre sujet en montrant que les cabinets de consulting en technologie disposent de la maîtrise d'un savoir et d'un savoir-faire qui les met dans une posture de pouvoir par rapport à ceux qui ignorent ces technologies et ces pratiques. Leur pouvoir de recommandation est d'autant plus accru par rapport à ceux qui ne détiennent pas ce savoir-faire et cette culture technoscientifique.

#### b. Un triple pouvoir heuristique, pragmatique et normatif ?

En formulant le postulat que ces consultants en technologie prétendent fournir une intelligence et aider à l'action à partir de « grosses données » (récoltées sur le web et dans les bases de données internes aux entreprises), nous distinguons qu'ils exercent différents types de pouvoirs. En reprenant les travaux de Thomas Grignon<sup>166</sup> sur les logiciels de surveillance de l'Internet à travers le cas de Google Analytics, on peut analyser les dispositifs technologiques qu'utilisent ces consultants comme « équipements de l'expertise » qui opèrent une redéfinition des professions, des pratiques professionnelles (dans les entreprises) et « l'institutionnalisation progressive de savoirs, savoir-lire et de savoir-faire<sup>167</sup> » à travers un dispositif de traitement des données « à la neutralité régulièrement revendiquée ». Nous pouvons qualifier ces pouvoirs de heuristique (c'est-à-dire qu'ils donnent une interprétation du réel), pragmatique (ils aident à la décision et à l'action concrète) et normatif (ils induisent une norme des pratiques professionnelles).

Concernant leur pouvoir heuristique, ces consultants en technologie<sup>168</sup> font en effet la promesse d'offrir une certaine connaissance du réel, à partir d'« énormes volumes de données<sup>169</sup> »

---

<sup>165</sup> Op. cit., p. 230

<sup>166</sup> « L'expertise communicationnelle au prisme de ses instruments. L'exemple de Google Analytics », cité dans « Le communicateur bousculé par le numérique : Quelles compétences à transmettre ? », COUTANT Alexandre, DOMENGET Jean-Claude (pp. 5-14)

<sup>167</sup> Ibid.

<sup>168</sup> Cf. Annexe 8. Support de présentation de Novagen – slide p. 7

<sup>169</sup> cf. Annexe 8. Support de présentation de Novagen – « Applications » slide p. 7

en croisant notamment les techniques et compétences « data-science » et « machine-learning ». Ils proposent même un savoir qui s'actualiserait en temps réel – en s'appuyant sur l'architecture « big data » Spark<sup>170</sup>, qui implique un certain savoir-lire, un certain cadre d'interprétation. On peut dire que ces ingénieurs-consultants induisent un sens de lecture propre au fonctionnement des interfaces. Luc Dall'Armellina précise que le « voir en relation » « se présente comme une modalité majeure du voir dans les interfaces<sup>171</sup> ». Il se produit des « opérations sémio-cognitives » quand on s'approprie une information *via* une interface, un écran (ici les tableaux de bords, diagrammes, histogrammes, tableurs de données entre autres). Il souligne que « cette modalité du voir relève d'une écologie de l'écran, dans laquelle la place, la forme, le rôle, la fonction, les relations et comportements de chaque « objets » composent ensemble un environnement, un milieu informé<sup>172</sup> ». C'est ce cadre d'interprétation qui invite à lire les « grosses données » de l'entreprise « en relation », à chercher des corrélations, à les interpréter comme un « environnement ». À la p. 150, il parle de ces médias comme « structures d'accueil complexes d'une information fluide alimentée en flux continu relevant d'un temps-réel fragmenté aux accès multi points et multi-supports hébergés dans des « nuages<sup>173</sup> ». Cette définition est éclairante pour notre sujet, puisqu'elle permet de lire les « technologies big data » comme des médias qui traitent les informations des entreprises en « flux continu », de manière fragmentée, et en temps réel (notion qu'on retrouve dans le discours de Novagen qui propose d' « analyser vos données en temps réel »). Ils sont des médias en cela qu'ils proposent une médiation de la réalité à travers leur « raison algorithmique<sup>174</sup> ». Lorsque Novagen promet une visualisation industrialisée et temps réel, il propose à ses clients de s'emparer d'un savoir-lire en temps-réel et « industrialisé », c'est-à-dire comme offrant un processus de savoir-lire automatisé, si on se réfère à Luc Dall'Armellina qui définit la « lecture industrielle » comme une activité d'automates qui scannent, crawlent, indexent, associent, proposent des contenus à un lecteur en fonction des informations que le robot a collectées sur lui ». Il précise qu'elle est accompagnée d'une « lecture humaine » car « seule, elle serait sans effets, c'est parce qu'elle est hybridée et interprétée – on pourrait dire performée – par une lecture humaine que son nom prend tout son sens<sup>175</sup> ». On a là une activité d'interprétation de l'information qui est propre à ces dispositifs technologiques, qui imposent d'interpréter à partir d'une activité de lecture par des « robots », et de venir compléter cette activité d'interprétation par

---

<sup>170</sup> Ibid., slide p. 5 « Notre approche »

<sup>171</sup> DALL'ARMELLINA Luc, « Voir en relation. Habiter l'information ? », *Études de communication*, 2013, mis en ligne le 01 décembre 2013, consulté le 14 mars 2017. URL : <http://edc.revues.org/5419> (p. 143)

<sup>172</sup> Op. cit., p. 143

<sup>173</sup> Ibid.

<sup>174</sup> Notion travaillée par Eric Sadin dans *La vie algorithmique*

<sup>175</sup> Op. cit., DALL'ARMELLINA Luc, p. 149 : « La lecture industrielle prend son origine dans une activité d'automates qui scannent, crawlent, indexent, associent, proposent des contenus à un lecteur en fonction des informations que le robot a collectées sur lui. Mais seule, elle serait sans effets, c'est parce qu'elle est hybridée et interprétée – on pourrait dire performée – par une lecture humaine que son nom prend tout son sens »

une interprétation humaine qui va donner du sens – au tableau de bord, au diagramme, à l’histogramme. Ces cabinets de conseil en technologie, en proposant des « technologies big data » qui lisent les « grosses données » des entreprises en appliquant leur « rationalité algorithmique », exercent un pouvoir heuristique, dans la mesure où elles imposent aux entreprises cette manière de lire leurs données, souvent en « temps réel », donc avec une certaine immédiateté de l’interprétation. Elles impliquent une activité de lecture « hybridée » caractérisée par une interprétation par la machine et par l’humain, sur un mode de l’instantanéité. Cette hybridation de l’interprétation des données et son interprétation immédiate se voit dans l’expression de « visualisation industrialisée et temps réel » sur le site de Novagen<sup>176</sup>.

Concernant leur pouvoir pragmatique, ces consultants en technologie proposent des outils d’aide à l’action, et sont souvent suivis d’une action ou une décision stratégique. Cela va être par exemple ajuster sa chaîne de production, s’adresser à une certaine catégorie de consommateurs, proposer un parcours client spécifique, ou encore ajuster une stratégie de réputation suite à une cartographie de communautés suite à une surveillance d’évènements sur le Web<sup>177</sup>. Autre exemple, dans son support de présentation, Novagen aide à repérer des « matrices d’opportunités<sup>178</sup> » (« Segmentation des besoins et opportunités identifiés ») et à prioriser ces opportunités d’action pour l’entreprise, préconise les « architectures big data » à adopter (« Préconisation Architecture »). Par ces actions de priorisation et de préconisation, les consultants-ingénieurs de Novagen influencent un certain cadre d’action. En témoigne aussi le slogan d’Affini-Tech sur la première page de son site, « de l’idée au service, nous accélérons votre business avec vos données !<sup>179</sup> », qui illustre sa volonté de se positionner comme un « accélérateur » du « business » de ses clients, comme celui qui va « transformer » ce business. Ses consultants veulent produire une série d’actions, avoir un impact sur les décisions des entreprises avec lesquelles ils travaillent. Ce pouvoir de recommander des actions *via* leurs dispositifs technologiques « big data » se traduit chez Affini-Tech, dans son exemple de projet « Un moteur de recommandation pour augmenter l’engagement des lecteurs<sup>180</sup> », qui implique un sens de l’action clair pour son client. Ce dernier étant un média en ligne, le dispositif de traitement des « streams de données » et l’application d’algorithmes de « machine-learning » ont eu un effet sur l’action de ce média, qui a modifié la distribution de son contenu éditorial en fonction des recommandations données par Affini-Tech<sup>181</sup>. Ainsi se manifeste

---

<sup>176</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

<sup>177</sup> Cf. Annexe 8. Support de présentation de Novagen – slide p. 7 « Surveiller des évènements sur le web (réputation, évolution des prix...) »

<sup>178</sup> Ibid., slide p. 9

<sup>179</sup> « Simplifions vos Data : De l’idée au service, nous accélérons votre business avec vos données ! Combinant Data-Science et expertise technologique, nous vous aidons à Imaginer, Expérimenter et Exécuter des services qui transforment votre business. » (cf. Annexe 1. e. Site web Affini-Tech – Page d’accueil)

<sup>180</sup> Cf. Annexe 1. e. Site web Affini-Tech - Projets

<sup>181</sup> Ibid., Donner « le meilleur contenu » éditorial à ses lecteurs

le « pouvoir pragmatique » de ces consultants-ingénieurs qui s'appuient de la sorte sur leurs « technologies big data », lesquelles fiabilisent – en apparence – et légitiment la recommandation qu'ils donnent à leurs clients. Ce pouvoir pragmatique s'illustre aussi chez le cabinet Novagen, dans son support de présentation, dans lequel il précise sa démarche d'accompagnement <sup>182</sup> : « La mise en œuvre de vos projets, en respectant un rythme itératif », ou encore : « Formation + Action : Vous faire un projet Big Data » mais aussi « Feuille de route » sont autant d'injonctions à produire de l'action, à agir dans l'entreprise, à mettre en œuvre des décisions pour celle-ci. Lors de notre entretien, les fondateurs du cabinet Novagen ont formulé ce même pouvoir pragmatique qui invite à l'action, ce positionnement de leurs technologies comme aide à l'action sur le terrain : « Imaginons, vous avez des équipes terrain, des points de vente, vous pouvez leur donner des informations, des alertes en temps réel. Là aussi, dans le *data lake*, vous aurez structuré vos données, vous aurez mis les processus de transformation et les algorithmes de valorisation, ils vont multiplier après les use cases (les cas d'application) <sup>183</sup> ». Novagen se donne pour mission d'aider les « équipes terrain » en leur donnant des « cas d'application » qu'ils peuvent appliquer de manière concrète, dans leur quotidien.

Concernant leur pouvoir normatif, comment ces acteurs de l'industrie du traitement de données, *via* leurs dispositifs technologiques de « big data », influencent-ils les pratiques des entreprises ? En s'appuyant sur la notion des « programmes » de Bruno Latour <sup>184</sup>, on peut voir que ces outils font des postulats d'usages en entreprise, c'est-à-dire que ces technologies « font faire » et ne font plus faire certaines choses. Par exemple, un algorithme « big data » qui va faire une « analyse sémantique » d'un corpus de textes sur le web, implique un « anti-programme », le fait que l'activité d'analyse des bases documentaires soit déléguée à un robot, et que l'homme n'opère plus cette activité mais ajuste sa pratique, en fonction des tâches effectués par l'algorithme. Quand Novagen propose dans un projet de « big data » en entreprise <sup>185</sup>, de « Croiser des données 'froides' ou en temps réel » ou encore « Mettre en place des analyses 'intelligentes' grâce au croisement des techniques de Data Science / ML », ce sont autant de manières d'appréhender les données de l'entreprise, les travailler, les traiter, les analyser. Ces dispositifs de « machine-learning » et de « data-science » vont forcer les acteurs de l'entreprise à adapter leurs pratiques et à travailler différemment (comme par exemple devoir croiser leurs données, ou évaluer en temps réel leurs campagnes marketing). Ce « pouvoir normatif » va aussi s'exercer lorsque Novagen accompagne les chargés d'études dans leurs pratiques professionnelles. Comme le déclare Novagen lors de notre entretien : « On peut faire part de notre savoir-faire et faire monter en compétences les

---

<sup>182</sup> Cf. Annexe 8. Support de présentation de Novagen – slide p. 11

<sup>183</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>184</sup> Notion vue en cours de Sociologie des usages et des techniques avec Jérôme Denis, citant les *Petites leçons de sociologie des sciences* de Bruno Latour

<sup>185</sup> Cf. Annexe 8. Support de présentation de Novagen – slide p. 7

équipes. Les équipes peuvent nous solliciter sur une thématique d'analyse sémantique, on fera des travaux, on les partagera avec le client, pour qu'il se les approprie ensuite<sup>186</sup> ». Ce terme d'« appropriation » par le client suppose qu'il y a une instance extérieure qui a conçu un dispositif (ici une technique d'analyse sémantique « big data »), on retrouve ce geste d'appropriation qui passe d'un acteur social à un autre. Dans l'appropriation, il y a aussi le fait que celui qui s'approprie l'objet qui a été inventé le reconnaît une valeur comme légitime, lui attribue une valeur. Ici on peut supposer que si Ipsos s'« approprie » ces dispositifs d'analyse sémantique « big data », il leur reconnaît une valeur d'efficacité et une certaine performance pour améliorer sa pratique professionnelle. Ainsi, le dispositif de moteur sémantique, reconnu comme légitime pour aider le métier des chargés d'étude dans leur quotidien, fait l'objet d'un processus de reconnaissance de son autorité. Se voyant « approprié » dans les métiers au sein des entreprises, il devient ainsi une nouvelle pratique et une norme professionnelle. Ainsi ce geste d'appropriation entre un cabinet de conseil en technologie et une entreprise peut se voir comme la manifestation d'un pouvoir normatif qui s'exerce entre ces deux acteurs socio-économiques.

Jouissant d'un pouvoir rhétorique de recommandation lié à leur ethos algorithmique, les consultants en « technologie big data » influencent des pratiques, comme en témoigne leur recommandation sur le fait de « désiloter » les compétences et les métiers, de croiser les données, changer les approches de travail en interne. Par exemple Affini-Tech, sur son site, propose de « Construire une plateforme analytique universelle pour supprimer les silos dans l'entreprise<sup>187</sup> », ou lors de notre entretien, les fondateurs de Novagen qui recommandent d'entreprendre « un changement organisationnel au sein des entreprises, parce qu'elles marchent souvent en silo (vous avez ceux qui sont spécialisés vente, spécialisés du marketing, etc.) et en fin de compte on regroupe toutes ces données dans un même endroit, ça oblige une réflexion beaucoup plus transversale, et avec une valeur ajoutée qui est assez extraordinaire<sup>188</sup> ». Sous leurs recommandations (« désiloter » les bases de données internes et les croiser entre elles), ils influencent des pratiques professionnelles, des schémas de pensée, des manières de travailler (une campagne marketing, une maintenance des équipements industriels ou le fait de collaborer entre différents corps de métiers). Lorsqu'Affini-Tech propose de « Construire un Datalab pour expérimenter sur données et concevoir des modèles prédictifs<sup>189</sup> », les consultants-ingénieurs influencent les pratiques des entreprises selon cette notion de prédiction, cette nouvelle manière de penser la stratégie en

---

<sup>186</sup> Cf. Annexe 2. Entretien avec Novagen (« Ce sont par exemple des départements ou des équipes de ce type qu'on peut accompagner »)

<sup>187</sup> Cf. Annexe 1. g. Site web Affini-Tech – Projets – « Consolider un datalake »

<sup>188</sup> Cf. Annexe 2., Entretien avec Novagen : « « En rapprochant des données dans un même lieu, qui concernent des métiers différents et des problématiques différentes, en les rassemblant, ça permet d'être beaucoup plus fin aussi bien sur l'analyse des ventes que sur la fraude, que sur les risques »

<sup>189</sup> Cf. Annexe 1. g. Site web Affini-Tech – Projets

entreprise, qui est susceptible de se transformer en norme dans l'organisation. Pour de la maintenance prédictive, ils vont utiliser des techniques de « machine-learning » « pour anticiper des évènements industriels », appliquer des « algorithmes prédictifs sur des flux d'évènements en temps réel<sup>190</sup> » : ainsi une Direction Stratégique d'une entreprise qui travaille avec ces cabinets va être invitée à davantage anticiper les défaillances de son industrie, à prévoir des plans de maintenance industriels, à penser, à écrire la stratégie de l'entreprise sur ce mode de l'anticipation et de prévision des évènements.

Ainsi ces manières de concevoir l'efficacité et la performance se trouvent-t-elles intégrées à un mode de management, une philosophie d'entreprise, qui vont encourager les métiers à se « désiloter », qui vont aussi encourager à innover *via* les données, et surtout à davantage intégrer ces « calculateurs » au sein de l'entreprise. Ainsi, nous mettons en évidence ces différents pouvoirs qui ont trait à la « raison algorithmique » des consultants ingénieurs en traitement des données. Ils ont le pouvoir de recommander et d'influencer des cadres de lecture, d'écriture, et d'action.

### c. Des « scribes » de l'univers : une nouvelle élite savante ?

Dans quelle mesure pourrait-on dire que ces consultants ingénieurs sont de nouveaux « scribes » du réel, de l'univers des entreprises ? Cette notion de « scribe », explicitée par Emmanuël Souchier dans son article « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture<sup>191</sup> », est éclairante pour notre sujet puisqu'elle permet de mettre à jour les pratiques des consultants en technologie et leur positionnement comme savants, voire comme élite savante au sein d'une société qui valorise une compréhension du réel par la « data » et les données numériques, qui prône un modèle d'appréhension du monde selon des critères de calcul, d'algorithmes, de croisement et d'enrichissement de données hétérogènes – les décideurs économiques et politiques souhaitant maîtriser un réel qui leur échappe, un monde qui se dit et se lit sur le mode du changement, que ce soit le changement du marché ou des comportements des consommateurs fragmentés et versatiles. Ainsi la volonté des acteurs économique-politiques de maîtriser une réalité fuyante et changeante induit que ces acteurs prônent la « raison algorithmique » de ces technologies de traitement des données comme autant d'outils pour appréhender, évaluer, mesurer afin d'avoir l'impression de mieux le maîtriser un réel qui leur échappe.

---

<sup>190</sup> Ibid., *Projet – Maintenance Prédictive* : « « Utilisation combinée de plusieurs techniques de Machine Learning, pour anticiper des évènements industriels. Application d'algorithmes prédictifs sur des flux d'évènements en temps réel. »

<sup>191</sup> SOUCHIER Emmanuël, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, de la production à l'interprétation* (pp. 213-229)

Dans ce contexte, les consultants en technologie aideraient ces décideurs à comprendre la complexité de notre société en réseau qui produit des flux de données en continu, ce « web » qui produit des données massives, brutes : « Le Web est pour nos contemporains pareil au « texte » des sumériens. Un univers de « signes » ou de « traces » que l'on doit agencer, organiser, donner à voir. Un « texte complexe » certes, mais un texte à lire et à interpréter, un monde à découvrir<sup>192</sup> ». Ils prétendraient rendre intelligible un système de signes complexes – ici l'univers des « data » et des flux continus de données, ils promouvraient ainsi leur capacité à lire, interpréter et traduire cet univers complexe au travers de leurs « technologies big data », ce qui leur conférerait un certain pouvoir, à l'image des scribes égyptiens qui représentaient une élite sociale car capables de déchiffrer et de poser sur leur écriteau les signes de la réalité qu'ils traduisaient – ici, les consultants en « big data » seraient les seuls capables de déchiffrer les données brutes et les retranscrire de manière intelligible, laissant croire à une compétence rare de pouvoir accéder à ce *savoir via* les données, difficilement compréhensible par tout un chacun. En cela, on peut les considérer comme des « scribes » contemporains, et la société les érigerait ainsi en savants, en une élite savante qui dispose de ce pouvoir d'entendement des « technologies big data », le pouvoir de faire parler les données. Emmanuël Souchier précise d'ailleurs ce statut particulier des « scribes » dans l'histoire de l'écriture : « De tous temps, dans toutes les sociétés où l'écriture s'est déployée, les 'scribes' ont joué un rôle prépondérant dans le rapport au pouvoir et dans la configuration même de l'outil qui leur conférait ce pouvoir (Martin, 1996). Comment s'étonner que les nouveaux scribes soient à leur tour happés par le pouvoir que représente « l'artefact politique » qu'ils modèlent ?<sup>193</sup> ». Emmanuël Souchier compare ici les ingénieurs-cartographes à des scribes qui détiennent un pouvoir puisqu'ils élaborent des « artefacts », ces outils qui sont utilisés par des acteurs ou des lieux où s'exerce le pouvoir et s'élaborent les stratégies, là où se font des choix importants. Dans cette même logique, on peut considérer nos consultants ingénieurs en « data » comme ces scribes dépeints par E. Souchier, puisqu'ils créent eux aussi des « artefacts », c'est-à-dire des construits (ici informationnels et organisationnels), des produits créés par l'homme à travers des procédés de création et de transformation successifs. D'ailleurs, Magali Ollagnier-Beldame désigne les inscriptions numériques comme étant elles-mêmes des artefacts, car « externes à l'humain », des « construits » dans un support<sup>194</sup> ». Ces artefacts sont ensuite mobilisés par des acteurs qui détiennent un pouvoir dans la société, des entités dirigeantes dans les sphères économiques et politiques (que ce soit dans les entreprises ou les institutions) qui vont s'emparer de ces artefacts et faire des choix stratégiques qui vont avoir un impact sur une organisation, une entreprise, une sphère de la société.

---

<sup>192</sup> Op. cit., p. 214

<sup>193</sup> Op. cit., p. 219

<sup>194</sup> OLLAGNIER-BELDAME Magali, « Interagir dans un monde de plus en plus réflexif : processus cognitifs et traces numériques – mémoire, interprétation et rapport au temps », *Traces numériques : de la production à l'interprétation*, CNRS Editions, 2013 (pp. 129-142)

Ces consultants-ingénieurs, traducteurs d'un univers complexe fait de « data » chaotique et déstructurée, se verraient érigés à la posture de savants, voire de démiurges, dans la mesure où ils produisent une connaissance qui aide à « gérer la complexité croissante de notre société et a fortiori celle de demain<sup>195</sup> ». Cette connaissance a une valeur pour les entités dirigeantes, et en structurant cette masse de données informe et en lui donnant une matérialité ordonnée.

