

# Tasmane

## DATA / CONFIANCE

PAS DE TRANSFORMATION DIGITALE SANS DATA...  
... MAIS PAS DE DATA SANS CONFIANCE !



### LA PAROLE

Thibaut MIDON

Directeur général de Tasmane

Franck BRECHET

Manager chez Tasmane

LA LETTRE • 02

NOVEMBRE 2020

## ÉDITO

Depuis le 25 mai 2018, nous avons tous été inondés de messages nous avertissant de la mise en place du règlement européen sur la protection des données à caractère personnel (RGPD). L'Union européenne s'est en effet dotée d'un cadre de régulation pour le stockage et l'exploitation des données des citoyens et consommateurs.

Non par plaisir de complexifier l'activité des entreprises, la Commission a souhaité mettre en place les conditions de la confiance des citoyens dans la « face numérique » de l'économie. Il s'agit d'éviter des scandales préjudiciables aussi bien pour les citoyens que le développement des entreprises.

Donc beaucoup a été fait pour sécuriser les citoyens le partage et l'emploi par les entreprises de leurs données personnelles entre entreprises et citoyens, mais par ailleurs en fait-on assez pour que les entreprises, en interne ou entre partenaires, disposent des conditions qui leur permette de libérer tout le potentiel de tous types de données, d'où qu'elles viennent et non pas seulement des données personnelles ?.

LA PAROLE

## THIBAUT MIDON

DIRECTEUR GÉNÉRAL TASMANE

## FRANCK BRÉCHET

PARTNER TASMANE

### QUELQUES CONSTATS...

#### L'effet silo frappe toujours !

Pour faire fonctionner son écosystème interne de données, de plus en plus au cœur de ses modèles d'affaires et de son innovation, une entreprise doit considérer la confiance comme une pierre angulaire de ses initiatives.

Avouons-le avec un minimum d'objectivité : si l'échange et la collaboration sont des valeurs portées haut et fort dans les entreprises, la réalité possède souvent une toute autre allure. Bien que les organisations soient nécessairement le siège de nombreux échanges de données (en interne comme avec l'externe), force est de constater que les silos sont naturels et nos réflexes sont souvent peu portés en première option sur le partage.

*« Utiliser la liste des sites industriels européens du contrôle de gestion ? C'est possible mais nous ne maîtrisons pas leurs mises à jour et il y a des différences avec celle du service logistique ... ». « Nous ne nous baserons pas sur les données d'un autre service... ». « L'information, c'est le pouvoir, je la garde ! Et la partager avec d'autres va nous apporter des ennuis ! ».*

Ainsi fleurissent entrepôts de données, pseudo-référentiels et autres solutions temporaires (qui s'éternisent !), complexifiant le SI et augmentant sa facture. Malgré l'existence de méthodes et d'actions permettant de garantir la fiabilité des données, la structure même de la plupart des entreprises reste souvent un obstacle majeur. Du fait des silos organisationnels, une même information (ou groupe d'informations) se retrouve fréquemment dans plusieurs services avec des données différentes, devenant un frein aux processus internes et générant des erreurs ou des pertes – souvent sous-estimées. En plus des problèmes de doublons et plus largement de cohérence de l'information, la mise en commun de données d'entités différentes de l'entreprise n'est pas systématique, alors même que les meilleurs acteurs du numérique montrent à quel point cette démarche est porteuse de valeur ajoutée.

#### La valeur est dans l'écosystème

Outre les échanges internes, l'ouverture et la gestion de l'information au sein d'écosystèmes sont devenues primordiales. En premier lieu, nous pensons logiquement aux partenaires principaux que sont les fournisseurs et les clients. Mais dans un environnement économique toujours plus sophistiqué, les échanges d'informations avec les partenaires, les startups, voire les coopérateurs, sont tout aussi vitaux.

En effet, le partage des données dans la « chaîne étendue » est générateur de valeur ajoutée, par exemple dans les situations suivantes :

- Dans le secteur agricole et agroalimentaire, l'échange de données est à la base de l'organisation de « Filières », dont l'enjeu principal est la traçabilité.

