



DÉMOCRATISER LA DONNÉE POUR CRÉER DE LA VALEUR

LES CLÉS DE LA RÉVOLUTION DATA-CENTRIC

ÉDITO



MARION GARDAIS

VP Head of Data & AI
Center of Excellence
South Central Europe

La démocratisation de la donnée, enfant de la révolution *datacentric*

Quand les entreprises ambitionnent de développer l'usage de la donnée, elles ont bien souvent le réflexe de partir de la technique. On sélectionne des outils, on bâtit une infrastructure, on forme des spécialistes... et tout cela est mis à la disposition des métiers qui sont priés d'en faire le meilleur usage. C'est typiquement cette logique qui a présidé à la mise en œuvre de quantité de data warehouses et datalakes où dorment désormais, sans risque d'être dérangées, des masses gigantesques de données inutilisées.

Parce que cette approche a clairement montré ses limites et parce que les entreprises font aujourd'hui face à des enjeux plus pressants que jamais, il n'y a d'autre choix que de renverser totalement la perspective : partir des besoins des métiers, se demander dans quelle mesure et comment la donnée peut y répondre, et mettre en place l'ensemble des ressources technologiques, humaines et organisationnelles adéquates - ce que nous appelons la plateforme de données.

La mise en œuvre de la plateforme de données est donc avant tout une révolution copernicienne dans la façon dont est abordée la data au sein de l'organisation. Pour autant, les fondamentaux restent les mêmes ; ils sont seulement envisagés sous un autre angle.

On aura donc toujours à traiter les trois enjeux clés qui conditionnent la capacité à produire de la valeur à partir des données : la mise sous contrôle du patrimoine de données, considéré comme l'un des actifs les plus précieux de l'entreprise ; le déploiement d'un socle technique tirant pleinement avantage du cloud et de ses innovations ; et la mise en place des méthodes, de la culture et de l'organisation qui maximiseront la productivité et les bénéfices des projets data.

Indissociables, ces trois grands chantiers doivent tenir compte des investissements, parfois lourds, déjà consentis et de la maturité de l'organisation sur les divers sujets abordés. Et c'est en les menant de concert, au fil d'une feuille de route elle-même tracée au prisme des besoins métiers, que l'on va créer les conditions et la dynamique de la transformation. Le terme de ce parcours sera une entreprise réellement et totalement centrée sur la donnée, c'est-à-dire où les métiers l'intégreront naturellement à leurs manières de réfléchir et d'agir, sauront spontanément y chercher la réponse à leurs questions, et s'en considéreront coresponsables.

À sa modeste échelle, cette publication vise à vous présenter comment atteindre cet objectif de démocratisation ultime de la donnée, vers lequel la vocation et l'expertise de Capgemini sont de vous accompagner.



OLIVIER PERHIRIN

CTO Insights & Data Capgemini
South Central Europe

SOMMAIRE

LES DONNÉES UN ACTIF PRÉCIEUX À METTRE SOUS CONTRÔLE

4

Data Trust :

comment renforcer la confiance dans les données — 5

YANN ORTODORO

Du Big Data au Right Data :
les plateformes à l'heure
de l'écoresponsabilité — 8

NICOLAS CLAUDON | NICOLAS YDDER

Avec l'EIM, **les données
non structurées** ne sont
plus les oubliées des
stratégies data — 14

LUCIEN GUILLERMIN

Data Mesh : la donnée
décentralisée enfin à l'heure
de la maturité ? — 11

ARNAUD ROVER | WILLIAM LESGUILLIER

Écosystèmes de données :
tirer parti des données
externes — 16

MARION GARDAIS

LE SOCLE TECHNOLOGIQUE TIRER PARTI DU CLOUD ET DE SES INNOVATIONS

18

**Modernisation des
plateformes :** bâtir une
infrastructure de données
grâce au cloud — 19

OLIVIER PERHIRIN | FLORENT BOUCHENOT

Data et FinOps,
un mariage nécessaire — 24

STÉPHANE MINOZZI | THOMAS SARRAZIN

Migration des données :
accélérer et réussir le
passage vers le cloud — 22

AUGUSTIN BARRIÈRE | MATHIEU JOLY

**Intégrer les données en
temps réel :** des avancées
qui ouvrent de nouveaux
horizons au décisionnel — 26

NEJMEDDINE REZGUI | ARNAUD ROVER

CULTURE, MÉTHODES, ORGANISATION MAXIMISER LES BÉNÉFICES DES PROJETS DATA

30

Culture data : tirer
le maximum des projets
et favoriser l'adoption
par les utilisateurs — 31

ISABELLE TACHET | GUILLAUME GÉRARD

Industrialiser l'IA :
comment produire vite
et bien des modèles
profitables — 37

GUILLAUME GÉRARD

AI Factories : démocratiser
l'usage de l'IA et des
Analytics — 34

KEVIN DUVAL

La Business Intelligence
à l'heure du self-service — 41

SYLVAIN CORNILLE-AVOT



LES DONNÉES

UN ACTIF PRÉCIEUX À METTRE SOUS CONTRÔLE

Avant d'être exploitable et de pouvoir produire toute la valeur qu'elle recèle, la donnée est une ressource brute. Comme un minerai, elle doit être localisée, extraite, nettoyée, raffinée, enrichie et préservée pour pouvoir être correctement façonnée. De la rigueur de ce processus dépendent sa qualité et, en définitive, la précision et la fiabilité des résultats obtenus. Prendre conscience de l'importance de ces phases amont fait partie intégrante de la révolution datacentric. La donnée est l'actif sur lequel l'entreprise va bâtir son futur et elle doit avoir une complète maîtrise de son cycle de vie.

Data Trust :
comment renforcer la confiance dans les données — 5

YANN ORTODORO | DATA TRUST OFFER LEAD

Du Big Data au Right Data :
les plateformes à l'heure de l'écoresponsabilité — 8

NICOLAS CLAUDON | CTO INSIGHTS & DATA FRANCE
NICOLAS YDDER | MANAGING DATA ARCHITECT

Data Mesh :
la donnée décentralisée enfin à l'heure de la maturité ? — 11

ARNAUD ROVER | PRINCIPAL ARCHITECT DATA
WILLIAM LESGUILLIER | PRINCIPAL ARCHITECT DATA & AI

Avec l'EIM, les données non structurées
ne sont plus les oubliées des stratégies data — 14

LUCIEN GUILLERMIN | EIM OFFERING LEADER

Écosystèmes de données :
tirer parti des données externes — 16

MARION GARDAIS | VP HEAD OF DATA & AI CENTER OF EXCELLENCE SOUTH CENTRAL EUROPE



DATA TRUST : COMMENT RENFORCER LA CONFIANCE DANS LES DONNÉES



YANN
ORTODORO

Data Trust Offer Lead

Dans des environnements de plus en plus complexes, avec un volume exponentiel de données, il devient urgent pour les organisations de répondre aux questions autour de la qualité, de la consommation et pilotage de la data. Comment développer une approche qui favorise l'adoption et la confiance dans la donnée ?

Lorsqu'on travaille sur des données, des questions reviennent régulièrement: Ce que je fais est-il correct ? Mon analyse se base-t-elle sur des données fiables ? Est-ce que je transmets la bonne information ? Avoir confiance dans les données, c'est n'avoir aucun doute sur la réponse à ces questions. L'enjeu ne réside donc seulement sur la qualité des données, mais bien sur la perception de cette qualité et sa maîtrise. Chez Capgemini, c'est ce que nous appelons la **Data Trust**.

Renforcer la confiance dans les données contribue à lever certains obstacles organisationnels qui entravent souvent les projets data. Les certitudes permettent d'aller plus vite, d'accepter plus facilement de partager ses données et/ou d'utiliser les données fournies par d'autres, et de prendre le bon niveau de précautions. Dans certains secteurs sensibles ou très réglementés, on voit souvent de « surqualités », des contrôles excessifs, qui brident les initiatives data de peur de commettre un impair. En libérant les utilisateurs de leurs doutes et de leurs craintes, la Data Trust est un levier de productivité, de collaboration et d'innovation tout en assurant un respect des règles et contraintes en vigueur.

DATA TRUST : DE QUOI PARLE-T-ON ?

Pour prendre une analogie, lorsqu'on est en voiture, on ne peut rouler à une certaine vitesse que si l'on a confiance : confiance dans son véhicule et dans les indicateurs qui décrivent son fonctionnement (compteur, jauge) ; confiance dans ses propres capacités de conducteur ; confiance enfin dans le fait que les autres usagers de la route respecteront un certain nombre de règles. Cet exemple nous montre trois choses qu'il est essentiel de garder à l'esprit lorsqu'on aborde la Data Trust.

Premièrement, c'est une notion dans laquelle la technologie joue un rôle majeur, mais qui reste avant tout humaine. Deuxièmement, c'est une notion subjective, qui dépend du profil de l'utilisateur, du contexte et des conséquences potentielles d'une erreur. Troisièmement, c'est une notion bilatérale: la confiance, c'est celle que l'on reçoit, mais aussi celle que l'on donne.

Autrement dit, il ne suffit pas d'améliorer la qualité des données pour améliorer la confiance : il faut aussi convaincre les parties prenantes que cette amélioration est utile et réelle (en revanche, si la qualité n'est pas au rendez-vous, la confiance s'envolera instantanément).

Pour faire progresser la confiance dans les données, il faudra donc travailler sur trois aspects nécessaires et complémentaires :

1 DES OUTILS

Ils permettront d'améliorer de façon visible et démontrable la qualité et la maîtrise des données : data lineage, Master Data Management...

2 DES FORMATIONS ET DES INFORMATIONS

Pour les utilisateurs afin d'appréhender les enjeux et les ambitions associés aux données, leurs origines, possibilités, mais aussi limites, et comment en maîtriser l'usage.

3 DES RÈGLES ET DES PROCESSUS COLLECTIFS

Leur respect évitera à chacun de se reposer systématiquement les mêmes questions.

Aujourd'hui, ces trois aspects peuvent s'appuyer sur des solutions technologiques abouties, qui utilisent notamment l'intelligence artificielle pour automatiser de nombreuses tâches de contrôle et de mise en qualité des données.

STIMULER L'ADHÉSION DES UTILISATEURS EN ALLIANT VISION À LONG TERME ET APPROCHE ITÉRATIVE/AGILE

Le chantier de la Data Trust est donc très vaste et il est possible de se sentir quelque peu découragé par sa complexité. Ou, au contraire, convaincu de son importance, on peut être tenté de se lancer dans une vaste entreprise de normalisation de la donnée à l'échelle de l'entreprise. Le risque est alors de s'enfermer dans une multitude de règles et de processus que les utilisateurs auront du mal à appliquer. Ces nouvelles exigences viendront perturber leur travail quotidien et la Data Trust sera pour eux davantage un frein que l'accélérateur qu'elle devrait être. Si l'ambition doit être de mise, mieux vaut donc procéder de façon pragmatique et par petits pas, en actionnant les bons leviers sur des périmètres choisis.

Une approche Data Trust réussie devient un actif organisationnel et métier précieux. Cela se traduit en premier lieu par une confiance renouvelée dans l'usage de la donnée au quotidien, la réduction des temps de traitement et de coûts, une prise de décision plus rapide et de meilleure qualité... Sans compter la possibilité de créer de nouvelles sources de revenus via les écosystèmes de données, tout en étant moins exposé au risque de non-conformité.