### 3. L'ambition démiurgique de leur ethos algorithmique : prétendre ordonner le chaos informationnel<sup>196</sup> par les technologies de traitement des données

#### a. Ces consultants en technologie exerceraient plus qu'un pouvoir performatif : un acte démiurgique ?

Plus que de donner à voir sous nos yeux, ces dispositifs de traitement des « grosses données » prétendent créer une réalité à partir d'un chaos de données. Plus qu'une dimension performative, on pourrait supposer que cet acte revêt une dimension démiurgique car on retrouve cette notion de création, à partir d'un chaos. Il va s'agir de créer une « signifiante » à partir du chaos, à partir de cette « technologie intellectuelle », qui positionne leurs « technologies big data » à l'égal de Dieu, puisqu'elles se donnent pour ambition de créer une « intelligence », une nouvelle réalité à partir d'une masse de données hétérogènes, massives. Il va s'agir de créer des corrélations de signification, de créer du sens à partir d'opérations techniques effectuées par les « data-scientist » et les ingénieurs de la « big data ». Il apparaît ainsi que ces acteurs du traitement de la donnée se donnent pour mission d'ordonner un chaos existant, postulant a priori que l'entreprise est dans une situation de désordre avec ses données. Novagen propose ainsi d'« Epurer, nettoyer, reformater, agréger les données<sup>197</sup> », présupposant une nécessité de venir structurer les données des organisations. On se retrouve presque en face d'un récit de la création, avec un objet technique considéré comme tout-puissant, capable de créer une information structurée, nettoyée et intelligible et utilisable par l'homme. Quand Novagen affirme sur son site web, « Le croisement des innovations et des explorations est créateur de valeur<sup>198</sup> », ou dans son support de présentation,

---

<sup>195</sup> Pascal Robert, dans *Communication & langages* (« Qu'est-ce qu'une technologie intellectuelle ? », 2000, pp. 97-114) évoque l'algorithme de Bell comme une « technologie intellectuelle » (p. 98), de même que l'écriture et la liste offrent une « intelligence gestionnaire » (p. 100).

<sup>196</sup> Expression utilisée par Béatrice Galinon-Mélénec à propos de l'utilisation de la cartographie du Web dans *Traces numériques. De la production à l'interprétation* : « Pour faire émerger un sens pertinent dans le chaos informationnel résultant de l'accessibilité permanente à un nombre de plus en plus grand de données, faut-il avoir recours à une cartographie du Web ? » (p. 10)

<sup>197</sup> Cf. Annexe 1. g. Site web Affini-Tech – Conseil IT (« Data Analysis »)

<sup>198</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

« Création d'indicateurs et de tableaux de bords », « De nouveaux services à créer<sup>199</sup> », on retrouve là cette notion de création. Les ingénieurs, par leurs « technologies big data », créent des « artefacts » c'est-à-dire des produits créés de toutes pièces (il y a aussi une dimension de création et de production) qui vont ensuite être utilisés dans le commun des mortels, circulant entre les hommes. L'information donnée à voir *via* ces dispositifs peut en cela être considérée comme étant le produit de la création de l'ingénieur des données qui se positionnent à l'égal de Dieu, produisant des objets signifiant à partir d'un chaos originel. Par là même, on pourrait dire que ces ingénieurs des données font un acte analogue à celui du démiurge. Cette dimension démiurgique se retrouve aussi dans le fait qu'il dispose d'une marge de manoeuvre libre dans la mesure où personne – ou alors une minorité – ne va remettre en question les artefacts qu'il donne à voir. Les tableaux de bord qu'il présente aux entreprises, les histogrammes, les tableurs de données, sont autant de produits de sa création qui font naturellement autorité, qui sont considérés comme allant de soi et existant de manière évidente, comme si elles faisaient partie d'un ordre du monde qui a été pensé et créé de manière cohérente et qui a un sens, car créé par un démiurge tout-puissant qui pense ce même ordonnancement du monde et donne une matérialité aux choses qu'il crée. Ici, les consultants-ingénieurs en données organisent des informations hétérogènes et déstructurées, les structurent selon un certain modèle d'ordonnancement, et leur donnent une matérialité (le schéma, l'histogramme, le tableur numérique de données). Emmanuel Souchier, évoque cette ambition du cartographe à donner forme au chaos et à « créer la « signifiante » à l'égal des dieux », parce qu'il formule son acte intellectuel comme étant « une grille logique méticuleuse où se trouve assujettie une partie du Web devenu maintenant plus signifiant<sup>200</sup> ».

L'activité d'interprétation des consultants-ingénieurs revêt une dimension démiurgique dans la mesure où ingénieur-cartographe (ici ingénieur en traitement de données) et devin ont chacun « une marge de manoeuvre analogue qui le laisse seul maître à bord », précisant que le « cartographe est maître des procédures<sup>201</sup> », ce qui marque à la fois « la puissance démiurgique de leur pouvoir et les limites de leur « scientificité ». En effet l'ingénieur en « data » est le seul à choisir les modèles de calculs qu'il va appliquer, à créer des algorithmes génératifs et ajuster les paramètres d'entrée, à identifier si les algorithmes fonctionnent bien ou non<sup>202</sup>, à repérer les failles, à sélectionner les résultats qu'il va présenter aux entreprises ». Par exemple la « data-scientist »

---

<sup>199</sup> Cf. Annexe 8. Support de présentation de Novagen – slide p. 14

<sup>200</sup> SOUCHIER Emmanuel, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, De la production à l'interprétation* : « Un geste démiurgique » (p. 226)

<sup>201</sup> Op. cit., p. 228

<sup>202</sup> Cf. Annexe 2. Entretien avec Novagen : « là aussi il y a une expertise humaine qui est liée à la compréhension du modèle mathématique qui se cache derrière les algorithmes. Et ensuite il faut pouvoir itérer, confronter les résultats qu'on obtient en confrontant les données à des algorithmes, avec les phénomènes observés : est-ce que la traduction est bonne ou mauvaise, et selon l'état, on va entrer encore dans des cycles, on va ajuster les paramètres d'entrée, ajuster notre stratégie de prévision, et à la fin aboutir à un résultat »

d'Ipsos, Charlotte Zapfel, explique ce procédé de choix d'opérations vectorielles qu'elle est la seule à faire : « Ici à Ipsos, je fais du *topic modelling*, qui a pour but de modéliser un corpus de texte pour qu'il soit utilisable par la machine. Je dois vectoriser des textes, c'est-à-dire créer une représentation par des vecteurs (qui sont des objets mathématiques très flexibles *via* des équations). J'applique donc des opérations vectorielles ». C'est elle qui est maître à bord puisqu'elle décide seule des « champs sémantiques » à représenter – montrant là une certaine « puissance démiurgique », elle est « maître des procédures » et dispose d'une marge de manoeuvre dans son activité<sup>203</sup>. « Bref, il est le seul à appliquer le « tison » informatique qui lui permettra de créer des « effets variés » sur la forme des cartes qui seront ensuite interprétées. » Emmanuël Souchier ajoute que « ce microcosme est marqué du sceau de son pouvoir » (Ghitalla, 2008)<sup>204</sup>. Ainsi, la « data-scientist » d'Ipsos marque du sceau de son pouvoir les artefacts qu'elle produit. Pour préciser cette dimension démiurgique qu'on retrouve dans la posture particulière des consultants ingénieurs, Emmanuël Souchier ajoute : « Il s'agit bien ici de « donner corps » à un monde invisible, de lui donner « forme ». L'acte est démiurgique, il fait apparaître ; c'est un acte de pouvoir, il donne à voir et régule la signification<sup>205</sup> ».

Enfin, il convient de souligner que leur pouvoir démiurgique de créer du sens à partir d'une masse de données brutes et informes, s'accompagne aussi d'un phénomène de croyance en leur « rationalité algorithmique » : « Ce pouvoir démiurgique – consistant à donner du sens, à établir des analogies comme dans la Chine ancienne ou l'antique Sumer, les technologues du Web le tirent également de leur croyance en la techno-science qui constitue leur objet<sup>206</sup> ». Ainsi, si les ingénieurs des données peuvent être comparés à des démiurges, c'est aussi parce qu'ils tirent leur pouvoir de leur « raison algorithmique » qui est au fondement de leur posture communicationnelle. Notre société étant marquée par un système de représentations et de croyances au sein duquel la « techno-science » est hautement valorisée, cela peut expliquer la dimension démiurgique que peuvent revêtir ces ingénieurs.

## b. L'influence des « signes » envoyés par la « raison algorithmique » dans le monde de l'entreprise

Nous avons remarqué une préoccupation des acteurs de l'entreprise à prendre en considération les « signes » que l'ingénieur en données donne à voir aux acteurs de l'entreprise,

---

<sup>203</sup> Pour illustrer cette posture de « maître à bord », elle ajoute : « Devant le client, je les ai présentés par une table de synonymes, par ordre croissant de similarité, et ça a permis de nourrir le quali et l'animation de focus groups ». Le choix de ce critère (la similarité) témoigne de cette marge de manoeuvre dans les choix qu'elle opère, en effet elle aurait pu présenter cette table de synonymes selon un autre critère de classification (ordre décroissant de différence, corrélations entre des sous-thématiques, ou encore présenter une table d'antonymes). Cf. Annexe 3. Entretien avec Charlotte Zaepfel

<sup>204</sup> Op. cit., SOUCHIER Emmanuël, p. 229

<sup>205</sup> Op. cit., p. 225

<sup>206</sup> Op. cit., p. 230

phénomène qu'on pourrait comparer à celui du demiurge créateur qui envoie des « signes » afin que ceux-ci soient pris en compte par les hommes – en tout cas ceux qui croient en la légitimité de son existence. Ces « signes » envoyés par l'ingénieur des données ou ses technologies, une fois identifiés comme tels sont interprétés par les hommes, qui s'en emparent pour déterminer quel comportement adopter. Ainsi les décisions prises par les sphères managériales des entreprises qui ont recours à ces cabinets de conseil en données, se feraient en fonction de ces « signes » envoyés par le technologue demiurge. Pour Ipsos, cela va par exemple consister à identifier les « signaux faibles » pour aider à leur tour leurs propres clients. Notre entretien avec Agnès Gilbert, « Chief Digital Officer » à Ipsos, montre qu'en utilisant le « topic modelling » avec le pôle « Data Sciences », leur réponse à une problématique va s'ajuster différemment en fonction des résultats – sous forme de graphes de mots par exemple – que donnent à voir les logiciels de « topic modelling ». Ils servent à « faire remonter les sujets sur un corpus varié, à partir d'algorithmes qui établissent des corrélations. On va travailler le mot avec un moteur d'analyse sémantique sur le web, qui isole les occurrences et analyse les récurrences et les fréquences d'associations entre les mots<sup>207</sup> ». Ainsi la recommandation d'Ipsos (qu'ils appellent des « insights ») va être influencée par le graphe qui va donner à voir des « signes » spécifiques (les thématiques d'un corpus de textes numériques, leurs associations sémantiques). Ipsos va ensuite élaborer ses « insights » en fonction des signes qui sont remontés par les algorithmes de « topic modelling ». On constate ici l'influence de la « raison algorithmique » qui est prise en compte par les acteurs de l'entreprise, qui à leur tour exercent une « raison algorithmique », du fait de leur croyance même en cette rationalité.

Dans le monde de l'entreprise, il semble ainsi que les hommes croient au fait qu'ils doivent prendre en considération les signes de cette rationalité technoscientifique pour déterminer un comportement qu'ils doivent adopter. Dans un article de *Stratégies* intitulé « Les instituts domptent la data », certains interprètent ces « signes » comme un ensemble d'actions à suivre pour enrichir les compétences des différents métiers. Ces signes technologiques sont vus comme « une ouverture du champ des possibles, qui selon François Baradat, de Kantar « plonge le marché des études dans une époque inédite, avec une accélération phénoménale des pratiques de notre métier, tout autant vertigineuse que passionnante<sup>208</sup> ». On voit ainsi la croyance à l'œuvre que les acteurs de l'entreprise (ici les instituts d'étude) doivent suivre un ordre du monde lié à la « raison algorithmique », caractérisé par l'attente en de « multiples innovations qui attendent le marché des études ». Nous avons vu que ces « signes » issus de la « raison algorithmique » présentés par les consultants-ingénieurs sont certes accueillis favorablement, mais ils semblent aussi être craints, par ceux justement qui croient peu en cette « raison algorithmique ». Dans l'article « Les instituts domptent

---

<sup>207</sup> Cf. Annexe 4. Entretien avec Agnès Gilbert

<sup>208</sup> Delphine Le Goff, « Les instituts domptent la data », *Stratégies*, n°1900, 13 avril 2017 (p. 25 à 27), trouvé sur l'intranet d'Ipsos (*Ipside*)

la data<sup>209</sup> », un acteur exprime cette inquiétude : « Soulagement : les instituts n'ont pas été engloutis par la *data*. Il existait une grande crainte de la *big data*, censée remplacer nos métiers<sup>210</sup> », ou encore : « Beaucoup ont prédit la mort de ce métier. Je préfère l'idée d'une extension du domaine de la lutte pour les instituts<sup>211</sup> ». On voit là que les signes sont interprétés différemment selon la croyance des acteurs (y voir un mort des métiers, ou pour d'autre une opportunité d'étendre leurs compétences). On voit aussi que ces technologies sous-tendent des craintes de la part des acteurs dans l'entreprise, du fait de ce pouvoir démiurgique qui peut avoir un effet créateur ou destructeur.

Nous voyons ici l'influence que peuvent avoir les « signes » dans les artefacts que présentent les ingénieurs des données auprès des acteurs de l'entreprise, qui vont modifier leurs pratiques professionnelles, créer de nouvelles normes de leurs métiers en interne (intégrer les données brutes et massives dans le métier du chargé d'études), élaborer des recommandations stratégiques pour leurs clients, suivre un plan d'actions en s'éclairant à partir de leur croyance en la « raison algorithmique ».

### c. La prétention démiurgique à maîtriser et prévoir les événements : des simulacres

Les ingénieurs technologues aident à déchiffrer une réalité abstraite, qui paraît complexe et désordonnée. Cette image s'inscrit dans l'imaginaire d'Internet qui figure des flux de données continus et chaotiques, à l'image d'une toile où l'on s'embourberait, rendant nécessaire la présence d'une entité qui nous aide à suivre notre propre chemin le sens sur cette toile chaotique. Les consultants-ingénieurs en « data » façonnent ainsi des outils technologiques qui rassurent et donnent l'impression de maîtriser un réel complexe, abstrait, en donnant l'impression de soumettre ce réel à ces dispositifs qui l'organise et lui donne un éclairage. Plus encore que de faire valeur de connaissance, le traitement et l'analyse des « traces » numériques fait valeur de maîtrise d'un monde. Le dispositif technologique rassure et réduit les incertitudes des décideurs stratégiques (Direction Générale ou Marketing), c'est en cela qu'on peut les considérer comme des « simulacres » de maîtrise du monde. Par une interprétation des données de l'entreprise et du web, les « technologies big data » prétendent, comme les cartes, fournir une intelligence à base de données et se constituer comme des outils d'intelligence économique qui aident à connaître le cours des événements pour s'y adapter le plus pertinemment possible. On pense à Novagen qui

---

<sup>209</sup> Ibid.

<sup>210</sup> Op. cit., Luc Balleroy, directeur général d'Opinion Way

<sup>211</sup> Op. cit., François Baradat, directeur marketing de Kantar TNS

propose des « outils de Business Intelligence, temps réel et sur-mesure », pour « Maîtriser son Marché par le traitement Analytique des Données du Web<sup>212</sup> ».

Nous pouvons dire que ces consultants ingénieurs élaborent un simulacre de la maîtrise du cours des évènements, faisant croire qu'ils offrent aux décideurs économiques la possibilité d'avoir la main sur le cours des choses, de pouvoir, de leur transmettre ce pouvoir presque démiurgique de maîtriser les évènements du monde, voire de prédire ceux-ci. On retrouve dans leurs dispositifs technologiques des principes de contrôle, par exemple le fait de présenter un diagramme témoigne d'un procédé de contrôle (celui du cadre interprétatif défini par les contours du diagramme). En effet, si l'on définit le « diagramme comme geste de riposte à la crise et tracé d'un plan de composition<sup>213</sup> », on peut l'analyser comme un geste intellectuel qui voudrait rompre un système chaotique, le soumettre à son contrôle. Le geste du tracé délimite l'objet, délimite l'artefact numérique et est par essence un « construit », externe à l'homme. Nous aurions à faire à une lecture du réel cantonnée à ce cadre interprétatif réduit, donnant l'impression de maîtrise car on fait rentrer des éléments du réel dans ce tracé, dans ce cadre qui fait croire aux dirigeants qu'ils ont une vue synthétique du marché, des ventes, du chiffre d'affaires. On retrouve chez Affini-Tech des expressions qui traduisent cette rationalité qui aspire à une maîtrise du cours des évènements, telles que « Construire une vue 360° cross-canal » ou « Détection d'anomalie avec des algorithmes de classification ». Concernant le simulacre de leur capacité à prévoir les évènements, on peut l'identifier à travers un champ sémantique lié à la prédiction, avec « Application d'algorithmes prédictifs sur des flux d'évènements en temps réel », « Prédire la prochaine action dans la relation client » ou encore « Utilisation combinées de plusieurs techniques de Machine Learning, pour anticiper des évènements industriels<sup>214</sup> ». Ainsi ce ne sont pas des outils de maîtrise, mais des simulacres car des outils soumis au contrôle d'une certaine rationalité, car ils donnent seulement l'impression d'avoir une prise sur les évènements. Or on ne peut avoir de maîtrise du cours des choses, la réalité étant par définition contingente et non calculable en avance. On assisterait à la mise en place d'un système de contrôle pour faire croire à ce contrôle, avec des expressions qui seraient proférées telles des invocations pour invalider, conjurer un monde fait d'incertitudes et d'imprévisibilité. On ne se retrouve pas tant devant une réelle maîtrise des évènements que devant un simulacre de cette maîtrise. Car la maîtrise du réel relève d'une prétention, le réel est trop complexe pour avoir la prétention de s'en saisir.

Nous voyons bien que ces acteurs de l'industrie du traitement de la donnée opèrent un processus de construction de sens, qui découle d'un rapport au monde particulier, de la « raison

---

<sup>212</sup> Cf. Annexe 1. b. Site web Novagen – Data Factory

<sup>213</sup> BERGEN Véronique, « Saisies du simulacre chez Baudrillard et chez Deleuze », *Lignes*, 2010, n° 31, 192 p. (URL : [www.cairn.info/revue-lignes-2010-1-page-37.htm](http://www.cairn.info/revue-lignes-2010-1-page-37.htm))

<sup>214</sup> Cf. Annexe 1. g. Site web Affini-Tech – Projets

algorithmique » que nous avons définie. Concernant la posture de devin ou de savant dont sont investis les consultants-ingénieurs, on peut penser à l’aveu du philosophe par rapport au savoir – Socrate disait « La seule chose que je sais, c’est que je ne sais rien ». Le vrai savant serait-il précisément celui qui reconnaît qu’il ne sait pas, plutôt que celui qui prétend détenir la connaissance ? Il semble plutôt que ces acteurs de la donnée soient plutôt des « démiurges », puisqu’ils sont dans la posture de celui qui croit qu’il peut avoir un impact sur le cours du monde. Cette prétention démiurgique caractériserait la « raison algorithmique », et nous avons tenté de déconstruire celle-ci en mettant en avant les simulacres qu’elle construit.

Nous remarquons donc que l’activité de traitement des « traces » numériques correspond à une certaine forme d’ordonnement du monde, fondée sur la croyance en l’application de modèles de calculs pour appréhender un univers (le marché économique, les comportements des consommateurs, le diagnostic d’un chaîne de production industrielle). Les consultants-ingénieurs, adressant leurs outils de valorisation des données aux entités dirigeantes, entraînent une circulation de leurs dispositifs technologiques dans l’entreprise. L’univers gestionnaire s’empare des univers ingénieur et technologique, et élabore ses discours stratégiques en justifiant des choix par l’argument technologique et de la performance.

### III. LES DISCOURS DES DÉCIDEURS SUR LA PERFORMANCE SE RÉINVENTENT PAR LA CIRCULATION DE LA « RAISON ALGORITHMIQUE » AU SEIN DES ENTREPRISES

Nous avons vu que les « technologies big data » promues par les cabinets de conseil en technologie sont un objet qui se trouve légitimé auprès des entités dirigeantes. Les sphères décisionnelles des entreprises ont tendance à les considérer comme producteurs d'une « intelligence » du marché, et révélateurs de phénomènes économiques et sociaux. Il semble considéré comme un outil qui fait valeur de connaissance et d'aide à la décision, qui est susceptible d'apporter de « nouveaux gisements de valeur<sup>215</sup> ». Ces dispositifs de traitement des données constituent un capital stratégique pour l'entreprise, qui va pouvoir ajuster sa stratégie à partir de la connaissance produite à partir des données traitées et valorisées.

Nous allons analyser comment les discours sur les « technologies big data », promus par les cabinets-conseil en technologie, se trouvent mis en circulation au sein des sphères managériales des entreprises, faisant l'objet de procédés d'appropriations par ceux-ci. Nous allons analyser comment circule la « raison algorithmique » que nous avons examinée jusqu'à présent. Comment les décideurs d'entreprise s'approprient-ils ces outils de traitement des données et construisent-ils à leur tour un discours sur ces mêmes technologies ? Sous l'influence des discours technocentrés des consultants en technologie, il semble que les entités dirigeantes s'emparent de ces mêmes discours pour à leur tour légitimer et justifier leurs choix stratégiques au sein de l'organisation.

#### 1. Une appropriation par les acteurs économiques : des outils de gouvernement au pouvoir ?

Comment la circulation des discours de ces consultants-experts entraîne une modification des discours et des pratiques au sein de l'entreprise, parmi les sphères managériales ? Comment le discours de la performance se construit-il à partir de cette « raison algorithmique » ? Il semblerait que la circulation de ces outils technologiques qui traitent les « big data » sous-tende des enjeux de pouvoir entre les acteurs de l'entreprise.

Ils seraient utilisés comme outils de gouvernement, c'est-à-dire que les chiffres sont utilisés pour ajuster et légitimer le discours des acteurs dans l'entreprise. Ils confèrent une posture

---

<sup>215</sup> Cf. Annexe 1. f. Site web Affini-Tech – Nos Services

rationnelle, basée sur la « raison algorithmique » que nous avons analysée, et par là même donnent du pouvoir, viennent appuyer un plan d’actions, des choix stratégiques. Les prises de décisions se voient justifiées à travers les données et constituent des arguments à part entière dans le discours du manager et dans celui du salarié – des enjeux de pouvoir semblent s’y jouer.

#### a. L’influence de la « raison algorithmique » sur les choix stratégiques des discours des décideurs : des technologies intégrées à un discours général sur la performance économique

Nous constatons une multiplication des discours technocentrés des acteurs organisationnels et institutionnels, qui souhaitent intégrer davantage de dispositifs technologiques au cœur de leurs stratégies d’entreprise. Les « technologies big data » circulent comme un objet de discours parmi les sphères décisionnelles de l’entreprise (marketing, direction générale...), pour venir justifier des actions et des décisions.

