Valorisation des données dans les grandes entreprises

Maturité, pratiques et modèle

cigref
réussir
le numérique

DEVELOPPER LA CAPACITE DES GRANDES ENTREPRISES
A INTEGRER ET MAITRISER LE NUMERIQUE

Valorisation des données dans les grandes entreprises

Maturité, pratiques et modèle

Novembre 2016

LE MOT DU PILOTE

Les données constituent, et vont constituer davantage encore durant la décennie qui vient, un levier essentiel de la transformation numérique des entreprises et de leurs conditions de création de valeur.

En conclusion de ses travaux sur l'Entreprise 2020 confrontée au défi du numérique, le CIGREF avait fait de la donnée l'un des neuf enjeux, intitulé *« Valoriser les données et créer la confiance »*. Un certain nombre de défis y étaient associés, tels que : l'exploitation du capital données, la garantie de leur qualité, la promotion de l'éthique dans leur usage, leur protection et enfin, la mise en place d'un management global de la donnée.

Même si l'exploitation des données a déjà fait l'objet de nombreux développements visant notamment à améliorer la prise de décisions dans nos entreprises, la révolution internet, et plus largement la révolution numérique, ont engendré l'ère de l'information. Auparavant, cette dernière était rare et difficile à trouver, elle est dorénavant pléthorique, largement accessible, mais sa collecte et son analyse restent cependant toujours de véritables challenges. Ne pas savoir exploiter toutes ces données aujourd'hui disponibles, et ne pas se donner les moyens de le faire, met l'entreprise en danger, dans un contexte où les concurrents, eux, le feront. Il suffit de voir les termes habituellement associés au *Big Data* : infobésité, tsunami, avalanche de données, etc… rien que des catastrophes ! Alors l'enjeu est clair : l'entreprise doit optimiser l'usage de ses données, pour en tirer de la valeur et en faire son nouvel « or noir ».

Au-delà de la masse de données disponibles, la véritable nouveauté réside dans la démultiplication des potentiels de croisement de données, dans cette perspective de création de valeur. De plus, dans un monde en pleine accélération, l'usage en temps réel des données devient une variable essentielle de la prise de décision opérationnelle. Enfin, les réseaux sociaux donnent lieu à un champ immense de données non structurées et encore insuffisamment maitrisées par les entreprises. En outre, le numérique exacerbe les risques sur les données de l'entreprise, et induit l'obligation d'une protection des données renforcée, et d'une attention particulière aux obligations réglementaires, notamment sur les données personnelles, pour créer la confiance des clients et veiller à la réputation de l'entreprise. La donnée est au cœur de la transformation de nos entreprises, de leur développement face à la concurrence, de l'évolution de leurs modèles d'affaire, de l'apport de services aux clients, de leur efficacité opérationnelle. Que ce soit pour anticiper et répondre aux menaces, ou pour tirer profit des opportunités, la donnée est une source de valeur considérable, mais encore largement sous-estimée… C'était là tout l'objet des travaux du groupe qui a réuni, depuis l'automne 2015, une vingtaine d'entreprises membres.

Le présent rapport s'adresse aux décideurs, tous d'ores et déjà convaincus de l'actif stratégique que leurs données représentent, de leur richesse, sans pour autant toujours être en mesure de l'évaluer. Les grandes entreprises qui ont collaboré à ces travaux ont contribué à faire progresser

cet objectif d'évaluation de la valeur des données, par leurs expériences partagées, et par leur créativité dans les réflexions autour des approches possibles de valorisation. Nul doute que le lecteur y trouvera matière à renforcer sa maturité sur ce sujet critique pour l'avenir de nos organisations…

- Renforcer la maturité sur un sujet, encore largement exploratoire, était justement l'un des principaux enjeux des travaux du groupe. D'où le questionnaire élaboré ensemble, qui comporte 80 pratiques (dont certaines reprises de la méthode *« Data Value Modelling - DVM® »* développée par le Pr Bounfour de l'Université Paris Sud XI) et devrait permettre aux entreprises d'auto-évaluer leur maturité.
- Autre production originale du groupe de travail, un modèle de valorisation d'un service de mise à disposition de données à son écosystème : toutes les entreprises, quelles qu'elles soient, développent ce type de service à valeur ajoutée pour les personnes ou entreprises intéressées. Proposer une méthode pour en évaluer la pertinence, économique mais aussi réputationnelle ou réglementaire, apparait comme une avancée significative.
- D'autres éléments contribuent directement à la valorisation des données, comme la gouvernance mise en place au sein de l'entreprise, ou encore les architectures techniques et les solutions déployées. Ces aspects font l'objet d'un chapitre spécifique.
- Enfin, le livrable contient toute une partie consacrée à des témoignages sur des pratiques d'entreprises qui ont accepté de faire part de leurs initiatives et de leurs succès. Elles sont autant de références qui viennent confirmer, au travers de cas concrets, que la donnée devient cœur de métier, que sa contribution au développement et à la transformation de l'entreprise est une réalité, dont la mesure objective reste néanmoins à faire.

A l'heure où les travaux du groupe s'achèvent, matérialisés par un livrable riche de vrais apports opérationnels sur ce thème neuf et ambitieux de la valorisation des données, je tiens à remercier vivement tous les participants des entreprises membres du CIGREF pour leur contribution active et créative, et les intervenants extérieurs qui nous ont fait bénéficier de leur expertise, sans oublier l'équipe du CIGREF qui a animé les réflexions et rédigé ce rapport.

Bonne lecture.

Régis DELAYAT
Vice-président du CIGREF - Pilote du groupe de travail
DSI de SCOR

Le CIGREF, réseau de Grandes Entreprises, a été créé en 1970. Il regroupe plus de cent très grandes entreprises et organismes français et européens de tous les secteurs d'activité (banque, assurance, énergie, distribution, industrie, services...). Le CIGREF a pour mission de développer la capacité des grandes entreprises à intégrer et maîtriser le numérique.

TITRE DU RAPPORT : VALORISATION DES DONNEES DANS LES GRANDES ENTREPRISES – MATURITE, PRATIQUES ET MODELE

EQUIPE DU CIGREF

Jean-François Pépin, Délégué général
Sophie Bouteiller, Directrice de mission
Flora Fischer, Chargée de programme de recherche
Josette Watrinel-Herniou, Secrétaire de Direction
Frédéric Lau, Directeur de mission
Marie-Pierre Lacroix, Chef de projet
Josette Leman, Assistante de direction

REMERCIEMENTS

Le groupe de travail « Valorisation des données » a été piloté par **Régis DELAYAT, DSI de SCOR et Vice-président du CIGREF.** Nous remercions les personnes qui ont participé aux échanges et contribué à la rédaction du présent document :

Eric Alix, La Poste
Sophie Allaire, Total
Cyril Bauer, L'Oréal
Mehdi Ben Hmida, Edenred
Irinel Béta, ADP
Christophe Binot, Total
Loris Bosa, RTE
Ahmed Bounfour, Université Paris Sud
Jean-Christophe Brun, Carrefour
Franck Denié, Pôle emploi
Brigitte Declerck, AGIRC-ARRCO
Christophe Digue, EDF
Olivier Djomby, Essilor
Olivier Druel, SCOR
Alain Dubost, Michelin
Bernard Dupré, RATP
Georges Epinette
Christine Eymard, Pôle emploi
Michèle Fort, SNCF
Marie-Noëlle Gibon, La Poste
Jérôme Gueydan, Orange
Christian Guy, Lisi Aérospace
Gérard Hatabian, EDF
Lionel Heytens, Société Générale
David Huck, Air France KLM
Michel Hutin, Arkema
Bruno Lalande, Renault
Yann-Maël Larher, Total
Gauthier Lemasne, Air France KLM
Michael Lepoutre, Auchan
Ludovic Letort, Nexans
Christian Maitre, Orange
Gérard Michalik, Michelin
Jean-Luc Raffaëlli, La Poste
Arnaud Renard, RTE
Yao Sogadzi, Fondation de France
Olivier Sorba, Lagardère
Marie-Christine Tuton, RTE
Stéphane Vial, carrefour
Michael Ymele-Leki, BPCE

Ce rapport a été rédigé par Sophie BOUTEILLER et Flora Fischer (CIGREF), avec la contribution des membres du CIGREF.

POUR TOUT RENSEIGNEMENT CONCERNANT CE RAPPORT, VOUS POUVEZ CONTACTER LE CIGREF AUX COORDONNEES CI-DESSOUS :

CIGREF, Réseau de Grandes entreprises
21, avenue de Messine 75008 Paris
Tél. : + 33.1.56.59.70.00
Courriel : contact@cigref.fr

Sites internet :
http://www.cigref.fr/
http://www.entreprise2020.fr/
http://www.histoire-cigref.org/
http://www.questionner-le-numerique.org
http://www.entreprises-et-cultures-numeriques.org

Droit de propriété intellectuelle

Toutes les publications du CIGREF sont mises gratuitement à la disposition du plus grand nombre, mais restent protégées par les lois en vigueur sur la propriété intellectuelle.
Est autorisée la copie du titre et d'extraits de 500 caractères, suivis chacun de la mention « Source : » assortie de l'url de la publication CIGREF. Toute autre reprise doit faire l'objet d'une autorisation préalable auprès du CIGREF cigref@cigref.fr.

SOMMAIRE

1. Introduction 1
   1.1. La donnée, un actif stratégique de l'entreprise… 1
   1.2. …cependant mal géré et trop peu valorisé 1
   1.3. La valorisation des données, comment aborder le sujet ? 3
      1.3.1. Ce qui a fait rapidement consensus dans le groupe de travail 4
      1.3.2. Ce qui fait encore débat et/ou qui pose question 5
2. Maturité des grandes entreprises en matière de valorisation des données 6
   2.1. Présentation de l'outil d'auto-évaluation CIGREF 6
   2.2. Maturité globale des entreprises membres du CIGREF 7
3. Illustration d'une démarche de mise à disposition de données 10
   3.1. Comment attribuer une valeur à une donnée ? 10
   3.2. Exemple de service : « Mise à disposition de données à l'écosystème » 11
      3.2.1. Le modèle 11
      3.2.2. Explication du modèle 12
   3.3. Quelques prérequis à la mise en œuvre de ce service 14
4. Quelle gouvernance pour valoriser les données ? 15
   4.1. Insuffler la « culture donnée » 15
   4.2. Repenser l'organisation et responsabiliser les acteurs 16
   4.3. Gérer le cycle de vie des données pour mieux innover 19
   4.4. Faire de l'éthique un facteur de différenciation 22
   4.5. S'outiller 24
5. Valorisation des données : cas d'usage et exemples de pratiques 27
   5.1. Faire évoluer les modèles d'affaires 28
      5.1.1. Usages et approches : exemples sectoriels 28
      5.1.2. Démarches de valorisation 30
      5.1.3. Défis identifiés 31
   5.2. Améliorer la connaissance des clients, de l'écosystème, des partenaires 32
      5.2.1. Développer une vision 360° des clients 32
      5.2.2. Intégrer le temps réel 34
      5.2.3. La qualité des données et l'approche « référentiels » 35
      5.2.4. Défis identifiés 35
   5.3. Créer de la valeur et la partager 36
      5.3.1. La redistribution 36
      5.3.2. Une approche possible du partage de la valeur 37
      5.3.3. Défis identifiés 39
6. Conclusion 40
7. Annexe – Référentiel de pratiques de l'outil d'auto-évaluation 42
8. Bibliographie 45

1. INTRODUCTION

1.1. La donnée, un actif stratégique de l'entreprise…

La donnée est devenue l'or noir de toutes les entreprises, quelle que soit leur taille. Selon Gartner, les entreprises qui auront su anticiper avec le *Big Data* tireront un avantage compétitif de 25% ou plus sur leurs concurrents à l'horizon 2019. Et selon une étude *International Data Corporation* (IDC – 2014), 100% des grandes entreprises achèteront de la donnée et les transactions se développeront de plus en plus d'ici 2019. Il semble donc évident qu'un nombre croissant d'organisations va chercher à monétiser ses données, soit en les vendant à court ou moyen terme, soit en en optimisant l'usage pour développer son *business*.

L'enjeu relatif aux données a été formalisé et décrit par le CIGREF en 2014 dans son ouvrage *Entreprise 2020 : enjeux et défis à l'ère du numérique*. La nécessité d'inventer la relation entre valorisation et données est portée par le développement du *Cloud Computing* et du *Big Data*, qui sous-tendent les grandes tendances technologiques.

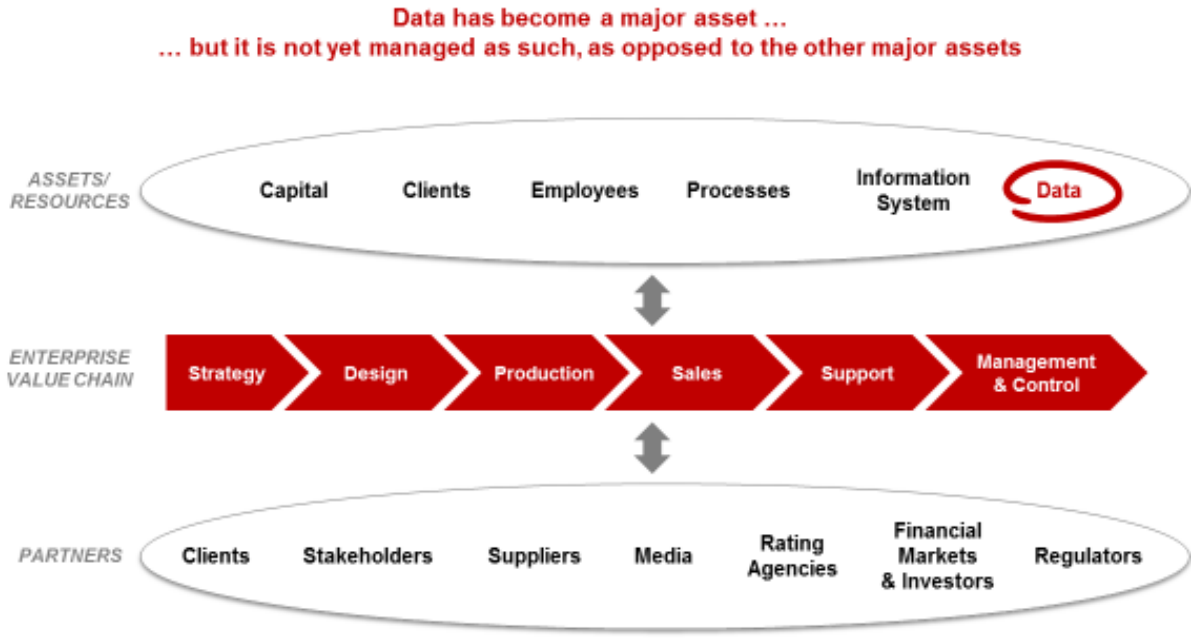
Elles posent par ailleurs les questions cruciales pour l'entreprise de l'exploitation et de la protection des données, et notamment des données personnelles. L'enjeu ici est de créer les conditions de la confiance dans l'exploitation des données, car il n'y a pas d'économie numérique sans confiance.

L'apparition de nouvelles lignes de produits et de nouveaux services, basés sur l'exploitation et la valorisation des données, crée des opportunités et fait évoluer – de manière plus ou moins profonde – les modèles d'affaires des entreprises. Dans un contexte où le risque de désintermédiation est majeur, la maîtrise de la relation client, la personnalisation de masse, la co-conception avec les consommateurs ou l'écosystème seront fondamentales dans la réussite des entreprises en 2020.