*La coopérative Terrena a par exemple lancé Agrimatrice, une interface publique où les consommateurs peuvent consulter « l'ensemble des événements survenus dans l'histoire du produit ». Cette capacité à s'ouvrir et à échanger est bénéfique pour l'ensemble des acteurs, voire indispensable lors de détection de problème sanitaire.*

- De nombreuses entreprises et organisations ouvrent une partie de leur patrimoine de données à des partenaires, le plus souvent des startups ou des universitaires, afin d'explorer et de générer plus de valeur que les seules équipes internes ne peuvent imaginer.

*La Caisse nationale de l'assurance maladie (Cnam) et l'École polytechnique ont par exemple signé fin 2014 une convention de partenariat de R&D qui visait à favoriser le développement des technologies du big data appliquées au domaine de la santé*

- Le secteur des transports est fortement impacté par l'ouverture des données. La législation sur l'open data a généré du « combustible » pour tout un ensemble d'acteurs offrant des services basés sur des données en provenance des opérateurs historiques ou régulés. Dans le domaine public, la tendance est d'ailleurs désormais de considérer la data comme un élément facilement « consommable », afin d'accompagner cette ouverture et de favoriser la création de valeur.

### **Un effet « boîte noire » qui peut inhiber la consommation des données mises à disposition**

Une mauvaise manie est souvent d'attribuer la confiance à la donnée elle-même, en se limitant naïvement à sa fiabilité et sa qualité. Cette situation traduit souvent le point de vue des personnes qui en ont la gestion (dans la DSI ou chez les métiers) : la qualité intrinsèque de la donnée serait suffisante : « dans les cases, les valeurs sont bien les bonnes ».

Nous avons encore peu le réflexe de nous décentrer vers les usages : qui manipule, quand et pourquoi ? Comment la donnée est-elle exposée, à quel moment et à quelle granularité ? Finalement, qu'est-ce qui est le plus important : les données, ou les processus qui les supportent et les exposent ? Les deux sans aucun doute, bien que nous oublions régulièrement le second. Avec des processus non maîtrisés, le traitement à lui seul peut être source d'une perte de confiance dans la donnée.

*Le cas de Kobe Steel, entreprise majeure de l'industrie sidérurgique nipponne, est édifiant sur ce point. Florissante jusqu'en 2017, un scandale éclate suite à la découverte d'une falsification de nombreuses données produites, distribuées à plus de 500 partenaires industriels, dont des fleurons européens. Il y a fort à parier que les standards d'échanges étaient respectés (formats, flux, etc.). Pourtant, des pratiques douteuses des gestionnaires de ces informations, et une direction peu scrupuleuse tant que des marchés étaient à prendre, ont abouti à une mise en difficulté majeure de l'entreprise : crédibilité réduite à néant, chute vertigineuse de son action en bourse et départ de son dirigeant.*

Outre l'absence ou le manque de maîtrise des processus autour des données, la confiance est aussi mise à rude épreuve par la tendance à « l'algorithmisation » des entreprises et de la vie courante. Sur la base de données avec un certain niveau de fiabilité, des algorithmes complexes proposent des résultats qui se traduisent de plus en plus souvent en décisions automatiques. Ces dernières se révèlent de fait opaques et sujettes à critiques : un algorithme est moralement neutre, mais pas ses usages.



Cet effet « boîte noire », comme on l'appelle parfois, est amplifié par la sophistication des calculs. Il n'y a pas si longtemps, les calculs complexes et leurs usages étaient réservés à des populations restreintes, ayant souvent participé à leur conception. Les technologies liées au big data ont changé la donne. L'algorithmisation de masse engendre une déconnexion grandissante entre des calculs toujours plus complexes à expliquer et des populations beaucoup plus larges.

C'est d'autant plus important qu'une tendance logique est de réinjecter les résultats de ces algorithmes dans des applications opérationnelles (outil métier, application client...)

- de nombreuses technologies dites d'intelligence artificielle correspondent à cette situation.