Ces technologies ont la prétention d’améliorer la performance des acteurs économiques, d’optimiser leurs métiers, prétentions inhérente à la « raison algorithmique » qui souhaite structurer les flux de données pour en faire des « leviers business <sup>216</sup> ». Cette prétention de rendre l’entreprise plus performante se retrouve chez Novagen : « tout ce qu’on aura acquis comme donnée et qu’on aura préparé au travers de *stratégies data lake*, stratégies de réconciliation des données, vont donner des avantages très compétitifs dans les années à venir. On va mieux connaître ses clients, on va être plus efficace, faire des prédictions sur les ventes beaucoup plus fines, on peut optimiser un métier industriel sur la maintenance des chaînes de production <sup>217</sup> ». Eric Sadin souligne ce phénomène d’injonction à la performance propre à la « raison numérique » : « Les flux de data « s'érigeant comme l'instance primordiale de l'intelligibilité du réel [...] visant les plus hautes optimisation, fluidification et sécurisation des existences et des sociétés <sup>218</sup> ». Cette rationalité circulant dans l’univers de l’entreprise, elle incite ses acteurs à s’emparer de cette injonction de performance. Les acteurs intégreraient cette rationalité à leurs pratiques professionnelles, concevant progressivement leurs métiers à travers le filtre de l’amélioration continue, l’optimisation, l’efficacité commerciale. Erik Bertin, dans son article dans *Influencia*, note qu’avec une telle rationalité, « toute conduite, toute action est optimisable, du point de vue du calcul et de la modélisation. L’algorithme est la clé d’accès à un savoir cumulatif sur soi et sur la société, dans un but d’amélioration ». On va soumettre l’univers de l’entreprise à des modèles de calcul et à des algorithmes en vue d’optimiser

---

<sup>216</sup> Cf. Annexe 1. e. Site web Affini-Tech – Page –d’accueil: « De l’idée au service, nous accélérons votre business avec vos données ! », et « Le savoir-faire de nos Data-Engineer et de nos Data-scientists vous éclairera sur les leviers business, et l’intégration des plateformes nécessaires »

<sup>217</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>218</sup> SADIN Eric, *La vie algorithmique, critique de la raison numérique*, Paris : Éditions L’Échappée, 2015, 278 p. (Quatrième de couverture)

les pratiques, et ce dans un discours des managers sur la performance de l'entreprise, nécessaire pour avoir des « avantages compétitifs » et développer son activité. Par exemple, on va avoir à faire à la mise en place de normes qui poussent à intégrer des méthodologies qui permettent de mesurer, de croiser les données, de mesurer les « traces numériques », pour les évaluer et améliorer les pratiques et les métiers, et à terme, rendre l'entreprise plus efficace et performante. À titre d'illustration, dans le magazine interne d'Ipsos, quand Agnès Gilbert parle d'aller « chercher de nouvelles compétences/ressources », pour « répondre aux briefs de manière différente, plus rapide et plus efficace<sup>219</sup> », on remarque ici l'idéal d'efficacité et de performance économique à l'œuvre dans les discours des entités dirigeantes. Le philosophe Hermut Rosa souligne ce phénomène selon lequel nous serions en quête de croissance perpétuelle. On voit donc l'opérativité des idéologies technocentrées, qui promettent, grâce au traitement informatique de données massives, d'améliorer un état existant, la santé économique d'une entreprise, l'efficacité de ses procédures internes, de ses métiers, ou encore la performance globale de son modèle économique.

Des éléments de langage se construisent autour de la « raison algorithmique » et fondent le discours des décideurs sur la stratégie d'entreprise. Nous identifions des discours sur la performance de l'entreprise qui portent sur l'amélioration d'un ROI, la mesure de l'efficacité de campagnes marketing, ou encore l'amélioration de la connaissance client. C'est d'ailleurs ce qui compose l'offre de services des cabinets de conseil en technologie<sup>220</sup>. On constate que se formulent des injonctions à la performance, avec des stratégies « data-driven », c'est-à-dire guidées par la croyance des l'efficacité économique du traitement des données. Ce phénomène de formulation d'injonctions à la performance *via* l'intégration des technologies de « big data », peut s'éclairer par le concept d'Antoinette Rouvroy de « gouvernementalité algorithmique<sup>221</sup> » qui impose son « régime de vérité numérique », expliquant cette tentation des acteurs économiques à la « mise en nombre » du monde social, cette tentation à mesurer les phénomènes du monde selon les spectre des données, en vue de les optimiser. Dans « Gouvernementalité algorithmique et perspectives d'émancipation<sup>222</sup> », Antoinette Rouvroy et Thomas Berns révèlent cette tendance qui touche tous les acteurs de la société : « Les gouvernements les récoltent à des fins de sécurité, de contrôle, de gestion des ressources, d'optimisation des dépenses... ; les entreprises privées recueillent quantité de données à des fins de marketing et de publicité, d'individualisation des offres, d'amélioration de leur gestion des stocks ou de leurs offres de service, bref, en vue d'accroître leur efficacité

---

<sup>219</sup> Cf. Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017, « Dossier spécial inno : portraits. Agnès Gilbert, Chief Digital Officer, une ressource à partager ! » (p. 9)

<sup>220</sup> Cf. Sites web de Novagen et d'Affini-Tech

<sup>221</sup> ROUVROY Antoinette, BERNIS Thomas, « Gouvernementalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », *Réseaux*, 2013, n° 177 (pp. 163-196)

<sup>222</sup> Op. cit., p. 168

commerciale et donc leurs profits... ; les scientifiques collectionnent les données à des fins d'acquisition et d'amélioration de connaissances... ».

Nous voyons ainsi que cette « gouvernamentalité algorithmique » caractériserait les choix stratégiques des décideurs économiques et politiques, sous l'effet des discours technocentrés qui affirment un « pouvoir des données ». Les dirigeants élaborent donc des discours orientés sur la performance de l'organisation, qui contiennent des injonctions à devenir plus performant grâce au traitement des données.

#### b. Valoriser ses données pour survivre dans un environnement économique compétitif : le traitement de données au cœur des stratégies d'entreprise

L'appropriation des discours autour des « technologies big data » par les entreprises passe aussi par la conviction des entités dirigeantes que la « raison algorithmique » aide à développer l'efficacité économique et commerciale d'une entreprise, d'autant plus dans un environnement compétitif. Quand l'article « Les instituts domptent la data » paru dans *Stratégies*, évoque « l'apparition de nouveaux concurrents », il rejoint cette thématique de l'adaptation à un contexte concurrentiel *via* l'intégration de plateformes et de méthodologies liées à la « data ». Dans un contexte de compétitivité économique, ces technologies de traitement de données massives font partie des discours des décideurs stratégiques (que ce soit la Direction Générale ou Marketing), constituent des éléments de langage à part entière, prenant souvent la forme d'injonctions à la performance comme nous l'avons vu dans la sous-partie précédente, édictant de nouvelles normes à adopter pour améliorer la performance. On constate la mise en œuvre de stratégies d'entreprise qui mettent la « data » au cœur de leur plan stratégique, et imposent ces outils et méthodologies de travail reposant sur le traitement et l'analyse des « big data ». Ils prônent ce modèle, pensant qu'il faut à tout prix valoriser ses données pour ne pas « être mort » sur le marché. L'entretien avec Novagen révèle cette préoccupation des décideurs : « J'ai travaillé par exemple avec PSA et Renault, les grands groupes se disent, « si on ne le fait pas, si on n'investit pas à fond, on est mort » ; et ils ont commencé à le faire dans des volumes significatifs. Pareil pour les grands groupes bancaires, qui peuvent se faire concurrencer très largement d'un mois sur l'autre, et avec des questions réglementaires, assurances... Tous les acteurs se disent qu'ils sont potentiellement menacés. Donc ils ont initié des démarches big data<sup>223</sup> ». Aujourd'hui il semble qu'être une structure performante passe par ces outils de « big data », pour mieux comprendre un marché devenu complexe, des consommateurs aux multiples faces. Les technologies de traitement des

---

<sup>223</sup> Cf. Annexe 2. Entretien avec Novagen

« grosses données » se trouvent considérées comme des outils de décision et d'aide à l'action incontournables pour les entreprises qui veulent maîtriser les incertitudes du marché d'aujourd'hui<sup>224</sup>.

Ce phénomène de forte légitimation des acteurs de l'industrie de la donnée se manifeste par la place centrale des postes liés au traitement des données, la création de postes spécifiques en interne, le recrutement de « data-scientist » en entreprise (chez Ipsos) ou de Responsables de « projets big data » (comme chez Airbus<sup>225</sup>), et dans les cabinets conseil (la page « recrutement » des sociétés de traitement de données sont remplies d'offres d'emploi liées au traitement des « big data »). Notre entretien avec Charlotte Zaepfel nous a aussi révélé qu'elle ressentait cette injonction de la « data », a remarqué que les sociétés sont avides de ces profils de « data-scientists » qui savent manier le langage de programmation des « big data » et envisage donc se former en conséquence : « des langages comme Hadoop et Spark, auxquels il faut que je me forme car c'est très demandé sur le marché par beaucoup de boîtes », [...] « il faut d'ailleurs que je me mette à une formation pour Hadoop et Spark pour connaître ce langage « big data », puisque c'est là où il y a des perspectives d'embauche !<sup>226</sup> ». Cette légitimation à l'œuvre des acteurs spécialisés en « technologies big data » se manifeste aussi par une collaboration croissante avec des partenaires spécialisés dans le traitement des données, comme Ipsos qui noue des partenariats avec Weborama et Big Fish<sup>227</sup> ; ou le fait qu'Ipsos soit en discussions avec le cabinet-conseil Novagen. Dans « Les instituts domptent la data », on voit bien ce phénomène d'hybridation des pratiques professionnelles à l'œuvre entre les acteurs de l'entreprise et les acteurs des traitements des données massives qui vendent du logiciel « big data » : « Autre mot-clé : l'hybridation. La *data* n'est plus l'ennemie, mais plutôt un complément. Et les pure-players pas forcément des rivaux, mais des partenaires. Nous manions les données structurées et non structurées, sans remettre en cause notre cœur de métier »<sup>228</sup>. Par ailleurs, les entités dirigeantes veulent montrer que leur entreprise est « à la page », notamment en prenant l'exemple des grands acteurs du web et de la donnée qui font valeur de modèles de performance économique et de succès commercial. Novagen, lors de notre entretien, fait remarquer le pouvoir que donne le fait de disposer de larges corpus de données : « Les niveaux de maturité sont différents. Déjà on voit que les entreprises du web se sont construites là-dessus et

---

<sup>224</sup> L'offre de services de Novagen entre en cohérence avec ce contexte, quand ils proposent des « outils de Business Intelligence » pour « maîtriser son marché » (cf. Annexe 1. b. Data Factory – Ce qu'apporte le Data Lab)

<sup>225</sup> « Un Datalab chez Airbus pour diffuser le big data dans tous les métiers », *Usine Digitale*, 10 mars 2016, <http://www.usine-digitale.fr/article/un-datalab-chez-airbus-pour-diffuser-le-big-data-dans-tous-les-metiers.N383312>

<sup>226</sup> Cf. Annexe 3. Entretien avec Charlotte Zaepfel

<sup>227</sup> Cf. Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017, « Dossier spécial inno : portraits. Agnès Gilbert, Chief Digital Officer, une ressource à partager ! » et article de *Stratégies* « Les instituts domptent la data » (« Nous nouons ainsi des partenariats avec des acteurs comme Weborama, avec qui nous avons élaboré une solution, Big Fish, un outil de *datamining* qui permet de dresser la carte sémantique d'une marque sur le net. »)

<sup>228</sup> Cf. Annexe 7. « Les instituts domptent la data », *Stratégies*, 13 avril 2017, n° 1900

ont des avantages compétitifs extraordinaires. Vous prenez Google, ils ont Waze, ils connaissent tous les comportements, donc quand on est un acteur automobile par exemple, Google est capable de piloter des voitures, de faire l'optimisation du trafic. Si moi je ne m'y mets pas, je suis mort<sup>229</sup> ». Ils font le postulat que le niveau de maturité d'une entreprise dépend désormais du fait qu'elle ait intégré, ou non, des « technologies big data » dans son modèle économique.

### c. L'utilisation politique : des outils de gouvernement qui assoient la légitimité des décideurs

Autre procédé d'appropriation de cette « raison algorithmique » par les acteurs de l'entreprise, le fait qu'ils se dotent d'un pouvoir *via* ces « outils big data », en montrant qu'ils connaissent et maîtrisent cette culture technologique, et détiennent une solution pour améliorer la performance de l'entreprise, pour constituer une figure auctoriale qui fait autorité à partir de cette « raison algorithmique ». Dominique Cardon, dans *À quoi rêvent les algorithmes : Nos vies à l'heure des big data*, met en lumière les « technologies big data » que nous devrions considérer comme des « outils de gouvernement »<sup>230</sup>. Les algorithmes sont « plus que de simples outils techniques<sup>231</sup> », c'est-à-dire qu'ils font partie d'un projet politique. Ils sont au cœur des discussions des Directions Stratégiques des entreprises et ont une place de plus en plus importante dans les comportements des acteurs sociaux (professionnels, politiques, économiques). Pourrait-on considérer ces technologies comme un « nouveau pouvoir statistique »<sup>232</sup> ? De la même manière que les statistiques ont été l'outil du gouvernement au pouvoir, les « big data » aident à gouverner, sont des outils qui asseyent le pouvoir, le légitiment, témoignent de l'efficacité d'un mode de management ou de gouvernement à travers la preuve des données, comme l'explique Alain Desrosières<sup>233</sup> à propos des usages de l'outil statistique par l'Etat « néolibéral » à partir des années 1970. Il note aussi « l'ambivalence des usages qui en sont faits, tantôt outil scientifique d'administration de la preuve ou outil politique d'action gouvernementale<sup>234</sup> ». Ainsi nous pourrions analyser ces « technologies big data » comme un objet ambivalent, mobilisé dans différents contextes et lieux de pouvoir, tour à tour « outil scientifique d'administration » ou objet politique. Alain Desrosières précise la nature de ce gouvernement par la statistique, qu'on peut aujourd'hui appliquer à la

---

<sup>229</sup> Cf. Annexe 2. Entretien avec Novagen

<sup>230</sup> CARDON Dominique, *À quoi rêvent les algorithmes ? : Nos vies à l'heure des big data*, Paris, Seuil, La République des idées, 2015, 105 p. (Quatrième de couverture)

<sup>231</sup> Ibid.

<sup>232</sup> ROUVROY Antoinette, BERNS Thomas, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques », *Multitudes*, 2010, n° 40 (pp. 88-103)

<sup>233</sup> SCHLEGEL Vianney, « Alain Desrosières, Prouver et gouverner. Une analyse politique des statistiques publiques », *La Découverte*, 2014, 264 p. », *Revue Projet*, 2014, n° 342 (pp. 90-91)

<sup>234</sup> Ibid.

problématique des « grosses données » : « La « mise en nombre » du monde social effectuée par la science comme par l'action gouvernementale, au-delà d'une traduction chiffrée supposée objective, est aussi une « mise en forme » de ce monde, la création d'une réalité nouvelle (chapitre viii). Elle comporte une dimension éminemment politique, dans la mesure où elle offre une prise à l'action et soumet cette réalité nouvelle à l'intervention<sup>235</sup> ». C'est précisément cette « traduction chiffrée supposée objective » qui permet de légitimer les choix stratégiques d'une Direction Générale ou d'une Direction Marketing (par exemple couper le budget sur telle action marketing du fait d'un croisement de corpus de données qui démontre que telle campagne marketing a été inefficace). On voit ainsi ce « nouveau pouvoir statistique » qui s'exerce dans les entreprises, ces dispositifs technologiques étant brandis comme des médiateurs neutres et objectifs, détenant une vérité sur la situation économique de l'entreprise. C'est la théorie que défendent Antoinette Rouvroy et Thomas Berns dans « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps "numériques" ». Ils alertent sur ce mode de gouvernement « par les traces », qui se fait faussement passer pour « inoffensif » et dont tous ses sujets sont asservis sans s'en rendre compte : « La puissance d'un tel gouvernement réside donc dans son caractère essentiellement « inoffensif » : c'est aussi à ce titre qu'on doit comprendre que ses sujets sont gouvernés par les traces qu'ils laissent. Nulle naïveté ou paresse des sujets dans cette nouvelle forme de servitude volontaire<sup>236</sup>. On pense alors aux opérations techniques et mathématiques qui sont mises en avant pour justifier et légitimer les décisions au sein de l'organisation. Lorsqu'un algorithme donne à voir un certain résultat, la scientificité de son processus va être mise en avant par le dirigeant d'entreprise, se servant de cet outil technologique comme un outil de gouvernement.

On voit ainsi comment entre en circulation la « raison algorithmique » des acteurs de l'industrie du traitement des données. L'imaginaire de l'objet technique tout-puissant circule et se trouve mobilisé dans des modes de management et de gouvernement. La technicité de l'outil semble être utilisée pour légitimer les décisions des sphères managériales. Plus encore, on pourrait même l'analyser comme un objet qui a une portée religieuse voire messianique, dans la mesure où il cristallise des croyances et attentes fortes, et impose un certain dogme idéologique.

---

<sup>235</sup> Ibid.

<sup>236</sup> Op. cit., p. 93 (paragraphe « Gouverner par les traces »)

## 2. Une dimension messianique croissante : d'un ethos algorithmique à un ethos messianique ?

Plus qu'un outil de gouvernement ou un moyen pour réaliser une fin, nous pouvons aussi analyser les « technologies big data » à l'aune des croyances qu'elles cristallisent dans les organisations. L'expression « transformation numérique » qui circule au sein des entreprises<sup>237</sup>, témoigne d'une croyance des acteurs de l'entreprise en l'avènement d'un nouveau modèle, d'un nouveau mode de management, qui a vocation à changer les pratiques des acteurs économiques, politiques et sociaux. Il apparaît intéressant d'analyser les dispositifs technologiques de « big data » à l'aune de leur dimension messianique, puisqu'ils nourrissent un sentiment d'espoir, la croyance en une technologie qui viendrait rendre le monde meilleur, tel un messie, il promettrait de « sauver » l'entreprise et de venir à la rescousse de ses métiers.

### a. Un régime de croyance qui se manifeste par une symbolique messianique et religieuse : la « conversion numérique »

On peut examiner la circulation de la « raison numérique » au sein de l'entreprise comme la manifestation d'un nouveau régime de croyance autour de la « data », au sein duquel on pourrait identifier des « messies » de la « big data » – consultants-ingénieurs, et entités dirigeantes croyant dans le pouvoir de la technologie – qui évangéliseraient ces outils technologiques auprès d'un plus grand public, qui serait ici le reste des métiers dans l'entreprise. Ces « messies » promettraient une meilleure efficacité et performance des pratiques, de meilleurs modes de management, de collaboration et de gouvernance *via* cette « data ». Cette attente d'une technologie qui va venir transformer la réalité de l'entreprise est sous-jacente dans l'article de Stratégies, « Les instituts domptent la data » : « Des compétences qui vont demander à être encore enrichies devant les multiples innovations qui attendent le marché des études ». On retrouve dans ce discours un registre lexical lié à l'attente d'un changement en faveur d'un progrès – ici, un progrès concernant les instituts d'études et le marché. Ces « messies », sauveurs de l'organisation par la technologie et le numérique, peuvent s'incarner dans l'entreprise par le « Chief Data Evangelist » ou le « Chief Data Officer », à l'image d'Agnès Gilbert qui a vécu la création de ce poste éponyme à Ipsos en début d'année 2017. On retrouve une symbolique messianique dans son discours, quand elle

---

<sup>237</sup> Phénomène observé lors de nos expériences professionnelles précédentes, et notamment la plus récente, à Ipsos. L'expression se retrouve dans « Les instituts domptent la data » (« Le secteur a été chahuté par la transformation digitale ») et dans le magazine interne d'Ipsos (p. 9, « pour que les sociétés ne ratent pas le virage numérique », « réussir la transformation numérique » (cf. Annexe 7)

évoque sa « mission » au sein d'Ipsos : « Ma première mission est d'accompagner les LBL dans la problématisation de la transformation numérique qui a complètement bouleversé le rapport des gens au monde<sup>238</sup> ». On retrouve dans son discours la thématique de la conversion, qui dans l'imaginaire religieux, désigne le « fait d'adopter une nouvelle religion en abandonnant sa religion, ses convictions antérieures<sup>239</sup> ». Dans le *What's Up Mag*, le magazine interne d'Ipsos, Agnès Gilbert affirme que son poste a été conçu pour « pour que les sociétés ne ratent pas le virage numérique » pour « réussir la transformation numérique en digitalisant certaines fonctions traditionnellement off-line ». Dans cette même logique de convertir tous les acteurs de l'entreprise au numérique et aux technologies intégrant la « data », Bruno Schumtz insiste sur le fait d'« accélérer la diffusion des innovations méthodologiques et technologiques (*topic modelling*, mesure passive, géolocalisation...)»<sup>240</sup> », comme objectif stratégique pour l'année à venir. Cette affirmation montre une volonté d'imposer des valeurs – ici technologiques et d'innovation – dans une logique verticale, à la manière de ceux qui viennent prêcher la bonne parole pour convaincre de transformer leur modèle de société existant.

Cette thématique liée à la « conversion numérique » est travaillée par Milad Doueïhi dans *La Grande Conversion numérique*<sup>241</sup>. Il utilise le concept de « conversion » pour signifier le « changement radical de perspective » et la « remise en cause fondamentale et profonde de tous les éléments constitutifs d'une vision du monde »<sup>242</sup>. En comparant la révolution numérique et l'histoire des religions, il donne un éclairage sur ce concept de « conversion » : « Du point de vue du religieux, il exprime un changement radical de perspective ; une remise en cause fondamentale et profonde de tous les éléments constitutifs d'une vision du monde. Sous l'angle technique, il désigne l'opération par laquelle des éléments provenant de l'ancien monde, le monde analogique, sont transmutés, convertis dans le nouveau monde, numérique celui-là ». L'auteur précise : « La révolution numérique est donc pour l'essentiel un processus civilisateur contribuant à l'émergence d'une nouvelle culture ; la culture numérique ».<sup>243</sup> Le concept de « conversion » permet ici de lire les processus de circulation de la « raison algorithmique » à l'œuvre dans les organisations.

---

<sup>238</sup> Cf. Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017, « Dossier spécial inno : portraits. Agnès Gilbert, Chief Digital Officer, une ressource à partager ! »

<sup>239</sup> D'après le CNRTL, « Fait d'adopter une nouvelle religion en abandonnant sa religion, ses convictions antérieures. Conversion des protestants, des juifs au catholicisme. » Ou « Fait d'adopter des idées, des opinions, un système politique ou philosophique; adhésion à ces nouvelles idées. Conversion au communisme, au fascisme, à la démocratie, à la révolution; conversion démocratique, révolutionnaire. » ([www.cnrtl.fr/definition/conversion](http://www.cnrtl.fr/definition/conversion))

<sup>240</sup> Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017. Bruno Schumtz de Ipsos Connect, dans « Objectifs 2017 » : « continuer à dynamiser et structurer notre effort commercial », « accélérer la diffusion des innovations méthodologiques et technologiques (*topic modelling*, mesure passive, géolocalisation...) » (p. 3)

<sup>241</sup> Source consultée sur « La Grande Conversion numérique, par Milad Doueïhi », *Homo Numericus*, 7 avril 2008. URL : <http://www.homo-numericus.net/article282.html>

<sup>242</sup> Il y (dans le résumé de son ouvrage) évoque notamment les « changements de l'environnement numérique au cours des vingt dernières années » (ibid.)