1.2. …cependant mal géré et trop peu valorisé

Mais pour arriver à créer de la valeur à partir des données, encore faut-il que celles-ci soient de qualité : la qualité des données est la condition de base à la valorisation, et pour cette raison, la gestion des données dans l'entreprise est fondamentale. Sur ce dernier point, nous vous invitons à vous reporter à un rapport du CIGREF, *Enjeux business des données : comment gérer les données de l'entreprise pour créer de la valeur ?*, publié en octobre 2014. Ce rapport propose une méthodologie de gestion des données, avec des exemples concrets de mise en œuvre et des bonnes pratiques, ainsi qu'une ouverture sur une démarche de valorisation des données. Il est assorti d'un outil d'auto-évaluation de la maturité des entreprises en matière de gestion des données qui montre que si les fondamentaux sont bien maîtrisés dans les entreprises (opérations et sécurité), un long chemin reste à faire sur la qualité, l'architecture, les métadonnées, la gouvernance et… la valorisation ! Et aujourd'hui encore, force est de

constater que les entreprises valorisent encore assez peu leurs données, notamment du fait d'un manque de compétences, d'investissements dans les outils, et de stratégie dédiée à la donnée et sa valorisation[1].



Source : *Enjeux Business des données*, CIGREF (2014)

Le défi pour l'entreprise est aussi de promouvoir des règles éthiques qui favorisent l'émergence d'un système de confiance, garantissant une création de valeur et une innovation responsables et respectueuses des clients, à partir de l'exploitation des données. Car s'il s'agit désormais de faire du *business* autour de la donnée, où qu'elle se situe, et de créer de nouveaux produits et services à partir de son exploitation, cela impose d'être agile dans la valorisation et la sécurisation de cet actif stratégique, tout en ayant une gestion responsable et éthique, qui favorise la confiance. Cette dimension a d'ailleurs fait l'objet d'un groupe de travail CIGREF et d'un rapport publié en octobre 2015, sur *L'Economie des données personnelles : les enjeux d'un business éthique*.

[1] Au démarrage de ce groupe de travail, et pour initier la réflexion, le CIGREF a réalisé une analyse de plusieurs études qui mettaient en évidence ces constats (études : *Data & Analytics Trends 2015* – Deloitte, *Databerg 2015* – Veritas Technologies LLC, et *Seizing the information advantage*, septembre 2015 – PwC).

1.3. La valorisation des données, comment aborder le sujet ?

Pour évaluer les grandes tendances en matière d'usages des données dans les grandes entreprises, le CIGREF a réalisé une courte enquête auprès de ses membres au lancement du groupe de travail, fin 2015. Sur la trentaine d'entreprises qui ont répondu, il est apparu que :

- Les principaux usages faits de la donnée par les différents Métiers et Fonctions de l'entreprise visent en priorité à améliorer la prise de décision, à analyser la rentabilité, à renforcer la connaissance client, à améliorer les services aux clients, mais aussi à élaborer de nouveaux produits et services ;
- Une grande majorité d'entreprises a mis en place une application *Big Data*, et utilise très largement des données externes ;
- Plus de 50% des entreprises répondantes ne mettent pas leurs données à disposition du public et n'ont pas de politique d'*Open Data* ;
- Les bénéfices tirés de l'usage des données visent à fidéliser les clients et améliorer la qualité pour développer l'avantage concurrentiel ;
- Les entreprises n'ont globalement pas élaboré de méthode de mesure de la valeur de leurs données ;
- Les principaux freins à la mise en place d'une démarche de valorisation des données sont de nature culturelle et organisationnelle ;
- La question de la valorisation des données se positionne d'abord au niveau des Directions Métiers, puis du Comex.

Ainsi, si l'objectif initial du CIGREF, à travers ce groupe de travail, était de proposer une méthode pour chiffrer la valeur des données, cette courte enquête a permis de constater que cette ambition était largement prématurée par rapport au niveau de maturité dans les entreprises, d'autant que les discussions au sein du groupe ont montré que la valorisation des données fait débat. Certes, toutes les entreprises sont conscientes et convaincues des enjeux *business* des données, et pour une grande partie d'entre elles, elles tirent d'ores et déjà de la valeur de leurs données, internes et externes. Cependant, si elles les utilisent déjà beaucoup, elles le font très insuffisamment en regard du volume de données aujourd'hui disponibles, et ne savent pas y associer une valeur, voire n'ont pas fait l'exercice.

1.3.1. Ce qui a fait rapidement consensus dans le groupe de travail

De nombreuses affirmations qui ont fait consensus se sont retrouvées au cœur des réflexions :

- **Le temps réel** peut être fondamental dans la valorisation d'une donnée à un moment t : une même donnée peut avoir une très grande valeur pour un Métier ou une Fonction dans l'entreprise à un moment t, puis plus aucune au moment t+1.
- La valeur des données varie selon les usages qu'en font les Métiers et Fonctions, il y a une richesse évidente à **« transversaliser » les données** dans l'entreprise, où les silos perdurent encore trop souvent.
- **L'approche culturelle** est fondamentale pour intéresser tous les acteurs à la valeur des données : au sein d'une même entreprise par exemple, certains métiers très digitaux ont une culture de la donnée très développée alors que d'autres métiers, plus traditionnels, sont très immatures et pourtant pas moins concernés.
- La **gouvernance** est incontournable : pour être efficaces dans leurs démarches de valorisation, les entreprises doivent définir et mettre en place une gouvernance adaptée. Dans le groupe de travail, il est apparu que dans certaines entreprises, des initiatives autour de la mise à disposition et de l'usage de la donnée existent, mais elles sont souvent isolées, il n'y a pas de gouvernance globale. D'autres ont créé un département « Data » (*Data Science*, architecture, gouvernance) qui travaille à partir du cycle de vie de la donnée pour identifier la valeur, mais elles rencontrent de vraies difficultés dans l'outillage et la méthodologie pour simplifier la collecte et l'analyse. D'autres enfin ont mis en place des équipes pluridisciplinaires qui rassemblent selon les cas *Chief Digital Officer, Chief Data Officer, Chief Data Quality Officer, Digital Data Officer, Data Protection Officer* et *CIO* pour casser les silos et diffuser la culture du partage (favoriser une continuité dans l'utilisation de la donnée au sein de l'entreprise, partager et convaincre les Métiers de s'ouvrir à l'*Open Data*).
- Il manque une **approche méthodologique et outillée** pour extraire plus de valeur des données. L'utilisation des données pour créer de nouveaux produits et services est complexe car s'il existe une quantité énorme de données stagnantes dans les différents Métiers partout dans le monde, on ne sait pas les récupérer facilement, ni les analyser. On sait que l'information existe mais comment capter les données qu'on ne voit pas dans nos systèmes mais dont on pressent pourtant qu'elles ont un potentiel de valeur ? Beaucoup d'entreprises sont encore en phase d'apprentissage et multiplient les *Proof of Concept* (PoC) sur la prise de décision ou le pilotage de l'entreprise, mais le mode industriel est encore rare : peu d'entreprises du groupe de travail ont développé un niveau de maturité significatif en la matière. Le passage des PoC à **l'industrialisation** reste difficile.
- La valorisation pose la question des **engagements de l'entreprise vis-à-vis des Métiers et de l'extérieur :** la donnée est un vecteur par lequel une entreprise peut opérer sa

transformation (personnalisation, reconnaissance des clients, anticipation des besoins) mais pas dans n'importe quelles conditions (loyauté des algorithmes, responsabilité dans l'usage des technologies). *Privacy* et création de valeur vont de pair.

- La **question de l'écosystème** (établir un réseau avec des partenaires autour des données) est fondamentale dans l'approche valorisation.
- Au-delà de la compréhension du modèle organisationnel, les entreprises ont **besoin d'indicateurs** objectifs et quantifiables pour prioriser les investissements et les projets structurés autour de la donnée.

1.3.2. Ce qui fait encore débat et/ou qui pose question

D'autres réflexions font encore débat, notamment :

- **Sur l'objet même de la valorisation :** pour certains participants, valoriser les données revient à valoriser des services embarquant de la donnée ou de l'information, alors que pour d'autres, ce qui importe est bien la valorisation de l'actif « donnée » à l'état brut. Quelles données veut-on valoriser et dans quel but (*Big Data*, *Open Data*) ? A quels coûts ?
- **Sur le caractère stratégique de la valorisation en fonction de la nature de l'entreprise (*BtoB/BtoC*) :** la dimension stratégique de la valorisation des données ne fait pas consensus. Elle serait un sujet stratégique pour les entreprises en *BtoC* mais pas nécessairement pour les entreprises en *BtoB*. Pourtant, les entreprises historiquement en *BtoB* cherchent justement à se rapprocher de leurs clients finaux et à développer « l'intimité client » (*BtoBtoC*), et celles en *BtoC* déploient d'importants efforts pour conserver et renforcer cette proximité avec leurs clients. La donnée est l'énergie qui supporte les initiatives permettant aux entreprises d'innover pour rester au plus proche du consommateur.
- **S'agissant de la gouvernance :** ce n'est pas tant la gouvernance des données qui importe (qui reste cependant fondamentale), selon certains, que la gouvernance de la valorisation des données. L'enjeu pour les entreprises est de fédérer les initiatives, de partager les bonnes pratiques et d'animer une communauté autour de la valorisation des données. Les entreprises auraient surtout besoin de réfléchir à la gouvernance de la valorisation, et de définir une méthodologie. Par ailleurs, la gouvernance de la donnée est un sujet souvent porté par les Métiers, sur lequel la Fonction SI doit être un support (gouvernance, technologies, agilité…) face à des Métiers qui ont besoin d'aller vite.

2. MATURITE DES GRANDES ENTREPRISES EN MATIERE DE VALORISATION DES DONNEES

A la lumière des premiers échanges du groupe de travail, est rapidement apparu le besoin pour les participants de donner un moyen à chacune des entreprises membres de s'auto-évaluer et de faire progresser leur maturité en matière de valorisation des données. C'est pourquoi, le premier chantier a porté sur l'élaboration d'un outil d'auto-évaluation, basé sur un référentiel de 80 pratiques identifiées par les participants, et quelques-unes reprises de la méthode *« Data Value Modeling - DVM® »* développée par le Pr Bounfour (Université Paris Sud XI) présentée ci-après.

2.1. Présentation de l'outil d'auto-évaluation CIGREF

L'orientation choisie par le groupe de travail a été de créer un **référentiel de pratiques, assorti d'un outil d'auto-évaluation,** permettant à chaque entreprise de se positionner et d'évaluer sa maturité en matière de valorisation des données.

La grille de maturité est repartie en 6 axes structurants :

1. *Business Models*
2. Connaissance (clients, écosystème)
3. Transversalité et efficacité opérationnelle
4. Stratégie et gouvernance
5. Ethique et conformité
6. Outils et solutions

Chacune des pratiques est notée sur une échelle de 0 à 3 (0 = inexistant, 1 = débutant, en phase d'initialisation, 2 = Avancé, en cours de mise en œuvre, 3 = Confirmé, dispositif opérationnel et/ou en phase d'amélioration). Le radar obtenu en fin d'évaluation permet de positionner les entreprises sur chacune des briques d'une démarche de valorisation des données.

Ce référentiel de pratiques, disponible en annexe du présent document, a été élaboré dans le cadre des réunions du groupe de travail et à partir des témoignages de divers acteurs. Le groupe de travail s'est également inspiré du modèle *« Data Value Modelling - DVM® »* développée par le Pr Bounfour (Université Paris Sud 11), dont certaines pratiques sont reprises dans l'outil CIGREF.

Au-delà de l'évaluation de la maturité, l'objectif de l'outil est de sensibiliser les directions Métiers à la valeur de la donnée. Pour le groupe de travail, il s'agissait de **proposer un modèle concret sur la valeur des données, impliquant autant la Fonction SI que les Métiers.**

La méthode « Data Value Modeling - DVM® »

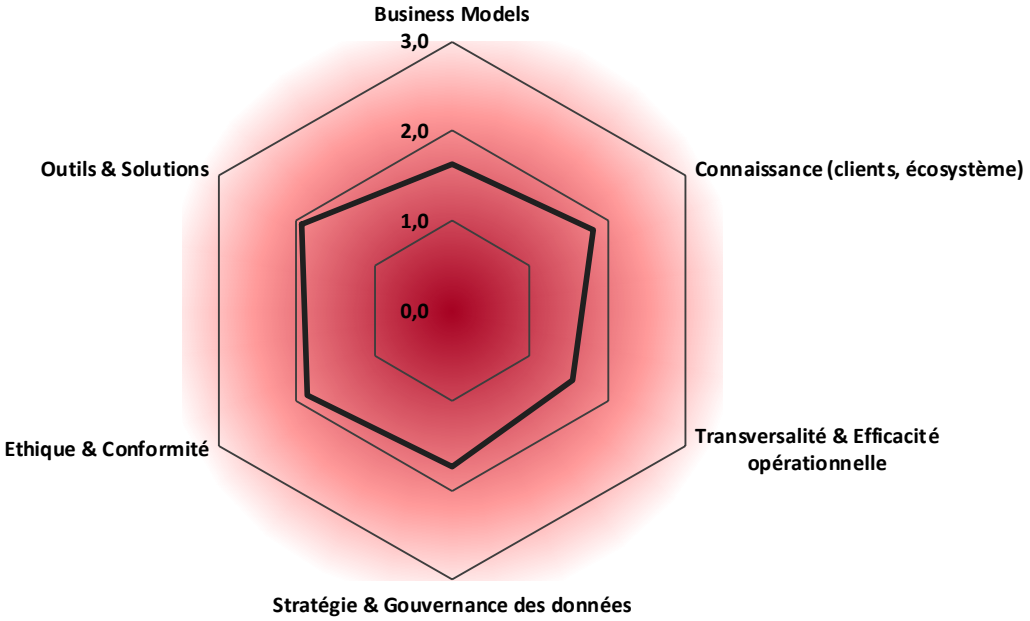
Cette méthode a un caractère exhaustif puisqu'elle est basée sur les résultats du programme ISD (dans la cadre de la Fondation CIGREF) et sur les travaux du Pr Bounfour relatifs au capital immatériel. Elle a pour objet de faire migrer les données structurées et non structurées vers la catégorisation des 3 types d'actifs (propriétaires/conjoints/non propriétaires) dans une perspective de création de valeur, et explore 6 dimensions :

- *Potentiel stratégie /* business models
- *Connaissance client / partenaires*
- *Partage de la valeur*
- *Transversalité*
- *Temps réel/décision*
- *Ethique des usages/RSE*

La méthode comporte 25 niveaux de pratiques et 5 niveaux de maturité. Elle a été construite en collaboration étroite avec une entreprise membre du CIGREF, et a mobilisé principalement la DOSI, le Chief Data Officer, *les responsables de projets, les responsables internet, les contrôleurs de gestion et le marketing.*

Cette méthode peut s'appliquer à tous secteurs car elle est suffisamment générique. Une des forces du modèle est le benchmarking *qu'il permet de réaliser entre entreprises pour montrer aux Directions Générales et aux directions Métiers comment l'entreprise se positionne à un moment donné en matière de valorisation des données.*

2.2. Maturité globale des entreprises membres du CIGREF



Source : CIGREF, Enquête 2016

Le questionnaire d'auto-évaluation a été soumis à l'ensemble des entreprises membres du CIGREF, une trentaine de réponses ont été collectées. Le radar ci-dessus est le résultat consolidé de ces différentes auto-évaluations du niveau de maturité en matière de valorisation des données. Il donne une première idée du niveau de maturité globale des grandes entreprises.