Pour valider la réussite des projets et démontrer leur bien-fondé, il sera important de définir des indicateurs qui soient suffisamment explicites pour les utilisateurs et qui mettent en perspective bénéfices métiers obtenus avec le temps et le coût de la gestion de la donnée. Une mesure comme le temps nécessaire pour

enregistrer un nouveau client et le valider, qui a à la fois un intérêt métier et un sens data, peut constituer un bon point de départ. L'objectif est de s'assurer que l'on met en place un dispositif viable et créateur de valeur, et non un système qui produise des données certes irréfutables, mais à grand frais.

« Il ne suffit pas d'améliorer la qualité des données pour améliorer la confiance : il faut aussi convaincre des parties prenantes de cette amélioration.

En revanche, si la qualité n'est pas au rendez-vous, la confiance s'envolera instantanément. »

Renforcer la confiance dans la donnée est un travail sur le long terme, qui doit sans cesse être remis en perspective par le métier, car le moindre faux pas peut suffire à ébranler des certitudes difficilement acquises. C'est une démarche qui exige un investissement et un engagement de tous dans la durée et la première tâche est sans doute de convaincre (régulièrement) toutes les parties de son importance capitale pour la réussite de la stratégie de l'entreprise.



À RETENIR

En levant les doutes, la Data Trust est un levier de productivité, de collaboration et d'innovation.

La Data Trust n'est pas seulement une notion technologique, mais fondamentalement humaine.

La Data Trust est un chantier de long terme mais à mener pas à pas, et qui doit sans cesse démontrer sa valeur pour convaincre.



DU BIG DATA AU RIGHT DATA : LES PLATEFORMES À L'HEURE DE L'ÉCORESPONSABILITÉ



NICOLAS
CLAUDON

CTO Insights & Data France



NICOLAS YDDER

Managing Data Architect

Maîtriser son patrimoine de données, c'est aussi limiter son expansion débridée. À rebours d'une culture de l'accumulation à tous points de vue insoutenable, il faut adopter une approche plus sélective. Bonne pour l'environnement, cette frugalité l'est aussi pour l'entreprise.

En matière de données, on a longtemps considéré que plus on en collectait, mieux on se portait. Après tout, cela pouvait toujours servir. Et avec l'essor du Big Data et du cloud, on n'avait aucune raison de se limiter puisque les infrastructures pouvaient supporter sans sourciller des volumes gigantesques. Cependant, on se rend de plus en plus compte que ces masses de données, qui peuvent se chiffrer en pétaoctets, sont rarement utiles, voire contre-productives tant il devient rapidement impossible de s'y repérer. Résultat : 80 % des données ne sont pas utilisées¹, et ont peu de chances de l'être un jour.

Or, même inutilisées, ces données nécessitent des transferts, des manipulations, des traitements, de l'espace de stockage, des mesures de sécurité... Et comme tout cela se passe à très, très grande échelle, l'empreinte environnementale qui en résulte est loin d'être négligeable. À titre d'exemple, au sein de l'Union Européenne, les datacenters représentaient déjà 2,7 % de la demande d'électricité en 2018 et ce chiffre atteindra 3,21 % en 2030 si la trajectoire actuelle se poursuit². Au moment où l'IT est pointée du doigt pour son impact écologique grandissant, lutter contre cette infobésité, aussi

inutile que nuisible, et adopter une politique de frugalité des données, apparaît comme un axe prioritaire en matière d'éco-responsabilité.

UN ENJEU CULTUREL ET STRATÉGIQUE

Réduire la quantité de données, c'est avant tout rompre avec l'habitude et la facilité qui consistent à tout conserver. Du Big Data exhaustif, il faut passer au Right Data : privilégier la qualité à la quantité en ne collectant et conservant que ce qui est vraiment utile. Mais choisir, trier, sélectionner, c'est un effort et un changement, donc avant tout un enjeu culturel. Il faut faire prendre conscience à tous les acteurs concernés qu'en dépit de son caractère « immatériel », l'accumulation de la donnée est un gaspillage préjudiciable à l'entreprise et à l'environnement. Et qu'il appartient à chacun d'œuvrer pour le réduire : à l'IT, chargée des infrastructures ; à l'organisation data, chargée du patrimoine de données et de sa mise en valeur ; aux métiers, qui sont les seuls à pouvoir évaluer la valeur de chaque donnée et donc l'intérêt ou non de la posséder.

1. Source : [Data Act: measures for a fair and innovative data economy](#)

2. Source : [Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market](#)

UNE NÉCESSAIRE GOUVERNANCE

D'un principe, la frugalité des données devient donc un objectif commun. Cependant, tout le monde n'aura pas le même point de vue sur ce qu'il faut conserver, à quelles fins et dans quelles conditions. Il est donc nécessaire de mettre en place une gouvernance pour définir une politique, des lignes de conduite, des rôles, et arbitrer les différences d'appréciation. L'instance de gouvernance pourra être épaulée, au niveau technique, par une Design Authority, équipe composée par des architectes et des décideurs métier qui édictera des règles, pilotera leur déploiement et veillera à leur application rigoureuse et homogène.

Une des raisons de l'inflation galopante des données est que personne n'est aujourd'hui responsable d'en maîtriser les volumes. Dans le

cadre de la gouvernance, il est donc capital que quelqu'un endosse ce rôle. Il lui appartiendra de s'assurer que la donnée consommée aura l'impact business optimal en fonction de son empreinte environnementale, voire de mettre son veto si cette dernière apparaissait démesurée au regard des bénéfices envisagés. Pour prendre de telles décisions, il conviendra de mettre en place des indicateurs plus fins que le volume des données et leur coût de stockage. Quant au pilotage, il s'effectuera de préférence au niveau des portefeuilles de projets afin de tenir compte de l'hétérogénéité des situations.

« En incitant à s'interroger en permanence sur le sens et l'utilité de ce que l'on collecte, une politique de frugalité renforce la maîtrise du patrimoine de données, ce qui contribue aussi à réduire les coûts et les risques associés. »

DES MESURES OPÉRATIONNELLES

Concrètement, la frugalité des données passe par un certain nombre de mesures opérationnelles dont beaucoup peuvent commencer à être mises en œuvre sans attendre que soit posé le cadre de gouvernance. Ces actions sont la priorité pour obtenir des premiers gains significatifs et amorcer un changement de perception :

LE STOCKAGE

De nombreuses données ne prêtent pas à débat et pourraient être rationalisées en tirant simplement parti des fonctionnalités des plateformes comme la déduplication pour optimiser les dispositifs de stockage, avec des retours positifs en lien avec l'approche FinOps.

L'ACCÈS

Autre axe prioritaire et relativement facile à mettre en place, la différenciation du mode de stockage au fil du cycle de vie des données avec des accès « chauds » pour les données les plus immédiatement nécessaires, « tièdes » pour les analyses et les reportings moins fréquents, puis de l'archivage de précaution, « froid », et enfin de l'archivage à fins purement historiques, sur bande.

LA FORME

Au niveau de l'infrastructure de données, la DSI dispose également de plusieurs leviers techniques pour limiter les volumes comme la compression (à condition que les gains ne soient pas absorbés par des décompressions trop fréquentes), la sérialisation binaire (optimisation du stockage de l'objet) et la virtualisation des données (pour éviter des répliquions superflues).

ADOPTER L'APPROCHE DU CYCLE DE VIE

Toutefois, le changement le plus important doit intervenir au niveau des projets et des produits data, où l'on doit désormais se focaliser sur l'information, sur ce que l'on a besoin de savoir et non sur les données qui permettent de le savoir.

De cette manière, on pourra faire, tout au long du cycle de vie, des choix économes en données sans que ce soit au détriment des résultats métier : utiliser, s'il en existe, des données tierces plutôt que collecter et posséder les siennes ; filtrer les données à la source et les prétraiter pour ne remonter que ce qui a du sens (edge computing) ; choisir des algorithmes pré-entraînés ou peu gourmands en données (few-shot learning, zero-shot learning) ; déterminer le seuil de précision/pertinence juste nécessaire et ne pas prolonger les calculs au-delà ; ne conserver que les résultats et pas les données brutes qui ont permis de les obtenir (ou bien n'en conserver que des échantillons représentatifs) ; se contenter, quand c'est possible et pertinent, de résultats agrégés, de moyennes, plutôt que des chiffres détaillés... À noter que toutes ces mesures devront être documentées et traçables au cas où l'on pourrait avoir à rendre compte des divers tris et suppressions réalisés.

Si elles coordonnent leurs efforts, les organisations métiers, IT et data disposent donc de très nombreux leviers pour réduire les volumes de données et leur empreinte environnementale. Surtout, en prenant le réflexe de s'interroger en permanence sur le sens et l'utilité de ce que l'on collecte, cette politique de frugalité renforce la maîtrise du patrimoine de données, ce qui, par ricochet, contribue à réduire les coûts et les risques associés. Pour l'heure, aucune réglementation n'oblige à modérer ses quantités de données, mais les entreprises auraient, à tous points de vue, le plus grand intérêt à l'anticiper.

La frugalité des données est donc une préoccupation transverse. Elle dépasse par conséquent le cadre du seul Green IT pour s'inscrire dans celui, plus large, de la stratégie RSE de l'entreprise tout en recoupant les transformations en cours vers une organisation davantage fondée sur la data. Aussi, le changement doit être porté et impulsé au plus haut niveau, et figurer parmi les objectifs stratégiques en matière d'environnement.

À RETENIR

Les masses de données conservées « au cas où » sont rapidement inutilisables tout en occasionnant une empreinte environnementale considérable.

Une politique plus frugale repose sur une gouvernance transverse car les points de vue différents sur les données à collecter et conserver impérativement.

De très nombreux leviers techniques et opérationnels tout au long du cycle de vie des données offrent des gains rapides et significatifs.



DATA MESH : LA DONNÉE DÉCENTRALISÉE ENFIN À L'HEURE DE LA MATURITÉ ?



ARNAUD ROVER
Principal Architect Data

Le concept du data mesh a séduit comme solution pour plus de collaboration. Cependant l'approche s'est heurtée à la réalité terrain des systèmes historiquement silotés, et à des équipes tiraillées entre la décentralisation demandée et le modèle centralisé de base sur lequel tout a été construit. Est-ce encore possible de retrouver un équilibre entre centralisation et autonomie dans la transformation data-driven ?

La data est un matériau avec lequel on construit des produits, des services, des automatismes, afin de créer davantage de valeur. Parallèlement, la digitalisation des activités et des échanges, la dématérialisation et l'Internet des objets multiplient les sources de données, aussi abondantes que variées. Bref, on a à la fois beaucoup plus de matière première et beaucoup plus de façons de l'exploiter.

premières initiatives ont souvent créé des déceptions. Guère enthousiastes à l'idée de démanteler ce qu'elles avaient si patiemment construit, les équipes data centrales, à qui était généralement confié le projet, ont commencé à bâtir des architectures compliquées à base de référentiels, de passerelles, et d'outils intermédiaires entre les sources de données et les produits finaux. Approches qui se sont souvent révélées inefficaces, mais qui n'étaient pas à proprement parler du Data Mesh.