<sup>243</sup> « La Grande Conversion numérique, par Milad Doueïhi », *Homo Numericus*, 7 avril 2008. URL : <http://www.homo-numericus.net/article282.html>

L'importance des discours sur la « data » et la « révolution numérique » dans les sphères stratégiques des entreprises peut s'analyser comme la mise en place d'une certaine vision du monde, qui serait en train de s'imposer à tous les acteurs de l'entreprise. Si l'on pousse la logique du régime de « conversion numérique, tous (organisations, décideurs, salariés) doivent être « convertis » à ce nouveau régime de croyance, sous peine de se voir exclus de la sphère économique, d'être considérés comme appartenant à un « ancien monde » qui n'aurait plus lieu d'être.

#### b. Une stratégie d'évangélisation *via* des discours idéologiques sur le « big data »

Plus encore qu'une croyance en la donnée qui fait autorité dans les sphères managériales, nous assisterions à une « évangélisation »<sup>244</sup> de cette idéologie au sein des organisations. L'utilisation récurrente et la circulation du terme « transformation numérique » témoignent de cette stratégie qui vise à convaincre un grand nombre d'acteurs organisationnels d'intégrer la technologie dans leurs pratiques professionnelles et dans leur manière de concevoir leurs manières de travailler. Il s'agit de convaincre le plus grand nombre d'acteurs possibles d'adopter ces technologies et d'en devenir des adeptes. À l'égard du métier de « Chief Digital Evangelist », Amaia Errecart fait remarquer que les discours d'escorte autour du numérique et des « technologies big data » comportent cette dimension messianique, avec une « prégnance du registre religieux qui se fait jour autour du numérique », qui sous-tend des imaginaires particuliers. L'utilisation du terme « évangéliste » « participe de la construction d'une véritable rhétorique messianique autour du numérique, laquelle relève d'une conception de la technique comme enchantement », ici l'enchantement de l'entreprise. L'intitulé du métier « Chief Digital Evangelist » contient la notion d'enchantement par la technologie (ici, « big data ») et suppose une vision de la technique comme objet extérieur à la société, à la manière d'un messie qui amène un monde meilleur. Un tel intitulé de poste revêt aussi une dimension salvatrice, renfermant l'imaginaire que la technique serait capable de venir soigner les maux des hommes. On peut penser à l'ONG Bayes Impact qui se donne une mission quasi salvatrice de venir « construire le futur du service public » en s'appuyant sur « la puissance de logiciels complexes et de la big data », à l'image de sa page d'accueil<sup>245</sup> et de sa page de recrutement qui invite à les rejoindre « pour changer le monde avec les données<sup>246</sup> ».

---

<sup>244</sup> Le CNRTL définit le terme « évangéliser » ainsi : « Annoncer l'Évangile de Jésus-Christ; convertir à la foi chrétienne. Synon. christianiser » ([www.cnrtl.fr/definition/évangéliser](http://www.cnrtl.fr/definition/évangéliser))

<sup>245</sup> [www.bayesimpact.org/fr](http://www.bayesimpact.org/fr)

<sup>246</sup> « Les logiciels et la big data sont d'incroyables leviers pour le progrès social. Ils peuvent permettre à un petit groupe de personnes d'avoir un impact social positif à grande échelle. Rejoignez notre petite mais super équipe pour construire de meilleurs services sociaux et servir des millions de personnes » ([www.bayesimpact.org/fr/careers](http://www.bayesimpact.org/fr/careers))

Comme le souligne Milad Doueïhi, la notion de « transformation digitale », semble être « la vulgate moderne<sup>247</sup> », autrement dit « une idéologie, un courant de pensée vulgarisés, à l'usage du plus grand nombre<sup>248</sup> ». Agnès Gilbert affirme d'ailleurs dans notre entretien qu'« il y a un enjeu de formation au digital et à la data ». Les entités dirigeantes semblent opérer une stratégie d'évangélisation autour des pratiques liées à la « data » et au « big data », participant à la diffusion d'une idéologie autour des données. L'ouvrage de Milad Doueïhi invite à conscientiser la manière dont « une technologie, essentiellement collective, modifie radicalement la vie de chacun, le lien social même, mobilisant nos repères les plus tangibles ». Elle constitue une idéologie, « Avec ses faiblesses, ses aveuglements, ses richesses et ses promesses<sup>249</sup> ». Nous passons ainsi de la croyance à la diffusion d'une idéologie, avec une stratégie qui consisterait à convaincre les acteurs économiques qu'ils doivent changer leur manière de fonctionner, de travailler, qu'ils doivent opérer un changement de paradigme en dédiant leurs pratiques aux technologies numériques – et liées à la « data ». La promotion de cette idéologie, de ce paradigme autour des technologies de la « data » (penser qu'on peut extraire de « l'intelligence » à partir de corpus de données massives et hétérogènes) montre un glissement entre les cabinets de conseil en technologie qui promeuvent leurs « technologies big data », et les acteurs stratégiques de l'entreprise qui à leurs tours ont intégrés ces discours et élaborent une stratégie pour diffuser ces technologies au sein de leur organisation. On assiste par exemple à la redéfinition des rôles et des métiers en interne, avec une valorisation et centralité des postes liées à la donnée. La création de poste de « Chief Data Officer » et le recrutement d'une « Data Scientist » à Ipsos témoigne de ce phénomène de centralité des valeurs liées à la donnée. Le magazine interne d'Ipsos *What's up Mag* consacre à ce titre un article-portrait à Charlotte Zaepfel<sup>250</sup>.

On voit ainsi comment opère la croyance en une « technologie big data » qui viendrait « enchanter » l'univers de l'entreprise et ses métiers, avec des rôles qui sont construits autour de l'imaginaire des données. Les entités dirigeantes, convaincues de ce rôle enchanteur de la technologie, délaissent-elles cependant une posture davantage critique à l'égard de celles-ci ?

### 3. Les limites des « technologies big data » : un manque de recul sur les pratiques

Justement, parce que ces technologies constituent un outil de gouvernement et sont mobilisées par les acteurs stratégiques de l'entreprise, il convient d'observer de manière critique

---

<sup>247</sup> Cf. Résumé de l'ouvrage *La Conversion numérique*

<sup>248</sup> Définition de « vulgate » selon le Larousse en ligne

<sup>249</sup> Cf. Résumé de l'ouvrage de Milad Doueïhi

<sup>250</sup> Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017, « Dossier spécial inno : portraits. Charlotte Zaepfel : Data Scientist »

celles-ci. Emmanuel Souchier, en parlant de la nécessité de porter un regard critique à l'égard de l'« écriture du Web », nous rappelle d'analyser « de manière critique, c'est-à-dire en pleine connaissance de cause, l'environnement technique de la pensée, en montrant que les instruments dont elle se saisit sont aussi ceux par lesquels elle est elle-même saisie, que les dispositifs matériels par lesquels elle se fabrique sont aussi des dispositifs qui la fabriquent, auxquels elle se plie ou avec lesquels elle se débat » (Chatelain et Jacob, 2004 : 15-16)<sup>251</sup>. C'est cette posture critique qu'il conviendrait d'adopter face aux « technologies big data », qui font partie d'un « environnement technique de la pensée », c'est-à-dire qu'elles sont un objet dont les acteurs économiques et politiques se saisissent, travaillent, réfléchissent avec, et qu'il convient dès lors d'en connaître les rouages.

#### a. Une « raison algorithmique » relativement peu mise à l'épreuve

On peut analyser le régime de croyance de la « raison numérique » à travers la notion de « simulacre » de Baudrillard. Cela nous invite ainsi à penser la société à la lumière de ses régimes de croyances, qui ont tendance à primer sur le réel et font qu'on n'interroge pas de manière réflexive ce modèle de pensée. Les acteurs économiques et sociaux seraient guidés, se fieraient à leurs représentations avant de se fier au réel – si réel il y a. Le simulacre est de l'ordre du virtuel, de la pensée, il souligne le pouvoir des représentations sociales et culturelles au sein d'une société<sup>252</sup>. L'intérêt de ce sujet est donc bien de déconstruire ces « simulacres » qui opèrent dans les sphères économiques, sociales et politiques.

Soulignons que la « raison algorithmique » est relativement peu mise à l'épreuve, comme le fait remarquer Antoinette Rouvroy. Elle parle de « l'opérationnalité sans épreuve du comportementalisme numérique »<sup>253</sup>. Les outils de traitement des données ne seraient pas suffisamment mis à l'épreuve, et nous aurions tendance à les considérer comme un savoir acquis, naturel, qui va de soi (et l'aspect juridique s'avère délaissé). L'auteure déconstruit la prétention des « big data » à faire valeur de savoir puisque « le « monde réel, dans sa complexité, dans sa globalité, nul ne peut en rendre compte de manière satisfaisante<sup>254</sup> ». Le fait même de prétendre à comprendre le réel *via* ces technologies algorithmique devrait alerter notre raison critique.

---

<sup>251</sup> SOUCHIER Emmanuel, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, De la production à l'interprétation*, CNRS Editions, 2013 (pp. 213-229) (p. 231)

<sup>252</sup> BERGEN Véronique, « Saisies du simulacre chez Baudrillard et chez Deleuze », *Lignes*, 2010, n° 31, 192 p. URL : [www.cairn.info/revue-lignes-2010-1-page-37.htm](http://www.cairn.info/revue-lignes-2010-1-page-37.htm)

<sup>253</sup> ROUVROY Antoinette, « Pour une défense de l'éprouvante inopérationnalité du droit face à l'opérationnalité sans épreuve du comportementalisme numérique », *Dissensus*, Revue de philosophie politique de l'ULg, N°4, Avril 2011, pp. 127-143 (p. 127)

<sup>254</sup> *Ibid.*, p. 127

Antoinette Rouvroy ajoute que « Cette question de la mise à l'épreuve me semble particulièrement épineuse et urgente relativement à une « gouvernementalité algorithmique<sup>255</sup> », soulignant la nécessité de s'interroger sur la nature de ce savoir issu de ces outils, qui offrent un savoir-lire réducteur – réduit à ses propres cadres interprétatifs (le tracé du diagramme, le modèle de calcul appliqué, la tableur – or la réalité n'est pas réductible à une série de calculs ou de bugs informatiques. En effet, ce travail est une invitation à porter un regard critique et distancié sur ces technologies qui circulent dans le monde économique et social : « Portons donc notre regard au-delà de la fascination créée par la découverte de nouvelles potentialités technologiques – et de leurs implications plus ou moins directes -, afin de nous intéresser à la geste intellectuelle que nos contemporains élaborent pour comprendre le monde<sup>256</sup> ». Elle nous invite à porter un regard réflexif qui dépasse l'effet de fascination exercé par ces technologies. Les décideurs, les acteurs économiques et politiques doivent s'interroger sur ce régime de croyance (à tendance technocentrée), à questionner leurs pratiques pour mettre en perspective leurs décisions dans un cadre interprétatif plus large, considérant ces technologies à l'aune de leurs limites, de leur histoire.

Il apparaît nécessaire de produire un savoir réflexif sur ces pratiques liées aux « technologies big data ». Nous pouvons penser au rôle que jouent les chercheurs de produire cette réflexion, pour aider à mieux intégrer ces technologies avec le corps social<sup>257</sup>. Béatrice Galinon-Méléneq, dans *Traces numériques*, s'interroge sur ce phénomène de croyance qui occulte une rationalité critique à l'égard des technologies de traitement de « traces numériques », des algorithmes et de l'Intelligence Artificielle : « Au moment où la société contemporaine se voue au culte de la mémoire et de la mémorisation, apparaît une montée en visibilité des traces. De plus en plus nombreuses, sont-elles pour autant plus pertinentes pour cerner le réel ?<sup>258</sup> » Il convient en effet de s'interroger sur la pertinence de la notion de trace et de données. La valeur des « big data » semble résider dans le fait qu'elles sont massives, qui ont tendance à nous faire considérer les « technologies big data » comme pertinentes pour cerner le réel, pour appréhender et comprendre une réalité. Or cette masse de données est-elle vraiment pertinente pour produire une connaissance à partir d'un tel corpus, que certains qualifient de « chaos informationnel<sup>259</sup> » ? Dans « Gouvernementalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », Antoinette Rouvroy précise que ce qu'on appelle « donnée », qui

---

<sup>255</sup> Op. cit. p. 129

<sup>256</sup> SOUCHIER Emmanuël, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, de la production à l'interprétation* (p. 231)

<sup>257</sup> Cf. société et technique ont toujours été imbriquées selon Bruno Latour, nous devons cesser de penser que la société est dans une posture passive, réceptrice, face à une technologie qu'on a tendance à se figurer comme « émettrice, active », et venant de l'extérieur pour venir changer la société. Les deux sphères sont en réalité imbriquées l'une dans l'autre et s'analysent de manière plus nuancée et complexe.

<sup>258</sup> GALINON-MELENEC Béatrice, ZLITNI Sami, « L'Homme-trace, producteur de traces numériques », *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013, pp. 7-10 (p. 9)

<sup>259</sup> Op. cit. p. 10, chapitre « L'Homme-trace, producteur de traces numériques »

est le matériau de départ des « technologies big data » et au fondement du modèle économique des acteurs de l'industrie du traitement des données, est finalement peu questionné : « qu'il s'agisse de conserver la trace d'un achat, d'un déplacement, de l'usage d'un mot ou d'une langue, chaque élément est ramené à sa nature la plus brute, c'est-à-dire être tout à la fois abstrait du contexte dans lequel il est survenu et réduit à « de la donnée ». Une donnée n'est plus qu'un signal expurgé de toute signification propre – et c'est bien sûr à ce titre que nous tolérons de laisser ces traces, mais c'est aussi ce qui semble assurer leur prétention à la plus parfaite objectivité : aussi hétérogènes, aussi peu intentionnées, tellement matérielles et si peu subjectives, de telles données ne peuvent mentir !<sup>260</sup> ». L'analyse que fait Antoinette Rouvroy des « données » pointe du doigt la prétention d'objectivité des données, justement lié à leur caractère abstrait, qui fait qu'on peut les façonner selon une certaine logique d'interprétation. Une « donnée » n'existe pas en soi et ne prend son sens que par les acteurs qui la façonneront et y appliqueront leur subjectivité. C'est ce phénomène que nous invitons à « mettre à l'épreuve », à déconstruire, pour empêcher d'être aveuglé par la manière dont elles opèrent dans notre société.

#### b. Ne donner à voir qu'une traduction algorithmique du monde, une limite intrinsèque à la « raison algorithmique »

Il convient enfin de souligner que les acteurs de l'industrie de la donnée ne donnent à voir qu'une version algorithmique du monde qu'ils analysent, ayant aussi opéré un processus de sélection des données qu'ils traitent, analysent et donnent à voir. Les dispositifs technologiques de « big data » révèlent leurs limites dans la mesure où ils ne donnent à voir qu'une réalité inscrite dans un cadre interprétatif donné. Ils offrent une certaine version du réel, en donnant une vue synthétique, et procèdent de processus de simplification de l'information. Yves Jeanneret déconstruit cette prétention des acteurs de la donnée à faire valeur de connaissance en analysant les diagrammes, comme des mises en formes de la réalité qui imposent un savoir-lire réducteur et sont des « processus de construction de sens<sup>261</sup> » effectués par une raison algorithmique qui organise l'information selon une certaine vision du monde. Déjà, il convient de s'interroger sur ce que nous donnent à voir les outils de traitement des données numériques, comme Béatrice Galinon-Méléneq le rappelle dans *Traces numériques. De la production à l'interprétation* : « Que nous donnent finalement à voir ces traces numériques ? Nous sommes sensibilisés à leur présence. Mais comment se construit cette présence ? Qui la décide ? Quelle est la valeur de cette présence ? A notre sens, s'y joue non

---

<sup>260</sup> ROUVROY Antoinette, BERNS Thomas, « Gouvernamentalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », *Réseaux*, 2013, n° 177, pp. 163-196. (p. 169)

<sup>261</sup> JEANNERET Yves, « Les chimères cartographiques sur l'Internet. Panoplie représentationnelle de la 'traçabilité' sociale », *Traces numériques*, pp. 235-255

pas l'exactitude de la reproduction d'un réel pluridimensionnel, mais plutôt sa réduction à des traits identifiables par le plus grand nombre<sup>262</sup> ». En effet les « technologies big data », du fait qu'elles analysent des « traces numériques » massives et variées dont la production semble infinie, sont promues comme capables de donner à voir un monde dans sa globalité, complexe. Or l'interprétation du réel qu'elles en font n'est pas « pluridimensionnel », mais réduit au cadre d'interprétation de ces mêmes technologies, réduit au traitement informatique de ces « données » numériques, réduit au fonctionnement des algorithmes génératifs et analytiques qui sont appliqués, et réduit à une représentation, un format qui soit intelligible auprès du plus large public possible. Il convient aussi de souligner les processus de sélection qui sont à l'œuvre par les acteurs du traitement des « traces numériques », qui exercent leur propre subjectivité et fixent leurs propres cadres interprétatifs : « Un processus commun de sélectivité émerge : tout n'est pas pris en compte et traité par l'informatique. Tout n'est pas perçu par le lecteur et l'interprète des traces. De chaque côté, des tris s'opèrent qui produisent des effets de visibilité différenciés, allant de l'invisible à l'impression d'évidence existentielle et interprétative<sup>263</sup> ». Le processus de traitement des « big data » est une activité interprétative sélective, qui opère des tris, donne à voir certaines informations plutôt que d'autres, et opère des catégorisations des éléments du monde qu'elle analyse. Béatrice Galinon-Méléneq souligne que ce qui pose problème, c'est le fait que les résultats que donnent à voir ces acteurs de l'industrie de la donnée sont présentés comme révélateurs d'une « intelligibilité du monde ». Or rappelons que « l'interprétation et le jugement leurs sont indexés. Les catégorisations qui soutiennent souvent l'intelligibilité du monde ne sont que des commodités cognitives et culturelles pour parler le monde. Elles ne sont pas le monde<sup>264</sup> ». L'auteur dénonce cette prétention à traduire le monde à l'aide du traitement des « données » numériques, qui sont promues comme neutres et érigées en valeur de vérité, mais dont les rouages de leur fonctionnement complet sont occultés. En effet ces acteurs de l'industrie de la donnée ne vont pas se risquer à admettre les limites de leur activité d'interprétation propre à leur « raison algorithmique », afin de garantir la survie de leur activité commerciale. Comme le rappelle Emmanuël Souchier, « le processus fonctionne selon une boucle de croyance, car même de mauvaise foi – et à défaut de cynisme -, je dois croire en mon activité, puisqu'elle me fait vivre. Quant à la communauté, elle doit aussi participer de cette croyance afin que l'activité perdure socialement<sup>265</sup> ».

---

<sup>262</sup> GALINON-MELENEC Béatrice, ZLITNI Sami, « L'Homme-trace, producteur de traces numériques », *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013 (p. 8)

<sup>263</sup> Op. cit., p. 10

<sup>264</sup> Ibid.

<sup>265</sup> SOUCHIER Emmanuël, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, de la production à l'interprétation* (p. 225)

Il convient de s'interroger sur le type de connaissance que nous offrent ces acteurs de l'industrie du traitement des données numériques : « Quel degré de la connaissance produit la mise en visibilité de la connexion des traces personnelles des individus ? Est-ce une révélation ? Ou une méconnaissance ? Pour faire émerger un sens pertinent dans le chaos informationnel résultant de l'accessibilité permanente à un nombre de plus en plus grand de données, faut-il avoir recours à une cartographie du Web ?<sup>266</sup> ». Il convient, pour les décideurs économiques et politiques, de garder en tête que ces technologies de traitement des « données » ne sont pas des dispositifs neutres mais déjà investis d'une subjectivité, d'une certaine activité interprétative qui opère des constructions de sens – effectués par une « raison algorithmique » qui organise l'information selon une certaine vision du monde, souvent technocentrée. En effet, « quels biais instaure la « traduction » de propriétés mathématiques en objets visuels, produits à partir de données elles-mêmes supposés rendre compte du « social » en ce qu'elles sont notamment qualifiées de « traces d'usage » ?<sup>267</sup> ». Il convient effet de s'interroger sur le biais interprétatif que suppose l'activité de traduction des « traces numériques », qu'on postule être une expression du « social », mais dont on omet de préciser qu'elles ont été successivement traduites en propriétés mathématiques puis en objets visuels.

Enfin, Antoinette Rouvroy souligne avec justesse ce mode de « gouvernementalité » exercé par les « données » et les algorithmes, constatant une multiplication des usages en entreprise des pratiques de « data-mining » et de « profilage » - surtout pour le secteur du commerce : « les sujets, « individualisés », se prêtent entièrement à leur propre gouvernement statistique, par les traces qu'ils laissent, par la répétition et/ou le fléchissement de ces traces, sans qu'aucune instance extérieure de surveillance active, aucune règle, aucune médiation – sinon une intermédiation purement technologique – ne soient désormais plus nécessaires. C'est à ce titre aussi qu'on peut considérer que ce gouvernement statistique maintient et continue de supposer la réflexivité de ce qui est gouverné, alors même que cette réflexivité parvient à se passer d'une véritable intentionnalité qui réclamerait un questionnement sur les fins<sup>268</sup> ». Etant donné qu'on délègue le pouvoir d'analyser et de choisir à ces technologies, on ne prend plus la peine de les questionner la manière dont elles gouvernent. On postule que les « traces » numériques sont capables de gouverner car elles s'autorégulent, suggérant que nous n'avons pas besoin d'une instance extérieure pour les surveiller. Le gouvernement par les données aurait une réflexivité propre à son système (elles seraient neutres, objectives). Pensant ainsi que les « données » s'autorégulent d'elles-mêmes et sont réflexives, c'est ainsi faire l'écueil de ne pas les considérer de manière réflexive. Antoinette Rouvroy dénonce cette

---

<sup>266</sup> Op. cit., p10

<sup>267</sup> Op. cit., p. 225

<sup>268</sup> ROUVROY Antoinette, BERNS Thomas, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques », *Multitudes*, 2010, n° 40, pp. 88-103 (p. 97)

vision technocentrée du monde, qui serait, grâce aux « technologies de big data », désormais réductible à des bases de données. Elle souligne les limites épistémologiques de ce « comportementalisme numérique », cette rationalité algorithmique, qui « se fonde sur la pure observation des données, complètement coupée du monde ».