Avec une moyenne de 1,6 (sur 3), les axes « Transversalité & Efficacité opérationnelle » et « *Business Model* » représentent les plus faibles taux de maturité des entreprises. Les principales difficultés liées à la valorisation des données en entreprise reflètent ainsi certains freins ayant trait au (à la) :

- Partage des données en interne et à l'externe ;
- Accessibilité des données ;
- Décloisonnement des Métiers ;
- Optimisation de la qualité des produits et/ou des services ;
- Capacité à prendre des décisions éclairées, à partir de l'analyse de données (structurées et non structurées) en temps réel ;
- Capacité à développer des partenariats pour faire évoluer le *business model* et se différencier de la concurrence grâce à la valorisation des données.

Par contre, le radar montre que les points les plus forts de la maturité sont les axes « Ethique & Conformité » et « Outils & Solutions », avec une moyenne de 1,9 chacun. Il existe donc, pour une majorité d'entreprises répondantes, des stratégies visant à lier usages responsables des données avec la vision stratégique. Il semble essentiel pour les entreprises membres du CIGREF de faire de l'éthique un élément de différenciation, en garantissant à leurs clients finaux le respect de leur vie privée. La confiance est vue comme un argument de vente, l'éthique ne s'opposant pas au *business* : *« privacy is good for business »*[2]. Parallèlement, les entreprises s'outillent pour explorer et expérimenter les différents usages possibles des données, *via* notamment la mise en œuvre de systèmes et d'architecture *Big Data* mais aussi d'*APIs.*

Entre ces deux extrêmes se trouvent les axes « Connaissance clients, écosystème » (1,8 de moyenne) et « Gouvernance » (1,7 de moyenne), qui reflètent ainsi un niveau de maturité moyen, laissant donc une marge d'amélioration certaine. Ceci illustre bien les difficultés constatées dans le groupe de travail autour de la gouvernance des données et de leur valorisation, encore balbutiantes dans nombre d'entreprises.

Globalement, ce radar montre que les entreprises sont bien conscientes de la valeur potentielle des données qui sont stockées en masse dans leurs systèmes d'information et en dehors. Mais il montre également qu'un long chemin reste à parcourir avant que les

[2] Interview d'Ann Cavoukian, Commissaire à l'information et à la protection de la vie privée de l'Ontario : http://www.theatlantic.com/magazine/archive/2001/03/the-reinvention-of-privacy/302140/

organisations ne deviennent, à l'image des GAFA, des entreprises *« Data Centric »* où la donnée est un élément structurant de la culture de l'entreprise et où elle est gérée comme un actif clé.

3. ILLUSTRATION D'UNE DEMARCHE DE MISE A DISPOSITION DE DONNEES

Pour donner une illustration concrète de la forme que peut prendre une démarche d'ouverture des données, et de la valeur qu'une entreprise peut en tirer, le groupe de travail a travaillé sur un cas d'usage qui a permis d'**identifier les axes sur lesquels les entreprises doivent faire porter leurs efforts** en matière de valorisation des données.

3.1. Comment attribuer une valeur à une donnée ?

Pour le groupe de travail, deux difficultés majeures se posent en matière de valorisation des données :

- Définir la **valeur des données en interne,** avant de pouvoir exprimer comment la combiner et créer de la valeur supplémentaire ;
- Identifier les **données qui n'ont pas de valeur aujourd'hui mais qui pourraient en avoir demain :** on est un peu ici dans une démarche de prophétie/spéculation.

Objectiver la valeur de la donnée : les 10 variables clés de Dawex (www.dawex.com)

Dawex est une plateforme BtoB *d'achat, de vente et d'échange de données pour les entreprises qui souhaitent acquérir des données sans intermédiaire et/ou générer des revenus en les monétisant. Membre de Cap Digital, Dawex a été créée en mars 2015 avec le soutien financier de BPIFrance. Dawex joue le rôle de tiers de confiance entre l'acheteur et le vendeur et sécurise la transaction de jeux de données des entreprises sur plusieurs plans : technique, réglementaire et contractuel.*
Dawex a présenté au CIGREF les 10 critères qui permettent, selon cette start-up, de définir la valeur d'une donnée :

1. *Notoriété de l'offreur : réputation et confiance*
2. *Complétude : se suffit à elle-même*
3. *Rareté : peu d'organisations sont capables de la fournir*
4. *Fiable : le reflet de la réalité*
5. *Organisée : présentation stable, utilisable facilement*
6. *Fréquence et délai : rafraîchissement*
7. *Pérennité : sera-t-elle encore produite dans l'avenir ?*
8. *Volume, dont l'historique*
9. *Compréhensible et documentée*
10. *Dépendance : données cruciales pour l'acquéreur*

… / …

Quelques points et questions à vérifier pour vendre ou acheter des jeux de données :

- *Conseils au du vendeur : Quelles données puis-je vendre ? A quel prix ? Comment packager mes offres de données ?*
- *Conseils au de l'acheteur : Quelles données acquérir ? Auprès de quelles sociétés, à quel prix ? Comment démarcher des sources de données ? Comment organiser le partage de données ?*
- *Conseils techniques et préparation de la donnée : Quelle est la qualité requise pour vendre une donnée ? Quels processus pour stabiliser et pérenniser ma production ou l'acquisition de données ?*
- *Accompagnement sur les règles de transfert de données à caractère personnel, en France, en Europe, et dans le reste du monde.*

3.2. Exemple de service : « Mise à disposition de données à l'écosystème »

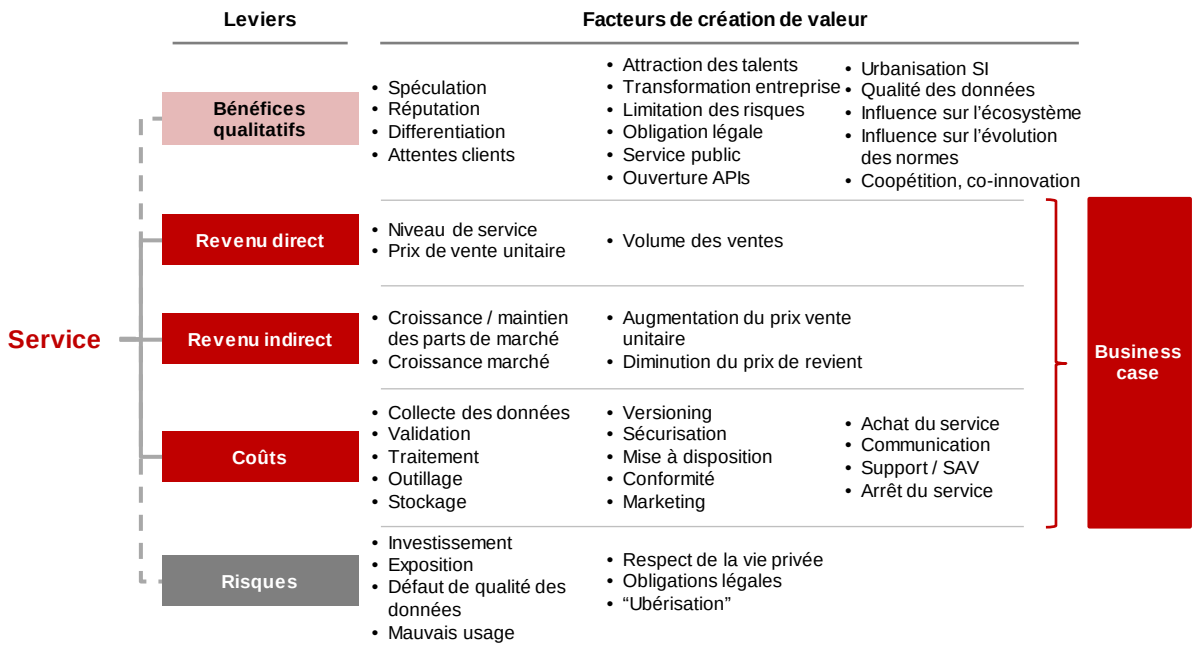
3.2.1. Le modèle

Le groupe de travail a identifié quelques **cas d'usages « génériques »** qui auraient pu être étudiés : ressources humaines ou finances. Cependant, ces cas d'usages ne sont pas apparus comme les plus intéressants.

Le cas d'usage « **Mise à disposition de données aux clients/à l'écosystème »** a été jugé comme étant le plus pertinent à traiter, car toutes les entreprises sont confrontées à cet objectif de partager des données et à leur valorisation possible aujourd'hui. Ce cas d'usage a donc été le second chantier du groupe de travail, avec pour objectif **d'objectiver la valeur des données, et de formaliser quelques bonnes pratiques**, au travers d'un modèle de valorisation qui prenne en compte des critères tant quantitatifs que qualitatifs.

Il est important de noter ici que ce modèle « Mise à disposition des données à l'écosystème » concerne :

- Des données que l'on ouvre aux clients, tant en interne qu'en externe ;
- Des entreprises qui se situent autant en *BtoB* qu'en *BtoC*.



Modèle de mise à disposition des données à l'écosystème (source : CIGREF, Octobre 2016)

3.2.2. Explication du modèle

Dans ce cas d'usage, le groupe de travail a structuré les différentes étapes clés nécessaires à une démarche de valorisation d'un service basé sur la mise à disposition de données à son écosystème.

Trois critères apparaissent comme fondamentaux : il s'agit de définir l'objectif du service, et d'identifier les éléments du revenu attendu (direct et indirect) et les éléments de coût.

- **1er critère : l'objectif ➔** savoir à quel objectif répond le service : en quoi crée-t-il des opportunités ? Veut-on développer un nouveau *business* ? Répondre à une obligation de marché ou à un objectif de fidélisation client ? Veut-on développer le *business* traditionnel ? Challenger la qualité et la production des données ? Développer l'innovation avec l'écosystème (gain par réciprocité) ? Dans l'esprit de l'*Open Innovation*, cherche-t-on à travailler plus facilement avec les *start-ups* (la donnée devient une monnaie d'échange qui facilite le travail) ?
- **2ème critère : le revenu ➔** pour déterminer la valeur du service, il faut travailler sur le revenu que l'on espère en tirer : quel type de revenu direct ou indirect ce service peut-il générer ?
    - Direct : on peut proposer différents niveaux de services, établir un prix de vente unitaire, un volume des ventes, etc.
    - Indirect : ce service peut permettre de développer les parts de marché, mais on peut aussi jouer sur le prix de vente du cœur de métier pour équilibrer les coûts engagés pour ce service.

- **3ème critère : le coût ➔** s'agissant des coûts, on peut choisir de créer et vendre tout le service, de la collecte au marketing, ou alors d'acheter une partie du service à l'extérieur (du traitement à la mise à disposition) et de le refacturer. On peut également décider de ne pas vendre ce service : si l'on développe le *business* traditionnel, le service sera clairement rentabilisé.

Ce service s'accompagne des **bénéfices qualitatifs, mais aussi de risques** comme on peut le voir sur le modèle.

- **S'agissant des bénéfices,** il peut s'agir d'efficacité, de conformité, de préservation des actifs, d'éléments en termes d'image pour l'entreprise, etc. Autant de bénéfices qui ne sont cependant pas directement quantifiables et qui se mesurent généralement *ex-post*. Par ailleurs, la mise à disposition des données personnelles nécessite des précautions, c'est pourquoi l'aspect *privacy* est important à intégrer.
- **S'agissant des risques,** il s'agira du risque « de ne pas faire » (« ubérisation » par d'autres acteurs extérieurs/des concurrents), mais aussi du risque d'utiliser des données de mauvaise qualité (qui conduirait à une destruction de valeur) ou encore du risque d'un mauvais usage des données, notamment personnelles (qui mettrait l'entreprise hors la loi et conduirait à une perte de clients). Il y a également un risque financier de non-conformité, d'où la dimension « limitation du risque » dans le modèle : avec le règlement européen sur la protection des données personnelles, les sanctions financières reposeront sur le chiffre d'affaires annuel mondial (jusqu'à 4%).

Dans le modèle, « (revenu direct + revenu indirect) – coûts » permet de déterminer le *business case* d'un tel service. On peut imaginer différents niveaux d'offre du service : offre *freemium* (service gratuit à un premier niveau) / offre *platinum* (pour accéder à d'autres services, l'offre devient payante). Il s'agit d'une démarche classique de calcul de ROI cependant, quand il faut acheter une infrastructure, le prix unitaire doit être valorisé pour l'ensemble des services. Il faut donc travailler sur la **mutualisation possible du service** pour valoriser les services associés.

Par ailleurs, il existe une dimension spéculative dans une telle démarche de type *Open Data* : on ne sait pas nécessairement quel service sera viable, et s'il y a des chances qu'il génère du revenu, personne ne sait avec certitude quelle forme il prendra. Il est donc difficile d'apprécier le revenu réel que l'on pourra tirer, d'où la nécessité de l'évaluer même si cette estimation est approximative et pleine d'incertitudes.

3.3. Quelques prérequis à la mise en œuvre de ce service

Pour mettre en œuvre un tel service, il faut faire preuve de **pédagogie et développer la « culture donnée » dans l'entreprise :** savoir comment parler de la donnée aux Métiers et les acculturer, expliquer le besoin d'ouvrir les Métiers pour qu'ils partagent leurs données en interne et casser les silos de traitements de données. Il s'agit de mutualiser et de « transversaliser » les données, et donc de repenser l'organisation interne pour favoriser la fluidité de la collecte et du traitement des données. Cette réorganisation interne pose la question cruciale de l'identification des besoins en compétences.

La mise en œuvre d'un tel service impose aussi de lever des freins :

- **Juridiques et culturels** d'abord, avec la peur du mauvais usage des données par les autres utilisateurs. Il y a une réelle crainte d'être pénalisé en ouvrant ses données ; il faut donc définir des règles du jeu, pour que l'intérêt commun soit préservé.
- **Organisationnels** ensuite. Il faut mettre en place des règles de gouvernance des données, partagées par les différentes lignes de métiers, impliquant notamment un partage des responsabilités sur la qualité et le respect de la vie privée.

4. QUELLE GOUVERNANCE POUR VALORISER LES DONNEES ?

4.1. Insuffler la « culture donnée »

Parler de valorisation des données et mettre en place une gouvernance associée amènent à faire le postulat que le *top management* considère que les données sont un actif stratégique de l'entreprise. La donnée brute a une valeur, mais celle-ci aura tendance à augmenter avec les traitements effectués, la rendant d'autant plus stratégique et appelant la mise en place d'une gouvernance adaptée, qui induit une classification, des règles, des propriétaires (présents partout dans l'entreprise), en bref, un processus de validation des données. Ceci suppose, en amont, de construire un dialogue avec les Métiers pour définir ce qu'est, selon eux, une donnée qui a de la valeur. Le besoin d'usage de la donnée est guidé par le Métier (besoin fonctionnel).

Exemples de pratiques

- Questionner les Métiers sur les données, en définissant les 10 questions clés à leur poser pour qualifier la valeur.
- Identifier avec les Métiers ce que l'on veut faire de la donnée, se projeter dans son usage, organiser la transversalité pour casser les silos, lever les incohérences dans le traitement du cycle de vie de la donnée et responsabiliser les propriétaires.
- Faire le tri dans la diversité des données exploitées.
- Identifier les usages de la donnée : de façon macroscopique, on peut distinguer trois types d'usages :
  - Finalité de communication,
  - Finalité business,
  - Finalité études.
- Mettre en place un dispositif de contrôle de la qualité de la donnée tout au long de son cycle de vie.