WILLIAM LESGUILLIER
Principal Architect Data & AI

Beaucoup plus et, très vite, beaucoup trop ! La traditionnelle gestion centralisée des données, parfaitement adaptée à la BI, se révèle impraticable à l'ère de l'IA. Les data lakes monolithiques deviennent si vastes et si complexes qu'il n'est plus humainement possible de savoir précisément ce qu'ils contiennent. Chacun n'utilise que ce qu'il connaît, la collaboration est très difficile, et les croisements de données, où se situent pourtant la valeur, sont impossibles. Résultat : l'immense majorité des données collectées restent en friche, attendant un cas d'usage qui ne viendra jamais.

UN DÉFI : DÉCENTRALISER SANS CRÉER DE NOUVEAUX SILOS

Le défi est donc de décentraliser la donnée pour la rendre plus maniable sans pour autant créer de nouveaux silos. Il faut en particulier veiller à préserver deux capacités clés. Premièrement, la collaboration, pour que les équipes puissent enrichir et exploiter le patrimoine de données avec un effort de coordination minimal. Deuxièmement, l'extensibilité, pour que ce patrimoine, qui ne cesse de s'enrichir, puisse continuer à croître sans que cela pèse sur l'agilité nécessaire pour bâtir les nouveaux usages.

Dans cette perspective, l'approche Data Mesh, qui propose de scinder le monolithe de données en plusieurs domaines orientés métier, a d'emblée suscité beaucoup d'intérêt, car elle est apparue comme la réponse parfaite à ces enjeux. Malheureusement, les

UN IMPÉRATIF : RESPECTER LES GRANDS PRINCIPES DU DATA MESH

C'est pourquoi les expériences mitigées ne condamnent pas le Data Mesh, mais soulignent au contraire l'importance de bien appréhender le concept. Rappelons que le Data Mesh est la déclinaison pour la donnée du Domain Driven Design (DDD), une approche de la conception de système qui postule que la valeur d'un système réside dans la réconciliation transverse de ses composantes. Le DDD vise donc à maximiser le découplage entre les composantes technologiques du système ainsi qu'entre les équipes qui les créent, les fabriquent et les maintiennent. Autrement dit, chaque domaine est géré à sa manière par ceux qui le connaissent le mieux et tout le monde se retrouve au niveau des interfaces.

Le Data Mesh, qui peut se définir comme une approche sociotechnique décentralisée pour gérer et accéder aux données analytiques à l'échelle¹, DÉCLINE LA PHILOSOPHIE DU DDD EN QUATRE GRANDS PRINCIPES :

- > UN DÉCOUPAGE PAR DOMAINES
- > UNE GOUVERNANCE FÉDÉRÉE
- > UNE PRÉSENTATION SOUS FORME DE PRODUITS
- > UNE PLATEFORME EN SELF-SERVICE

La réussite de la mise en œuvre du Data Mesh passe impérativement par une application rigoureuse de ces quatre principes. Si l'un d'eux manque à l'appel, la décentralisation des données sera bancal, incomplète, et n'apportera pas les bénéfices escomptés.

CONCRÈTEMENT, CELA SE TRADUIT PAR QUATRE AXES DE TRANSFORMATION :

1 ÉVOLUER VERS UN URBANISME PLUS DÉCENTRALISÉ, PORTÉ PAR LES CAS D'USAGE.

Le patrimoine de données est découpé en domaines dont la gestion est confiée à des équipes indépendantes. C'est à elles qu'il appartient de faire vivre leur domaine, d'en maintenir la qualité et l'homogénéité, et de proposer des produits data clés en main, faciles à appréhender par les non spécialistes extérieurs. Bien entendu, la condition est de préserver l'interopérabilité des domaines. Pour cela, on pourra recourir à une méthodologie commune de modélisation comme Data Vault, qui permet de s'accorder sur des clés de croisement tout en laissant une grande liberté par ailleurs. On débutera de préférence par les domaines les plus matures et dont le retour sur investissement apparaît le plus immédiat.

2 METTRE EN PLACE UNE GOUVERNANCE FÉDÉRÉE

Pour maintenir la cohérence du patrimoine commun, il faut que chacun puisse faire entendre sa voix. C'est à ce niveau que l'on décidera de la création et du découpage des domaines, des règles d'interopérabilité, de l'outillage commun...

3 METTRE EN PLACE UNE PLATEFORME EN LIBRE-SERVICE

On l'aura compris, le Data Mesh n'est pas une technologie, mais un principe de gestion et de développement collectif. Il doit s'appuyer sur un outillage adapté, capable d'autoriser cette souplesse tout en préservant la cohésion d'ensemble et la capacité à passer à l'échelle. Aujourd'hui, on trouve essentiellement de tels outils, à la fois robustes et faciles à prendre en main, chez les grands acteurs tels Snowflake et Google Cloud Platform (BigQuery). Dans la logique de maintien de la cohésion d'ensemble, une approche "Infrastructure as Code" (IaC) garantie la standardisation de l'infrastructure des nœuds du Data Mesh et facilite leur maintenance.

4 FAIRE ÉVOLUER LES PARTIES PRENANTES INTERNES

Il s'agit sans doute du chantier le plus délicat. En premier lieu, l'organisation data centrale va voir son rôle profondément remis en question alors même qu'elle a en général beaucoup investi ces dernières années pour se renforcer. Elle ne doit pas vivre le Data Mesh comme un désaveu ou un déclassement, mais prendre conscience de son importance clé au cœur du dispositif : c'est elle qui peut accompagner les métiers dans leur propre évolution, diffuser les compétences nécessaires, animer la gouvernance fédérée, maintenir le référentiel cœur, ou encore porter les enjeux transverses d'outillage, d'interopérabilité et de sécurité. Côté métier, il faudra sensibiliser, réorganiser, former et impliquer tous les acteurs qui ont une appétence pour la donnée mais qui ne sont aujourd'hui que très peu associées à sa gestion (par exemple, les actuaires dans le secteur de l'assurance).

1 Source: Data Mesh: Delivering Data-Driven Value at Scale, Zhamak Dehghani

Le Data Mesh n'est pas nécessaire pour toutes les entreprises. Tant que le modèle centralisé répond aux exigences métiers en termes de délai et de qualité de donnée, il n'y a pas de raison de changer. En revanche, dès que les produits data deviennent trop nombreux et qu'il est question de créer des leviers opérationnels fondés sur des données, une approche décentralisée est indispensable. Là, le Data Mesh est la meilleure façon de la mettre en œuvre.

En somme, le Data Mesh est l'enabler du data-driven, et c'est pourquoi ce sont bien les métiers qui doivent en être à l'origine. Ce sont eux qui doivent constater leur impossibilité de développer leurs projets data avec la structure centralisée existante ou même de façon artisanale, en « shadow IT ».

C'est une transformation à part entière, qui demande l'acculturation des équipes sur les nouvelles possibilités qu'apporte la data, et qui passe par un travail de découverte des cas d'usages en collaboration avec l'IT, ainsi que par un transfert progressif des compétences data vers les métiers et les équipes de développement du SI opérationnel.

Les métiers doivent être les demandeurs du Data Mesh, qui doit être implémenté par eux, avec eux et pour eux. Car une chose est claire : personne d'autre qu'eux ne pourra extraire la valeur des données inexploitées.

« Les métiers doivent être les demandeurs du Data Mesh, qui doit être implémenté par eux, avec eux et pour eux. »

À RETENIR

Il devient de plus en plus difficile de faire face à l'explosion des besoins et des sources de données avec un modèle de gestion centralisé.

Approche décentralisée orientée métier, le Data Mesh est la réponse idéale, mais les déceptions initiales soulignent l'importance de bien en respecter les principes.

Les métiers doivent être à la manœuvre de cette transformation qui va leur accorder un rôle et des responsabilités élargies.



AVEC L'EIM, LES DONNÉES NON STRUCTURÉES NE SONT PLUS LES OUBLIÉES DES STRATÉGIES DATA



LUCIEN
GUILLERMIN

EIM Offering Leader

Bien que constituant une large part du patrimoine de données, les données non structurées sont trop souvent négligées. En unifier la gouvernance et les intégrer à sa stratégie data permet de mettre ce précieux gisement d'informations sous contrôle et d'en exploiter la richesse.

Aujourd'hui, les entreprises axent en priorité leurs stratégies data autour des données dites « structurées » que contiennent leurs bases et leurs applications métiers. À juste titre puisque ces données constituent un gisement majeur de connaissances que les nouveaux outils d'analyse avancée permettent d'exploiter. Cependant, les données structurées ne constituent qu'une partie émergée de l'iceberg informationnel des organisations : selon une étude du cabinet IDC1, plus de 90% des données se trouvent en effet sous des formes non structurées (documents, vidéos, images, ...), et négliger cette masse considérable constitue à la fois un risque et une perte d'opportunités pour l'entreprise.

DES DONNÉES HÉTÉROGÈNES, UNE GESTION FRAGMENTÉE

Outre son volume, la principale caractéristique de cet ensemble, qui traverse tous les métiers et tous les systèmes d'information, est son extrême hétérogénéité. Production, formats, usages, valeur, durée de vie, système mis en œuvre... les données non structurées ont peu de choses en commun. C'est d'ailleurs ce qui explique largement que

ce patrimoine informationnel soit si souvent laissé en jachère, ou bien géré et exploité localement, de façon très silotée. Chaque acteur se charge, sans coordination ni partage, des données qu'il produit et qu'il utilise, s'appuyant pour cela sur des outils disparates, peu intégrés au reste du SI, et parfois obsolètes.

Le revers de cette approche fragmentée est qu'elle rend plus difficile une réponse adéquate aux enjeux qui concernent tout autant les données non structurées que structurées : sécurité, conformité, qualité, exploitabilité, maîtrise des coûts et des volumes... Or avec la dématérialisation qui va de pair avec l'essor des modes de travail et de collaboration actuels, ces enjeux se posent avec de plus en plus d'acuité.

La mise en conformité par rapport aux exigences en matière de protection des données personnelles en constitue un très bon exemple. Dans une masse informationnelle hétérogène et silotée, une non-conformité peut en effet résulter de l'incapacité d'identification des informations se rapportant aux personnes physiques. Grâce à l'extraction automatisée d'informations, il est toutefois possible d'adresser cette difficulté tout en permettant la valorisation de contenus a priori « monolithiques ».

POUR UNE GOUVERNANCE GLOBALE

Pour tenir compte de l'ensemble des contraintes qui pèsent sur les données non structurées, il s'agit donc de mettre en place une approche permettant d'aborder de façon systématique les questions de sécurité, de conformité (RGPD, fiscale, métier...), de qualité, de traçabilité et de conservation.

Au cœur d'un tel dispositif, les organisations ont tout à gagner en mettant en place une instance de gouvernance pour définir les règles, les rôles et l'outillage qui permettront de garantir qu'à chaque étape du cycle de vie des données seront prises les mesures et les précautions adéquates, et cela en dépit de la diversité des situations et des formats.

Outre qu'une telle politique permet de limiter les angles morts, et donc de maximiser la maîtrise des risques, c'est aussi un levier de réduction des coûts car cela permet de prendre en compte aisément et avec plus d'agilité les nouveaux contenus numériques qui ne cessent d'émerger, et de gagner en productivité sur des traitements nécessaires pendant longtemps réalisés de façon artisanale.