Nous avons vu que se construit un discours stratégique qui valorise ces technologies comme outils pour améliorer la performance de l'entreprise et de ses acteurs. Il apparaît que les entités dirigeantes se servent des dispositifs technologiques liés à la « data » pour asseoir leur parole, qui leur conférerait un certain pouvoir rhétorique face aux autres acteurs de l'organisation, du fait qu'on ne questionne pas les résultats justifiés des « technologies big data ». En cela ils peuvent s'analyser comme des outils de pouvoir et de gouvernement, dans la mesure où ils confèrent une certaine légitimité à la parole du dirigeant.

Nous avons aussi examiné la circulation de la « raison algorithmique », qui passe entre les mains de différents acteurs : des cabinets de consulting en technologie aux décideurs économiques, lesquels s'emparent de ces outils et des éléments de langage associés, afin de justifier leurs décisions (cela peut être un plan stratégique, une décision pour les métiers dans le département marketing...). Se trouvant intégrés à l'écosystème de l'entreprise, les technologies liées à la « data » et à la « big data » paraissent exercer une certaine influence sur la configuration des métiers dans les organisations, des pratiques professionnelles, et de la vision stratégique de l'action.

Nous pointons aussi du doigt un régime de croyance propre à cette « raison algorithmique » que l'on retrouve dans les sphères managériales, et qui a créé des « simulacres » opérants. Les acteurs agissent avec la conviction que ces technologies les aident à maîtriser les événements, à connaître voire à prédire ceux-ci. Nous avons mis en lumière les effets de croyance qui opèrent dans le monde économique et social, avec des rôles qui sont construits autour de l'imaginaire des « données » – telle que la place centrale donnée au rôle du « data-scientist » ou la nouvelle place accordée au « Chief Data Officer » dans les organisations.

# CONCLUSION

## Bilan des réflexions soulevées

Nous avons vu que les « technologies big data » sous-tendent une certaine rationalité et une vision du monde qui sont à l'œuvre – celle d'un monde interconnecté traversé d'énormes flux de « données », qui permettraient de comprendre la totalité des phénomènes qui nous entourent, d'appréhender leur complexité voire de prédire certains événements. Les consultants-ingénieurs spécialisés dans la valorisation des « données » exercent une « raison algorithmique » en cohérence avec la logique de ces dispositifs « big data ». Ils élaborent leurs discours d'escorte en s'appuyant sur des procédés discursifs qui s'ancrent dans des imaginaires numériques forts, dans une vision technocentrée du monde, dans une « rationalité occidentale » et une « raison computationnelle », dans un système de croyances qui témoigne d'une grande confiance aux valeurs de la technoscience.

Nous avons aussi vu que ces technologies de traitement des « big data » se posent comme des outils de révélation de phénomènes sociaux et symboliques, avec la prétention de nous donner accès à un savoir social qui préexisterait, et serait rendu visible par ces outils. Nous avons vu les prétentions que la technologie condense ainsi que son pouvoir de recommandation face aux entités dirigeantes : prétendre ordonner un « chaos informationnel », en interpréter le monde à la manière des « barû » sumériens, se positionner comme des « savants » qui prétendent faire valeur de connaissance et d'aide à la décision.

Nous avons aussi analysé l'opérativité des discours des consultants-ingénieurs en « big data » et comment ils font autorité auprès des entités dirigeantes. Ils entraînent une circulation de leurs discours dans les organisations, discours technocentrés qui font l'objet de processus d'appropriation dans les organisations. À la manière de la statistique, les « technologies big data » sont utilisées comme outils de gouvernement pour justifier de certaines décisions stratégiques, justifier un plan d'action, ou un certain discours sur la performance. Nous constatons l'élaboration d'un discours sur la performance de l'entreprise qui prend appui sur cette « raison algorithmique », technocentrée et promue par les cabinets de conseil en technologie – par les acteurs de l'industrie de traitement des données. Nous avons analysé comment la « raison algorithmique » est opérante dans l'univers gestionnaire de l'entreprise. Les entités dirigeantes, semblent souscrire à ce système de croyances, supposant qu'en mettant en place des « architectures big data », ils vont pouvoir mieux gérer leurs chaînes de production, leurs procédures en interne, agir de manière plus efficace ou encore mieux connaître leurs clients ou un type de population. Nous avons mis en évidence

l'influence de ces discours technocentrés sur les « technologies big data », et l'opérativité d'un système de croyance qui croit en la techno-science comme salvateur de l'organisation. Outre cette dimension politique, ces technologies revêtent une dimension presque messianique avec des personnes chargées de leur évangélisation, et de former à cette « révolution numérique ».

À travers ce travail de recherche, nous voyons comment ces acteurs de la « donnée » parviennent à construire un *ethos* algorithmique, c'est-à-dire fondé sur la rationalité du nombre et du calcul, sur une « raison computationnelle<sup>269</sup> » : la promotion d'une certaine vision du monde, technocentrée, qui attribue un certain pouvoir à la machine, aux ordinateurs, aux algorithmes. Nous avons constaté les limites d'une telle vision technocentrée qui ne remet pas en perspective les algorithmes, présuppose leur infaillibilité et questionne peu leur légitimité. Nous invitons, par ce travail de recherche, à questionner ces « technologies big data », les discours d'escorte qui les entourent, et à prendre conscience du fonctionnement de l'industrie du traitement des données. Gardons à l'esprit que ces acteurs de la donnée procèdent d'une vision du monde particulière et ont tendance à appréhender le réel à travers le prisme de la technique et de la science, de voir de monde comme une série de « bugs » ou de calculs – or le monde est en ensemble complexe, difficilement saisissable.

Nous rappellerons cette comparaison d'Emmanuel Souchier entre l'ingénieur contemporain et le devin de la Chine ancienne : « De Sumer ou de la Chine ancienne à Internet, la pratique de « lecture - écriture » qui préside à « l'univers des dieux » ou à celui du « numérique » est passée d'un monde de croyances religieuses ou de superstitions à un monde de pratiques marchandes nourries par le messianisme scientifique et technique qui préside aux destinées et à l'introduction des « nouvelles technologies » dans notre société<sup>270</sup> ». Il convient de penser ces nouvelles technologies à l'aune de notre histoire anthropologique, et de la place que notre société accorde à la science et à la technique, en les observant avec recul.

### Vérification des hypothèses

L'intuition de départ postulant que les cabinets de conseil en « big data » construisent un *ethos* algorithmique qui prend appui sur des imaginaires technoscientifiques se confirme, puisque l'étude révèle que les consultants travaillent leurs discours – sur leurs sites web et leurs supports de présentation – pour légitimer leur offre de services auprès des entités dirigeantes et auprès de leurs prospects. Ces consultants utilisent des éléments de langage en cohérence avec cet *ethos* algorithmique, ce qui les positionne comme experts dans le secteur de l'industrie du traitement des

---

<sup>269</sup> Cf. Travaux de Bruno Bachimont évoqués en cours de *Prospective digitale* avec Stéphan-Éloïse Gras

<sup>270</sup> Cf. GALINON-MENELEC Béatrice, *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013, 290 p. (p. 215)

données. Comme l'a évoqué Thomas Grignon dans son cours de « Cartographie du web »<sup>271</sup>, l'expert n'est pas seulement celui qui a une compétence et savoir-faire, mais il est aussi celui qui est reconnu comme ayant une compétence. Afin que sa parole soit écoutée, il doit travailler sa légitimité, comme s'il était dans un jeu de représentation de sa propre expertise. En effet, il n'y a pas d'expertise sans une reconnaissance de cette connaissance experte. C'est cette crédibilité qui se travaille au quotidien ; ainsi le spectacle de la technologie qui fait tourner les algorithmes est utilisé comme appuyant l'effet de fascination des cabinets de conseil auprès de leurs clients.

Notre analyse du corpus confirme également que les consultants-ingénieurs en technologie consolident leur posture d'expert par le fait qu'ils recommandent des actions spécifiques à leurs clients, par le fait qu'ils portent à leur connaissance un savoir qu'ils présentent comme un révélateur, *via* les « technologies big data ». Pour nuancer cette hypothèse, nous pourrions dire que plus qu'un expert, le consultant-ingénieur en données se positionne comme un médiateur, souvent unique, entre des « données » déstructurées et un réel qu'il est possible d'appréhender, voire de prédire. Nous avons même vu qu'il existe un rapport de pouvoir entre le consultant et les décideurs économiques qui font appel à lui, puisqu'il semble le seul à détenir ce savoir-faire technologique, cette maîtrise des « technologies big data », cela exerçant un certain effet de fascination.

Enfin, l'hypothèse selon laquelle les consultants en technologie font circuler leurs discours dans les entreprises, s'est confirmée, en montrant que les discours et les pratiques des acteurs économiques s'imprègnent de cette « raison algorithmique » et la réutilisent dans leurs discours. Nous avons vu qu'elle s'exerçait de l'intérieur, avec un effet de croyance fort de la part des dirigeants stratégiques. Elle implique en effet une rationalité qui prône une performance qu'on peut atteindre par le biais de ces technologies, qui proposent de mesurer et d'évaluer des phénomènes sociaux et économiques. Ainsi a-t-on vu que la donnée chiffrée n'est pas seulement un outil ni un moyen pour réaliser une fin, mais relève d'une croyance qui fait autorité.

### Principales difficultés méthodologiques rencontrées

La principale difficulté a d'abord été de cadrer notre sujet de recherche, le terme « big data » étant riche, à la fois technique et investi de discours médiatiques, économiques, politiques, en constante redéfinition par différents acteurs. L'étape de lectures flottantes s'est avérée longue, et nous avons dû opérer une sélection des articles et ouvrages les plus pertinents pour ce sujet.

La définition de notre corpus d'analyse a elle aussi été fastidieuse. Souhaitant au départ analyser un grand nombre de start-ups spécialisées dans le « big data », nous avons dû réviser le

---

<sup>271</sup> Cette année dans le Master Communication, Médias et Numérique

cadrage de ce corpus pour l'adapter aux contraintes temporelles et d'accès au terrain que nous avons. Ainsi nous avons préféré réduire l'analyse à deux cabinets de conseil, pour en produire une analyse plus fine.

Nous avons aussi fait face à une richesse documentaire et de pistes de réflexion, et avons dû réduire l'analyse à l'angle de recherche choisi. Nous avons dû opérer des choix et ne pas céder à la tentation de l'exhaustivité.

### Perspectives de prolongement et ouvertures

En explorant ces différents axes de réflexion, il serait par ailleurs intéressant de s'interroger sur la question de l'éthique liée au traitement des « traces » numériques que laissent les individus sur le Web. La question d'établir une éthique voire une jurisprudence à l'égard de ces nouvelles technologies se pose, dans la mesure où les acteurs de l'industrie de la donnée exploitent ces « traces » qui ont été laissées de manière non intentionnelle par les individus.

Nous pourrions aussi élargir l'analyse à la question de la place de l'automatisation dans nos sociétés contemporaines, de ce que nous concédons de déléguer à la technologie et aux machines. Cette question des actes de « délégation », travaillée par Bruno Latour, serait intéressante pour examiner quel rapport nos sociétés entretiennent avec l'intelligence artificielle et les algorithmes. Quels processus automatisés ou en voie d'automatisation sont davantage acceptés et quels autres sont rejetés ? Sur quels critères ? Cette question de l'automatisation dans nos sociétés serait un angle d'analyse intéressant pour étudier le rapport homme-machine, les imaginaires autour des technologies ou encore le modèle de la confiance de nos sociétés. Face à une tendance à la robotisation de nos sociétés, comment se structurent alors les discours des concepteurs des robots et de l'intelligence artificielle ? Quels discours élaborent-ils pour légitimer leurs technologies ?

# BIBLIOGRAPHIE

## OUVRAGES

ALSAFAR Ali. *Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012. Linguistique*. Université Paul Valéry - Montpellier III, 2014. URL : <https://tel.archives-ouvertes.fr/tel-01146590>

BABINET Gilles, *Big Data, penser l'homme et le monde autrement*, Le Passeur éditeur, cop. 2015, 1 vol., 248 p.

CANDEL Etienne, « Textualiser les interfaces. Opérativité et épistémologie d'une requalification », Volume 1, *Soutenance de HDR*, 7 décembre 2015, 337 p.

CARDON Dominique, *À quoi rêvent les algorithmes : Nos vies à l'heure des big data ?*, Paris, Seuil, La République des idées, 2015, 105 p. (Quatrième de couverture)

DELORT Pierre, *Le Big Data*. Presses Universitaires de France, « Que sais-je ? », 2015, 128 p. URL : <http://www.cairn.info/le-big-data--9782130652113.htm>

GALINON-MENELEC Béatrice, *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013, 290 p.

KRIEG-PLANQUE Alice, *Analyser les discours institutionnels*, Paris, Armand Colin « ICOM », 2012 (chapitre 6. Polyphonie, dialogisme, interdiscours : l'ouverture du discours (pp.186-210))

MAYER-SCHÖNBERGER Viktor, CUKIER Kenneth, *Big data: la révolution des données est en marche*, R. Laffont, 2014, 296 p.

SADIN Éric, *La vie algorithmique, critique de la raison numérique*, Paris : Éditions l'Échappée. DL 2015, 278 p. (quatrième de couverture)

## ARTICLES SCIENTIFIQUES (SIC)

BERGEN Véronique, « Saisies du simulacre chez Baudrillard et chez Deleuze », *Lignes*, 2010, n° 31, 192 p. URL : [www.cairn.info/revue-lignes-2010-1-page-37.htm](http://www.cairn.info/revue-lignes-2010-1-page-37.htm)

DALL'ARMELLINA Luc, « Voir en relation. Habiter l'information ? », *Études de communication*, 41, 2013, mis en ligne le 01 décembre 2013, consulté le 14 mars 2017. URL : <http://edc.revues.org/5419>

DE CHANAY Hugues, « Patrick Charaudeau. — Le discours politique. Les masques du pouvoir. Paris : Vuibert, 256 p. », *Corpus*, 4, 2005, mis en ligne le 05 septembre 2006, consulté le 23 août 2017. URL : <http://corpus.revues.org/320>

EFFEUILLAGÉ, « La trivialité pensée par Yves Jeanneret ». URL : <http://effeuillage-la-revue.fr/portfolio-item/la-trivialite-pensee-par-yves-jeanneret/>

GALINON-MELENEC Béatrice, ZLITNI Sami, « L'Homme-trace, producteur de traces numériques », *Traces numériques. De la production à l'interprétation*, CNRS Editions, 2013 (pp. 7-10)

GANASCIA Jean-Gabriel, « Les big data dans les humanités », *Critique*, 2015, n°819-820, pp. 627-636

GRAS Stéphan-Eloïse, « Éthique computationnelle et matérialisme numérique : l'apport des Software Studies », *Critique* 2015/8 (n° 819-820), pp. 667-679

GRIGNON Thomas, « L'expertise communicationnelle au prisme de ses instruments. L'exemple de Google Analytics » dans « Le communicateur bousculé par le numérique : Quelles compétences à transmettre ? », COUTANT Alexandre, DOMENGET Jean-Claude (pp. 5-14)

JEANNERET Yves, « Les chimères cartographiques sur l'Internet. Panoplie représentationnelle de la 'traçabilité' sociale », *Traces numériques, De la production à l'interprétation*, CNRS Editions, 2013, pp. 235-255

KUÇURADI Ioanna, « Rationalité' » et « rationalités » dans le cadre du débat modernisme – post-modernisme », *Diogène*, Presses Universitaires de France, 2003, n° 202, 192 p. (pp. 12-20)

LATOUR Bruno, « Les Vues de l'esprit », *Culture et technique, Réseaux*, volume 5, n°27, 1987, Questions de méthode. pp. 79-96. DOI : 10.3406/reso.1987.1322

OLLAGNIER-BELDAME Magali, « Interagir dans un monde de plus en plus réflexif : processus cognitifs et traces numériques – mémoire, interprétation et rapport au temps », *Traces numériques : de la production à l'interprétation*, CNRS Editions, 2013 (pp. 129-142)

OLLION Étienne, BOELAERT Julien, « Au delà des big data. Les sciences sociales et la multiplication des données numériques », *Sociologie*, 2015, Vol. 6 (pp. 295-310). URL : <http://www.cairn.info/revue-sociologie-2015-3-page-295.htm>.

ROBERT Pascal, « Qu'est-ce qu'une technologie intellectuelle ? », *Communication et langages*, n°123, 2000, Dossier : Les médias en Russie (pp. 97-114) DOI : 10.3406/colan.2000.2992

ROUVROY Antoinette, BERNS Thomas, « Gouvernamentalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », *Réseaux*, 2013, n° 177, pp. 163-196. DOI 10.3917/res.177.0163

ROUVROY Antoinette, BERNS Thomas, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques », *Multitudes*, 2010, n° 40, pp. 88-103

ROUVROY Antoinette, « Pour une défense de l'éprouvante inopérationalité du droit face à l'opérationalité sans épreuve du comportementalisme numérique », *Dissensus*, Revue de philosophie politique de l'ULg, N°4, Avril 2011, pp. 127-149 URL : <http://popups.ulg.ac.be/2031-4981/index.php?id=963>

SCHLEGEL Vianney, « Alain Desrosières, Prouver et gouverner. Une analyse politique des statistiques publiques », *La Découverte*, 2014, 264 p. », *Revue Projet*, 2014, n° 342 (pp. 90-91)

SOUCHIER Emmanuël, « Voir le Web et deviner le monde. La 'cartographie' au risque de l'histoire de l'écriture », *Traces numériques, De la production à l'interprétation*, CNRS Editions, 2013 (pp. 213-229)

TÉTU Jean-François, « Yves Jeanneret, *Penser la trivialité*. Volume 1 : *La vie triviale des êtres culturels* », *Questions de communication*, 2009, mis en ligne le 17 janvier 2012, consulté le 19 novembre 2016. URL : <http://questionsdecommunication.revues.org/176>

WAHL Iodie, « Hartmut Rosa, Accélération. Une critique sociale du temps », *Lectures, Les comptes rendus*, 2010, mis en ligne le 16 avril 2010, consulté le 23 août 2017. URL : <http://lectures.revues.org/990>

## DIVERS

HOMO NUMERICUS, « La Grande Conversion numérique, par Milad Doueïhi », 7 avril 2008. URL : <http://www.homo-numericus.net/article282.html>

## SÉMINAIRE

Séminaire « Big data et anticipation : vers une gouvernementalité algorithmique ? », organisé le 16 avril 2015 à l'Espace éthique/IDF. Disponible sur YouTube, vidéo publiée le 22 mai 2015 par Espace Éthique. URL : <https://www.youtube.com/watch?v=RUEkEsB6FJA>

## ARTICLES DE PRESSE

ANDERSON Chris, « The End of Theory: the Data Deluge makes the Scientific method obsolete », *Wired*, 23 juin 2008

BERTIN Erik, « Data, l'onction du nombre », *Influencia*, 16 octobre 2016. URL : <http://www.influencia.net/fr/actualites/media-com,data-trends,data-onction-nombre,6740.html>

MADDYNESS « Le big data est-il le nouvel eldorado des entreprises européennes ? », 5 juin 2015. URL : <https://www.maddyness.com/innovation/2015/06/05/big-data-eldorado/>

MADDYNESS, « Infographie. Le big data et son influence sur le quotidien des consommateurs en un coup d'œil », 23 juillet 2015. URL : <https://www.maddyness.com/innovation/2015/07/23/big-data-vouchercloud/>

LE FIGARO, « Lectures, Le mythe du progrès », Rémi Soulié, 21 février 2011. URL : [www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php](http://www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php)

LE MONDE, « Hartmut Rosa : « Plus on économise le temps, plus on a la sensation d'en manquer », Anne-Sophie Novel, 1<sup>er</sup> avril 2016. URL : [www.lemonde.fr/m-perso/article/2016/04/01/hartmut-rosa-plus-on-economise-le-temps-plus-on-a-la-sensation-d-en-manquer\\_4893818\\_4497916.html](http://www.lemonde.fr/m-perso/article/2016/04/01/hartmut-rosa-plus-on-economise-le-temps-plus-on-a-la-sensation-d-en-manquer_4893818_4497916.html)

L'USINE DIGITALE, « Un Datalab chez Airbus pour diffuser le big data dans tous les métiers », Usine Digitale, 10 mars 2016. URL : [www.usine-digitale.fr/article/un-datalab-chez-airbus-pour-diffuser-le-big-data-dans-tous-les-metiers.N383312](http://www.usine-digitale.fr/article/un-datalab-chez-airbus-pour-diffuser-le-big-data-dans-tous-les-metiers.N383312)

#### PRINCIPAUX SITES INTERNET MOBILISÉS

[www.alliancy.fr](http://www.alliancy.fr)  
[www.cnrtl.fr](http://www.cnrtl.fr)  
[www.c-radar.com/lexique-de-lopen-data](http://www.c-radar.com/lexique-de-lopen-data)  
[www.journaldunet.com](http://www.journaldunet.com)  
[www.maddyness.com](http://www.maddyness.com)  
[www.la-philosophie.com/positivisme-auguste-comte](http://www.la-philosophie.com/positivisme-auguste-comte)

#### CORPUS D'ÉTUDE

[www.novagen.tech/conseil-it](http://www.novagen.tech/conseil-it)  
[www.affini-tech.com](http://www.affini-tech.com)

# ANNEXES

## Annexe 1. Corpus d'étude : Sites internet des cabinets de conseil en technologie (Novagen et Affini-Tech)

### NOVAGEN

- a. Site web Novagen – *Page d'accueil* (www.novagen.tech)
- b. Site web Novagen – *Data Factory* (www.novagen.tech/data-factory)
  - *Data Factory Présentation*
  - *Pourquoi une Data Factory ?*
  - *Ce qu'apporte le Data Lab*
  - *Notre Méthodologie*
- c. Site web Novagen – *Conseil Data* (www.novagen.tech/conseil-data)
- d. Site web Novagen – *Conseil IT* (www.novagen.tech/conseil-it)
  - *Spécialités des consultants & Engagement de résultats*
  - *Externalisation, nos domaines d'intervention* (développement applicatif, AMOA, Data Analysis, Data Science)

### AFFINI-TECH

- e. Site web Affini-Tech – *Page d'accueil* « Simplifions vos data » (www.affini-tech.com)
- f. Site web Affini-Tech – *Nos Services* « Imaginer, Expérimenter, Construire & Exécuter
- g. Site web Affini-Tech – *Projets*
- h. Site web Affini-Tech – *Technologies*
- i. Site web Affini-Tech – *Clients*

Annexe 2. Entretien avec Novagen, cabinet de conseil IT et Big Data

Annexe 3. Entretien avec Charlotte Zaepfel, « data-scientist » à Ipsos

Annexe 4. Entretien avec Agnès Gilbert, Chief Digital Officer à Ipsos

Annexe 5. Entretien avec Vincent Cottin, Développeur web PHP à Atsuké

Annexe 6. Magazine interne d'Ipsos – *What's Up Mag*, n° 1, avril 2017, « Dossier spécial inno : portraits. Agnès Gilbert, Chief Digital Officer, une ressource à partager ! » & « Charlotte Zapfael : Data Scientist »

Annexe 7. « Les instituts domptent la data », *Stratégies*, 13 avril 2017, n° 1900 (trouvé sur l'intranet d'Ipsos)

Annexe 8. Support de présentation de Novagen (slides Powerpoint en PDF)

## Annexe 1. Corpus d'étude – Sites internet des cabinets de conseil en technologie (Novagen et Affini-Tech)

[www.novagen.tech](http://www.novagen.tech), [www.affini-tech.com](http://www.affini-tech.com)

### NOVAGEN

#### a. Site web Novagen – Page d'accueil ([www.novagen.tech](http://www.novagen.tech))



CONSULTING ▾ DATA FACTORY CARRIÈRES ▾ ACTUALITÉS CONTACT

## NOTRE EXPERTISE DATA

Au service de création de Valeur Business pour les Entreprises

Machine Learning, IoT, Analyse sémantique ...