La donnée au cœur de la Transformation Digitale d'Orange

Placer la valorisation des données au cœur de l'entreprise implique des changements de culture et de mode de fonctionnement qu'il faut accompagner. Orange a mis en place un programme global de transformation digitale, avec un volet spécifique sur le « Big Data[3] » comprenant :

- *Un sponsorship appuyé du Président-Directeur Général, relayé par plusieurs membres du comité exécutif.*
- *Une acculturation de tous les métiers pour que l'approche s'enracine durablement, au travers d'un « Passeport Digital » et d'un « Visa Big Data » pour tous (examen*

[3] Le terme est à prendre sous son acception la plus large.

culturel accessible en ligne à destination des dirigeants et collaborateurs) ainsi que d'une « Journée de sensibilisation » pour les managers. En 3 ans, 80 000 collaborateurs ont obtenu leur « Passeport Digital », 3 000 l'ont complété avec le « Visa Big Data *» et 1 500 décideurs ont été sensibilisés.*

- *Une incitation répétée à l'exploration transverses des données afin d'identifier de nouveaux cas d'usage pertinents pour soutenir la promesse d'une expérience incomparable pour les clients.*
- *L'accompagnement agile de tous les Métiers qui le souhaitent, par une équipe transverse d'experts (notamment des* Data Scientists*) afin de co-développer les premiers cas d'usage et procéder ainsi au transfert opérationnel de compétences.*
- *L'accès aisé à des moyens et plates-formes communalisées, pour que chaque Métier accède à moindre coût aux meilleurs outils et que la « technique » ne soit pas un obstacle.*

Un changement culturel très fort impulsé par le Big Data *chez Orange a été l'acceptation du « droit à l'échec » (*try fast, fail fast*) pour libérer les énergies et s'autoriser à progresser pas à pas (*think big, start small, act fast*) dans ce monde plein de promesses mais recelant nombre d'inconnues. Les impulsions du* top management *sur ce point ont été libératrices pour les collaborateurs, notamment face à la culture « planification et qualité » des opérateurs de télécom.*

Au bilan, ces mesures ont permis aux différents Métiers de se forger une conviction par l'exemple, sans besoin préalable d'investissements majeurs, sur la base de projets ou PoC simples qui ont donné d'excellents résultats. Elles ont en outre permis de tester et mûrir progressivement les modes de travail agiles ainsi que la collaboration renforcée entre les équipes Métier et techniques, démontrant ainsi à la fois leur efficacité et leur compatibilité avec la culture d'entreprise.

Cette transformation s'est diffusée rapidement au sein du Groupe et touche désormais tous les métiers : à titre d'exemples, 40 projets opérationnels « de conviction » ont été ainsi co-développés depuis début 2016 (sans compter ceux développés en propre par chaque Métier déjà convaincu).

4.2. Repenser l'organisation et responsabiliser les acteurs

En termes d'organisation, les solutions sont diverses, allant de la mise en place d'un département *Data* chargé d'organiser la transversalité, les flux d'échanges, la modélisation et fournir aux Métiers un terrain de jeu favorable aux expérimentations, à la désignation des propriétaires chargés d'assurer la qualité du référentiel de données, en passant par la nomination d'un *Chief Data Officer* (CDO). Cela dit, seul, le *Chief Data Officer* ne peut rien faire : il doit avoir des interlocuteurs dans les Métiers, voire un département dédié, avec des *Data Scientists* notamment, pour faciliter la diffusion des pratiques dans l'entreprise, ainsi qu'un *sponsoring* fort au niveau du *top management*. Gartner a mis en évidence dans une

publication récente la dimension éminemment politique du poste et a proposé une *roadmap* pour les 100 premiers jours du CDO dans l'entreprise.

Focus sur le positionnement des* Data Scientists *: exemple d'organisation chez Orange

Chez Orange, une équipe centralisée de Data Scientists *(DS) a pour objectifs de :*

- *Co-développer avec les Métiers les premiers cas d'usage afin d'assurer la diffusion du savoir-faire, pour que ces Métiers puissent ensuite le faire par eux-mêmes,*
- *D'apporter une aide méthodologique et algorithmique dans l'analyse des données (c'est cet aspect qui relève de la « science »)*
- *De développer des outils automatisant les analyses, afin que les DS ne soient justement plus indispensables aux métiers (outils de* Machine Learning*, typiquement).*

Ce modèle permet notamment d'avoir à la fois une équipe centralisée pour ces aspects et des équipes spécialisées par Métier pour les analyses à faire au quotidien.

Certaines entreprises disposent d'équipes mutualisées de *Data Scientists,* dont le rôle est de développer des premiers cas d'usages, à partir de l'exploitation de toutes les données de l'entreprise, sans aboutir nécessairement à de nouveaux *business models.*

Les principes fondateurs de la* Data Governance *chez Air France KLM

La DSI a beaucoup œuvré pour expliquer, convaincre, et accompagner les Métiers, avec l'objectif de tirer de la valeur, et savoir ce qu'il faut mettre en place en termes de process pour garantir le niveau de qualité, l'hygiène de la donnée. C'est le Métier qui doit assumer ses responsabilités sur ce point-ci.

Il n'y a pas aujourd'hui de Chief Data Officer *chapeautant l'ensemble des métiers du Groupe, c'est un choix délibéré : Air France KLM n'en a pas ressenti la nécessité, d'autant que la stratégie est de diffuser une approche* bottom-up*. Cependant, si l'entreprise évolue vers une approche « data-centrique » par domaine, il sera sans doute pertinent de désigner des* Data Stewarts*. Un* ***CDO n'aura du sens que si l'approche « data-centrique » existe déjà dans chaque domaine****. Là, il y aura une vraie nécessité* ***d'incarner cette gouvernance transverse de la donnée*** *(via un CDO ou autre).*

Deux domaines Métiers, particulièrement sensibles au Big Data*, ont développé leurs premières applications :*

- ***La connaissance client*** ➔ *amener des données en provenance de plusieurs systèmes opérationnels distincts dans une plateforme* Big Data *(les mettre en un endroit unique et les restituer à chaque acteur qui est à proximité du client).*
- ***La maintenance aéronautique*** *et en particulier* ***la maintenance prédictive*** ➔ *exploiter les données qui sont soit dans les systèmes de maintenance, soit dans les*

avions. La maintenance des flottes d'avions produit des champs de données permettant de faire de la maintenance prédictive.

Il y a un important travail de pédagogie à faire pour sensibiliser les équipes Métiers. Il est également nécessaire de faire se rencontrer les Data Stewarts *afin que la « culture donnée » se diffuse.*

Pour qu'une démarche de valorisation des données puisse se mettre en place et que la « culture donnée » se diffuse dans l'entreprise, il semble indispensable qu'un responsable (*CDO* ou autre), haut placé dans la hiérarchie :

- Orchestre la démarche à tous les niveaux de l'entreprise et dans tous les Métiers,
- Ait la responsabilité globale de la gouvernance et de la stratégie sur les données.

La gouvernance des données mise en place chez SCOR

Une initiative pour la mise en place d'un Data Management *global a été engagée au sein de la DSI, pour structurer les différents volets : gouvernance, architecture, référentiels et* Master Data*, méta-données, qualité,* Data Analytics *et* Big Data*, protection des données, développement,* Database Operations*.*

Par ailleurs, une organisation spécifique a été mise en place pour garantir la qualité des données, à la base pour répondre aux attentes des régulateurs, mais de fait, à des fins opérationnelles :

- *Au niveau Groupe, un* Chief Data Quality Officer *(CDQO) pilote l'assurance qualité des données, et s'appuie sur des relais dans chaque ligne de* business*.*
- *Dans chaque Métier, des* Business Data Governance Lead *(BDGL) pilotent, pour le Métier, la vérification de la pertinence des données tout au long de leur cycle de vie. Le BDGL s'appuie, pour chacun des processus Métiers, sur des* Process Owners *et des* Data Quality Engineers*.*
- *Une organisation opérationnelle est en charge des contrôles effectifs :* Control Performers*,* Control Assessors *et* Control Approvers*.*

Ce dispositif est complété par un Data Protection Officer *(DPO), garant de la protection des données personnelles devant le régulateur.*

Il faut également disposer de comités *ad-hoc*, qui se réunissent régulièrement et qui rassemblent différents acteurs (propriétaires des référentiels, représentants du Métier, *Data Scientists*, responsables SI, etc.) pour examiner des indicateurs d'un tableau de bord de pilotage et mesurer la qualité des données, la cohérence du catalogue et sa complétude, l'ouverture des référentiels, et pour faire émerger et exprimer les besoins. Une gouvernance adaptée doit aussi s'interroger sur les rapports Investissement/Coût/Valeur (coûts/gains – dépenses ou pertes évitées) dans ces comités *ad-hoc*.

La sensibilisation du management, des salariés, des dirigeants chez EDF

La sensibilisation des managers, des salariés, des dirigeants est un levier essentiel pour atteindre l'ambition du projet ***CAP2030 d'EDF*** *sur les données, à savoir « des données [qui]* ***sont gérées comme un actif stratégique majeur, mieux partagées dans le Groupe, et créent de la valeur*** *: nouveaux services, optimisation des processus ».*
Deux initiatives parmi d'autres :

- *Un engagement personnel des membres du COMEX sur la base d'une campagne de communication vers tous les salariés. Un slogan* ***"la donnée ne s'use que si l'on ne s'en sert pas"*** *porté par un des Directeurs exécutifs et présent sur des affiches largement diffusées dans les locaux et sur le Réseau Social d'Entreprise VIVRE EDF* Online
- *Une* ***Politique de la Donnée*** *adoptée au niveau COMEX qui commence par* ***10 questions simples*** *à destination des managers :*
  1. *Je connais les données produites par mon entité*
  2. *Je sais comment classer les données utilisées par mon entité et d'ailleurs je connais les classements pertinents (DCP, ICS, OIV) et ce qu'ils induisent comme traitement*
  3. *Je sais quelles données générées par mon entité peuvent être/sont partagées en interne et pourquoi*
  4. *Je sais quelles données mon entité récupère/souhaite récupérer (de l'interne de l'entreprise, de l'externe de l'entreprise) et sous quelle convention*
  5. *Je sais quelles données générées par mon entité peuvent être/sont publiées à l'externe et pourquoi*
  6. *Je sais qui est le garant de la qualité et de la sécurité des données utilisées par mon entité*
  7. *Je sais qui prend la décision de publier une donnée produite par mon entité*
  8. *Je dispose de règles écrites concernant la durée de conservation de mes données*
  9. *Je dispose de règles écrites concernant l'usage qui peut être fait des données qui sont traitées par les équipes sous ma responsabilité*
  10. *Je connais les outils disponibles pour partager les données en interne et vers l'externe*

4.3. Gérer le cycle de vie des données pour mieux innover

Définir une hygiène de la donnée, basée sur des processus et des méthodes de validation de la donnée, permet de faire vivre des référentiels garants de la qualité des données qu'ils contiennent, et ce tout au long de leur cycle de vie.

La question qui se pose alors, en termes de gouvernance notamment, est celle de l'ouverture des référentiels de données pour pouvoir alimenter/constituer le *Data Lake* (ou « lac de

données ») : on peut imaginer qu'en bout de chaine du processus, un observatoire compétent (ou une autorité) certifie la donnée sortante vers les parties prenantes.

En synthèse, gouverner les données peut être représenté comme suit :

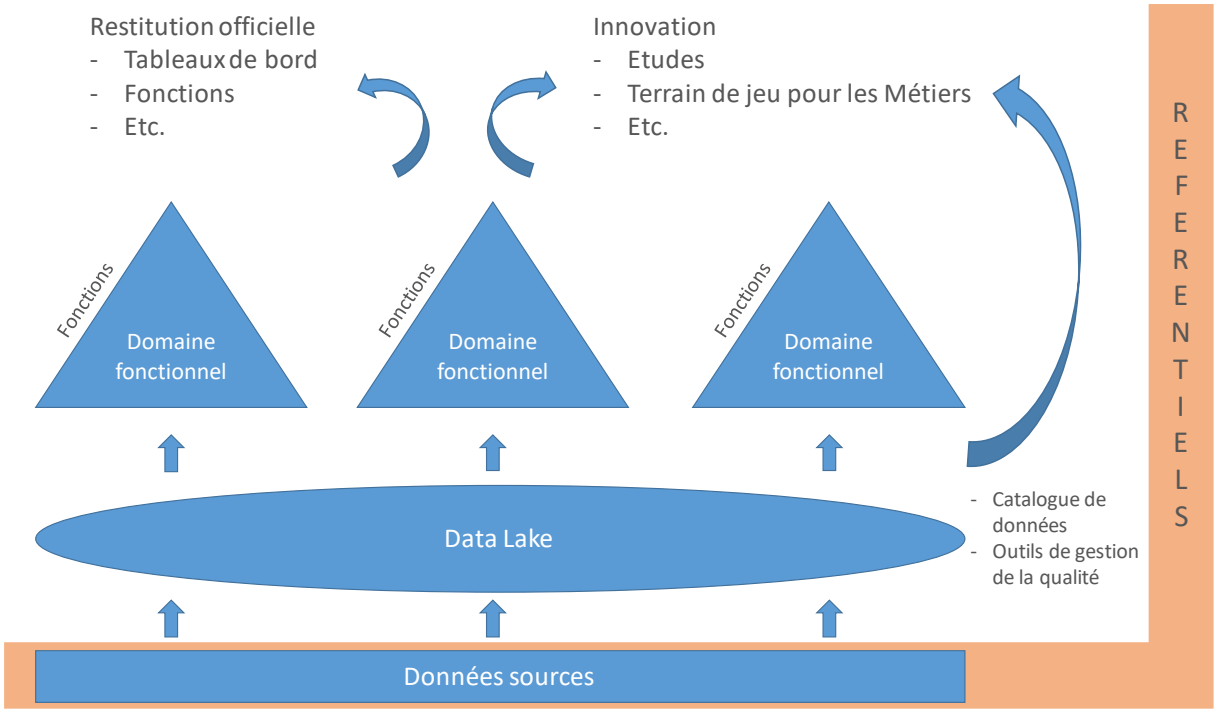


Schéma de gouvernance des données (source : CIGREF, Octobre 2016)

Explication du schéma

- Brique de base « Données sources » : stocker toutes les données dans un *Data Lake* pour une exploitation immédiate ou différée (principe : à ce stade, il s'agit de tout stocker et d'évaluer plus tard ce qui est effectivement exploité, en fonction des usages identifiés avec les Métiers) ;
- Brique intermédiaire « *Data Lake* » :
  - Définir des règles et mettre en place des outils pour raffiner les données ;
  - Organiser le *reporting* ;
  - Faire du *Data Mining*.
- Tout au long du processus, il faut garantir la qualité de toutes les données, même si elles n'ont pas toutes le même degré de criticité. On peut imaginer une gouvernance minimale sur les données contenues dans le *Data Lake* et responsabiliser chaque Métier sur la qualité de ces données pour aboutir à un catalogue de données de qualité et, créer le terrain de jeu propice avec les Métiers pour créer de la valeur à partir d'usages divers qui peuvent être la communication, les études ou le *business*.