L'EIM, UNE APPROCHE UNIFIÉE CRÉATRICE DE VALEUR

Par la gouvernance et les cas d'usage, cette gestion globale des contenus non structurés (Enterprise Content Management, ECM) se rapproche dans les faits de celle des données structurées (Data Management, DM). À terme, il apparaît pertinent de favoriser la rencontre de ces deux mondes autour de la gouvernance informationnelle. Il s'agit de bâtir une architecture de l'information unifiée autour de la notion de connaissance, indépendante des formats sous-jacents.

Reposant sur un cycle de vie commun – capture/production, gestion, valorisation –, une telle approche permet de mutualiser les capacités (gouvernance, outillage...), d'harmoniser les règles, et de démultiplier les possibilités de croisement et d'analyse pour faire émerger la connaissance de l'ensemble des sources disponibles. Et de créer ainsi davantage de valeur à partir des données... quelles qu'elles soient.

« Une gouvernance incluant les données non structurées permet non seulement de maximiser la maîtrise des risques, mais devient aussi un levier de réduction des coûts dans la prise en compte avec plus d'agilité des nouveaux contenus numériques, et de gagner en productivité sur des traitements réalisés pendant longtemps de façon non-industrielle. »

À RETENIR

Bien que majoritaires dans les systèmes, les données non structurées restent sous-exploitées en raison d'une gestion et d'un outillage cloisonnés.

Mettre en place une gouvernance globale permet de réduire les coûts et les risques associés à ce patrimoine, et de pouvoir en exploiter toute la richesse.

À terme, l'EIM a vocation à rapprocher données structurées et non structurées dans une même approche pour en démultiplier la valeur.



ÉCOSYSTÈMES DE DONNÉES : TIRER PARTI DES DONNÉES EXTERNES



MARION GARDAIS

VP Head of Data & AI
Center of Excellence
South Central Europe

En encadrant le partage de données, le Data Act sera pour les entreprises européennes une opportunité majeure. L'accès à des données tierces via les écosystèmes collaboratifs de données leur ouvrira de nouveaux champs d'innovation et leur permettra de traiter des enjeux globaux comme leur bilan carbone.

Avec le Data Act, dont le projet a été présenté en février 2022, la Commission européenne affiche ses ambitions en matière d'exploitation des données. Alors que près de 80 % des données industrielles générées dans l'Union européenne restent inutilisées, le Data Act vise à harmoniser les règles d'accès aux données produites au sein de l'Union européenne afin de faciliter les échanges entre entreprises européennes, et favoriser ainsi l'innovation et la co-innovation. En clarifiant le cadre régissant le partage des données, ce texte ouvre aux entreprises un nouveau territoire d'opportunités avec la perspective d'en tirer de réels avantages compétitifs.

Dans le cadre de leur transformation « data-centric », les entreprises constatent qu'elles auraient souvent beaucoup à gagner soit à enrichir leurs données avec de sources externes, soit à ouvrir leurs données à des tiers : clients, fournisseurs, partenaires, consortiums, organismes de recherche... Reconnaisant cette tendance et son potentiel pour l'économie européenne, le Data Act entend encadrer et faciliter de tels partages, et notamment la constitution d'écosystèmes collaboratifs de données.

QU'EST-CE QU'UN ÉCOSYSTÈME COLLABORATIF DE DONNÉES ?

Dépassant le périmètre de l'entreprise, ou même de l'entreprise étendue, un écosystème collaboratif de données est une plateforme qui permet à ses participants de valoriser leurs données d'une façon qu'il leur serait impossible de mettre en place seuls ou via une multiplicité de partenariats bilatéraux. Réunissant des acteurs d'un même secteur d'activité, de secteurs connexes ainsi que des institutions, il organise, entre ces partenaires, un partage de données qui préserve leur sécurité et leur confidentialité ; il établit, pour l'utilisation de ces données, des règles transparentes et conformes à la réglementation ; enfin, il garantit un partage équitable de la valeur qu'elles auront permis de créer.

Un écosystème collaboratif de données donne la possibilité de faire émerger des produits, des services ou des expériences client qui ne peuvent exister qu'en rapprochant des informations d'horizons différents. Skywise, la plateforme de données développée par Airbus pour le secteur aérien en fournit un très bel exemple. Skywise Health Management agrège les alertes et les messages de maintenance, les hiérarchise, rapproche chaque dysfonctionnement de sa procédure de dépannage et fournit l'historique de maintenance du système, ce qui aide les compagnies aériennes à identifier et déterminer leurs priorités en matière d'entretien à partir d'informations dont, seules, elles n'auraient jamais pu disposer.

UN PRINCIPE DIRECTEUR : LA MAÎTRISE DE LA DONNÉE

Dans le cadre d'un écosystème collaboratif de données, il ne s'agit pas de partager ou de transférer ses données à une autre entreprise, mais plutôt de mettre des indicateurs ou des informations à la disposition de la communauté au travers, par exemple, d'une marketplace. Ce que l'on ouvre, ce sont donc des données qui ont été extraites, raffinées et contextualisées pour leur donner davantage de valeur (tout en protégeant, le cas échéant, le détail confidentiel de leurs origines). Par exemple dans le cadre du calcul de l'empreinte carbone d'un produit, le lieu de fabrication de l'emballage est une information qui n'a guère de valeur en elle-même, mais qui prend tout son sens dans la perspective d'une démarche de consommation locale. Seul un écosystème collaboratif de données permet d'avoir accès facilement à ce type de connaissance.

Les écosystèmes de données collaboratifs n'en sont qu'à leur début car les freins à la mise en commun des données sont encore nombreux, mais leur potentiel est immense. Ils sont incontestablement un moyen d'accélérer l'innovation en réponse aux grands enjeux actuels. Dans la recherche pharmaceutique, par exemple, le projet européen Melloddy vise à améliorer les modèles prédictifs d'apprentissage automatique en se basant sur les données de dix entreprises pharmaceutiques, sans exposer d'informations sensibles. Cette collaboration pionnière va permettre d'accélérer la mise au point de médicaments et d'améliorer les résultats pour les patients en permettant, pour la première fois, de mener des recherches dans les bases de données décentralisées et hautement confidentielles des bibliothèques chimiques des membres du consortium.

« Un écosystème collaboratif de données donne la possibilité de faire émerger des produits, des services ou des expériences client qui ne peuvent exister qu'en rapprochant des informations d'horizons différents. »

LE SCOPE 3 DU BILAN CARBONE, LA KILLER APP DES ÉCOSYSTÈMES COLLABORATIFS DE DONNÉES

Lorsqu'on quantifie les gaz à effet de serre émis par une entreprise, on les répartit en trois catégories: les émissions directes (Scope 1), les émissions indirectes liées à l'énergie (Scope 2) et tout le reste (Scope 3). Le Scope 3 couvre donc les émissions générées par l'activité de l'entreprise, comme l'achat de marchandises ou de services, mais aussi toutes celles qui ne sont pas directement liées à la fabrication de ses produits. Il englobe aussi toutes les autres étapes du cycle de vie de ces derniers, depuis l'approvisionnement en matière première jusqu'à son utilisation par

les clients et les impacts de sa fin de vie. Le Scope 3 est donc extrêmement vaste et aucune entreprise ne saurait envisager d'en avoir une vision exhaustive et précise – et encore moins de définir et piloter des actions pour atteindre la neutralité carbone – sans un partage de données avec l'ensemble des acteurs concernés. Pour une entreprise, la gestion du Scope 3 rend donc de fait indispensable la mise en place d'un écosystème collaboratif de données couvrant la chaîne de valeur de ses produits et services.

À RETENIR

Le Data Act présenté par la Commission européenne cherche à encadrer les échanges de données entre entreprises afin de favoriser ces partages, vecteurs d'innovation.

En particulier, les écosystèmes collaboratifs de données permettront d'exploiter en toute confiance des informations tierces jusqu'à présent inaccessibles.

Le bilan et la réduction des émissions carbone de Scope 3 nécessitera la mise en place d'écosystèmes collaboratifs de données, dont ce sera l'une des applications phare.



LE SOCLE TECHNOLOGIQUE

TIRER PARTI DU CLOUD ET DE SES INNOVATIONS

Sans être le seul facteur en jeu, la technologie reste au cœur de la transformation data-centric. Des plateformes data cloud à l'IA, les nouveaux outils facilitent considérablement la collecte, la gestion et l'exploitation des données, et ouvrent des possibilités inédites en matière de cas d'usage et d'innovation. Mais, pour que ces solutions soient adoptées, maîtrisées et qu'elles apportent tous les bénéfices espérés, leur mise en œuvre doit être alignée sur la stratégie data.

**Modernisation des plateformes :
bâtir une infrastructure de données grâce au cloud** — 19

OLIVIER PERHIRIN | CTO INSIGHTS & DATA SOUTH CENTRAL EUROPE
FLORENT BOUCHENOT | COE CLOUD ADVISOR

**Migration des données :
accélérer et réussir le passage vers le cloud** — 22

AUGUSTIN BARRIÈRE | SENIOR DATA ARCHITECT
MATHIEU JOLY | SOLUTIONS ARCHITECT

Data et FinOps, un mariage nécessaire — 24

STÉPHANE MINOZZI | CTO OFFICE INSIGHTS & DATA FRANCE
THOMAS SARRAZIN | GLOBAL FINOPS OFFER LEADER

**Intégrer les données en temps réel : des avancées
qui ouvrent de nouveaux horizons au décisionnel** — 26

NEJMEDDINE REZGUI | MANAGING SOLUTION ARCHITECT
ARNAUD ROVER | PRINCIPAL ARCHITECT DATA



MODERNISATION DES PLATEFORMES : BÂTIR UNE INFRASTRUCTURE DE DONNÉES GRÂCE AU CLOUD



OLIVIER PERHIRIN

CTO Insights & Data South Central Europe

Un socle technique solide avec les bonnes technologies et la bonne gouvernance est essentiel pour devenir une organisation data-centric. Quelles sont les actions à mener à la lumière des opportunités offertes par le cloud ?

Un grand nombre d'entreprises disposent de masses considérables de données, souvent de data warehouses et de data lakes, et il ne leur resterait donc plus qu'à puiser dans ce gisement de richesses. Mais ce n'est pas si simple...

Pendant longtemps, la donnée a surtout servi à établir des rapports, à fournir des chiffres indicatifs de l'état de l'activité. Or, dans l'entreprise data-centric, on consomme la donnée de bien d'autres façons. Elle alimente de nouveaux business models (fondés sur l'usage, par exemple), elle devient un élément clé de différenciation (au travers de la personnalisation, de l'expérience client, des moteurs de recommandation...), elle nourrit la R&D et l'innovation, elle est indissociable des stratégies de développement durable. Les données deviennent un produit de l'entreprise, et à ce titre leur pertinence, leur fréquence de mise à jour et exhaustivité sont autant de critères qui permettent de qualifier leur potentiel.

Or, il est très difficile de mettre en place tous ces nouveaux usages avec des infrastructures essentiellement conçues pour le reporting. Avec ces systèmes verticaux, cloisonnés, les équipes ne peuvent pas accéder aux données aussi rapidement et efficacement qu'elles

le souhaiteraient. Pour développer de nouveaux services et modèles d'analyse, il leur faudrait pouvoir croiser des données issues de différents métiers, partager ces connaissances dans toute l'entreprise et souvent au-delà. Il faudrait aussi industrialiser la science de la donnée, faciliter l'utilisation et la réutilisation des algorithmes, avoir un cadre de sécurité et de gouvernance à la fois souple et rigoureux... Quand tout ceci n'est pas possible, elles ne peuvent que multiplier des POC sans lendemain et chercher le salut dans le « shadow IT », qui finalement ne fait qu'ajouter des obstacles et silos supplémentaires.