[En savoir plus](#)

# CABINET DE CONSEIL NOVAGEN



## UNE SOCIÉTÉ DE SERVICE

Un savoir faire éprouvé sur les métiers de l'IT et du Web, au sein des équipes MOE et MOA, Directions Informatiques et Marketing



## UN CABINET CONSEIL EN BIG DATA

Une expertise de l'écosystème Big Data au service des besoins Métiers de nos clients PME-PMI / ETI / Grands Groupes. Nous partageons nos expériences, nos usages et bonnes pratiques



## UNE DATA FACTORY

Un LAB interne pour développer et industrialiser nos savoir-faire Big Data. Un savoir faire opérationnel pour multiplier les Insight Business

## NOS SERVICES

CONSEIL IT



CONSEIL DATA



b. Site web Novagen – *Data Factory* (www.novagen.tech/data-factory)

*Présentation*



CONSULTING ▾ DATA FACTORY CARRIÈRES ▾ ACTUALITÉS CONTACT

# DATA FACTORY

Agir comme une fabrique de solutions Data innovantes  
en usages et en technologies Smart Data – Big Data

Promouvoir les bénéfices des Architectures Big Data  
auprès des PME / ETI et Grands Groupes

## Pourquoi une Data Factory ?

### POURQUOI UNE DATA FACTORY?

1. Les nouveaux usages exigent de préserver une démarche exploratoire et expérimentale

2. Cela permet d'industrialiser l'innovation en des solutions prêtes à l'emploi

3. Le croisement des innovations et des explorations est créateur de valeur

4. Nos équipes peuvent naturellement tester de nouvelles approches et les faire mûrir

### CE QU'APPORTE LE DATA LAB



Maîtriser sa complexité par l'Analyse de Données textuelles de données massives



Maîtriser son Marché par le traitement Analytique des Données du WEB



Des outils de Business Intelligence, temps réel et sur-mesure

### Ce qu'apporte le Data Lab



#### Maîtriser sa complexité par l'Analyse de Données textuelles de données massives

Extraire de l'Intelligence de vos bases de données internes, de votre GED, Historique Documentaire...

- Moteurs de recherches (articles, email, documentations)
- Analyse sémantique (analyse de corpus scientifiques, bases documentaires)



#### Maîtriser son Marché par le traitement Analytique des Données du WEB

Vous voulez découvrir de nouveaux usages en vous ouvrant à des données externes et prêtes à l'emploi ?

- Tableaux de bords de performance commerciale
- Analyse de campagnes de communication
- Cartographie d'acteurs économiques



## Des outils de Business Intelligence, temps réel et sur-mesure

Une meilleure approche dans l'exploitation de vos données internes

- Croisement données internes et 3rd Party
- Visualisation industrialisée et temps réel
- Optimisation de votre patrimoine Métier

## Notre Méthodologie

# NOTRE MÉTHODOLOGIE

### BRIEF CLIENT

#### Identification des usages et besoins client

Partage des possibilités offertes par la technologie BigData

### USE CASE

#### Choix des Use Cases

Présentation Prototypage  
ROI des Projets  
Go / NoGo du projet

### DÉVELOPPEMENT DES PROJETS

#### Identification des usages et besoins client

Partage des possibilités offertes par la technologie BigData

### DÉPLOIEMENT

#### Vous décidez de la livraison

- Plug & Play
- Cloud
- Dans vos serveurs
- Dans nos serveurs



### c. Site web Novagen – Conseil Data (www.novagen.tech/conseil-data)

# NOTRE OFFRE CONSEIL BIG DATA

Vous avez des idées Data plein la tête mais ne disposez pas des expertises pour les faire naître ?

[En savoir plus](#)

## NOS SERVICES

Nous identifions avec vous les étapes et jalons de votre stratégie Data et Big Data

Nous accompagnons nos clients PME-PMI / ETI et Grands Groupes dans l'élaboration de cette feuille de route



### FORMALISATION DE VOTRE STRATÉGIE DATA

Sélection et priorisation des use cases  
Création d'architectures applicatives orientées Big Data



### RÉALISATION DE VOS PROJETS BIG DATA

POCs et Projets QuickWin  
pour apporter la preuve de la pertinence de nouvelles solutions,  
adaptées à vos problématiques et usages



### ACCOMPAGNEMENT DE VOS ÉQUIPES

Vous faire profiter des dernières innovations  
à travers des  
« Formations axées mise en œuvre »  
et le partage de nos savoir-faire

### NOTRE APPROCHE DE CONSULTING IN SITU

Nos consultants Data Analyst/Data Scientist interviennent et développent les projets au sein de vos équipes

Nos consultants accompagnent vos équipes :

- Etudes en amont
- Développement des Projets
- Formation et Coaching

En mission chez nos client, nos consultants continuent de bénéficier de formations pour développer leur nouveaux savoir-faire

[En savoir plus](#)



L'Offre Big Data ne crée pas de valeur en tant que telle  
la technologie n'a de sens qu'au service des besoins métiers.



### NOTRE APPROCHE DATALAB

Nous définissons ensemble un ou des Use Case.

Nos équipes DataLAB développent vos projets en Mode Quick Win.

Nous prenons en charge toute la chaîne de développement jusqu'au déploiement final.

[En savoir plus](#)

## LES ÉTAPES DU MODE PROJET



### BRIEF

Identification  
des usages et  
besoins client



### USE CASE

Choix d'un Use  
Case



### DEV. PROJET

Dév. Agile du  
prototypage  
des solutions



### LIVRAISON

Livraison à la  
carte

- Interne
- Cloud



### ADRESSES

#### Ile de France

Les Collines de l'Arche - Immeuble  
Opéra E 76

### SITMAP

Conseil IT  
Conseil DATA  
DataFactory

### FICHES METIERS

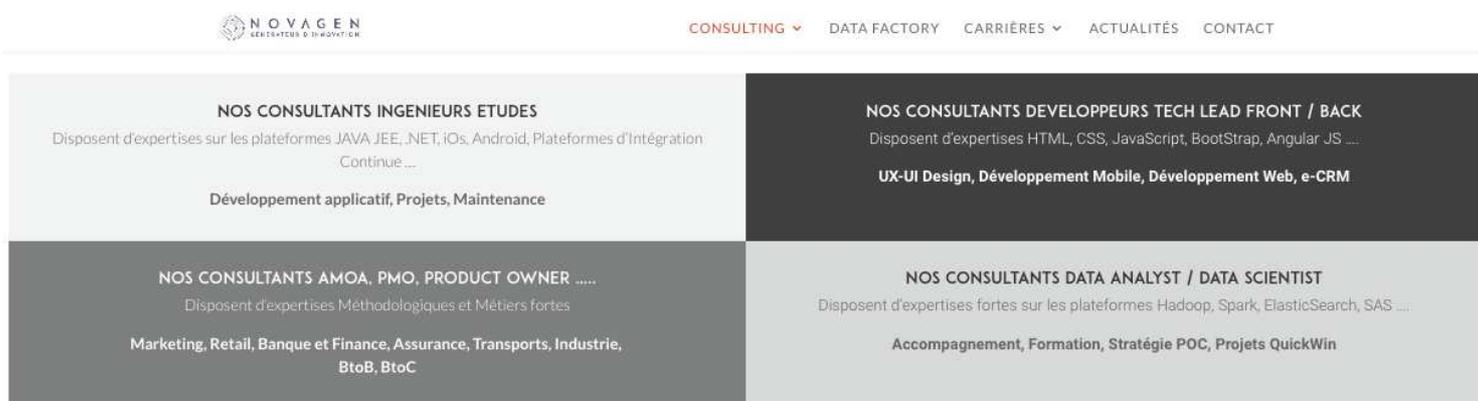
L'Analyste Confirmé  
Le CP MOE  
Le Consultant AMOA

### SUIVEZ NOUS



## d. Site web Novagen – Conseil IT (www.novagen.tech/conseil-it)

### Spécialités des consultants & Engagement de résultats



The screenshot shows the top navigation bar of the Novagen website with the logo on the left and menu items: CONSULTING, DATA FACTORY, CARRIÈRES, ACTUALITÉS, and CONTACT. Below the navigation bar, there are four main content blocks arranged in a 2x2 grid. The top-left block is titled 'NOS CONSULTANTS INGENIEURS ETUDES' and lists expertise in Java JEE, .NET, iOS, and Android. The top-right block is titled 'NOS CONSULTANTS DEVELOPPEURS TECH LEAD FRONT / BACK' and lists expertise in HTML, CSS, JavaScript, and Bootstrap. The bottom-left block is titled 'NOS CONSULTANTS AMOA, PMO, PRODUCT OWNER .....' and lists expertise in Marketing, Retail, and Finance. The bottom-right block is titled 'NOS CONSULTANTS DATA ANALYST / DATA SCIENTIST' and lists expertise in Hadoop, Spark, and ElasticSearch.



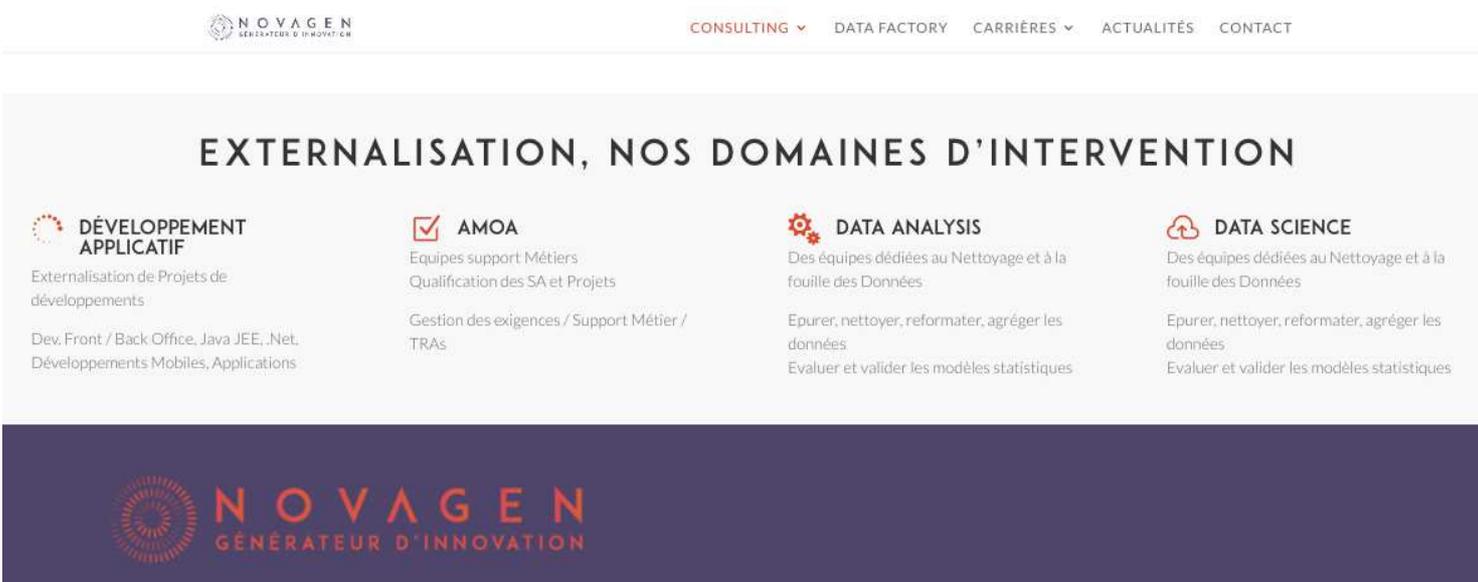
### ENGAGEMENT DE RÉSULTATS

**Connaissance approfondie des besoins clients et des offres de prestations à forte valeur ajoutée** du type Centre de Services, TMA, TRA, Forfaits...

**La Capacité à structurer, implémenter et piloter une offre de services complexe** sont des compétences que nous avons su déployer et proposer à nos clients grands Comptes.

Nous apprécions **l'agilité, la souplesse, le développement de solutions QuickWin**, à l'aide des dernières avancées technologiques

### Externalisation, nos domaines d'intervention (développement applicatif, AMOA, Data Analysis, Data Science)



The screenshot shows a page titled 'EXTERNALISATION, NOS DOMAINES D'INTERVENTION'. It features four columns, each with an icon and a title: 'DÉVELOPPEMENT APPLICATIF' (with a gear icon), 'AMOA' (with a checkmark icon), 'DATA ANALYSIS' (with a gear icon), and 'DATA SCIENCE' (with a cloud icon). Each column contains a brief description of the service. At the bottom of the page, there is a large dark blue banner with the Novagen logo and the text 'NOVAGEN GÉNÉRATEUR D'INNOVATION'.



[Home](#) [Services](#) [Projets](#) [Technologies](#) [Clients](#) [Blog](#) [Contact](#)

# Simplifions vos Data

De l'idée au service, nous accélérons votre business avec vos données !

Combinant Data-Science et expertise technologique, nous vous aidons à Imaginer, Expérimenter et Exécuter des services qui transforment votre business.

f. Site web Affini-Tech – *Nos Services* « Imaginer, Expérimenter, Construire & Exécuter »

## Nos Services



### Imaginer

Créer un business avec la donnée ?  
 Comment traiter mes données plus efficacement ?  
 Quels enjeux, Quelle roadmap ?



### Expérimenter

Tester et valider les intuitions.  
 Construire des POCs et pilotes.  
 Explorer les données, trouver de nouveaux gisements de valeur.



### Construire & Exécuter

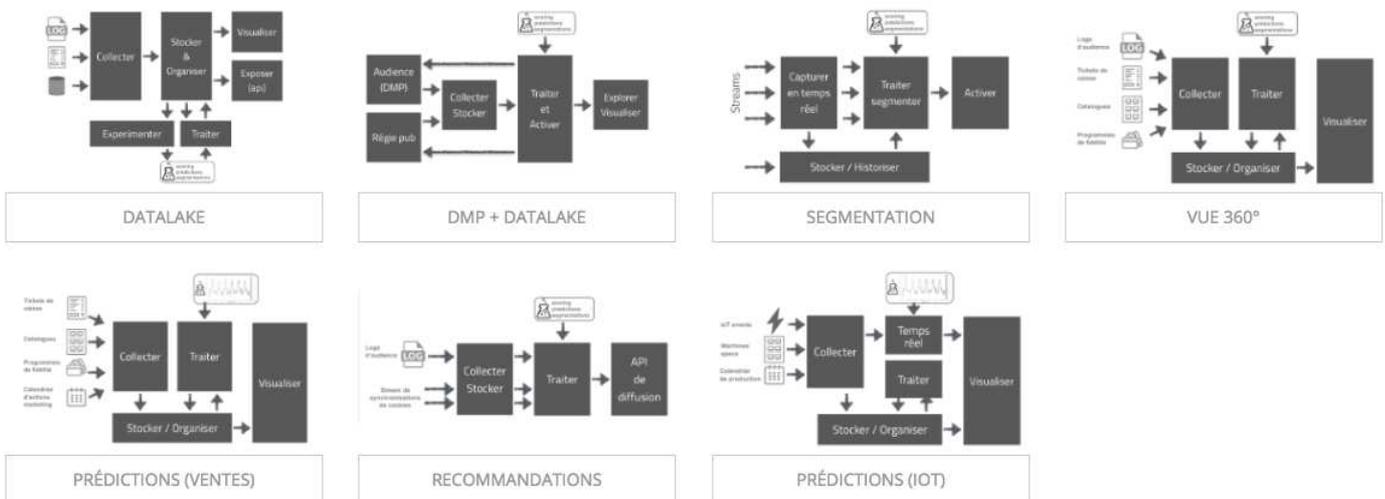
Construire un Datalake.  
 Exécuter des services pour vos business.  
 Vous accompagner au quotidien.

Nous mettons en œuvre des expertises technologiques sur des moyens modernes de traitement de la donnée. Le savoir-faire de nos Data-Engineer et de nos Data-scientists vous éclairera sur les leviers business, et l'intégration des plateformes nécessaires.

Nous nous attachons particulièrement à traiter ces sujets dans un temps très réduit, afin de vous permettre de les concrétiser rapidement.

g. Site web Affini-Tech – *Projets*

## Quelques projets

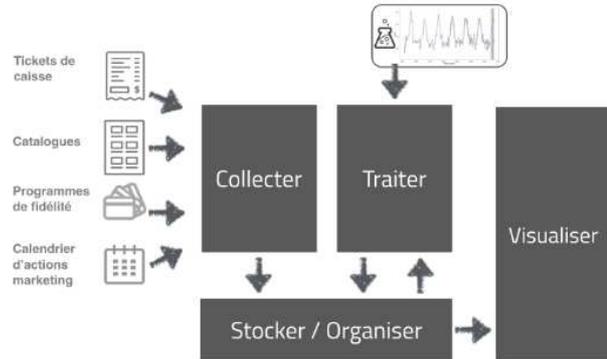


## Projets – « Prédiction des ventes »

### Prédictions des ventes

Prédire les ventes online et offline sur la base des historiques

- Utilisation de plusieurs modèles de Machine Learning
- Approches par séries temporelles
- Détections de Pics
- Amélioration continue des modèles
- Modèles d'attribution, pour intégrer les événements marketing



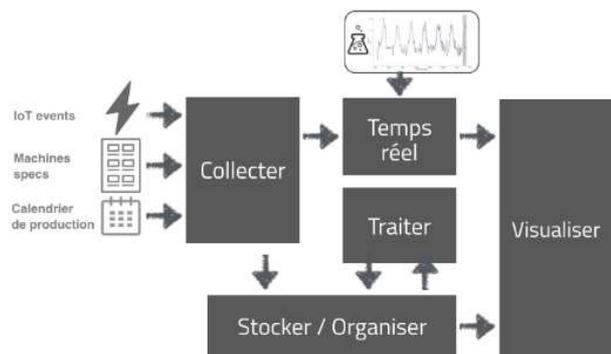
FERMER

## Projets – « Maintenance prédictive »

### Maintenance prédictive

Utilisation combinées de plusieurs techniques de Machine Learning, pour anticiper des événements industriels. Application d'algorithmes prédictifs sur des flux d'évènements en temps réel.

- Capturer de la data technique, commerciale, et des événements IIOT
- Approche statistique sur les cycles de fonctionnement.
- Détection d'anomalie avec des algorithmes de classification.
- Prédiction des défaillances et anticipation des plans de maintenances avec des moteurs de recommandations.
- Modélisation des équipements industriels pour de la simulation.
- Traitements batch et temps réel.