L'approche* Big Data *chez Orange

Les aspects techniques du Big Data *(plates-formes, volumes, passage à l'échelle, etc.) peuvent paraître complexes et un peu effrayants, au point qu'ils monopolisent parfois l'attention, au détriment des questions fondamentales de valorisation de données. Pour couper court à cette difficulté, l'approche d'Orange a consisté à régler de façon anticipée et décisive tous les aspects techniques du* Big Data *pour permettre ensuite aux Métiers de se concentrer sur la valeur à tirer des données.*

Pour cela, un modèle unique, partagé et « clé en main » d'architecture complète de Big Data *a été mis au point, exclusivement sur la base d'outils Open Source, puis progressivement déployé sur les plates-formes du Groupe et des filiales. Ce modèle, enrichi tous les 6 mois, englobe l'ensemble des préoccupations techniques : architecture, moteurs d'analyse de données, outils, interfaces, solutions, évolution capacitaire mais également exigences de sécurité, recommandations sur les modes de développement d'analyses et d'applications ou catalogue des compositions architecturales pour une bonne intégration dans le système d'information existant. Il est utilisé aussi bien pour des environnements d'exploration (pour les PoC ou le développement) que de production.*

Cette « pile technologique » a permis d'accélérer la diffusion des savoir-faire de valorisation de données, en assurant cinq aspects essentiels :

- *L'absence de perte de temps et d'investissements peu matures en cherchant des solutions auprès des différents éditeurs ou en ingénierie interne.*
- *Le fait que « la technique (ou l'IT) n'est pas un problème » et qu'il faut se focaliser sur les besoins, les usages et les opportunités Métier.*
- *La réutilisation aisée par les Métiers ou filiales des cas d'usage développés par les uns ou les autres (grâce à l'interopérabilité native des plates-formes).*
- *L'industrialisation aisée et le passage à l'échelle des cas d'usage.*
- *Ce dernier point est particulièrement important et il faut anticiper la montée en puissance des plates-formes ainsi que les ressources nécessaires : les premiers développements faits par chaque Métier (sous la forme de PoC, par exemple) amènent très rapidement beaucoup de valeur et, le plus souvent, un vif désir « d'en bénéficier tout de suite en production ». Si cette industrialisation ne se fait pas rapidement, la déception qui s'en suit est en général à la hauteur de l'effet Waouh initial.*

La Fonction SI doit être capable de créer des univers d'études pour les Métiers, pour leur permettre d'affiner les expressions des besoins, de faire des tests sur la base de jeux de données disponibles, dans des espaces de jeu adaptés : proposer des outils aux Métiers, avec un cadre léger pour leur permettre de « jouer » localement mais avec une gouvernance globale minimale, pour que les initiatives ne se dispersent pas trop dans tous les sens. Les

outils peuvent être des outils de *Data Vizualization*, de *Data Science,* de *BI* et de *Big Data* suivant l'expertise de l'utilisateur (utilisateur « lambda », *Data Scientists*, développeurs).

La* roadmap Data Management *d'Air France KLM

La Direction Data Management *d'Air France KLM a défini une* roadmap *autour de 5 blocs :*

- ***Une architecture* Customer Data** *: concevoir et mettre en place l'architecture IT pour qu'elle réponde aux besoins opérationnels ou R&D.*
- ***Une organisation* Data Centric** *: « évangéliser » les Métiers, leur faire comprendre l'intérêt qu'ils ont à utiliser les outils* Big Data*. Il s'agit également de définir dans chaque Métier des personnes qui s'approprient les données et les responsabilités associées (gestion des accès, qualité, protection de la vie privée…).*
- **Ethics as a driver** *: être garant de l'utilisation éthique des données clients. Les juristes sont impliqués dès le départ, pour rendre les services parfaitement transparents pour les clients.*
- ***Créer de la valeur à partir de la* data** *: une équipe IT en recherche opérationnelle forme des* Data Scientists *qui travaillent avec le* Business *pour développer de nouveaux modèles.*
- ***Qualité des données et* data management** *(projet de* Data Glossary*, de nettoyage de bases, etc.).*

Des cas d'usages sont mis en place comme par exemple un ***Moteur de recommandation de destination :*** *avant, les promotions étaient communes pour l'ensemble des clients. Aujourd'hui, en analysant l'historique des voyages clients, on obtient des résultats bien plus efficaces (clics et taux d'achat) que les promotions habituelles.*

4.4. Faire de l'éthique un facteur de différenciation

Certaines entreprises ont fait de l'éthique un avantage concurrentiel, avec des démarches relativement développées en la matière, accompagnées d'une communication externe robuste pour expliquer aux clients finaux que le respect de la vie privée des clients prime sur le *business*. Certaines entreprises ont ainsi défini des règles, en interne et en externe, qui dictent la posture à avoir vis-à-vis des données clients, comme par exemple refuser de commercialiser leurs données. D'autres ont élaboré des guides, avec la contribution active de leurs clients.

Mais il semble que sur ce sujet de l'éthique, les GAFA communiquent mieux auprès du public puisque selon une étude récente, les citoyens français font plus confiance aux GAFA qu'à leurs opérateurs nationaux en matière de protection des données personnelles. Il y a donc une vraie difficulté des entreprises à communiquer **vers l'extérieur et à faire savoir ce qu'elles font pour protéger les données personnelles de leurs clients**.

C'est une question urgente puisqu'aux Etats-Unis, en Russie, et dans de nombreux autres pays, les réglementations ou exigences commerciales se sont récemment durcies et imposent de stocker localement les données personnelles des ressortissants du pays, conduisant potentiellement à recréer des *Datacenters* localement.

La confiance est un argument de vente : dans certaines entreprises, il existe un *Data Protection Officer (DPO)*, dont un des rôles est de joindre l'éthique des usages au potentiel stratégique des données. Il faut faire en sorte que cette éthique soit une « culture de simplification de la complexité ». Mais l'idée selon laquelle l'éthique s'oppose au *business* est très prégnante. On développe toujours une « approche conformité », et non une « approche client », contre l'idée du *« privacy's good for business »*.

Cette confiance peut-elle être une marque de différenciation sur le marché ? On peut répondre oui, parce que les entreprises européennes n'ont plus le choix : elles sont obligées de se démarquer des GAFA. Et cela n'est pas gagné car nous progressons plus lentement que les GAFA du fait de nos règles plus contraignantes en Europe.

Exemples de pratiques :

- Développer une politique interne de gestion/exploitation des données dans l'entreprise et la faire connaitre auprès des clients finaux.
- Faire de la confiance un facteur de différenciation concurrentielle.
- Garantir la non-utilisation des données clients à mauvais escient.

Vers une gouvernance de la stratégie des données d'Orange

En 2014, Orange a défini 10 règles d'or pour formaliser sa posture d'opérateur de confiance, respectueux de la vie privée de ses clients. Cette première approche très accessible était bien entendu destinée à ses clients mais aussi et surtout au Groupe lui-même, afin que chaque collaborateur puisse facilement comprendre les enjeux et repérer au quotidien « la ligne rouge » à ne pas franchir dans la valorisation de données.
Cette étape a permis au Groupe d'approfondir les enjeux liés à la définition d'une stratégie des données, d'identifier les différents cas nécessitant un approfondissement de la posture, de mieux comprendre les difficultés engendrées par une valorisation holistique, tous Métiers confondus, des données de l'entreprise et de confirmer la nécessité d'une gouvernance de cette stratégie au plus haut niveau avec une capacité de résonnance sur l'ensemble du groupe.

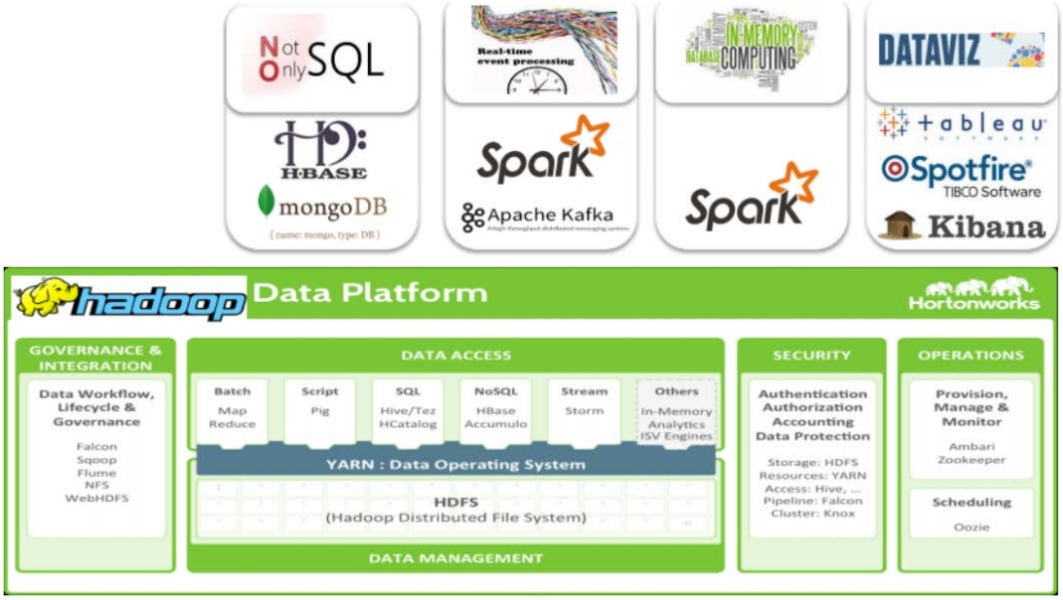
Au bilan, Orange a délibérément décidé de ne pas nommer de Chief Data Officer et a préféré mettre en place en 2016 un comité de Stratégie et de Gouvernance des Données, directement piloté par le Comité Exécutif, afin de conduire la déclinaison opérationnelle de la stratégie, de maîtriser l'impact des évolutions réglementaires (via un Group Data Protection Officer) et d'accélérer la valorisation transverse de ses données en interne.

4.5. S'outiller

Le *Cloud*, qui a baissé drastiquement les coûts de stockage, permet désormais de stocker des quantités importantes de données, pas nécessairement structurées, dans des plateformes telles que Hadoop : on peut y déposer plus de données, diverses et nombreuses, et diversifier et élargir les périmètres d'analyse.

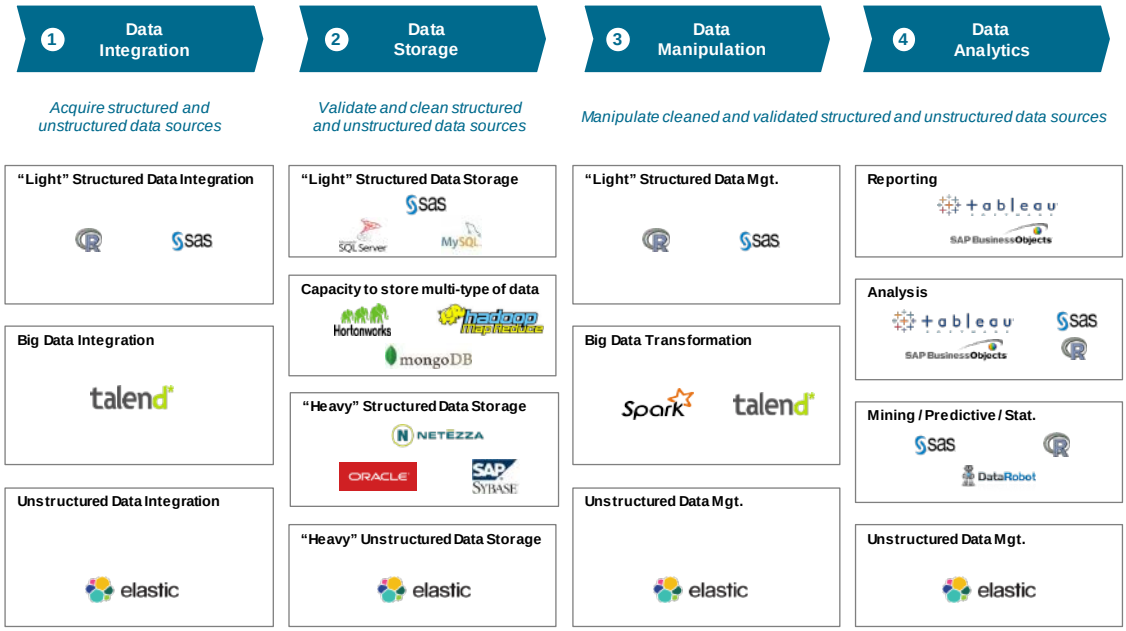
Les pans technologiques :

- Les technologies de *Data Sources* (*databases*, applications, émergence de places de marché des données, …) permettent de collecter les données ;
- Les technologies de *Data Wrangling* (enrichissement, traitement, référencement, sécurisation, …) permettent d'organiser et de gouverner les données ;
- Les technologies de *Data Applications* (*insights, models,* …) permettent d'analyser les données ; les modèles d'analyse permettent quant à eux d'orchestrer et d'opérationnaliser les traitements.



Exemple 1 d'outillage pour la gestion et l'analyse des données
(source : Air France KLM pour le CIGREF - Février 2016)

Par ailleurs, de nouvelles plateformes de gestion des données émergent, avec des bases de données analytiques et des applications, des outils d'*in-memory analytics*, *in-memory extreme processing*, les moteurs de recherche et NoSQL.



Exemple 2 d'outillage pour la gestion et l'analyse des données
(source : SCOR pour le CIGREF, Septembre 2016)

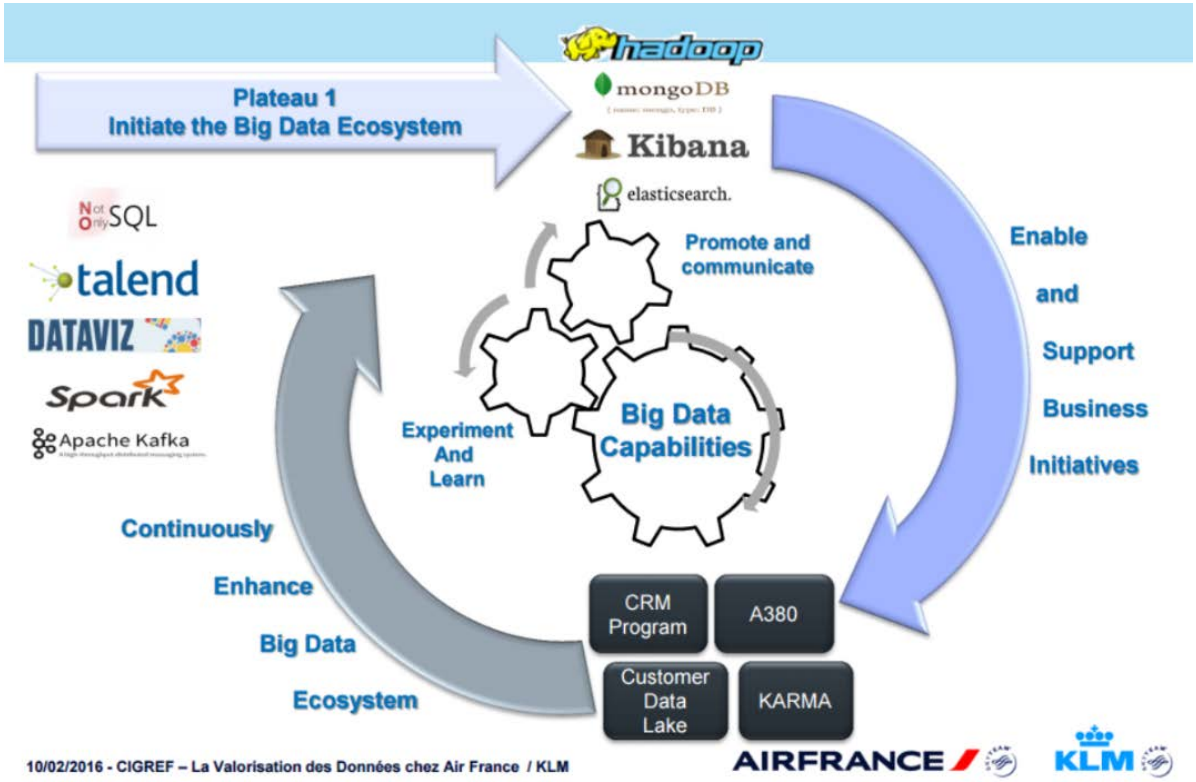
Le point sur Hadoop vu par Talend (***www.talend.com***)

C'est un peu le « couteau suisse » du traitement de données, qui a fédéré un environnement complet sur une plateforme capable d'intégrer le volume et la variété des données, ainsi que la vitesse des traitements. Il est à noter que les technologies, les acteurs de cet écosystème, mais aussi les besoins clients évoluent tellement vite que « choisir Hadoop » impose pour les entreprises de « remettre l'ouvrage sur le métier » en permanence, pour être à l'état de l'art.