« Flexible, puissant et sécurisé grâce aux investissements massifs des opérateurs, le cloud offre l'agilité, la scalabilité et les fonctionnalités nécessaires pour développer puis faire passer à l'échelle toutes sortes de cas d'usages data. »



FLORENT BOUCHENOT

CoE Cloud Advisor

UNE PLATEFORME RÉUSSIE SE MESURE PAR SON ADOPTION

Le succès de la plateforme se mesure en trois dimensions complémentaires. Premièrement, son empreinte économique. La modernisation de l'infrastructure, sa rationalisation, sa mutualisation, et la facturation à l'usage du cloud peuvent permettre de réduire de 20 % à 50 % le coût total de possession (à charge de travail constante). Deuxièmement, le nombre de cas d'usage que la plateforme aura permis de lancer. L'un des objectifs clés de la modernisation est de démultiplier la capacité à transformer les bonnes idées en solutions en production.

Enfin, la satisfaction des utilisateurs. Ce sont eux qui souffraient des limites des anciens systèmes, il est important de savoir s'ils adhèrent aux nouveaux. Quand ces trois aspects se combinent, la plateforme modernisée devient le moteur d'une transformation profonde et rapide de l'organisation vers un mode de fonctionnement résolument centré sur les données.

LE CLOUD, SOCLE TECHNIQUE DES PLATEFORMES À L'ÉTAT DE L'ART

Pour y remédier, le cloud est incontournable. Flexible, puissant et sécurisé grâce aux investissements massifs des opérateurs, il offre l'agilité, la scalabilité et les fonctionnalités nécessaires pour développer toutes sortes de cas d'usages, puis les passer à l'échelle. Avec la nouvelle génération d'outils data, le découplage et le partage de la donnée deviennent possibles dans un cadre maîtrisé, à moindre coût. La productivité est améliorée et la collaboration facilitée au sein d'équipes et de processus de plus en plus pluridisciplinaires.

Pour réellement devenir data-centric, les organisations doivent donc engager une démarche de modernisation de leurs plateformes data avec une double perspective : favoriser et industrialiser les nouveaux usages de la donnée ; tirer parti des spécificités techniques et économiques du cloud.

MODERNISER : OUI, MAIS COMMENT ?

Cette modernisation est un chantier de grande ampleur, vital pour l'entreprise, et à ce titre, il doit être initié au plus haut niveau par une approche stratégique. Il faut définir les conditions et les objectifs de cette rénovation afin d'en préciser les contours techniques (quels systèmes existants doivent être décommissionnés ?) et financiers (quels sont les retours attendus sur cet investissement conséquent ?).

Ceci établi, on peut définir l'architecture modernisée. Ce socle reposera sur deux éléments prépondérants : la sécurité et l'automatisation, essentielle car c'est elle qui permettra de créer des environnements au gré des besoins, et donc de recueillir les bénéfices opérationnels et financiers du cloud. En ce qui concerne le choix des technologies elles-mêmes, il faut être attentif à préserver sa liberté d'action et de décision vis-à-vis des fournisseurs (notamment les grands opérateurs). Il ne faut pas non plus sous-estimer le chantier technique. Certes, le cloud apporte beaucoup de facilités, mais les migrations de données, les intégrations techniques, la sécurisation, et la réalisation de tous les paramétrages nécessaires pour retranscrire l'intelligence des métiers dans la plateforme ne sont pas aussi immédiats que certains fournisseurs veulent bien l'affirmer. Sur ces aspects, être accompagné par un expert reste impératif.

En parallèle, il est fondamental de mettre en place une gouvernance appropriée. La nouvelle plateforme a vocation à découpler, industrialiser et démocratiser les usages de la donnée, donc à faire intervenir de très nombreux acteurs (data architects, data engineers, data scientists, data stewards, business analysts, développeurs, utilisateurs métiers...) tout au long de processus de production entièrement repensés. Pour obtenir les bénéfices attendus et les pérenniser, il convient de préciser les rôles et les tâches de chacun, et de mettre en place des instances de contrôle et d'arbitrage. Sans cela, la qualité des données collectées et des informations produites peut très vite dériver, et avec elle la confiance des utilisateurs.

À travers ces enjeux de gouvernance, transparaît enfin un très fort besoin d'accompagnement du changement. Avec la nouvelle plateforme de données, les acteurs métiers, data et IT voient leurs attributions changer. Il leur faut s'approprier de nouveaux rôles, de nouveaux outils, de nouveaux processus, et développer de nouvelles compétences. Qu'ils le fassent rapidement et sans réserve est capital pour la réussite de la plateforme. Cette conduite du changement s'amorce dès l'origine du projet, en impliquant les métiers : la plateforme doit être développée avec et pour eux, afin de faciliter son utilisation. Après tout, ce sont les métiers qui pourront tirer parti de la valeur ajoutée de ces plateformes.

A photograph of a modern, curved glass skyscraper at night. The building's facade is composed of many rectangular glass panels, some of which are illuminated from within, creating a warm glow. The building is set against a dark sky. In the foreground, a walkway with a metal railing runs alongside the building. Several people are walking away from the camera on the walkway. The overall atmosphere is urban and contemporary.

À RETENIR

Peu adapté aux nouveaux modes d'exploitation de la donnée, les systèmes « legacy » freinent le développement et le déploiement des cas d'usage.

Avec la nouvelle génération d'outils data dans le cloud, la productivité est améliorée et la collaboration facilitée au sein d'équipes et de processus de plus en plus pluridisciplinaires.

La modernisation n'est pas qu'un chantier technologique : la mise en place d'une gouvernance adaptée et la conduite du changement sont tout aussi déterminantes pour sa réussite.



MIGRATION DES DONNÉES : ACCÉLÉRER ET RÉUSSIR LE PASSAGE VERS LE CLOUD



AUGUSTIN
BARRIÈRE

Senior Data Architect

Les technologies proposées par les fournisseurs Cloud séduisent plusieurs organisations en quête de transformation data-driven. Comment maîtriser les enjeux lors des migrations vers le cloud ? Quels « pièges » déjouer ?

Le cloud n'offre pas seulement des ressources virtuellement sans limites pour absorber des volumes de données qui ne cessent de croître. Il apporte aussi une souplesse technique et une qualité de service inégalables. Ainsi, les métiers peuvent développer de nouveaux cas d'usage, les passer à l'échelle et créer plus rapidement de la valeur sans être bridés par les capacités restreintes d'un data center ou les procédures d'allocation de ressources de l'IT.

Toutefois, malgré l'intérêt manifeste de porter dans le cloud leur infrastructure de données, certaines entreprises y sont encore réticentes, arrêtées par la complexité apparente de la migration Or, il existe désormais des méthodes et des outils éprouvés pour maîtriser les coûts et la qualité de ces projets, et en réduire les délais. Ceci est à la portée de toutes les entreprises, quelles que soient leur taille, leur secteur d'activité et leur maturité sur la data, à condition de procéder avec rigueur.



MATHIEU JOLY

Solutions Architect

Il ne faut pas non plus négliger les avantages économiques du paiement à l'usage, avec, dans le cas particulier des données, la possibilité de découpler les ressources de stockage de la puissance de calcul... Ce qui change fondamentalement les approches à déployer pour une telle transformation.

Les grands providers cloud ajoutent à cela une expérience de longue date dans le domaine du Big Data. Très tôt confrontés à la croissance fulgurante de leur activité, ils ont développé pour leur propre usage des outils novateurs, capables de brasser des quantités considérables de données. Ces outils robustes font aujourd'hui partie de leurs offres. Ils s'appuient également sur un très riche écosystème de partenaires, où l'on retrouve les solutions les plus en pointe sur l'ensemble du cycle de vie de la donnée.

COMMENT SAVOIR OÙ J'ATTERRIS ? LES ÉTAPES DE LA DÉMARCHE DE MIGRATION

ANALYSER L'EXISTANT ET MAÎTRISER SON PATRIMOINE

Une fois que l'on a déterminé sa stratégie data, la première étape consiste à analyser l'existant. Cette phase concerne non seulement les données elles-mêmes (volumes, natures, formats, localisation...), mais aussi tout leur environnement : les traitements qu'elles subissent, les applications qu'elles alimentent, les connexions par lesquelles elles transitent, les règles internes qui les encadrent, les

réglementations auxquelles elles doivent se conformer (RGPD notamment), les dispositifs qui les sécurisent. Maîtriser sa migration, c'est maîtriser l'ensemble de son patrimoine pour minimiser ses efforts et éviter les mauvaises surprises.

DÉFINIR LA CIBLE, PRÉPARER LA MIGRATION ET L'EXÉCUTER

Après avoir précisé le point de départ, il s'agit de définir la cible, soit l'architecture de données et ses composants, ainsi que le mode opératoire de la migration et l'outillage qui sera utilisé. Par exemple, lors de la migration d'un entrepôt de données, il est indispensable d'établir

LES FACTEURS DE SUCCÈS

Les projets majeurs menés par Capgemini autour de la data nous ont permis d'identifier quelques-uns des facteurs clés de succès. En premier lieu, il est impératif de garder à l'esprit qu'une telle migration n'est pas qu'une simple montée de version : il faut adapter l'infrastructure de données à son nouvel environnement, et bien prendre en compte les spécificités de ce dernier.

Il y aura donc des ajustements à réaliser afin de garder le même comportement qu'à l'origine. Afin de se prémunir de leurs éventuels impacts en cascade, il est par conséquent essentiel d'avoir une vision d'ensemble du patrimoine concerné. C'est tout l'enjeu du soin qu'il faut apporter à la phase de préparation.

Un autre aspect déterminant concerne la mise en œuvre de la migration, qu'il faut industrialiser et automatiser au maximum, en plus d'une approche de cybersécurité par design. Tout cela est indispensable à la bonne réalisation du projet ainsi qu'à la viabilité de la plateforme data sur le long terme. On s'appuiera sur les outils d'automatisation évoqués plus haut, ainsi que sur les pratiques issues de DevOps : Infrastructure as code (IaC), pipeline d'intégration continue, tests automatisés, accélérateurs...

une cartographie complète pour les types de données, ainsi qu'une analyse des différences entre les deux entrepôts.

On s'appuiera volontiers sur des outils de transfert de données et de tests, qui permettent d'automatiser et de fiabiliser les tâches. Capgemini dispose également de ses propres outils d'analyse de code, qui permettent [d'accélérer la migration](#) et d'en abaisser le coût. Enfin, on définit la trajectoire de migration en planifiant et en lotissant les travaux, d'abord sur des périmètres limités avant leur généralisation.

« Une migration n'est pas qu'une simple montée de version : il faut adapter l'infrastructure de données à son nouvel environnement, et bien prendre en compte les spécificités de ce dernier. »

Enfin, en parallèle de la migration technique, il est important d'accompagner les conséquences du changement de plateforme et, en particulier, d'anticiper le passage au modèle de facturation à la demande. Pour suivre et piloter les dépenses, il est indispensable de mettre en place une démarche FinOps afin de contrôler l'évolution des usages et des coûts liée à ce changement technologique majeur. C'est aussi ce qui permettra d'accélérer l'innovation tout en maîtrisant ses dépenses.