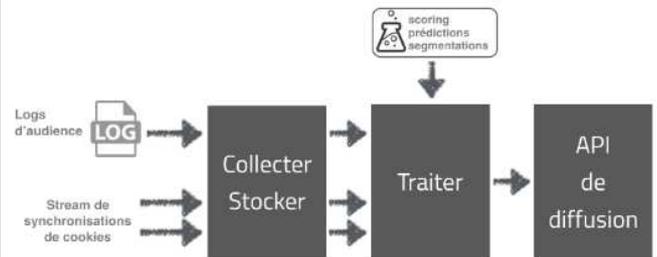
FERMER

## Projets – « Moteur de recommandation pour contenus éditoriaux »

### Moteur de recommandation pour des contenus éditoriaux

#### Un moteur de recommandation pour augmenter l'engagement des lecteurs (média en ligne)

- Collecte de données d'audience pour comprendre le lectorat
- Utilisation de métadonnées rédactionnelles
- Traitement en temps réel des streams de données
- Machine Learning
- Service distribuant pour chaque lecteur le meilleur contenu en temps réel



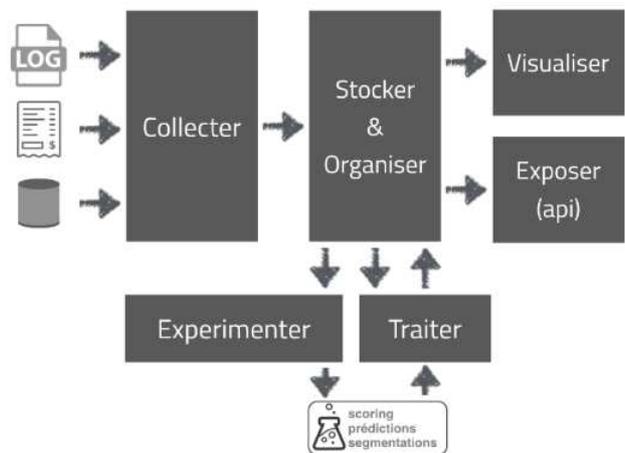
FERMER

## Projets – « Consolider un Data Lake »

### Consolider avec un Datalake

#### Construire une plateforme analytique universelle pour supprimer les silos dans l'entreprise

- Capturer des flux de données de toutes origines et formats
- Industrialiser les transformations de données pour les mettre en cohérence
- Intégrer des scénarios prédictifs
- Construire un Datalab pour expérimenter sur données et concevoir des modèles prédictifs
- Data-visualisation pour permettre des interactions riches
- Ouvrir l'écosystème avec des API



FERMER

## Projets – « Construire une vue 360° cross-canal »

### Construire une vue 360° cross-canal

Rapprocher les données de magasin et de site web pour identifier des parcours clients

- Disposer d'une vue globale sur chaque client et prospect.
- Développer un « SmartID »
- Comprendre les ventes en magasin, par l'étude des parcours web.
- Élaborer des stratégies cross-market
- Construire des expériences d'achat globales
- Prédire la prochaine action dans la relation client

FERMER

## Projets – « Segmenter des audiences en temps réel »

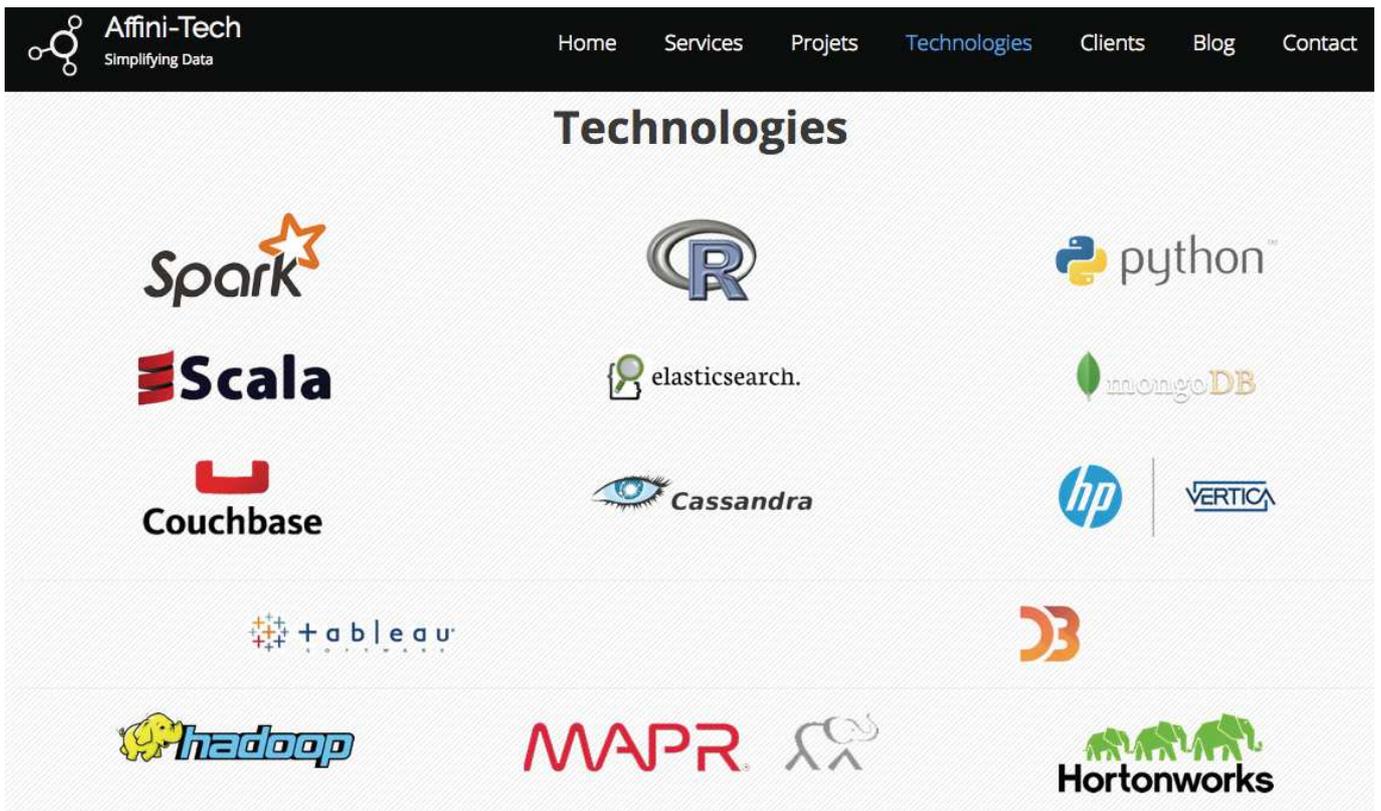
### Segmenter des audiences en temps réel

Étendre des segments pour une activation publicitaire plus efficace

- Collecter 500M d'items par mois en batch et en temps réel
- Augmenter la portée des segments définis (Algo / Machine Learning)
- Identifier des clusters d'intentionnistes
- Activation de la donnée sur des partenaires publicitaires
- Donnée segmentée et activable en quelques minutes après la collecte

FERMER

h. Site web Affini-Tech – *Technologies*



i. Site web Affini-Tech – *Clients*



## Annexe 2. Entretien avec Novagen, cabinet de conseil IT et Big Data

Retranscription de l'entretien avec Novagen, en présence de :

Olivier LEFIEVRE, CEO et Hubert STEFANI, Chief Innovation Officer

13 avril 2017

### **Pouvez-vous vous présenter en quelques mots, nous raconter ce que vous faites ?**

Nous sommes un cabinet de conseil qui travaille sur la valorisation des données. On est parti du constat, il y a quelques années que les technologies qui étaient mises en œuvre chez Google, Facebook, LinkedIn, entre autres, commençaient à s'ouvrir en mode open source et étaient finalement accessibles technologiquement. Ces technologies qui sont mises sur pied, améliorées, enrichies constamment pour leurs usages propres, pouvaient servir dans de multiples contextes.

Les contextes sont ceux auxquels ils font face tous les jours : quand on est Google, on aspire tout le web et on en fait une valeur ajoutée extraordinaire, c'est-à-dire la capacité à chercher en une fraction de secondes, à travers des millions de pages, l'information pertinente. Chez Facebook, c'est un meilleur profilage des personnes, la compréhension de leurs attentes, de leurs centres d'intérêts.

Et en fait ces grands groupes, il a fallu qu'ils inventent des nouvelles technologies. Ils inventent des nouvelles bases de données pour stocker des informations extrêmement différentes. Donc ils ont inventé tout un tas de choses : des bases de données qu'on appelle NoSQL qui permettent de stocker des informations d'une nature complètement différente de ce qui existait avant.

### **Qu'est-ce qui existait avant ?**

Avant on utilisait des bases de données structurées comme Oracle, mais ces technologies n'étaient pas capables de valoriser. On pouvait au mieux enregistrer l'information, mais on n'était pas capable d'offrir des moteurs de recherche ultra performants, ultra pointus, pertinents. Donc il y avait déjà une première barrière technologique à franchir, et eux l'ont franchie. Ils l'ont packagée et surtout, ils l'ont ouverte à la communauté.

On a donc eu l'opportunité de s'approprier ces technologies, ensuite il y a toute la valeur ajoutée : savoir prendre une problématique d'entreprise. Elle peut être grand public ou bien écouter des communautés sur les réseaux sociaux, écouter le web, et là on se rapproche de ce qui fait l'ADN de Google, Facebook, etc. Et puis c'est aussi des problématiques internes, par exemple mettre un moteur de recherche en interne sur sa propre banque d'informations, quand on est un grand groupe technologique, telco, industriel... On a des projets en pagaille, on a des soucis de communication, et là aussi c'est un champ formidable, c'est-à-dire se confronter à de la donnée peu structurée et rendre la communication interne plus efficace.

Nous, à titre d'exemple, on est capable de travailler en entreprise et de dire que les technologies, la big data, elles peuvent vous offrir l'avantage très compétitif en les appliquant à vos propres processus internes.

En fait quand on parle du big data, on l'associe assez souvent avec des gros volumes, donc ça c'est un aspect. Derrière le terme de big data, il y a un aspect qui est technologique, qui est en fait représenté par des nouvelles technologies qui sont accessibles, des nouvelles plateformes. Alors ça s'appelle Hadoop, Spark, Scala, avec des langages associés. Ce sont tous ces nouveaux outils, toutes ces nouvelles plateformes technologiques qui permettent de faire des choses qu'on ne pouvait pas faire avant. Donc il y a un aspect technique, le big data qu'on associe à de très gros volumes de données. Quand on a de très gros volumes de données, ces technologies permettent de faire tourner des algorithmes intelligents et de faire des traitements, là c'est de la *data-science*. On attaque des aspects mathématiques pour essayer de pouvoir mieux prédire un certain nombre de choses. Mais vous pouvez aussi très bien utiliser les technologies du big data pour faire du smart data. Le smart data ne nécessite pas de volumes colossaux mais c'est utiliser technologies pour pouvoir mieux travailler ces données internes. Je prends par exemple Ipsos, qui probablement stocke un nombre d'informations colossales depuis 10 ans, avec en interne dans les bases de données où il y a énormément d'informations, qui assez souvent correspondent à des enquêtes unitaires qui sont faites etc. Mais en mettant toutes ces informations dans un *data lake* et en utilisant les technologies du big data, il y a des croisements de données, des enrichissements de données, des usages et des apprentissages qui sont assez extraordinaires à faire.

**Concrètement quand on construit un « data lake », quelle est la part de l'expertise humaine, du conseil que vous donnez ? Qu'est-ce qui se passe concrètement quand on construit un « data lake » ?**

Quand on adopte la stratégie de mise en œuvre d'un data lake, et qu'on veut s'offrir une plateforme unifiée dans laquelle on va enregistrer toutes les données qui nous parviennent. Donc quand on est une entreprise, des données en interne, une base client, des applications, une base produits, une base référentielles, aussi des données qu'on va chercher à l'extérieur – un panéliste, d'autres informations qu'on irait aspirer sur le web et sur les réseaux sociaux.

Déjà, le fait que ça réponde à une première exigence qui est d'enregistrer beaucoup d'informations. Donc elle peut mixer à la fois des données structurées qui viennent de l'entreprise, elles ont passé déjà un certain nombre de tuyaux, de canaux qui font qu'elle est bien structurée et propre. Et puis d'autres informations qui font qu'elles le sont moins.

Donc déjà il y a une première mise en œuvre technique : on va faire cohabiter physiquement ces informations. Puis à l'intérieur de ça, on va avoir plusieurs niveaux de raffinement : il y a encore à la fois une approche vraiment stratégique (qu'est-ce qu'on va faire de ces données, comment on va les travailler, pour en faire une première version améliorée). Imaginons que vous vous confrontez à des forums, à des sites web, à des réseaux sociaux où vous voulez ressortir de l'information, peut-être la traduire automatiquement. Donc, on lance les premiers algorithmes. Imaginons que vous voulez reconnaître les noms d'entreprises, les noms de produits, de personnes... Vous allez faire jouer d'autres algorithmes. Vous voulez faire des indicateurs régulièrement sur les données qui vous arrivent brutes, pour détecter des tendances... Là encore, dans le data lake, vous allez partir d'une matière brute et vous allez en créer une nouvelle qui va être portée à la connaissance des décideurs. Et puis vous allez encore créer d'autres types d'informations, que l'on va réintroduire dans des applications (web, mobiles) pour être opérationnel. Imaginons, vous avez des équipes terrain, des points de vente, vous pouvez leur donner des informations, des alertes en temps réel. Là aussi, dans le data lake, vous aurez structuré vos données, vous aurez mis les processus de

transformation et les algorithmes de valorisation, ils vont multiplier après les use cases (les cas d'application).

Une solution technologique, c'est aussi toute une démarche de sélection de technologies, mise en œuvre, valorisation, accompagnement de la stratégie.

Je vais vous prendre un petit exemple pour un grand assureur français. Il avait rencontré plusieurs directions métiers. Et notamment la direction des ventes, elle travaillait sur sa base de données internes. Puis il a rencontré la direction de la fraude, et la direction des risques. Elle menait un travail pour aussi savoir quels étaient les risques qu'un client nous quitte. Par contre, elle travaillait sur sa base de données sur son périmètre. Mais on s'est rendu compte qu'en croisant les informations sur lesquelles travaillaient la direction des ventes, la direction des fraudes, des risques, du marketing, en réunissant toutes ces différentes sources d'information, on arrivait à gérer le risque beaucoup plus finement, on arrivait à générer du prédictif sur les ventes beaucoup plus facilement. En rapprochant des données dans un même lieu, qui concernent des métiers différents et des problématiques différentes, en les rassemblant, ça permet d'être beaucoup plus fin aussi bien sur l'analyse des ventes que sur la fraude, que sur les risques.

ça nécessite un changement organisationnel au sein des entreprises, parce qu'elles marchent souvent en silo (vous avez ceux qui sont spécialisés vente, spécialisés du marketing, etc.) et en fin de compte on regroupe toutes ces données dans un même endroit, ça oblige une réflexion beaucoup plus transversale, et avec une valeur ajoutée qui est assez extraordinaire.

### **Qu'est-ce que cette valeur ajoutée ? Qu'est-ce que le big data apporte au sein des entreprises (davantage de performance par exemple) ?**

On est là pour ajouter une valeur additionnelle, ça peut permettre de mieux faire certaines choses qu'on faisait avant, mais on est aussi là pour ajouter des opportunités. Ajouter une meilleure connaissance, ajouter une plus grande réactivité, et on s'ouvre à de tous nouveaux usages.

Quand le web est arrivé, on ne s'est pas dit 'ça va remplacer quelque chose', c'est venu ajouter une réactivité, un partage de l'information. Sur le big data, on est un peu sur le même ordre, entre deux magnitudes. Avec le big data, on va voir que ça va permettre de nouveaux usages, et on peut en parler avec les progrès de l'intelligence artificielle, tout ce qu'on aura acquis comme donnée et qu'on aura préparé au travers de stratégies data lake, stratégies de réconciliation des données, vont donner des avantages très compétitifs dans les années à venir. On va mieux connaître ses clients, on va être plus efficace, faire des prédictions sur les ventes beaucoup plus fines, on peut optimiser un métier industriel sur la maintenance des chaînes de production.

Nous on a beau être très au fait de ce qu'il se passe, des use case, quand on discute avec des collègues et des concurrents, chacun a des histoires incroyables à raconter autour de ce qu'on fait dans les données. On est vraiment dans un changement très profond de la façon dont on fonctionne.

### **Concernant le fonctionnement des algorithmes de valorisation des données, quelle est la place de l'automatisation et de la délégation à la machine ?**

Quand on met en place un algorithme de *machine-learning*, les grandes étapes qu'on fait, ce sont d'abord qu'on veut arriver à un objectif.

## **Comment mettez-vous en place les algorithmes ? Codez-vous des modèles de calculs spécifiques ?**

Ce qui change aujourd'hui c'est que avant, on codait une règle en dur. Imaginons que vous êtes une femme de plus de 50 ans et bien vous aurez tendance à consommer tel produit, on faisait vraiment des règles en dur. Et imaginons qu'on ait à vous pousser une publicité, on le faisait de manière très statique. Aujourd'hui on se connecte à plein d'informations, on les rassemble le plus possible, et on va utiliser les très grandes quantités d'informations pour les pousser dans un algorithme qui existe, et là l'expertise humaine c'est de bien choisir les données qu'on va faire rentrer. Il ne suffit pas de faire rentrer que l'on est une femme, tel âge, etc. mais il faut vraiment bien sélectionner les informations. On va aller chercher les informations de votre comportement (si vous avez tendance à utiliser telle application mobile, tendance à visiter tel site). Donc on peut rassembler beaucoup d'informations, mais encore faut-il l'appliquer, la structurer, la normaliser. Et une fois qu'on a fait tout ce travail qui est vraiment conséquent et qui nous mobilise beaucoup, là on peut le pousser dans un algorithme. Là aussi l'intervention humaine, c'est de bien choisir l'algorithme, selon qu'on veut reconnaître des objets dans une image, ou faire de la traduction, ou analyser un sentiment dans un texte... on ne va pas être sur les mêmes algorithmes. On ne va précisément pas être sur la même chaîne de traitement de l'information. Donc là aussi il y a une expertise humaine qui est liée à la compréhension du modèle mathématique qui se cache derrière les algorithmes. Et ensuite il faut pouvoir itérer, confronter les résultats qu'on obtient en confrontant les données à des algorithmes, avec les phénomènes observés : est-ce que la traduction est bonne ou mauvaise, et selon l'état, on va entrer encore dans des cycles, on va ajuster les paramètres d'entrée, ajuster notre stratégie de prévision, et à la fin aboutir à un résultat, celui qui est attendu.

Donc il y a beaucoup de choses qui sont de l'ordre de l'expertise, de l'intervention humaine. Une fois que ceci est au point, ça peut potentiellement remplacer un grand nombre d'actions humaines qui étaient auparavant mobilisées pour répondre à ce service.

Je vais vous prendre un tout petit exemple : si vous avez un historique des ventes sur dix ans, vous travaillez sur l'historique des ventes sur les cinq années les plus anciennes, vous faites tourner des algorithmes, et puis au bout de 1 ou 2 mois de travail, vous avez tuné un algorithme qui donne l'information juste à 99% sur les ventes de la sixième année, de la septième année, etc. Donc là vous avez trouvé un algorithme qui marche, puisqu'il est juste sur les années 6, 7, 8. On peut penser que avec cet algorithme, vous pouvez faire de la prédiction sur les années 11, 12, 13, 14, 15 qui arrivent demain.

Pour qu'on travaille sur des notions de data science et d'algorithmie, il faut une profondeur importante parce que plus il y a de données, plus les algorithmes peuvent être justes. Quand on tombe sur une combinaison d'algorithmes qui donnent des informations exactes sur le passé, on peut se projeter sur l'avenir. Donc à ce moment là, on va donner une vraie information à la direction des ventes et à la direction générale d'une entreprise, qui en fonction de paramètres qu'elle va faire évoluer, elle aura une vision juste du chiffre d'affaires qui peut être généré sur les 4 ou 5 années qui viennent.

**Si je comprends bien, quand vous construisez vos algorithmes, c'est un aller-retour permanent entre votre analyse et la justesse de l'algorithme ?**

ça marche de façon très itérative, très instrumentale. Si on a un risque d'erreur de 30%, les entreprises elles s'en fichent... Il faut arriver à 95, 96, à 98%, 99% et là ça devient intéressant, mais ça c'est un gros travail itératif. Pour du véhicule autonome il vaut mieux être à 100%.

### **Quels logiciels utilisez-vous (sont-ils en open source, est-ce vos logiciels propres...) ?**

On crée du logiciel en plus des technologies open source qu'on utilise. Il y a aujourd'hui un foisonnement extraordinaire de technologies, avec quelques incontournables (pour les citer, Hadoop, Scala, Spark) et quelques sélections audacieuses. Donc il y a déjà un premier savoir-faire à les sélectionner et imaginer comment elles vont collaborer ensemble, c'est tout un métier déjà. Et voir comment, au-dessus de cela, on va créer notre propre traitement spécial, notre sauce spéciale qui va faire qu'on va vraiment répondre à la problématique, qu'on va être pertinent, qu'on va être performant. A chaque fois qu'on pose un problème, il est multi tenants, il faut vraiment prendre en compte beaucoup de considérations, donc choisir quels logiciels ne répondent pas à toute la réponse. Ce sont des combinaisons, des sélections de savoir-faire et de développements spécifiques.

Ceux-là ce sont des profils qu'on appelle les « ingénieurs data ». Ce sont des ingénieurs informaticiens qui montent en compétences sur ces outils et plateformes big data. On différencie les ingénieurs qui s'occupent de faire cohabiter toutes ces briques technologiques, et puis vous avez les profils « data scientists » qui sont d'abord des mathématiciens. Parfois vous avez la perle rare qui a de très bonnes compétences mathématiques et technologiques.

### **Quelle est la différence entre une compétence technologique et une compétence mathématique ?**

Cela va être des études différentes. Vous avez beaucoup de data scientists qui ont fait de la recherche avec des masters en mathématique appliquée pure, les data scientist sont souvent des profils de ce type. Les data ingénieurs sortent des écoles d'ingénieurs et ont déjà appris beaucoup de choses sur les technologies du big data, et ont aussi pratiqué les algorithmes existants. Dans leur cursus ils ont aussi été formés aux mathématiques statistiques.

### **On lit aussi dans les revues et articles que le big data rompt avec la statistique traditionnelle. On ne calcule plus à partir d'une moyenne, d'une corrélation statistique, mais on identifie les signaux faibles dans une base de données. Qu'en pensez-vous ?**

Cela ne met pas au rebus toutes les approches statistiques classiques et cette approche va se poursuivre dans les entreprises. Elle va être enrichie avec les nouvelles technologies, elle permettra d'être plus fin dans un grand nombre de cas, elle va permettre d'intégrer de nouvelles dimensions de données dans les « pipelines », les chaînes de transformation habituelles ou techniques. La culture statistique va toujours être intéressante et va permettre d'étendre son champ d'application, et l'approche statistique va être utile aussi sur les nouvelles approches. Parce que dans les algorithmes, pour bien les alimenter, car ils ne sont pas structurés comme de simples moyennes ou écart type, mais elles sont utiles pour bien choisir et sélectionner les données qui vont entrer et ensuite valider les observations qui sont faites.

Pour le coup il y a des approches vraiment disruptives, en rupture avec ce qui se faisait auparavant, par les réseaux de neurones (le deep-learning), des choses comme ça qui rajoutent une couche de

magie ou en tout cas de changement de paradigme, qui font qu'aujourd'hui on peut vraiment cibler des comportements très complexes, ou avoir la victoire de Google au jeu de Go, ou encore l'histoire de la machine sur l'homme, ça laisse à penser qu'on change quand même de monde.

### **Pouvez-vous m'en dire plus sur les « réseaux de neurones », le deep-learning ?**

Le deep-learning est un sous-ensemble des réseaux de neurones, enrichi, qui profite des nouvelles innovations technologiques, et qui permet de construire des modèles de réseaux de neurones beaucoup plus complexes, et qui peuvent adresser des problématiques riches (jeux d'échec, jeux de go, jouer tout seul à un jeu vidéo, et à terme conduire une voiture tout seul). Ça sera très largement possible grâce à ces technologies.

### **Aujourd'hui, comment les entreprises accueillent les technologies big data ?**

Les niveaux de maturité sont différents. Déjà on voit que les entreprises du web se sont construites là-dessus et ont des avantages compétitifs extraordinaires. Vous prenez Google, ils ont Waze, ils connaissent tous les comportements, donc quand on est un acteur automobile par exemple, Google est capable de piloter des voitures, de faire l'optimisation du trafic. Si moi je ne m'y mets pas, je suis mort. J'ai travaillé par exemple avec PSA et Renault, les grands groupes se disent, « si on ne le fait pas, si on n'investit pas à fond, on est mort » ; et ils ont commencé à le faire dans des volumes significatifs. Pareil pour les grands groupes bancaires, qui peuvent se faire concurrencer très largement d'un mois sur l'autre, et avec des questions réglementaires, assurances... Tous les acteurs se disent qu'ils sont potentiellement menacés. Donc ils ont initié des démarches big data.

Mais il y en a d'autres qui ne l'ont pas encore fait mais qui se posent la question. On a rencontré des groupes du BTP, des entreprises moyennes dans la grande distribution, eux aussi vont s'intégrer, c'est en train d'arriver dans toutes les couches de l'économie. => III

En fait aujourd'hui, vous avez les très grands groupes qui investissent depuis 4/5 ans, et dont les investissements continuent, ils investissent très fortement et ils n'ont pas le choix. On avait le premier assureur français qui nous disait « notre principal concurrent potentiel c'est Facebook ou c'est Google, parce que le jour où ces groupes là se mettent à vendre des produits d'assurance, comme ils ont toute la Terre entière qui est connectée à eux, et comme ils ont énormément d'informations sur tous ces gens... », un assureur tel que Axa ou tel que Allianz, s'il veut dans dix ans rester numéro un sur son marché, il doit être presque aussi bon que Google dans sa connaissance client. Donc il y a des très grands groupes qui investissent énormément. Les compétences en data et en data science vont à deux endroits : dans les très grands groupes et dans les start-ups. Vous prenez Blablacar ou plein de start-ups dont on parle en France, Criteo, etc., ce sont des sociétés qui ont investi très fortement sur les données, sur le traitement intelligent de ces données (Blablacar avec de la géolocalisation extrêmement fine, Uber et tous les services offerts reposent sur ces technologies là). Entre ces très grands groupes et les start-ups, vous avez énormément d'entreprises qui n'ont pas forcément accès à ces compétences parce qu'elles sont encore rares, donc ils se posent des questions : « on y va, on n'y va pas, pourquoi, est-ce que c'est cher, est-ce que c'est pas cher, etc. ? ».