Et dans Hadoop, Spark semble être l'écosystème le plus dynamique du moment. Pour autant, il semble que le choix dominant aujourd'hui pour le Big Data *soit Hadoop, même si beaucoup de clients sont encore en phase de pilotes. La maturité est encore naissante sur Hadoop et l'enjeu réside dans l'expertise (il est difficile pour une entreprise de s'engager rapidement sur Hadoop du fait des expertises que ces technologies requièrent).*

Exemples de pratiques

- Pour tout projet de valorisation des données, il est nécessaire de définir un cap et un cadrage, les outils n'étant qu'un moyen au service d'une fin : attention aux boîtes noires (*deep learning*), il faut remettre de l'intelligence humaine dans l'intelligence des outils.
- Etre en capacité de resituer son *business* et de définir une stratégie par rapport à l'outillage.
- Adopter une démarche itérative, car le paysage technologique d'une part, et les besoins d'autre part, évoluent très vite.



Exemple d'outillage Big Data (source : Air France KLM pour le CIGREF - Février 2016)

5. VALORISATION DES DONNEES : CAS D'USAGE ET EXEMPLES DE PRATIQUES

L'analyse des données ouvre la possibilité de développer des services complémentaires. L'internet des objets ouvre par exemple de nouveaux champs pour le *business* : les données recueillies par les capteurs vont permettre de mieux développer les produits et de mieux répondre à la demande.

La finalité de la valorisation des données pour la SNCF

La valorisation des données permet notamment :

- *De faire de **l'optimisation des processus internes**. Ce qui apporte le plus de valeur au groupe SNCF est la supervision, l'identification des perspectives et l'optimisation pour améliorer les processus internes :*
  - ***Mieux superviser l'activité** : suivre la qualité des voies, le remplissage des trains au quotidien (là où les réservations ne sont pas obligatoires) ;*
  - ***Identifier des perspectives** : si on ne supervise pas bien les activités, il est difficile d'identifier les perspectives (meilleur ciblage des interventions de maintenance, adaptation de la capacité offerte par exemple) ;*
  - ***Optimiser** : optimiser certains processus internes (programmation d'opérations de maintenance grâce aux capacités prédictives, allocation de ressources, etc.) ;*
- *De faire de **la rentabilisation externe** par la monétisation et de réfléchir à la **métamorphose de l'entreprise** en créant d'autres activités liées aux données et à l'expertise métier associée (*Rail Solution *ou* Kisio Digital *par exemple)*

L'analyse des données permet aussi de réduire les coûts. Par exemple, beaucoup d'entreprises sont confrontées à la difficulté de contrôler leurs factures. Venant de nombreux fournisseurs (parfois plusieurs dizaines de milliers), étant très complexes à déchiffrer (certaines peuvent avoir plusieurs millions de lignes), ces factures sont bien souvent payées sans avoir été contrôlées. Diverses études (AT Kearney et Ardent Partner par exemple) montrent que la plupart des entreprises sont surfacturées de 10% en moyenne sur leurs achats, sans qu'elles s'en aperçoivent. Or, les récentes avancées en matière d'Intelligence Artificielle et de *« Machine Learning »*, fondées entre autres sur l'agrégation de toutes les données et expertises de l'entreprise dans une base de connaissance, permettent :

- Dans le domaine des Achats, d'automatiser les processus de suivi et de contrôle, depuis la lecture et l'extraction des données à valeur ajoutée (chaque ligne d'une facture est analysée et contrôlée) jusqu'au paiement ;
- Dans de nombreux autres domaines, d'automatiser toutes les tâches administratives afin d'optimiser les dépenses.

Cette approche digitale moderne de l'utilisation et de la valorisation des données est à la base de la transformation de l'entreprise. Elle s'inscrit également dans une démarche collaborative entre entreprises. En effet, des données non stratégiques peuvent être partagées dans un

environnement totalement sécurisé (entraide sur des sujets administratifs, sur des questions de règlementation, auto-apprentissage de règles de gestion, partage de connaissance d'une problématique rencontrée, échange de bonnes pratiques, etc…).

Utilisation de l'Intelligence Artificielle pour mieux contrôler les factures Energies chez ENGIE

ENGIE a ajouté de l'Intelligence Artificielle à sa plateforme de services pour le bâtiment intelligent « VERTUOZ by ENGIE », en intégrant la technologie et les plateformes d'un prestataire (Dhatim). Engie peut exploiter ainsi l'ensemble des données de facturation de ses clients utilisateurs (collectivités territoriales, résidentiel collectif, tertiaire), afin d'assurer un contrôle automatique de mise en cohérence des factures avec les consommations réelles des clients. Ces derniers peuvent suivre en temps réel la bonne application de leurs contrats, des règles d'acheminement, de la mise en œuvre des résiliations ou changements de puissance par exemple et générer des économies sur le paiement de leurs factures d'électricité, de gaz ou d'eau qui peuvent atteindre jusqu'à 10 %.

5.1. Faire évoluer les modèles d'affaires

5.1.1. Usages et approches : exemples sectoriels

Dans le secteur de la banque-assurance, les données sous-tendent des *business models* classiques qui reposent sur l'observation des marchés, mais avec la voiture sans conducteur par exemple, l'assurance sera incluse dans le service. C'est un axe de développement pour les *business models* mais aussi un facteur de risque. La cybersécurité est également un axe de développement des *business models* dans ce secteur : l'enjeu est de **valoriser les données afin d'affiner les connaissances et de mieux estimer les risques,** à partir d'historiques de sinistralité notamment.

Dans la distribution, l'usage des données, et du *Big Data* au sens large, permet de **faciliter le travail des opérationnels** (prévoir des stocks en parallèle, travailler à partir du traitement des informations en temps réel…). Il faut développer une vision 360° de tous les objets métiers, tant sur les produits, que sur les stocks.

L'évolution des *business models* autour des données se concrétise avec **l'apparition de nouveaux acteurs, qui militent pour la désintermédiation**. Pour certaines entreprises, la digitalisation vise avant tout à éviter *l'ubérisation*, et implique que ces entreprises se mettent en situation d'entrer directement en contact avec le client final, d'ouvrir de nouveaux services à tous les nouveaux utilisateurs.

Comment faire face à ces nouveaux concurrents, aux GAFA et NATU[4] notamment, qui captent de nombreux clients finaux des entreprises ? Les grandes entreprises doivent pouvoir offrir les mêmes interactions avec leurs clients que ces *pure players*. Il s'agira par exemple de développer de nouveaux services autour des objets connectés, d'y intégrer des éléments d'algorithmiques complexes, de miser sur le marketing pour apporter des services particuliers aux clients, d'échanger des informations au sens large, même sur la *supply chain*, de partager de l'information sur les magasins, en passant par les fournisseurs, et en intégrant l'immédiateté/le temps réel.

Dans le domaine des services, les transactions faites par cartes hébergent des **informations que certaines entreprises essaient de monétiser,** en créant de nouveaux *business models* basés sur l'interaction directe avec les clients finaux. Certaines entreprises sont ainsi plutôt dans les prémisses d'approches de valorisation des données.

Dans le **secteur des médias par contre, tous les métiers ont déjà un usage stratégique des données.** La vraie question pour ces acteurs est : comment réinventer le positionnement de l'entreprise par rapport aux GAFA ?

Quant au **secteur de l'énergie, on observe que les grands axes de valorisation reposent sur diverses pratiques :**

- Décloisonner les métiers, croiser les données internes ;
- Développer la démarche de transparence imposée par la loi : pour les clients et les régions, il y a un devoir de transparence et d'interactions avec les communautés, administrations locales ou collectifs associatifs ;
- Publier des données mais aussi des algorithmes à tout l'écosystème ;
- Développer les démarches également en Lab, se rapprocher des start-ups…

Exploiter les données pour mieux piloter les consommations d'énergie à la SNCF

La SNCF est l'un des premiers consommateurs d'énergie en France et cela représente des enjeux économiques très importants. L'objectif est d'arriver à 20% d'économie sur les consommations d'ici 2022 : comment la donnée peut-elle aider à progresser sur le sujet et à mieux piloter les consommations ?

La SNCF déploie aujourd'hui des compteurs sur l'ensemble des flottes, notamment avec les compteurs à bord des trains (un peu comme sur les Linky à domicile) pour avoir des données détaillées sur la consommation d'énergie. Cela permet de mettre en relation des données de consommation avec des données de production, en s'appuyant sur des outils de data visualisation *(train en circulation, train en préchauffage, etc.).*

[4] NATU : Netflix, Aibnb, Tesla, Uber

Cette thématique illustre également la métamorphose possible grâce à la valorisation des données SNCF ***: les trains consomment mais produisent aussi de l'énergie****. La SNCF travaille sur* ***la récupération de l'énergie*** *(réutilisation, stockage, …) avec des partenaires pour construire des écosystèmes énergétiques intelligents (la gare comme « hub » énergétique). Mais les investissements sont colossaux et pour faire cela, au-delà des partenaires, la première chose est de partager des données, de faire des maquettes numériques avec les partenaires (exemple : avec versaillessmartgare.com).*

Valoriser les données consiste à les transformer et à imaginer de nouvelles manières de faire (ou ne plus faire) les choses : la vraie valeur est dans cette transformation des métiers, ce n'est plus seulement un sujet technique.

Dans le secteur des télécoms, **la valeur des données est généralement poussée par l'exécutif**, à travers les questions **des plateformes et de la valeur que l'on peut tirer des données pour le *business*,** en fonctionnant par boucles de rétroaction.

5.1.2. Démarches de valorisation

La création/le développement de nouveaux *business models* à partir de l'exploitation des données peut aussi reposer sur la création de nouvelles alliances stratégiques, entre acteurs de secteurs différents : un opérateur avec une banque par exemple. La banque devient l'opérateur le plus naturel (*m-banking*, paiement avec téléphone portable, nouveaux systèmes de sécurité : possibilité de couper un distributeur automatique à partir de la géolocalisation du téléphone).

Les démarches de valorisation des données dans l'industrie semblent quant à elles encore assez embryonnaires. Certaines démarches récentes visent à identifier les avantages à tirer de la donnée, des technologies et à développer des *use cases*. Quelques projets stratégiques commencent à éclore, qui poussent au « dé-silotage » des données, tels que :

- ➔ L'optimisation du produit, par l'analyse fine des usages réels des clients, qui peut permettre de détecter les fonctionnalités réellement utilisées et celles qui ne le sont pas et dont on pourrait faire l'économie,
- ➔ La gestion de la qualité avec une détection anticipée des problèmes, une plus grande réactivité dans le traitement et un meilleur ciblage des opérations de rappel,
- ➔ La maintenance préventive et prédictive des installations industrielles, avec des gains potentiels importants sur les taux de disponibilité,
- ➔ Le pilotage des flux logistiques.

La finalité de ces démarches est de valoriser les gains rapides, mais il n'y a pas encore de réelle organisation autour de la *Data*. De façon générale, s'agissant de la maintenance préventive et prédictive, et du pilotage des flux logistiques, nous pouvons penser que l'on est à la veille d'une transformation majeure du monde des usines.

La maintenance prédictive sur l'A380 – Air France KLM

Il y a un réel ***enjeu stratégique autour de la propriété de la donnée*** *(opérateur, constructeur…). En vol, plusieurs centaines de milliers de capteurs engrangent des données sur l'A380 et celles-ci sont* downloadées *une fois que l'avion a atterri. Les données ainsi mises à disposition des opérateurs permettent de mieux voir comment tous les composants sont liés les uns aux autres. Et grâce à l'accès aux données, les opérateurs vont beaucoup plus vite dans la compréhension d'une panne. Cela améliore la réactivité et limite les impacts sur les opérations. En réalisant un PoC sur les circuits carburant (qui sont souvent la cause des retards sur les avions), les équipes ont ainsi déjà évité plusieurs annulations de vols.*

En plus des enjeux industriels cités précédemment, l'automobile voit apparaître un nouveau domaine lié à la multiplication des données captées, celui des services connectés. Le *business model* traditionnel persiste, mais d'autres émergent : celui des services connectés et celui de la mobilité *« as a service »*, qui eux sont totalement différents et encore largement à inventer. Le domaine des services connectés change la relation avec le client, car au lieu d'un *business* centré sur la vente, puis sur des retours ponctuels en atelier pour des opérations d'entretien, les services connectés supposent l'entretien d'une relation permanente avec le client. Cela nécessite de revoir le modèle de relations entre les acteurs qui est structuré depuis des décennies, entre le constructeur et son réseau de distribution : pour vendre des services connectés, il faut que le réseau y soit intéressé et formé. Il y a également un potentiel d'évolution dans le fait qu'en tant qu'opérateur de services, le constructeur est susceptible d'être en relation directe avec le client sans passer par l'intermédiation du réseau. Il y a également des questions dans la relation avec les équipementiers qui sont eux-mêmes susceptibles d'équiper leurs composants de capteurs : est-ce qu'ils doivent jouer leur propre jeu en proposant par exemple des services de maintenance, ou rechercher une coopération gagnant-gagnant pour offrir un service *via* le constructeur ?

5.1.3. Défis identifiés

- ➔ Développer des alliances stratégiques avec des partenaires dans d'autres secteurs ;
- ➔ Créer de nouveaux *business models* ciblant directement le client/bénéficiaire final ;
- ➔ Décloisonner les métiers, casser les silos en mettant en place une organisation *ad hoc* ;
- ➔ Eviter l'*ubérisation*, notamment pour les entreprises en *B2C* qui devront elles-mêmes développer un réseau de fournisseurs de services en *B2B2C* pour éviter de se faire « intermédier » par des opérateurs de type Uber ou GAFA ;
- ➔ Identifier les nouveaux acteurs et les tendances du marché ;
- ➔ Grâce au partage d'informations avec les partenaires, développer une vision 360° : clients, produits, stocks, etc. ;
- ➔ Capter les données en temps réel ;
- ➔ Exploiter les démarches Lab/start-up ;

- Organiser des boucles Métiers/IT - IT/Métiers, avoir des équipes mutualisées de Data Scientists … ;
- Croiser toutes les données.

5.2. Améliorer la connaissance des clients, de l'écosystème, des partenaires

Traditionnellement, les entreprises sont très orientées client, mais ces 3 statuts peuvent être considérés globalement.

5.2.1. Développer une vision 360° des clients

Avoir une vision transverse d'un client est un des enjeux que pose le numérique. Beaucoup d'entreprises cherchent à développer cette vision. Certaines déploient des programmes centralisés, qui visent à « transversaliser » les données entre les Métiers, *via* une entité dédiée ou une branche numérique quand elle existe. Dans ce cadre, le CDO (*Chief Data Officer*) peut être un atout pour l'entreprise, tant pour développer cette vision 360 que pour développer la connaissance de l'écosystème, car il a l'avantage d'avoir à la fois un pied à l'externe et un pied à l'interne, et peut traiter ainsi des problématiques et des besoins très différents.