Une migration data vers le cloud est un projet majeur qui nécessite de s'appuyer sur une plateforme fiable, des outils et des méthodes éprouvés et un conseil expérimenté, capable d'accompagner la démarche de bout en bout, ceci pour concilier maîtrise et vitesse d'exécution, et ainsi tirer au plus vite, et au maximum, les bénéfices de la plateforme cloud pour ses initiatives data.

À RETENIR

Ses nombreux atouts désignent le cloud comme l'environnement de choix pour les infrastructures de données, mais beaucoup jugent la migration réhibitoire.

Il existe des méthodes et des outils éprouvés permettant d'accélérer ces projets tout en maîtrisant les coûts et la qualité.

Soigner les préparatifs, industrialiser les opérations et anticiper les conséquences du changement sont les clés de la réussite.

DATA ET FINOPS, UN MARIAGE NÉCESSAIRE



STÉPHANE
MINOZZI

CTO Office
Insights & Data France



THOMAS
SARRAZIN

Global FinOps offer leader

À l'heure où les plateformes de données sont migrées vers le cloud, où y sont développées nativement, l'approche FinOps est nécessaire pour l'anticipation et la maîtrise de coûts. Quelles sont les bonnes pratiques à adopter et les pièges à éviter ?

Pour l'entreprise data-centric, son immense potentiel rend le cloud pratiquement incontournable. Dans le cloud, les plateformes de données ne sont limitées ni en volume, ni en puissance de calcul. Elles bénéficient sans cesse d'innovations et permettent, de surcroît, d'étendre l'utilisation de la donnée au sein de l'organisation à des acteurs à la fois plus nombreux et plus divers. Mais si les possibilités sont sans limite, les coûts également. Ils suivent mécaniquement l'évolution des usages, si bien qu'en l'absence de règles et de contrôles, le risque est grand de voir les factures exploser. Et c'est d'ailleurs ce que l'on constate chez certaines entreprises qui ont imprudemment bâti des infrastructures de données dans le cloud sans suffisamment prendre en compte les spécificités de sa tarification.

Pour éviter ce travers, les démarches FinOps ont pour vocation d'intégrer le paramètre financier dans l'équation des projets data. Concrètement, il s'agit de mettre en place les processus qui

permettront d'avoir de la visibilité sur les coûts et de réaliser des optimisations, sans pour autant imposer des restrictions drastiques aux utilisateurs. L'objectif est bien de maîtriser les coûts, pas de les réduire : une consommation du cloud qui préserve l'agilité et la capacité d'innovation des métiers tout en garantissant qu'il n'y aura ni gaspillage ni dérive des dépenses.

Le FinOps n'est donc pas un contrôle de gestion à posteriori, mais la prise en compte dans les projets de la dimension nouvelle qu'impose la variabilité des coûts des services cloud. Pour pleinement en tirer parti, ce n'est plus aux usages de s'accommoder aux ressources disponibles auparavant dans les datacenters, payées d'avance (y compris dans le cas d'une monétisation des données). C'est un changement complet d'état d'esprit dans les organisations et au sein de leurs DSI, qui doivent abandonner leurs schémas traditionnels basés sur des achats de capacités et des coûts fixes.

LE CAS D'USAGE ET LES APPLICATIONS CONCRÈTES PAR LES MÉTIERS SONT ROI

L'approche FinOps doit donc être étroitement intégrée à la démarche projet, et ce, dès l'origine. Partant des usages, on intervient à la phase de design pour que l'architecture soit conforme à la fois aux besoins métiers et aux principes technologiques et économiques du cloud. Pour ce qui est envisagé (reporting financier, reporting opérationnel, automatisation...), a-t-on besoin d'une base de données classique,

disponible 24h/24 et 7j/7 et dotée d'une puissance de calcul importante ou peut-on procéder autrement ? Quels sont les temps de réponse attendus ? Y aura-t-il des pics et des creux d'utilisation ? Quel sera le cycle de vie des données ? Faut-il un archivage ? De quel type ? Pour combien de temps ? Des réponses à ces questions dépendront les options d'architecture... et le coût de leur utilisation.

Pour mettre la bonne capacité en face du bon cas d'usage, la démarche FinOps doit donc être globale. Elle débute par le choix de la meilleure architecture selon le cas d'usage avec un double regard technique (performance, disponibilité...) et financier (modèle de facturation des services sollicités). Elle se prolonge par la mise en œuvre de bonnes pratiques de développement, comme

l'optimisation du requêtage, et par la mise en place d'un monitoring des coûts afin de les refactoriser convenablement ou encore d'émettre des alertes en cas de dérive. Enfin, on boucle le cycle d'amélioration continue en analysant ces indicateurs afin de comprendre l'origine des coûts et de mettre en place les optimisations techniques ou organisationnelles appropriées.

VOLUMES ET TRAITEMENTS, LES DEUX DIMENSIONS À CONSIDÉRER

Pour déterminer les coûts, deux aspects doivent être pris en compte : les volumes de données et les traitements/actions auxquels elles seront soumises. Or, les vieux réflexes incitent souvent à se focaliser sur les premiers en oubliant quelque peu les seconds, pourtant loin d'être négligeables. C'est ce qui va expliquer les différences de coûts significatives entre différentes typologies de gestion de données (data lake, base de données SQL classique, base de données NoSQL), ou encore le fait que des applications legacy peuvent devenir d'un emploi très coûteux lorsqu'elles sont transposées telles quelles dans le cloud (« lift & shift »).

Pour déterminer la meilleure architecture cible, il faut aussi connaître les catalogues de services des différents providers, qui peuvent être extrêmement vastes, et leurs modèles de facturation, très variés et parfois assez sophistiqués : en fonction du nombre de requêtes, des volumes de données échangées, du temps d'utilisation... Autant de paramètres qui n'avaient pas cours dans le datacenter et qu'il va aussi falloir prendre en compte en intégrant les bonnes pratiques FinOps avec, à la clé, une réduction de 30 % des coûts de consommation en moyenne. Ces gains peuvent aller jusqu'à des ratios de 1/10 lors du passage optimisé d'un système on premise vers le cloud.

UN FACILITATEUR ET UN ACCÉLÉRATEUR DE L'INNOVATION

Avec les plateformes de données, on quitte l'enceinte bornée du datacenter pour entrer de plain-pied dans l'univers sans limite du cloud. C'est, pour tous les acteurs concernés, un bouleversement culturel que les démarches FinOps permettent d'accompagner. En intégrant l'aspect financier, elles sont un accélérateur, et non un frein, à l'innovation et à l'optimisation, car elles préservent la flexibilité et la capacité à passer à l'échelle inhérentes au cloud tout en garantissant que les coûts resteront proportionnés à la création de valeur.

Les démarches FinOps sont là pour éviter toutes désillusions, souvent constatées par les

directions générales, lors de l'utilisation de la data dans le cloud. Bien intégré aux projets, le FinOps est un véritable facilitateur pour les entreprises qui veulent profiter d'un écosystème applicatif moderne, agile et modulaire pour devenir plus performantes, voire pour développer de nouvelles activités impossibles à mettre en œuvre avec un système d'information uniquement géré en mode capacitaire.


« Le FinOps n'est pas un contrôle de gestion à posteriori, mais la prise en compte dans les projets de la dimension nouvelle qu'impose la variabilité des coûts des services cloud. »

À RETENIR

Dans le cloud, les possibilités sont sans limites, mais pas les coûts, qui suivent l'évolution des usages et peuvent vite s'emballer si l'on n'y prend garde.

Une démarche FinOps permet d'intégrer l'aspect financier aux projets dès l'origine pour éviter que des gaspillages ou des dérives incontrôlées ne deviennent un frein.

En contribuant à bien choisir les architectures, en suivant l'évolution des coûts et en suggérant des optimisations, le FinOps maximise la capacité d'innovation de l'entreprise.



INTÉGRER LES DONNÉES EN TEMPS RÉEL : DES AVANCÉES QUI OUVRENT DE NOUVEAUX HORIZONS AU DÉCISIONNEL



NEJMEDDINE
REZGUI

Managing Solution
Architect

Les innovations actuelles des plateformes data dans le cloud pourraient considérablement faciliter la mise en œuvre de l'intégration des données en temps réel et ainsi ouvrir la porte à de formidables possibilités d'innovation.

Aussi bien stratégique que tactique ou opérationnelle, la prise de décision est elle aussi soumise à l'accélération d'un monde où tout va plus vite. Pour se déterminer, le décisionnaire doit avoir une connaissance précise de la situation à l'instant même où elle a évolué. Même si chaque métier définit son « temps réel » selon ses propres rythmes – de l'ordre de quelques microsecondes pour le trading haute fréquence à celui de

la minute pour les flux logistiques, par exemple –, les cas d'usage potentiels se multiplient dans tous les domaines. Qu'il s'agisse de détecter immédiatement la fraude, d'améliorer l'expérience client avec une personnalisation en direct ou de réagir sans délai à un incident opérationnel, la capacité à prendre des décisions en temps réel devient un enjeu majeur d'efficacité opérationnelle.

LE TEMPS RÉEL, UNE DISCIPLINE À PART



ARNAUD ROVER

Principal Architect Data

La mise en œuvre de systèmes temps réel est, encore aujourd'hui, une discipline à part, complexe et coûteuse. Il faut une lecture en continu des flux et les traiter au fil de l'eau, à l'aide de logiciels spécifiques (Spark Streaming, Flink...), ce qui nécessite des compétences de pointe et des infrastructures distinctes de celles des traitements par lot (batch). On peut certes contourner l'obstacle à l'aide de technologies traditionnelles en rapprochant les traitements (micro-batches), mais cette approche, lourde et peu satisfaisante, montre assez vite ses limites en termes de performance. Face aux coûts et aux difficultés, le temps réel reste donc généralement réservé aux cas d'usage où sa plus-value est indispensable.

Or, de nouvelles avancées technologiques pourraient bien changer la donne. Des éditeurs cloud tels que Snowflake, Databricks par exemple, vont proposer en 2023 de nouvelles approches visant à unifier et simplifier l'exploitation de données temps réel.

Concrètement, il en résulte des « pipelines déclaratifs » qui permettent aux utilisateurs finaux de construire leurs flux de données en se focalisant uniquement sur la logique métier et en n'ayant à spécifier que les niveaux de service souhaités (fréquence des rafraîchissements, performance, volumes...). Après l'exécution des traitements sur les données (ETL vers ELT), les plateformes Data Cloud vont occuper une place encore plus centrale,

prenant en charge en autonomie l'orchestration de ces traitements (au regard d'un niveau de service souhaité).

Le temps réel constitue la prochaine étape de cette révolution. On retrouve déjà sur le marché ces solutions qui vont permettre d'inclure les données temps aux pipelines déclaratifs grâce à un rafraîchissement incrémental automatique à

faible latence : Dynamic Tables (par Snowflake) et : Delta Live Table (par Databricks). En d'autres termes, les données batch et temps réel pourront être utilisées ensemble, sur la même plateforme, par les mêmes personnes et avec les mêmes technologies (SQL, Python ou autre selon l'appétence des équipes). C'est un pas de géant pour la démocratisation du temps réel !

LES BÉNÉFICES DES NOUVELLES APPROCHES CLOUD

Cette innovation constitue un pas de géant pour la démocratisation du temps réel car elle apporte trois avantages majeurs :

1

Développer des produits temps réel sera beaucoup plus simple et moins coûteux, puisque les infrastructures seront mutualisées avec les traitements batch. Ceci permettra d'abaisser considérablement les barrières techniques et financières qui limitaient jusqu'à présent les cas d'usage éligibles.