Comment s'y retrouver ? Comment comprendre les données, bien les utiliser ? L'explication du fonctionnement interne de la « boîte noire » sera primordiale à apporter pour installer et entretenir la confiance.

La plupart des secteurs ou des fonctions des entreprises sont ou seront soumis à ce travail d'explication. Prenons le cas de la banque ou de l'assurance : les modèles utilisés pour faire du scoring de risque, par exemple dans le cadre d'une décision d'octroi de crédit, se doivent d'être complètement explicables à un auditeur de l'organisme de régulation. Les modèles les plus évolués amènent à de très forts taux de performance pouvant conduire à remplacer la décision humaine, cependant ils ne sont pas facilement explicables face à un auditeur et les variables discriminantes souvent assez peu explicites. C'est pourquoi les organismes choisissent des modèles plus simples, moins performants - ce qui amène parfois à défausser le deep learning, censé être le must technologique.

*Un autre exemple du risque « boîte noire » est celui d'une société d'exploitation ferroviaire qui utilise ses données pour prédire et anticiper les problèmes de matériels, afin de fournir un meilleur service aux usagers. Le modèle prédictif produit est utilisé pour prendre des décisions opérationnelles, engager des moyens et ajuster les plannings de maintenance des équipes.*

Un modèle trop opaque pour ce type de cas d'usage permet-il aux managers opérationnels de comprendre les effets significatifs de l'ajout de telle ressource ou de l'utilisation de tel matériel au lieu d'un autre ?

En résumé : les données sont indissociables d'un contexte qu'il nous faut comprendre, à défaut de totalement le maîtriser. Plus les traitements et cas d'usages se complexifieront, plus les questions se poseront autour de la légitimité et de la capacité d'explication du traitement. Vous ne justifiez plus seulement de la disponibilité de vos données, vous rendez aussi des comptes sur l'ensemble de la chaîne en amont. Un maillon de cette chaîne fait défaut, et votre crédibilité peut être sérieusement mise à mal, surtout si les conséquences en aval sont importantes.

## **Quatre leviers clés au service de votre transformation digitale**

La **confiance** est au cœur du partage d'information nécessaire dans les processus, et de plus en plus fondamentale dans un monde de recours accru aux algorithmes.

Comment mettre en place concrètement cette confiance ? Comment l'entretenir ?

Comme toute problématique transverse, le management de la donnée est un sujet qui se travaille avec méthode. Nous proposons quelques pistes, sous forme de « mantras » pouvant orienter les pratiques.

## 1. Pas de nouveaux usages sans connaissance des données

Qui peut avoir foi en un général qui ne connaît pas le terrain de sa future bataille ? Bien des DSI et des directions métiers, lançant des initiatives digitales, se trouvent pourtant dans cette situation peu plaisante. Les entreprises ne connaissent majoritairement pas leur capital de données, et pas non plus toujours l'ensemble des flux (fonctionnels, applicatifs) qui leur sont liés. Les données font trop peu l'objet d'une description partagée et régulièrement mise à jour (sémantique, métadonnées). De plus en plus de spécialistes s'accordent à dire que les données constituent une « infrastructure essentielle » aux organisations. Or, normalisation, cartographie et langage commun sont les principes de base de toute gestion d'infrastructure. Souvent reléguée aux urbanistes du SI, rarement associée à une valorisation métier et économique, ne faisons pas de la donnée un cas à part.

La donnée possède par ailleurs cette caractéristique de « l'inversion » : dans la plupart des projets, les besoins fonctionnels sont définis puis la solution spécifiée et développée ; à l'inverse, les initiatives « data » débutent plutôt par des données disponibles et la recherche d'usages qui les exploitent. Comment être innovant, porteur de valeur quand on ne connaît pas son patrimoine informationnel ?