La donnée au centre de l'expérience client chez Air France KLM

*La vision d'Air France KLM est « **d'être la compagnie aérienne la plus** **customer centric** **d'ici 2020 ».** Pour cela, l'entreprise a considéré que sa différenciation pouvait se faire sur la personnalisation et l'interaction avec le client. Toutes les informations récoltées peuvent permettre à Air France KLM de proposer des services au client pour qu'il personnalise son vol, sans que cela soit intrusif.*

L'environnement « Data » d'un client est très vaste : données clients classiques (données personnelles, les données Frequent Flyers, ...) mais aussi des données externes, plus ou moins personnalisées, sur les médias sociaux, qui peuvent enrichir la connaissance client (Social data, Third party data…) des équipes Marketing notamment. Toutes ces dimensions peuvent être traitées par des technologies Big Data pour avoir des résultats pertinents. Le travail de récupération des données étant facilité en amont, il est possible pour les équipes de se concentrer sur la valorisation de ces données, sur les use cases et sur la mise à disposition des données aux Métiers.
L'ambition est de mettre la « customer intimacy » au cœur de chacune des interactions avec les clients. Il faut pour cela savoir restituer ces données sur les différents points de contact : aéroports, agents, points de ventes et d'informations, canaux digitaux (emailing, targeting, media sociaux, promotions, sites internet...) et également pour les clients eux-mêmes, par exemple via leur smartphone.

Dans la distribution, le *business* s'oriente naturellement sur la connaissance client, mais aujourd'hui, l'enjeu de qualification porte surtout sur **les données de produits**. Il y a là un enjeu de technologie pour mieux connaitre les produits, qui passe par tous les champs d'informations existants. Il faut croiser les informations avec les volumes de produits, et le volume de clients, afin d'obtenir des éléments pour offrir des nouveaux services aux clients.

Dans le domaine des services, la simplification du **parcours administratif des clients/usagers** sur internet est un enjeu, et l'exploitation des données peut être un levier : auto-remplissage de questionnaires à partir des données qui ont déjà été confiées. Ces démarches de services pragmatiques valorisent le parcours client.

D'autres entreprises témoignent être surtout dans une **démarche d'apport de nouveaux services aux Métiers,** grâce au développement d'une vision 360 des clients (en *BtoC* mais aussi en *BtoB*) sur la base de déploiement de systèmes de CRM.

Un cas d'usage emblématique ayant permis de convaincre et progresser rapidement : les* livebox *foudroyées chez Orange

La prévention des Livebox *endommagés par les impacts de foudre lors des orages s'est révélée particulièrement efficace pour déclencher la prise de conscience et propager au sein de l'entreprise l'intérêt pour la valorisation de données. Ce cas d'usage, construit par enrichissements successifs, a en effet permis d'illustrer de façon très pragmatique les différentes étapes de maturité dans la valorisation des données :*

- ***Etape 1 - Descriptive** : détecter* a posteriori *les* Livebox *foudroyées à partir de l'analyse des différentes données disponibles, pour proposer leur changement aux clients concernés.*
- ***Etape 2 - Exploratoire** : localiser les* Livebox *foudroyées et les comparer avec les cartes d'impact de foudre (météo).*
- ***Etape 3 - Prédictive** : identifier les motifs répétitifs (impliquant l'intensité des orages, les réseaux électriques, les modèles de* Livebox*, etc.) et prédire ainsi quelles sont les* Livebox *qui risquent d'être foudroyées.*
- ***Etape 4 - Répétable** : répondre efficacement à ce qui pourrait arriver en faisant appliquer une procédure spécifique «* Livebox *foudroyée » à chaque détection de ces motifs (procédure pour centre d'appels, par exemple, évitant d'avoir à répondre à des questions inadaptées au contexte).*
- ***Etape 5 - Continue** : automatiser ce qui pourrait arriver en prévenant automatiquement chaque client concerné dès détection des motifs récurrents, afin de lui conseiller de débrancher sa* Livebox.

Outre les gains financiers et opérationnels, ce cas emblématique et pédagogique a permis d'améliorer le Customer Net Promoter Score *de 40 points et a joué un rôle clé pour démontrer de façon simple et pragmatique l'intérêt majeur de la valorisation de données à travers le Groupe.*

5.2.2. Intégrer le temps réel

Les exigences en matière de temps réel sont différentes d'une entreprise à l'autre, car elles répondent à des besoins et à des objectifs distincts : traitement de réclamations clients, optimisation d'un réseau de distribution d'énergie ou d'un réseau de transport, analyse d'un sinistre et prise en charge d'un assuré, etc. Certaines entreprises font une **distinction entre « temps interactif » et « temps humain »,** qui peut aller de quelques secondes à quelques mois. **Le temps réel n'est par ailleurs pas une donnée systématiquement nécessaire :** pour des analyses de consommations, certains Métiers préfèrent travailler à partir de données statiques. Dans le domaine des médias, la même problématique se pose avec la publicité en temps réel sur les médias en ligne : les Métiers travaillent encore assez peu avec des outils de visualisation en temps réel.

L'*IoT*[5] et le temps réel : si l'on en croit les projections, il y aura 80 milliards d'objets connectés en 2020. Comment se préparer à cela ? Il y a d'abord un seuil technologique à franchir : le plus dur est de **transporter** les données, pas de les stocker, **au travers des relais**. Ensuite, la difficulté est de réussir à comprendre les informations générées par chacun des relais. Il n'est pas possible d'analyser chaque phénomène, donc il faut savoir arrêter la montée du flux quand on sait qu'on n'en aura pas besoin, ou savoir redescendre le traitement quand on en a besoin.

Pour certaines entreprises, **le temps réel est un élément essentiel dans la gestion de crises,** sur des cas de cybersécurité notamment, mais aussi pour analyser les dégâts d'une catastrophe naturelle ou pour gérer l'e-réputation sur les réseaux sociaux par exemple.

La donnée et le temps réel au cœur de la stratégie numérique de RTE

Les transitions énergétique et numérique font naitre de nouveaux comportements (les consommateurs deviennent des consomm-acteurs), de nouveaux usages (avec les objets connectés) et permettent l'optimisation de moyens locaux avec la création d'éco-quartiers et les smart-grids*. Le système électrique est plus ouvert, diversifié et décentralisé. Le numérique et la gestion des données sont au cœur des enjeux, et notamment la sécurité des SI.*

La stratégie numérique de RTE repose sur trois initiatives :

- ***Le Hub numérique,*** *lieu de rencontre de plateformes qui permet la gestion en temps réel de l'offre et de la demande d'électricité. Grâce à ce Hub, RTE propose de nouveaux services basés sur l'exploitation des données en temps réel.*
- ***Le* Mall *numérique,*** *socle des services aux clients : il regroupe l'ensemble des parties prenantes du réseau électrique. Il est à l'image d'un centre commercial dans lequel*

[5] *Internet of Things*

on ne trouverait que des données : l'idée est de mettre à disposition des données et des services (les produits sont donc les données), avec une grande variété de parties prenantes (clients utilisateurs du réseau, acteurs du système électrique, acteurs de l'économie numérique, start-ups, universitaires, écoles).

- ***Le Lab numérique,*** *guidé par le leitmotiv selon lequel « il faut être proche de l'extérieur pour rester dans la course ». Au sein du Lab numérique, RTE organise des concours de start-ups, développe des partenariats, lance des appels à compétences et R&D interne afin de mobiliser tous les ressorts de l'innovation numérique. Le groupe RTE travaille en collaboration avec un fond d'investissement (Truffle Capital) pour financer les entreprises innovantes.*

5.2.3. La qualité des données et l'approche « référentiels »

La qualité des données est fondamentale pour avoir une bonne connaissance client. Il existe encore de nombreux silos (CRM, *middle office…*) et de nombreux référentiels qu'il est difficile de mutualiser dans les entreprises. Il faut alors **initier des approches « référentiels »** pour aller vers une modélisation unique et fiable.

La valorisation des données oblige l'entreprise à revoir toute cette chaîne de valeur comprenant les données de fournisseurs, de divers intermédiaires, etc. Pour cela, **de nombreuses entreprises construisent des référentiels clients dont la finalité est la conception de services différents selon les clients.**

5.2.4. Défis identifiés

- ➔ Développer une vision transverse du client ;
- ➔ Connaître les données « produits » et « usages » pour construire des stratégies marketing et commerciales adaptées aux besoins des Métiers ;
- ➔ Améliorer la qualité des données grâce à une bonne gestion des référentiels ;
- ➔ Croiser les données en temps réel ;
- ➔ Accéder aux données des partenaires : achat, IoT… ;
- ➔ Garder un contact permanent avec le client ;
- ➔ Développer des applications pour remonter des informations en temps réel ;
- ➔ Adapter les applications traditionnelles au temps réel ;
- ➔ Veiller à l'image de marque de l'entreprise grâce au temps réel ;
- ➔ Prendre en compte l'interopérabilité – notamment avec les données d'IoT ;
- ➔ Prendre la cybersécurité comme *use case* pour travailler le temps réel.

5.3. Créer de la valeur et la partager

Le partage de la valeur permet d'améliorer l'écosystème dans son ensemble. Certaines entreprises soulignent la difficulté qu'il y a à partager la valeur : elles arrivent à signer des conventions mais c'est un long chemin, qui manque de souplesse.

5.3.1. La redistribution

Entre entreprises : Dans le domaine de l'énergie, où le tarif est régulé, le partage de la valeur pose les questions suivantes : est-ce que la valeur ajoutée est publique, distribuée de manière égale à chacun, ou personnalisée ? Le client est aussi une source d'optimisation sur le réseau (il peut libérer de l'espace réseau pour un autre client), mais le partage de cette information avec un fournisseur est aussi possible. Dans le domaine des services, certaines entreprises nouent des partenariats avec des acteurs de la banque pour développer leurs offres : chaque acteur trouve une valorisation à ce nouveau service proposé en commun.

Au client final : Il y a ici l'idée d'un contrat de confiance avec le client qui met à disposition ses données : que gagne-t-il en retour ? Une *start-up* s'est développée autour de cette idée : un client donne des informations qui sont exactes et il est rémunéré sous forme de monnaie d'échange (bitcoins par exemple) par les entreprises intéressées. Derrière l'échange de la valeur des données, il y a la construction de la confiance. Une entreprise propose certains de ses services par API, ce qui pousse les personnes à accéder à des données par le biais d'une sorte de « contrat de confiance ». L'objectif est ici est de créer un canal d'échange de données *via* une entreprise tiers de confiance.

A la collectivité : Derrière l'idée du partage de la valeur se trouve aussi celle de la création d'écosystèmes à l'échelle régionale : une entreprise qui a une relation étroite avec une région particulière et qui a des activités avec des *start-ups* et/ou des collectivités locales, se voit confier des données qui ne sont pas disponibles en *Open Data*. L'objectif ici est de détecter des entreprises partenaires intéressantes, déjà à un niveau local.

Le partage de la valeur : comment le Groupe SNCF fait vivre tout un écosystème

Le Groupe SNCF tire beaucoup de valeur des données disponibles à l'extérieur de l'entreprise : par exemple les données de constructeurs, les données de Météo France, des satellites, les données du marché de l'électricité. Pour l'entreprise, c'est une révolution industrielle qui dépasse la seule valorisation stricto sensu *: il est aussi question de faire de l'exploitation du trafic et de la mobilité personnalisée (clients et chargeurs), de la maintenance prédictive des infrastructures (équipements au sol et matériels roulants), ou encore de l'optimisation des consommations d'énergie.*

*Inversement, SNCF partage une partie de toutes les données disponibles et est l'un des précurseurs sur l'*Open Data *en France. Le site* https://data.sncf.com/ *met à disposition de*

nombreuses données sur des événements de sécurité, des données de réparations, sur des objets trouvés, le trafic en temps réel, des données LIDAR (scan en 3D de l'environnement qu'il y a autour des rails SNCF ; ces dernières peuvent être intéressantes pour des fournisseurs).

En partageant certaines données, la SNCF contribue à faire vivre un écosystème, même si elle sait que cela peut potentiellement la désintermédier.

5.3.2. Une approche possible du partage de la valeur

5.3.2.1. PARTAGE DES DONNEES ET CREATION DE VALEUR

Pour certains acteurs, le partage de la valeur est nécessairement pensé à partir d'une stratégie *business* : dans le secteur des télécoms, un opérateur peut refuser de vendre ses données clients. Dans le secteur de l'automobile, un constructeur peut au contraire accepter de partager des données anonymisées grâce à des algorithmes d'anonymisation. Mais cela est cependant susceptible d'affaiblir la valeur par rapport aux données en clair. **Ce qui crée de la valeur, c'est le croisement de données qui permettent de rattacher des informations (profession, catégorie socio-professionnelle, âge, statut, habitudes de consommation, …) à une personne, en temps réel, pour lui proposer des biens et services adaptés.**

Ainsi, au niveau d'un écosystème, dans le domaine des transports par exemple, les données d'un opérateur télécom, et celles mises à disposition en *Open Data* par des acteurs du transport ont *a priori* peu de valeur lorsqu'elles sont prises isolément. Mais **le croisement de ces données, issues de fournisseurs différents, peut créer de la valeur (optimisation du trafic en temps réel, information des usagers/clients, …)**.

Dans le domaine de la réassurance, tous les acteurs du métier se sont rassemblés pour dématérialiser les échanges : l'espérance de gain est une réduction de 30% des coûts administratifs associés, et cette valeur est partagée par l'ensemble des acteurs de la réassurance.

Pôle Emploi lance en 2016 un programme intitulé « Valorisation des données »

Fort de son positionnement central au sein du service public de l'emploi, Pôle Emploi a décidé d'exploiter les données collectées au sein de ses processus métier pour améliorer ses services auprès des demandeurs d'emploi et des entreprises.

Deux dimensions sont couvertes au sein du programme « Valorisation des données » :

- *L'analyse des données s'appuyant sur les technologies issues du* Big Data
- *Et l'ouverture des données à travers des initiatives* Open Data *et* Open API.

Des projets concrets ont été menés en 2016 par des auto-entrepreneurs fonctionnant en mode start-up et sélectionnés dans le cadre de challenges organisés par Pôle Emploi. Les services ainsi développés comme « La bonne boite » et « La bonne formation » utilisent des données du SI et ont été mis à disposition des demandeurs d'emploi à travers le site emploi-store.fr.

Sur le plan de l'ouverture des données, Pôle Emploi utilise son site institutionnel pole-emploi.org, pour mettre à disposition certaines de ses données en Open Data *(statistiques du marché du travail, le ROME, les offres d'emploi, …). Le site emploi-store-desv.fr (ESD) correspond au portail d'accès aux API mises à disposition par Pôle Emploi. Ces API sont aujourd'hui utilisées pour permettre la création de nouveaux services par des développeurs tiers. Ces services sont ensuite référencés par Pole Emploi et mis à disposition des usagers sur emploi-store.fr. L'ESD, c'est aussi une plateforme «* Open API *» qui sera bientôt opérationnelle pour favoriser l'interopérabilité du SI avec l'écosystème de Pôle Emploi (partenaires institutionnels ou privés).*

2016 a été l'année des premières mises en œuvre et a permis de démontrer l'importance d'une démarche de valorisation des données. 2017 doit être l'année du passage à l'échelle notamment dans le cadre du plan stratégique Pôle emploi 2020.

5.3.2.2. LES ENJEUX DU PARTAGE DE DONNEES EN INTERNE

Pour de nombreuses entreprises, le partage de la valeur se fait d'abord en interne : elle se partage entre les Métiers, autour de la **notion de « bien commun ».** Par ailleurs, la création de valeur peut se faire avec du *machine learning* mais la valeur n'est pas que dans la corrélation. Au-delà des machines, il y a aussi des **partages de valeurs fonctionnels** qui sont identifiés, comme la connaissance des applications des équipes entre elles par exemple.