2

La possibilité d'utiliser des technologies non spécialisées en temps réel n'est pas anodine car elles sont souvent abondantes dans les organisations. C'est donc une opportunité de conserver et de redéployer ces équipes sur de nouveaux projets créateurs de valeur.

3

Grâce au rapprochement technique et conceptuel entre temps réel et batch, il sera possible de mêler flux chauds et froids dans une même solution, ce qui était très complexe auparavant puisque les systèmes étaient différents et cloisonnés. Ceci permettra donc d'aider à la prise de décision en remplaçant aussitôt l'évènement (chaud) dans son contexte (froid). Par exemple, on pourra instantanément confronter une demande de transaction aux informations client nécessaires pour la valider.

Un autre avantage potentiel concerne l'empreinte environnementale des applications puisque la mutualisation des systèmes et le traitement incrémental limiteront les ressources nécessaires. Toutefois, il faudra tenir compte de l'effet rebond induit par la tentation de multiplier le recours au temps réel.

UNE NÉCESSAIRE MONTÉE EN MATURITÉ

De façon générale, si la levée de l'obstacle technique est une formidable nouvelle pour l'essor du temps réel, cela ne doit pas conduire à son utilisation tous azimuts. Techniquement, rien ne s'opposera à la production d'un reporting financier au fil de l'eau, mais cela aura-t-il un sens en matière de gestion ?

Il faudra donc toujours s'interroger sur la pertinence des nouveaux cas d'usage envisagés car le risque pourrait être de noyer les personnes sous un flot ininterrompu d'informations parmi lesquelles elles pourraient finir par ne plus avoir la capacité, ou la volonté, de discerner celles qui comptent vraiment. Les avancées technologiques en matière de temps réel devront donc s'accompagner d'une montée en maturité de l'organisation afin d'en exploiter avec discernement tout le potentiel.

« Les données batch et temps réel pourront être utilisées ensemble, sur la même plateforme, par les mêmes personnes et avec les mêmes technologies. C'est un pas de géant pour la démocratisation du temps réel. »



À RETENIR

Potentiellement un levier majeur d'efficacité opérationnelle, le temps réel reste limité à quelques cas d'usage en raison de ses contraintes techniques et de son coût.

À coups d'innovation, les grands acteurs de la donnée dans le cloud sont en train d'abattre la séparation entre batch et données temps réel.

Pouvoir exploiter des flux temps réel à moindre coût et les mixer avec des données froides va ouvrir d'immenses possibilités... à utiliser avec discernement.





CULTURE, MÉTHODES, ORGANISATION

MAXIMISER LES BÉNÉFICES DES PROJETS DATA

Si la technologie est nécessaire, elle n'est pas suffisante. À travers leur capacité à s'approprier les outils mis à leur disposition, à imaginer des cas d'usage fondés sur les données et à les mettre en œuvre à l'échelle, ce sont les utilisateurs qui détiennent la clé du succès de la transformation data-centric. Pour véritablement démocratiser la donnée et créer un maximum de valeur, l'entreprise doit s'attacher à diffuser la culture, à mettre place l'organisation et à définir les processus indispensables pour lever toutes formes de freins individuels et collectifs.

Culture data : tirer le maximum des projets et favoriser l'adoption par les utilisateurs ————— **31**

ISABELLE TACHET | LEAD DATA ARCHITECT
GUILLAUME GÉRARD | HEAD OF AI, DATA SCIENCE AND ANALYTICS FRANCE

AI Factories : démocratiser l'usage de l'IA et des Analytics ————— **34**

KEVIN DUVAL | MANAGING ARCHITECT DATA & AI
| OFFERING LEADER FOR AI FACTORIES AND DATA FOR NET ZERO

Industrialiser l'IA : comment produire vite et bien des modèles profitables — **37**

GUILLAUME GÉRARD | HEAD OF AI, DATA SCIENCE AND ANALYTICS FRANCE

La Business Intelligence à l'heure du self-service ————— **41**

SYLVAIN CORNILLE-AVOT | MANAGING APPLICATION CONSULTANT



CULTURE DATA : TIRER LE MAXIMUM DES PROJETS ET FAVORISER L'ADOPTION PAR LES UTILISATEURS



ISABELLE TACHET

Lead Data Architect

Une plateforme de données n'est un succès que si ses utilisateurs s'emparent sans réserve de ses possibilités. Obtenir leur adhésion et leur confiance est un chantier culturel, organisationnel et opérationnel d'une importance capitale.

Bien souvent, on constate de la part des métiers une forme de désenchantement vis-à-vis des solutions data existantes, comme les data warehouses ou les datalakes, dont les bénéfices n'ont que rarement rejoint les attentes. Si aujourd'hui les entreprises modernisent leurs plateformes de données, souvent dans le cloud, il convient de tirer les leçons de ces expériences mitigées et de mettre l'adoption au centre des préoccupations. L'enjeu n'est pas seulement de bâtir une plateforme de données, mais bien d'en maximiser l'utilisation par tous ceux qui, au sein des métiers, ont le pouvoir de transformer la donnée en connaissances, en décisions et en innovations.

L'ACCULTURATION DES UTILISATEURS : LE DATA MINDSET



GUILLAUME GÉRARD

Head of AI, Data Science and Analytics France

Pour utiliser un outil, il ne suffit pas de savoir s'en servir ; il faut aussi savoir à quoi il peut servir. Or, c'est souvent là que le bât blesse. Les utilisateurs sont formés à la manipulation des plateformes de données, mais insuffisamment à la compréhension des services qu'elles peuvent leur rendre. Pour favoriser l'adoption, il ne faut pas seulement leur inculquer le comment, mais aussi le quand et le pourquoi. Il est nécessaire d'investir dans la sensibilisation et l'évangélisation pour créer un terreau culturel, un « Data Mindset », qui les incitera à se tourner spontanément et à bon escient vers la donnée.

Il faut souligner que les réserves ne proviennent pas toujours d'utilisateurs finaux réfractaires au changement et à leurs nouveaux outils. On trouve aussi quantités de dirigeants et de managers sceptiques à l'utilisation de la data quand elle vient bousculer leurs conceptions et leurs pratiques de leur métier. L'indispensable chantier culturel pour démocratiser la data au sein de l'entreprise s'étend donc à tous les niveaux de l'organisation, à commencer par son sommet. Pour changer les attitudes vis-à-vis de la donnée, les dirigeants de l'entreprise doivent être les premiers convaincus que la data est la voie de l'avenir et soutenir ensuite visiblement les initiatives en ce sens.

CRÉER UNE ORGANISATION DÉDIÉE : LA DATA FACTORY

IT et métiers ayant de la donnée des perceptions très différentes, l'acculturation ne peut être menée que par une organisation dédiée, neutre, qui saura faire le lien entre le pourquoi et le comment de la data. Agent de dialogue, force de proposition et inlassable ambassadrice de la data, cette Data Factory sera la cheville ouvrière du changement culturel, aidant les métiers à découvrir toutes les possibilités de valorisation de la donnée au-delà de la seule production d'indicateurs : optimisation,

prédiction, automatisation, monétisation... Cette organisation data sera aussi chargée de mettre en place une gestion centralisée du patrimoine de données, d'installer une gouvernance afin d'encadrer et coordonner les initiatives data, de favoriser la diffusion des bonnes pratiques, et de créer une Design Authority, garante de l'homogénéité des choix techniques et du respect de règles générales (sécurité, conformité, RSE...).

GÉRER LA DONNÉE COMME UN PRODUIT : LE DATA-AS-A-PRODUCT

Une partie des réticences des utilisateurs à s'emparer de la donnée vient du fait qu'elle leur apparaît abstraite, volatile, sensible, obscure... Pour vaincre leurs hésitations, elle doit devenir pour eux un matériau beaucoup plus engageant. Elle doit être considérée comme un produit dont ils seront les consommateurs : un produit qui les séduira par la valeur qu'il peut leur apporter ; un produit qu'ils n'auront ni appréhension ni difficulté à utiliser ; enfin, un actif que l'entreprise est en capacité de produire avec les ressources dont elle dispose.

Assurée par la Data Factory, la gestion de la donnée sera donc guidée par une logique

produit, avec en arrière-plan, les besoins et les attentes des futurs consommateurs. Concrètement, les produits data devront réunir six caractéristiques clé : être « découvrables », accessibles, fiables, intelligibles, interopérables et sécurisés. Par exemple, la mise en place d'approches DataOps/MLOps permettra de s'assurer que les données sont et demeurent de bonne qualité tout au long de leur cycle de vie. Capables d'accéder facilement à des données dont ils seront sûrs de la signification, de la provenance, de la pertinence et de la sécurité, les utilisateurs seront plus enclins à les exploiter, à les partager et à collaborer.

ÉVALUER ET PRIORISER LES CAS D'USAGE

Le cadre et l'impulsion apportés par l'organisation data favorisant l'émergence de nouvelles idées, se pose rapidement la question de savoir lesquelles retenir et dans quel ordre les traiter. Il convient donc de mettre en place un processus d'évaluation et de sélection. Une méthode consiste à projeter les cas d'usage envisagés selon deux axes, l'un mesurant la faisabilité technique, l'autre, la création de valeur espérée. Associant nécessairement les fonctions IT et Data aux métiers, cette approche permettra d'identifier les quick wins (faciles et d'un bon rapport), d'écarter immédiatement les idées peu rentables (trop complexes ou d'un profit limité), et de programmer à plus long terme les idées

prometteuses mais délicates à concrétiser. Il en résultera une double feuille de route, des usages et des technologies, qui permettra d'optimiser l'allocation des efforts et des ressources. Aisément réalisables et visiblement profitables, les premiers projets contribueront aussi, par leur exemple, à diffuser la culture de la data et à engager la dynamique du changement.

« Les utilisateurs sont formés à la manipulation des plateformes de données, mais insuffisamment à la compréhension des services qu'elles peuvent leur rendre. »



À RETENIR

Pour maximiser l'utilisation des plateformes de données, la priorité est de **créer chez les utilisateurs un « Data Mindset »** qui les incitera à se tourner spontanément vers la data.

Chargée de ce chantier culturel, une organisation dédiée à la data s'occupera aussi de **gérer la donnée sous forme de produit pour la rendre plus appétante.**

En sélectionnant et en priorisant les cas d'usage, on pourra cibler des **quick wins visibles et emblématiques**, capables d'enclencher le changement.



AI FACTORIES : DÉMOCRATISER L'USAGE DE L'IA ET DES ANALYTICS



KEVIN DUVAL

Managing Architect Data
& AI
Offering leader for AI
Factories and Data For Net
Zero

Pour répondre aux grands enjeux des entreprises et fédérer toutes les équipes, les dirigeants doivent appliquer des décisions contextualisées, éclairées et intelligibles. Incontournables pour apporter ces éléments d'appréciation et d'explication, l'intelligence artificielle (IA) et l'analyse avancée des données doivent dépasser le morcellement des organisations et l'urgence du court terme.