## 2. Rien ne sert de cacher, mieux vaut exposer

Ainsi pourrait-on résumer une bonne pratique à appliquer en interne — en dehors des données sensibles bien entendu. Lorsque l'on veut limiter les cloisons des silos et faire émerger la valeur ajoutée, cette étape est indispensable. D'aucuns rétorqueront que l'exposition de celles-ci peut être compromise par leur qualité souvent moyenne, pas toujours maîtrisée. Prenons à contre-pied cette observation et osons rendre visible l'état des données lorsque les usages le permettent. Un nombre croissant d'entreprises prennent le parti de montrer les données telles qu'elles sont dans le SI opérationnel, par exemple via des interfaces de consolidation de sources différentes de type portail ou moteur de recherche. Si la qualité est un prérequis, l'objectiver l'est tout autant. Faisons confiance aux opérationnels (voire aux clients) pour se saisir des corrections, compléments à apporter plutôt que d'attendre LA solution (encore une !) dont la mise en place est rarement rapidement effective. Exposer les données, par exemple via des API, est souvent mobilisateur des métiers qui se sentent alors concernés par le capital qu'ils ont à gérer.

## 3. N'attendez pas pour expérimenter

Bien que les études de cadrage, les analyses de risques et de ROI aient leur nécessité, le meilleur facteur de confiance reste



le concret. La donnée se prête bien à l'itération avant l'industrialisation : pourquoi attendre le projet miracle ? À travers des moyens d'expérimentation ( « labs », « PoC », « PoV », etc.), les initiatives data doivent démontrer rapidement leur utilité : rassurer leur(s) promoteur(s), donner confiance aux partenaires en travaillant sur du tangible.

Le concret s'apporte aussi par petites touches, via des initiatives qualifiées parfois de « basiques ». Pourquoi ne pas débiter par des actions simplifiant la vie à tous ? Faire du « ménage » dans le SI legacy est souvent un bon début. Cela peut par exemple prendre la forme d'une mise à plat des nomenclatures les plus utilisées — ces tables lointainement cachées au fin fond de vos systèmes back-office dont tout le monde se sert mais dont peu se préoccupe de la gestion. Ou encore en standardisant certaines données basiques (adresses, numéros de téléphone, etc.).

#### **4. La data est l'affaire de tous**

Cela paraît une lapalissade de souligner l'importance d'embarquer les différents protagonistes sur une initiative touchant un même « domaine » de donnée (Client, Fournisseur, Produit, Tarif, etc.). Comment donner confiance lorsque qu'on oublie une partie des acteurs ? La plupart des organisations peine encore à mettre en place les outils d'une coordination transversale réellement exploitable, aboutissant à des avancées tangibles : pilotes, moyens, responsables, arbitrages. Deux parades sont souvent trouvées.

La première consiste à se décharger sur la DSI en lui donnant le rôle de décideur au lieu de celui d'animateur. Cette situation aboutit souvent à des choix de moins en moins partagés. Les projets traitant de la data sont par essence « métier ». En aucun cas la DSI ne peut porter seule l'initiative.

La seconde consiste à penser que la nomination d'un Chief Data Officer (CDO) est la recette miracle pour animer et embarquer les parties prenantes. Si la présence d'un CDO est une avancée souhaitable et à promouvoir, cette action ne suffit généralement pas à éclaircir un contexte déjà complexe.

Le travail transversal sur les données est freiné principalement par deux aspects. Tout d'abord il existe souvent des crispations sur la propriété des données : « à qui appartient le Client ? Au marketing ? Aux commerciaux ? À l'administration des ventes ? » Ensuite, ces initiatives ont pour conséquence une confrontation de points de vue sur un même concept de l'entreprise : « Qu'est-ce qu'un Produit ? Inclut-il les données d'emballage et d'étiquetage ? Inclut-il les données de R&D, de matières premières... ? ».

Songeons donc très tôt à instruire le sujet de la propriété des données et à utiliser les bonnes méthodes pour traiter les conflits de vision afin d'avancer sereinement avec de nombreux acteurs.

Mise en commun en interne, partage avec l'écosystème, mise sous-contrôle des processus, explication des traitements : et vous, qu'avez-vous engagé pour entretenir la confiance en vos données et réussir votre transformation digitale ?



RETROUVEZ TOUTES NOS PUBLICATIONS SUR:  
[WWW.TASMANE.COM](http://WWW.TASMANE.COM)