Par ailleurs, la culture *Open Source* peut présenter un avantage en matière de partage des données, puisqu'il fournit des codes accessibles et réutilisables par tous les autres Métiers. Et quand toutes les filiales d'un même groupe font le même métier, les collaborateurs ont tout intérêt à partager le code *Open Source*. Mais, en fonction du métier de l'entreprise, il peut y avoir une peur intrinsèque à chaque Métier qui est de « se faire voler son *business* par celui d'à-côté ». Ainsi, dans la réalité, les données mises à disposition par chaque acteur sont

rarement complètes, pour limiter le risque, ce qui justifie une **mise à disposition de données partielles, et fondamentales, pour mener les expérimentations**. Ceci permet de lever le frein au partage qui peut être par exemple la crainte de se faire voler un client, non pas la mise à disposition elle-même.

Optimisation des processus internes : exemple dans les Ressources Humaines chez CNP Assurances

Dans le domaine des Ressources Humaines, l'exploitation des données issues de la paie est un moyen efficace pour optimiser les processus et réduire les coûts, en suivant et contrôlant mieux les frais de déplacements, l'absentéisme, les cotisations sociales, les bulletins de paie, etc… Sur ces points, la mise en place de la Déclaration Sociale Nominative (DSN) représente une opportunité pour les entreprises de poursuivre leur transformation numérique, puisqu'elle repose sur la transmission unique, mensuelle et dématérialisée des données issues de la paie (soit 370 données par salarié) et la transmission dématérialisée de signalements d'événements aux organismes sociaux. C'est également par ce biais que le fisc et les employeurs vont partager les données nécessaires au prélèvement à la source de l'impôt sur le revenu à compter du 1er janvier 2018. D'où l'obligation de les contrôler avant de les transmettre aux organismes sociaux afin d'éviter tout risque de redressement. Le « Machine Learning » dans ce domaine permet de proposer des nouvelles règles de contrôles aux responsables de paie qui choisissent ou non de les activer. Grâce à cette approche, la machine auto-documente et auto-contrôle les règles de paie. La donnée se valorise comme génératrice de règles métiers et aide les responsables de paie à détecter des anomalies.

CNP Assurances contrôle ainsi ses Déclarations Sociales Nominatives avant envoi aux organismes sociaux. Au total, ce sont plusieurs centaines de milliers données qui sont contrôlées chaque mois, avec le prestataire Dhatim, et au-delà de la DSN, c'est la digitalisation de tout le contrôle de la paie de CNP Assurances qui est réalisée, ce qui permet in fine *d'en améliorer la qualité et d'en réduire le coût.*

5.3.3. Défis identifiés

- ➔ Créer des canaux d'échange de données ;
- ➔ « Désiloter » au sein de l'entreprise ;
- ➔ Créer un service en retour de la mise à disposition de données ;
- ➔ Identifier les besoins d'un écosystème régional ;
- ➔ Ouvrir les API à des partenaires/créer un portail de partenaires ;
- ➔ Mettre en commun la valeur entre plusieurs contributeurs ;
- ➔ Redistribuer la valeur aux clients qui partagent des données : « contractualiser » la fidélité client.

6. CONCLUSION

Si la donnée est bien reconnue comme un actif stratégique clé par les entreprises (évoquée régulièrement dans certains Comex), leur maturité en matière de valorisation des données reste pourtant très inégale. Cela soulève des enjeux importants, que ce rapport a tenté de mettre en évidence, en partageant quelques méthodes et bonnes pratiques :

- Le questionnaire d'auto-évaluation de la maturité des entreprises qui permet à chaque entreprise d'identifier ses points forts et ses axes d'amélioration en matière de valorisation des données ;
- Des méthodes et exemples de gouvernance et d'organisation autour de la donnée ;
- Un exemple de modèle générique présentant les différentes étapes clés nécessaires à une démarche de valorisation d'un service basé sur la mise à disposition de données à son écosystème ;
- Les quelques cas d'usage qui ont été partagés et qui sont présentés ici, s'agissant notamment de l'évolution des modèles d'affaire, de l'amélioration de la connaissance client et de l'écosystème, de la création et du partage de la valeur autour des données.

Le sujet de la valorisation des données, ses enjeux, mais aussi sa maturité encore insuffisante illustrent bien l'ampleur de la révolution engagée, à laquelle aucune industrie, aucune entreprise ne pourra échapper. Dans de nombreuses grandes organisations déjà, ces masses de données disponibles sont exploitées par les technologies de de l'Intelligence Artificielle, telles que le *Machine Learning* et le *Deep Learning*, en vue d'innover en permanence, pour optimiser l'usage et la capacité d'impact de la donnée dans le développement du *business* de l'entreprise :

- Gouvernance globale ;
- Partenariats ;
- Gestion du cycle de vie de la donnée ;
- Transversalité de l'organisation ;
- Méthodologies et outillages adaptés ;
- Ethique et conformité ;
- Nouveaux services au client ;
- Monétisation et partage de la valeur avec les partenaires ;
- Etc.

Mais toutes les innovations basées sur l'exploitation des données ne seront bonnes que si elles sont bien gérées, de qualité, partagées et que l'entreprise crée la confiance, sur la base de règles explicites et transparentes. Sur ce dernier point, le sujet le plus urgent à traiter aujourd'hui, auquel le CIGREF s'attellera dès cette année, est certainement celui relatif à la mise en œuvre du Règlement sur la protection des données (dit *GDPR : General Data Protection Regulation*). Les grandes entreprises ont certes depuis plusieurs années déjà,

nommé des CIL (Correspondants Informatique et Libertés) et mis en place des politiques de protection des données. Mais le GDPR crée des obligations renforcées en matière de protection des données personnelles pour les entreprises, notamment en matière d'analyse de risques, de mise en place de standards concrets liés à la *« privacy by design »*, de portabilité des données, etc. Toute la question pour les entreprises est de savoir comment traduire opérationnellement ce texte, en tenant compte des dispositifs déjà en place par ailleurs. Cette question doit d'ailleurs être traitée avec les fournisseurs, puisque s'agissant notamment des services génériques (messagerie, gestion des ressources humaines, relation client), ceux-ci font l'objet d'une offre logicielle large, dorénavant proposée de plus en plus souvent en SaaS, dans le cloud public, par les éditeurs.

Autant de sujets sur lesquels les dirigeants doivent faire progresser leurs organisations, et contribuer ainsi au développement de la culture de la donnée au sein de l'entreprise.

7. Annexe – Referentiel de pratiques de l'outil d'auto-evaluation

Axes	Critères d'évaluation (* signale une pratique issue du modèle DVM®)
Business Models	Des cas d'usages basés sur les données ont été déployés
	Les données sont utilisées pour se différencier de la concurrence au travers de nouveaux services
	De nouveaux *business models* et des services innovants ont été mis en place grâce à l'usage et l'analyse des données
	Des challenges développeurs sont organisés pour identifier de nouveaux services basés sur les données
	Les données (marketing, techniques, financières, RH ou autre) sont une composante essentielle et effective du développement des produits et services*
	Une forme de rétribution (*bitcoins* par exemple) est offerte au client qui partage ses données (contrat de confiance) : contractualisation de la fidélité client
	Des accords de partenariats sur le partage de la valeur existent (sous-traitants, co-traitants, fournisseurs…)
	Les objectifs du "temps réel" sont identifiés (traitements de réclamations clients, optimisation d'un réseau de distribution d'énergie/de transports, analyse de sinistre, gestion de crise, e-réputation, …)
	Des capteurs sont intégrés dans les produits de l'entreprise
	Les données sont analysées en temps réel
	Les données sont exploitées pour éclairer les décisions
Connaissance (clients, écosystème)	Des actions sont mises en place pour identifier les clients / les partenaires
	L'utilisation des données permet un accès au client final
	Une vision globale du client à 360° est développée/mise en place
	L'accès à des données de granularité plus fine est organisé et permet d'améliorer la connaissance (clients, partenaires, concurrents, …) et les processus industriels
	Des données sont collectées sur les attentes des clients et contribuent à personnaliser les services et l'expérience-client
	Au-delà des données clients, les données produits sont également exploitées (pour améliorer la connaissance des produits, des *process*, et *in fine* la connaissance des clients)
	Nous mobilisons les données pour simuler/prédire/anticiper les comportements des clients/écosystème*
	Les données sont exploitées afin d'identifier les acteurs disruptifs et prendre les mesures nécessaires pour ne pas se faire "disrupter"
	Des stratégies marketing et commerciales adaptées aux besoins des Métiers sont définies à partir de la connaissance des données clients, produits et/ou des usages
	Une surveillance des concurrents et des nouveaux entrants disrupteurs par l'analyse des données externes est en place
	Les données permettent une bonne identification des clients perdus/gagnés
	Les données externes à l'entreprise, issues de l'*IoT*, sont exploitées en interne
	Une plateforme de partage de données est alimentée au nom de l'intérêt commun et contribue à créer un nouvel écosystème
	Une veille active est assurée sur les réseaux sociaux en temps réel pour anticiper, prendre des décisions...
	Un service fondé sur les données mises à disposition a été créé par l'écosystème

Transversalité & Efficacité opérationnelle	Le décloisonnement des Métiers et des données associées est organisé
	La connaissance des outils d'analyse des données est diffusée de façon organisée au sein de l'entreprise
	La valeur de la donnée est partagée en interne
	Des ateliers et communautés de pratiques sont en place pour favoriser le travail collaboratif (mise en place d'espaces de co-working, création de FabLabs...)
	Les *datalabs* coopèrent
	La valeur des données est estimée/mesurée (monétisation, optimisation du trafic sur un réseau routier, répartition des consommations énergétiques en temps réel, meilleure information des usagers, etc…)
	Les données issues de capteurs et objets connectés internes sont collectées et analysées
	Nous mobilisons les données pour offrir des combinaisons de produits/services les plus adaptés*
	L'analyse des multiples données non-structurées internes/externes disponibles et accessibles est systématisée
	Des données sont collectées pour mesurer l'usage réel des produits et services
	Les données sont partagées avec des concurrents/partenaires pour améliorer la productivité de l'activité
	L'utilisation/la valorisation des données permet d'améliorer/d'optimiser la qualité des produits & services
	Les processus RH intègrent une perspective donnée, en particulier celle relative à la valorisation du capital humain de l'entreprise*
	Des analyses sont faites des bénéfices tirés de la mise à disposition de données
	Des alliances avec des partenaires permettent de renforcer et d'accélérer l'analyse des données
	L'analyse des données est reconnue comme un facteur d'accélération du *time-to-market*
Stratégie & Gouvernance des Données	La donnée est reconnue comme un actif clé au sein de l'entreprise et son management est organisé
	La donnée est un élément structurant de la culture de l'entreprise
	Il existe une stratégie et des politiques de valorisation des données
	La cohérence de la politique de gestion des données et sa traduction opérationnelle fait l'objet d'un suivi régulier au niveau de la DG*
	L'organisation, les rôles et les responsabilités sont clairement définis en matière de gestion des données
	Une gestion (prospective) des compétences requises autour des données est en place
	Un système d'exploration des enjeux autour des données est en place (Lab ou autre)
	Des cas d'usage tirés des expérimentations sont industrialisés
	Une démarche d'identification des sources de données non-structurées existe
	Le traitement d'informations en temps réel est fait à bon escient, et en réponse à un besoin Métier identifié
	Des organisations mixtes type IT-Métiers sont mises en place (organisation Agile, co-location des équipes...)
	Une politique d'ouverture des données est mise en œuvre en interne (*open data*/publication de données et/ou d'algorithmes)
	Une politique d'ouverture des données est mise en œuvre à l'externe (*open data*/publication de données et/ou d'algorithmes), à l'écosystème
	Un portail interne a été créé et les API ont été ouvertes (dans le cadre d'un partage de données entre différents acteurs d'un même Groupe)
	Un portail externe pour les partenaires a été créé et les API ont été ouvertes (dans le cadre d'un partage de données entre différents acteurs extérieurs à l'entreprise)
…/…	Les processus SI intègrent de manière routinière la dimension donnée, en particulier la dimension Création de Valeur*

Stratégie & Gouvernance des Données (suite)	Les exigences relatives à la qualité des données sont explicitées
	La vérification de la qualité des données est organisée (les retours sur les défauts de qualité sont gérés)
	Une logique d'amélioration continue est appliquée à la qualité des données (prise en compte de leur contexte notamment)
	Des approches "référentiels de données" sont mises en place pour garantir la fiabilité et la qualité des données
	Une démarche d'harmonisation des référentiels de données (internes et/ou externes) existe
	Les données clés sont définies et leurs attributs identifiés
	Des méthodes de préparation des données existent pour créer des ensembles de données cohérents
	Les données de référence sont utilisées en transverse par l'ensemble des systèmes
	Les risques liés à l'ouverture des données sont pris en compte
	Les contraintes réglementaires et juridiques sont connues et prises en compte (données à caractère personnel, droits de propriété intellectuelle, …)
	La sécurité des données fait l'objet d'un audit régulier
	Un processus de protection des données est en place et le niveau de protection adéquat est défini en fonction de la nature des données
	La gestion des droits d'accès interne, la contractualisation de gré à gré entre différentes directions d'une même organisation et/ou entre partenaires sont encadrés
Ethique & Conformité	Des règles éthiques sur l'usage et l'exploitation des données sont définies, partagées et appliquées en interne
	La confiance et l'éthique dans la gestion des données sont des facteurs de différenciation concurrentielle
	Les règles éthiques sont communiquées aux clients finaux et/ou partenaires, écosystème (en particulier sur l'exploitation et le traitement des données personnelles)
	Un système de gestion de crise est défini et activable en cas de fuites et/ou d'usages détournés des données
	Une démarche de *consumer empowerment* pour aider le consommateur à mieux maîtriser ses données personnelles est mise en place
	Un processus d'anonymisation des données personnelles est en place pour permettre leur exploitation
	Les règles de protection des données personnelles sont prises en compte dès la conception (*privacy by design*)
	Une démarche d'influence sur la définition et l'élaboration de normes et standards est effective auprès des régulateurs
	La localisation et l'accès aux données personnelles tiennent compte des obligations réglementaires (multi-)locales
	L'éthique n'est pas considérée comme une contrainte mais comme une source d'opportunités
	Les DPI (Droits de Propriété Intellectuelle) sont clairement identifiés et encadrés
Outils & Solutions	Des connecteurs, APIs, et autres *web services* ont été développés pour favoriser le partage des données
	Des systèmes/outils sont mis en place pour capter les données en temps réel
	L'interopérabilité est prise en compte, notamment avec les données de l'*IoT*
	Une architecture et/ou des applications *Big Data* sont en place

8. BIBLIOGRAPHIE

Rapports CIGREF

- *Entreprise 2020 à l'ère du numérique : enjeux et défis* (2014)
- *Enjeux business des données : comment gérer les données de l'entreprise pour créer de la valeur ?* (2014)
- *Economie des données personnelles : les enjeux d'un business éthique* (2015)
- *Gouvernance de l'Intelligence Artificielle dans les entreprises : enjeux managériaux, juridiques et éthiques* (2016)

Collection CIGREF

MIGLIETTI Charles – *Les données au cœur de la transformation numérique* (Nuvis, 2016)

BOUNFOUR Ahmed – *Futurs numériques – Transformation numérique : du lean à l'accéluction* (Eska, 2016)

CIGREF

21 avenue de Messine

75008 PARIS

Tel. : +33 1 56 59 70 00

cigref@cigref.fr

www.cigref.fr