**Ipsos possède un département data science, plus spécialisé dans le web sémantique...**

Ce sont par exemple des départements ou des équipes de ce type qu'on peut accompagner. On peut faire part de notre savoir-faire et faire monter en compétences les équipes. Les équipes peuvent nous solliciter sur une thématique d'analyse sémantique, on fera des travaux, on les partagera avec le client, pour qu'il se les approprie ensuite. On a aussi travaillé sur une application qu'on a développée dans le cadre de la campagne présidentielle qui fait une représentation des communautés sur Twitter en fonction du hashtag que l'on suit. Plus on en sait sur les métiers et sur la façon de travailler de nos clients, plus on peut être amené à bien réfléchir sur les solutions qui pourraient être intéressantes pour lui.

### **Quand vous présentez vos solutions big data à vos clients, que leur présentez-vous ? Est-ce des tableaux de bord, des graphiques... ?**

Dans les types de présentations qu'on fait, on leur présente déjà la démarche dans laquelle on s'inscrit, qui est une démarche d'accompagnement, en expliquant bien les types de prestations qu'on peut faire. On vend un savoir-faire et un service, donc la première des choses c'est de faire en sorte qu'un client ait confiance en nous, à la fois sur notre savoir-faire et à la fois sur notre démarche. On essaie de bien expliquer ce qu'on sait faire, comment on le fait. Les clients ont de toute façon des problématiques différentes, il y a certes des étapes incontournables, mais on va effectuer une étude de leurs besoins, imaginer ce qu'ils ont comme patrimoine de données actuel et à venir, et la finalité va être de rendre un service opérationnel. Cela passera vraisemblablement sur la capacité à faire des tableaux de bord, ou en tout cas des outils qui, connectés à la donnée, leur apportent une vraie valeur ajoutée. C'est dans ce cas là que souvent, la fin de notre présentation se termine par une démonstration live des tableaux de bord, de notre capacité à matérialiser tout ce qui est acquisition et valorisation des données. On montre que pour arriver à ce produit qui réponde à leurs attentes stratégiques, il y a une démarche d'accompagnement et de formation qu'il faut bien avoir à l'esprit.

### **Quel a été votre parcours ?**

*Hubert* : Pour ma part j'ai un profil ingénieur en informatique, j'ai travaillé sur les bases de projets informatiques et de développement web. Puis j'ai évolué vers des fonctions technico-commerciales. J'aime bien la dualité, être à la fois pertinent techniquement, avoir une large culture technique, et le mettre au service de produits. Depuis 6-7 ans, je suis entre les deux, à un moment les mains dans le code et quatre minutes après je discute avec une Direction stratégique. Il y a une vraie place pour les acteurs techniques au profil ingénieur qui savent vulgariser les enjeux du moment.

*Olivier* : Ce sont des profils ingénieurs, et pour maîtriser ces outils là, il y a vraiment une dimension ingénieur qui est absolument inévitable. Et la nécessité de maîtriser des langages, de coder et d'avoir toutes les bonnes pratiques et démarches de la gestion de projet, parce que quand vous voulez faire cohabiter toutes ces technologies open source les unes avec les autres sur des cas concrets clients, il faut travailler 20 à 30 jours, il faut y aller à fond. Après on travaille aussi avec des autres métiers qui n'ont pas un background ingénieur, mais qui ont des connaissances de là où doit aller la stratégie et qui ont la connaissance de mise en œuvre et d'opportunités qui sont sous-tendues par la big data.

## Annexe 3. Entretien avec Charlotte Zaepfel, « data-scientist » à Ipsos

Retranscription de l'entretien avec Charlotte Zaepfel, « Data Scientist » à Ipsos

4 Mai 2017

### **Peux-tu te présenter en quelques mots, nous raconter ton parcours jusque chez Ipsos ?**

J'ai suivi des études de linguistique informatique. Le nom de ma formation, c'est « traitement automatique du langage » (qui comprend des cours de formalisation du langage, c'est ce qui m'a attiré). On l'appelle aussi sémantique ou syntaxe formelle, et cela consiste à mathématiser, à modéliser, computer le langage pour le rendre systématique.

Ici à Ipsos, je fais du *topic modelling*, qui a pour but de modéliser un corpus de texte pour qu'il soit utilisable par la machine. Je dois vectoriser des textes, c'est-à-dire créer une représentation par des vecteurs (qui sont des objets mathématiques très flexibles *via* des équations). J'applique donc des opérations vectorielles. L'objectif est de récupérer les grands thèmes du corpus de texte, soit les « champs sémantiques » par *topic*, si l'on vulgarise. Ces opérations vectorielles calculent des scores (ou probabilités) d'affinité entre les mots (les mots les plus probables par *topic*), autrement dit ces opérations de la machine donnent une probabilité d'appartenance du mot au *topic*.

### **Comment fonctionnent les algorithmes que tu utilises ?**

Ce sont des algorithmes génératifs, qui s'améliorent en faisant différentes itérations (à l'inverse d'un algorithme analytique). Avant de générer l'*algo génératif*, on peut préparer le corpus et optimiser les *topics* en détectant les collocations entre mots, pour récupérer les thèmes d'un texte. L'ordinateur/la machine traite les mots comme des index, comme en statistique (sur la base de 0, 1, 0, 1...). Ce sont des opérations complexes de langage *processing*, qui permet la classification de documents (c'est la méthode Word 2 Vec de Mikolov).

### **Concrètement, comment fais-tu lorsque tu dois traiter des gros corpus de textes ?**

Concrètement, je récupère une centaine de fichiers .csv, je fais des tourner des *algos* pour traiter chaque document. Je ne stocke pas le corpus brut, je le traite d'abord, à la différence de la méthodologie big data qui travaille sur des corpus stockés et avec beaucoup de *processing* (type *deep learning*). Moi je traite ligne par ligne, manuellement, et je stocke chaque résultat. Je convertis chaque ligne en vecteur.

### **Quelle est la particularité de la *méthodologie big data* selon toi ?**

Pour faire du *big data*, il faut construire des architectures pour accéder rapidement à des données. Cela demande d'avoir une machine avec de grosses capacités de stockage de mémoire. La manière dont je comprends le *big data*, c'est créer des bases de données intelligentes et créer des langages de parallélisation. C'est de la programmation pure. Cela requiert une autre connaissance du langage de

la machine. Ce sont des langages comme Hadoop et Spark, auxquels il faut que je me forme car c'est très demandé sur le marché par beaucoup de boîtes.

### **Comment utilises-tu les algorithmes ? Sur quelles méthodologies t'appuies-tu ?**

Pour moi, un algorithme c'est une recette de cuisine, une suite d'instructions. En ce moment j'utilise Python (mais c'est un peu lent, bien qu'il y ait une communauté active avec laquelle on peut échanger), et j'utilise R pour vectoriser mes opérations. Je fais un script bash pour lier les deux.

La méthodologie que je suis, c'est du « word embedding », une technique de vectorisation des mots par le contexte. Cela marche un peu comme un neurone, avec des input et des output, et de l'apprentissage. Si on prend l'exemple de ham ou spam, la machine fait le calcul sur la base d'une fonction affine. C'est en donnant des exemples d'apprentissage à la machine que *l'algo génératif* se crée.

On crée et on utilise des algorithmes pour que la machine construise des modèles en lien avec l'achèvement d'une tâche. Au niveau des résultats, idéalement on voudrait que *l'algo* soit aussi bien que le serait un humain. Après l'objectif n'est pas de simuler l'humain, car la machine a un langage propre, et l'humain aussi, ce sont deux langages et deux fonctionnements différents. L'Intelligence artificielle est un autre mode de fonctionnement que l'humain.

### **Peux-tu nous donner des exemples de missions que tu fais chez Ipsos ?**

Chez Ipsos, on me demande de rechercher des protosynonymes. J'ai fait ça pour un client dans la cosmétique spécialisé dans les soins du visage, qui voulait connaître dans quels cas sont utilisés ses soins du visage. Donc j'ai fait ressortir les soins du visage nourrissant, hydratant, nettoyant... On a fait ressortir des clusters de sens (les synonymes) et a trouvé les vecteurs les plus proches qui sont utilisés dans des contextes similaires. Devant le client, je les ai présentés par une table de synonymes, par ordre croissant de similarité, et ça a permis de nourrir le quali et l'animation de focus groups.

### **Tu es donc spécialisée en linguistique programmatique. Quels sont pour toi les principaux enjeux dans ton métier ?**

La recherche est importante en linguistique informatique, c'est une discipline jeune, qui bouge très vite. La difficulté c'est d'écrire un rapport entre technique et vulgarisation. Il faut vulgariser pour le client, et en même temps ça donne plus de légitimité si j'utilise un langage technique. Il y a aussi une chose qui fait peur au client, c'est le fait que l'algo ait trouvé un résultat tout seul. Il met en doute la légitimité de l'algorithme utilisé, c'est pour ça qu'il est important de s'appuyer sur des modèles d'algos qui sont reconnus par la communauté scientifique. J'ajoute donc toujours quelques slides à la fin pour expliquer ça, et ajouter une caution scientifique à ce que je présente.

J'ai l'impression qu'on parle beaucoup de l'algorithme avec un grand A. Les chargés d'études pensent que ça va « à la place de ». Sauf qu'il faut que je travaille les vecteurs pour qu'ils fassent ce que je leur demande. J'ai d'ailleurs sorti un Manuel d'utilisation pour expliquer en quoi consiste mon périmètre d'action, et les opérations d'algorithmes que je fais. On fait beaucoup d'aller-retour entre les chargés d'études et moi. Je les aide en leur sortant des topics et des corpus avec des

verbatim au score d'affinité le plus élevé (par exemple quand on parle à 90% du topic A). Et je leur donne une hiérarchisation des topics.

**Quelle est la part de l'automatisation effectuée par la machine et de la compétence humaine ? Que penses-tu de ces deux compétences ?**

Pour moi la machine est au même niveau que l'homme. Ma position face à la question, c'est que je me place du côté de l'intelligence – machine : je fais du nettoyage, je programme l'algorithme, mais je me place peu du côté intervention humaine. J'interviens avant pour calibrer l'algorithme, mais pas pendant. Puis j'essaie d'injecter l'intelligence humaine au moment le plus opportun.

Mon objectif c'est de rendre la machine la meilleure possible, trouver l'hybridation optimale. Il y a des tâches où l'homme ou la machine sont l'un et l'autre meilleur. L'homme pour détecter les émotions, les intuitions de langage et de grammaire. D'ailleurs l'homme ne construit pas l'ordre syntaxique de la phrase, mais la machine oui. Elle peut même écrire des poèmes et des symphonies.

L'homme est limité par sa mémoire, et en cela, le big data fait mieux que l'homme, il traite et croise des données massives et hétérogènes. Il permet une simultanée à intégrer des valeurs, la capacité à croiser les données.

Après le « big data », ça fait vendre... C'est vrai qu'on est dans l'ère des données, de la révolution de l'accès à ces données. Le « web sémantique » aussi, c'est un mot qui fait vendre. Pour moi, big data ce n'est qu'un big word : il faut d'ailleurs que je me mette à une formation pour Hadoop et Spark pour connaître ce langage big data, puisque c'est là où il y a des perspectives d'embauche !

## Annexe 4. Entretien avec Agnès Gilbert, Chief Digital Officier à Ipsos

Retranscription de l'entretien avec Agnès GILBERT, Chief Digital Officier à Ipsos

11 mai 2017

### **Pouvez-vous nous raconter votre parcours ?**

J'ai fait Sciences Po, puis de l'information-communication, j'ai travaillé dans une boîte de production qui vend des modèles statistiques. Puis j'ai intégré Nielsen comme panéliste, puis la société Eri, et enfin Ipsos. Je suis aujourd'hui Directeur du Digital.

### **On voit beaucoup de définitions circuler autour du big data (marketing prédictif, machine-learning, structuration des données...). Selon vous, qu'est-ce que le big data ?**

Le big data c'est d'abord les 3V : des données variées, volumétrique, de la vitesse, avec de la récurrence et un rafraîchissement rapide des données. Cela va être souvent des données non structurées, non sollicitées, presque inutilisables en tant que telles, et qui doivent être traitées.

Il y a beaucoup de déchets et cela nécessite beaucoup d'intervention humaine. On a l'impression qu'il y a avant tout de l'automatisation, des processus automatisés, or en réalité l'humain intervient beaucoup derrière.

### **Quelle est la part de l'automatisation et de l'humain ?**

La machine ne peut pas aller sur le chemin de l'arbitrage, qui relève du domaine de l'expert. L'algorithme ne fait que répliquer ce que l'humain lui dit de faire. Même s'il y a une forme d'intelligence artificielle, avec un apprentissage de la machine, elle n'aura jamais de libre-arbitre. On ne comprend pas toujours le fonctionnement de l'algorithme, et là il devient une boîte noire pour les programmeurs.

### **Comment utilisez-vous le big data chez Ipsos ? Sur quelles méthodologies vous appuyez-vous ?**

Pour un institut d'études, l'enjeu va être de chercher à répondre à une problématique, souvent de la grande consommation. On doit d'abord faire le choix du corpus qu'on analyse, puis le choix du traitement. Par exemple le *topic modelling* va être la mécanique principale qu'on choisit ici, il sert à faire remonter les sujets sur un corpus varié, à partir d'algorithmes qui établissent des corrélations. On va travailler le mot avec un moteur d'analyse sémantique sur le web, qui isole les occurrences et analyse les récurrences et les fréquences d'associations entre les mots.

Il y a certes aussi du marketing prédictif, qui va consister à établir des corrélations en analysant un passif. S'il y a tant de chances A, alors il y aura tant de chances B. Dominique Cardon parle de cette « soit disant prédictivité » : nous sommes des êtres routiniers, on n'a pas besoin de faire tourner des algorithmes. En réalité, on n'apprend pas grand chose avec des modèles prédictifs.

On ne peut qu'être prédictif de quelque chose qui n'existe pas, on ne sait pas chiffrer la prédictivité. On n'a pas l'équivalent de modèles comme Basis ou Novaction, qui font des modélisations statistiques et de la prévision de probabilité de succès d'innovation en volume, en se basant sur les habitudes des consommateurs et des modèles adjacents, ils construisent ainsi des modèles prédictifs. Ce n'est pas très novateur. L'innovation va être de travailler sur un corpus de données hétérogènes, une volumétrie d'informations. Ce spectre de données élargi élargit les possibilités mais c'est très chronophage de la structurer.

Le traitement de corpus de mots mène à des chiffres, le Net fait qu'on change d'échelle et qu'on arrive à des corpus énormes. Pour aller au bout de l'*insight*, on fait une analyse qualitative, on ne s'arrête pas au quanti. L'intervention humaine se fait là, avec les verbatims qui donnent du contexte aux chiffres.

### **Comment donnez-vous à voir la donnée ?**

La mise en forme de ces *insights* est important, cela peut être *via* des graphes ou des mappings... Souvent les moteurs d'analyse sémantique ont des graphes qui cartographient la récurrence et la fréquence de mots, comme Big Fish qui fait des cartographies. Il est clair que la représentation graphique est plus parlante. La représentation dépend du sujet que l'on traite. Par exemple pour le cas d'un parcours d'achat, on va faire un tracking de la navigation et on va représenter l'univers de consommation.

### **Concrètement, comment se fait l'analyse pour le département « data-science » ?**

Pour ce qui est de l'analyse, c'est assez artisanal mais cela prend du temps. On fait l'analyse sur Excel avec des matrices de séquençage, qui vont séquençer les adresses URL, le temps passé sur le site... On va pouvoir faire l'analyse manuellement sur la base d'un échantillon faible de 120 pax. Mais vu la quantité d'informations à analyser, on ne peut la traiter qu'en passant par des algorithmes. A partir du moment où ça tient dans un fichier Excel, ce n'est pas du big data.

### **Aujourd'hui, quelle est la place du big data à Ipsos ? Quels sont les grands enjeux autour de la data ?**

Il y a un enjeu de formation au digital et à la data. Il faut avoir la curiosité d'esprit d'utiliser d'autres choses, d'autres outils pour résoudre et répondre à des questions auxquelles on ne peut répondre, pour compléter une méthode initiale d'analyse. Il va s'agir d'aller plus loin dans notre manière de travailler.

Alors que chez Ipsos on va d'abord recueillir de la donnée structurée à partir de questionnaires et de questions précises, le big data va consister à récupérer de la donnée passive, recueillir des informations précises mais sans poser de questions au préalable. La question à se poser c'est : est-ce que mon analyse est auto-suffisante ou la compréhension va-t-elle être améliorée avec le big data ?

Mon rôle, c'est de faire comprendre que le digital, la data, sont plus une opportunité qu'une menace, il faut faire preuve de pédagogie et ne pas faire peur. Tout le monde en parle, mais peu en font. Aujourd'hui il y a de nouveaux acteurs, comme Linkfluence, qui se positionnent comme des

instituts d'études, et beaucoup de sociétés comme Facebook et Google vendent et offrent des données. Mais peu de ces sociétés vendent de l'analyse, ils font seulement de la tendance. C'est facile de compter sur Internet, mais il est difficile de mesurer car il faut une question. Par exemple lors des élections présidentielles, la société Filteris a travaillé à partir d'une source non représentative, elle a juste fait de la volumétrie des mentions Twitter et y a associé des algorithmes avec une charge émotionnelle. L'erreur a été qu'elle ne s'est pas basée sur un échantillon représentatif pour en faire une donnée fiable. Pour les grosses boîtes, la vraie problématique est aussi de structurer les données. Or aujourd'hui, les gros producteurs de data ont une utilisation basique des données.

## Annexe 5. Entretien avec Vincent Cottin, Développeur web PHP à Atsuké

(Startup Web & Mobile)

18 juillet 2017

### **Présentation :**

Tout d'abord, je ne suis pas un grand spécialiste dans le domaine du Big Data mais j'ai des connaissances en réseau et en programmation qui me permettront de t'aider un peu à sa compréhension. Je suis particulièrement sensible à ce sujet de manière personnelle et je le côtoie très souvent en étant actuellement développeur web dans une startup.

### **Selon toi, qu'est-ce que le big data ? à quoi ça sert, pour faire quoi ?**

Je vois le Big Data comme le regroupement des données informatiques stockées sur l'ensemble des serveurs inter-connectés au réseau Internet. Un serveur c'est tout simplement le nom que l'on donne à un ordinateur qui a été spécialement configuré et optimisé pour réaliser des tâches bien précises comme par exemple l'hébergement de sites web ou encore le stockage d'informations dans des bases de données. Dans la majorité des cas, un serveur est continuellement connecté au réseau Internet ce qui le rend accessible en permanence à toute personne également connectée à ce même réseau et leur grand intérêt est de permettre aux utilisateurs d'utiliser les services qu'ils proposent. De nos jours les choses ont un peu dérivées et ces services récoltent de plus en plus de données sur leurs utilisateurs en usant de diverses méthodes. Les informations personnelles sont devenues très convoitées et font désormais partie d'un grand marché. Elles sont en général utilisées par tout type de vendeurs qui cherche à mieux cibler ses clients en utilisant leurs informations et en identifiant leurs comportements. Les grandes compagnies ont commencé à construire des centres de données (Data Centers) qui regroupent des milliers de serveurs configurés pour le stockage de masse en embarquant de nombreux disque dur. L'appellation Big Data est apparue à ce moment là pour désigner ces prestataires qui accumulent la vie privée des gens de façon continue.

### **Comment traites-tu les gros volumes de données dans ton métier ? Quels logiciels/outils tu utilises, sur quelles méthodologies t'appuies-tu ?**

On utilise le plus souvent des systèmes de gestion de base données (MySQL, Oracle) très utile pour stocker efficacement les données sans avoir des informations redondantes. De plus, ils permettent d'exécuter des recherches très complexes pour trouver des informations bien précises. Tout d'abord, on construit le schéma de la base en créant des tables avec les types de données que nous allons récupérer puis on effectue des opérations de lecture, d'insertion ou de suppression et des mises à jours pour actualiser la base.

### **ça marche comment un algorithme big data ?**

Je ne connais pas en détail le nom des algorithmes utilisés mais ils sont nécessaires pour réaliser des traitements sur les données récupérées. Il y a surtout des opérations d'indexation pour accélérer les recherches parmi toutes les informations pour pouvoir les recouper plus facilement par la suite.

**Quelle est la part de l'automatisation vs l'intervention humaine ? Qu'est-ce que vous déléguez à la machine ?**

En général, on délègue la majorité des tâches à la machine et on utilise des outils de monitoring pour contrôler le bon fonctionnement des opérations en temps réel.

**Avant de lancer les algorithmes, comment se fait la collecte des données ? De quelle base vous partez ?**

Dans le domaine du web, la méthode la plus utilisée est le "tracking" des clients qui consiste à récupérer un maximum d'informations sur les personnes connectées comme leur localisation (si activée), langage et version du navigateur, action clics. Ce sont les informations de bases que les sites récoltent à tout moment et les autres sont souvent fournies par les utilisateurs eux-mêmes, soit par le biais de formulaires les poussant à donner des informations personnelles ou par les actions réalisées sur la pollution visuelle qui les incitent à cliquer sur des pubs.

**Quel est le livrable que vous présenter au client, comment donnes-tu à voir les résultats des algorithmes big data ?**

Toujours dans mon domaine, lorsque l'on livre un site à un client, celui-ci se compose de deux parties : une visible par les utilisateurs (Front Office) et une deuxième (Back Office) destinée uniquement au client pour lui permettre de visualiser plus facilement les données. Cette dernière très souvent constituée de tableaux et de graphiques permet de mieux comprendre les chiffres et les informations récupérées.

**Selon toi, quelle importance ont ces technologies big data aujourd'hui dans les entreprises, et le secteur économique plus généralement ?**

D'un point de vue personnel, je pense que ces technologies sont de plus en plus envahissantes et continuellement à limite du droit à la vie privée. Les entreprises ont pris conscience du potentiel des données utilisateurs et c'est devenu le passage obligé dans le monde du web.

**Pour des raisons de droits, les pages 121 à 143 ont été retirées de la version diffusée en ligne.**