Alors que près de 9 entreprises sur 10 jugent que l'IA est une priorité stratégique, moins de la moitié des projets d'analytics avancée atteignent la production. Il y a donc urgence à industrialiser les méthodes et les process, et cela en dépit de trois freins majeurs qu'une étude du Capgemini Research Institute¹ a clairement mis

en évidence : un manque critique de compétences data confirmées sur le marché ; un déficit de conduite du changement pour favoriser l'acceptation et l'adoption des solutions fondées sur l'IA ; l'absence de modèles de gouvernance solide permettant de consolider les initiatives locales et de passer à l'échelle.

UN ENJEU ORGANISATIONNEL À ADRESSER

Ces freins ne sont pas technologiques, mais organisationnels et humains. L'enjeu est donc d'établir un modèle opératoire : cadre unifié et partagé qui permet de tirer le meilleur parti des compétences IT et data disponibles, de démocratiser l'usage de l'IA et de l'analytics auprès des métiers, et finalement d'avoir des outils dont décideurs et collaborateurs ont bien confiance dans leurs insights.

La solution est ce que nous appelons une AI Factory, c'est-à-dire un dispositif mutualisé à l'échelle de l'entreprise,

transverse à plusieurs organisations, permettant à la fois de catalyser et de canaliser les initiatives des différentes branches métiers. La mission de l'AI Factory est d'instaurer une culture, des concepts, des méthodes, des pratiques collaboratives et un outillage commun afin de favoriser l'éclosion des projets et de maximiser leurs chances de réussite à grande échelle.

¹ Source: The AI-powered enterprise: Unlocking the potential of AI at scale, Capgemini Research Institute, 2020

PLUS QU'UNE PLATEFORME TECHNOLOGIQUE, UNE STRUCTURE POUR RÉPONDRE AUX DÉFIS COLLECTIFS

Une AI Factory repose sur trois piliers classiques : les personnes, les processus, les outils. En ce qui concerne les personnes, l'objectif est de répartir au mieux les rôles, les responsabilités et les compétences (data scientists, data analysts, data engineers, développeurs professionnels et occasionnels...) pour éviter les doublons et les interférences, aplanir la collaboration entre IT et métiers, harmoniser les approches et offrir à chacun des perspectives individuelles.

Les processus qui irriguent cette organisation doivent permettre de diffuser l'expérience méthodologique et l'expertise technique, de faire remonter les connaissances et les besoins du terrain et de définir les étapes clés de l'industrialisation des projets. Par exemple, il est important d'établir quand et comment un citizen developer doit solliciter les équipes spécialisées pour transformer en solution d'entreprise une idée qu'il aura d'abord explorée avec des capacités exploratoires tel que l'AutoML.

Enfin, concernant l'outillage, l'AI Factory doit mettre en place un socle technologique intégré de production à l'échelle, depuis les pipelines de préparation de la donnée jusqu'au déploiement d'applications embarquant des modèles de machine learning. Ces outils doivent aussi contribuer à accroître la productivité des projets à travers l'industrialisation des méthodes DataOps et MLOps, ainsi que la réutilisation des développements et modules matures.

Selon la maturité de l'entreprise et ses usages, l'AI Factory s'appuie sur différents modèles de gouvernance : centralisée (les projets sont réalisés de bout en bout par une seule équipe dédiée pour toute l'organisation), décentralisée (une structure de ce type est créée dans chaque entité métier) ou « hub & spoke » (les équipes locales sont appuyées par un pôle transversal d'expertise et de coordination). Lorsqu'il est possible, ce dernier modèle se révèle optimal car il permet de concilier expertise et mutualisation, d'une part, et capacité d'initiative et connaissance métier, d'autre part.

QUELS AVANTAGES POUR LES PARTIES PRENANTES ET POUR QUELS RÉSULTATS BUSINESS ?

Le déploiement d'une AI Factory établit un modèle de production associant IT et métiers, permettant de créer de la valeur à la fois rapidement (les cas d'usage aboutissent plus vite), à moindre coût (mutualisation des efforts et capacités) et durablement (les modèles sont maintenus pour rester pertinents et performants). La plateforme IT sous-jacente est activée en mettant à disposition des données de qualité, des capacités exploratoires efficaces, et des capacités de déploiement en continu pour le passage à l'échelle.

En structurant la pratique de l'IA et en instaurant un langage commun, l'AI Factory contribue à créer une culture inspirant confiance dans la stratégie et des résultats intelligibles. Par ailleurs, les gains de productivité permettent aux data scientists de se libérer de tâches à faible valeur ajoutée et d'aborder des sujets fondamentaux qu'il leur fallait souvent reléguer au second plan, comme l'éthique, l'explicabilité des résultats, la gestion des données sensibles ou la maintenance des algorithmes.

Ce modèle d'AI Factory a d'ores et déjà fait ses preuves dans de nombreuses entreprises. Il a ainsi permis à un grand constructeur automobile européen de mettre plus de 100 modèles d'IA en production en seulement cinq mois, ou à un acteur majeur du secteur aéronautique de multiplier en trois ans ses cadences de production par plus de six ! Dans les deux cas, la clé a été de ne pas vouloir forcer la mise en place d'un modèle théorique, mais de chercher à tirer parti de façon pragmatique des forces et des expériences existantes, avec une approche itérative et collaborative de l'innovation.

« La mission de l'AI Factory est d'instaurer une culture, des concepts, des méthodes, des pratiques collaboratives et un outillage commun. »



À RETENIR

Les principales difficultés que rencontrent les entreprises pour faire aboutir leurs projets data ne sont pas technologiques, mais organisationnels et humains.

L'AI Factory est un dispositif de l'échelle de l'entreprise qui fédère, mutualise et organise les ressources et les compétences en matière d'IA.

En diffusant culture, outillage et bonnes pratiques, l'AI Factory favorise l'éclosion des projets et maximise leurs chances de réussite à l'échelle.

INDUSTRIALISER L'IA : COMMENT PRODUIRE VITE ET BIEN DES MODÈLES PROFITABLES



GUILLAUME
GÉRARD

Head of AI, Data Science
and Analytics France

Pour amplifier la dynamique du changement et éviter au contraire qu'elle ne s'essouffle, les initiatives data doivent aboutir plus rapidement et créer visiblement de la valeur. Cela passe par l'industrialisation de la production des modèles d'IA.

Désormais largement convaincues du potentiel de l'intelligence artificielle (IA), les entreprises investissent sur le sujet avec des ambitions claires et des objectifs précis. Pourtant, beaucoup constatent des résultats en deçà de leurs attentes. Selon une étude du Capgemini Research Institute¹, seulement 13 % des projets IA sont déployés avec succès et appliqués à grande échelle. Bien souvent, la raison principale tient aux difficultés rencontrées, ou sous-estimées, lors du passage à l'échelle des modèles.

¹ Source: "[The AI-powered enterprise: Unlocking the potential of AI at scale](#)", Capgemini Research Institute, 2020.

LES ÉLÉMENTS CLÉS DE SUCCÈS DE L'INDUSTRIALISATION

Pour tout projet d'IA, le passage à l'échelle est une étape d'autant plus délicate que la marche est particulièrement haute entre le prototype et la solution opérationnelle. Pour qu'un outil d'IA porte pleinement ses fruits, plusieurs aspects sont à prendre en compte dès le départ :

FACTEURS OPÉRATIONNELS

Au-delà de la pertinence et de la performance théoriques du modèle, il faut s'assurer que le temps de traitement, la fréquence de rafraîchissement ou encore la présentation des résultats seront compatibles avec les usages envisagés. Ces facteurs doivent être définis sous forme d'indicateurs de performance et être suivis tout au long du projet. Il faut aussi veiller à ce que les utilisateurs acceptent et adoptent la solution, si possible en leur donnant les moyens d'en comprendre et d'en expliquer les résultats.

FACTEURS TECHNIQUES

L'entreprise doit pouvoir mettre en œuvre les moyens IT que suppose la solution, par exemple pour traiter en temps réel de très forts volumes de données ou gérer des bases de données orientées graphe.

FACTEURS ENVIRONNEMENTAUX

L'empreinte environnementale des infrastructures, des données et des traitements, notamment leur consommation énergétique, ne doit pas être démesurée par rapport aux objectifs visés.

FACTEURS DE RENTABILITÉ

Il faut vérifier que le coût de ces infrastructures n'excèdera pas les bénéfices attendus, ce qui n'est pas toujours acquis. L'approche FinOps permet de traiter de ces problématiques en ayant une démarche maîtrisée de suivi des coûts. Dans ce calcul de ROI, on tiendra aussi compte des gains extra-financiers potentiels, par exemple en matière de productivité, de qualité de service, de satisfaction client et employé, ou d'impact environnemental.

FACTEURS RÉGLEMENTAIRES

Dans le cas où le RGPD s'applique aux données utilisées, il faut prendre les dispositions adéquates pour les anonymiser, les protéger, et obtenir le consentement des personnes.

FACTEURS ÉTHIQUES

Il est essentiel de s'assurer que le modèle et l'usage qui en sera fait sont bien compatibles avec les valeurs de l'entreprise et qu'ils resteront dans les limites qu'elle se fixe.

DÉRIVE DES DONNÉES ET DÉRIVE DES USAGES

À ce jeu de contraintes, s'ajoute une particularité propre à l'IA : son besoin fondamental de suivi. Du jour où elle est lancée, une IA peut en effet voir ses performances se dégrader peu à peu jusqu'à passer au-dessous d'un seuil minimum d'acceptabilité. Un modèle étant conçu pour répondre à une question précise à partir d'un ensemble de données, sa réponse perdra en pertinence sitôt que les données ou la question évolueront.

Le premier cas, la dérive des données, est très courant car, on choisit généralement pour le POC un jeu de données stable, restreint et de bonne qualité, ce qui ne reflète pas toujours les conditions réelles. Pour maintenir la qualité des résultats, il faudra donc être capable de réentraîner le modèle sur des données plus à jour et plus représentatives.

Plus insidieux, le second cas, la dérive des usages, tient à l'évolution des attentes des utilisateurs. Alors qu'elle correspondait parfaitement à leur besoin à l'origine, l'IA peut, par exemple, ne plus leur suffire, le contexte ayant changé, ou bien continuer à attirer leur attention sur des problèmes désormais réglés. L'application ne leur étant plus utile, ils s'en détournent progressivement.

Ces phénomènes de dérive, communément appelés "drift", sont tout à fait naturels et souvent inévitables. Ils doivent par conséquent être anticipés, détectés et réglés au fil de l'eau. C'est pourquoi l'IA doit être considérée et gérée comme un produit, en constante évolution. Contrairement à la plupart des solutions IT, le jour de la mise en production ne marque donc pas la fin du projet, mais seulement celle du premier cycle d'un processus d'apprentissage continu.



UNE APPROCHE INDUSTRIALISÉE DU CYCLE DE VIE

Pour adresser les multiples contraintes susceptibles d'entraver le passage à l'échelle d'un modèle d'IA, puis corriger au fur et à mesure ses éventuelles dérives, une approche industrialisée et orientée produit couvrant l'ensemble de son cycle de vie est nécessaire. Cette démarche comporte cinq étapes :

1 LA DÉFINITION DU PROBLÈME

dans un premier temps, il est nécessaire que métiers et équipes data aient une même compréhension des résultats attendus. Pour cela, le métier doit exposer la problématique à traiter aux data scientists, lesquels émettent alors un premier diagnostic de faisabilité. Cette première étape est également l'occasion de définir les différents KPIs à suivre tout au long du projet.

3 L'INDUSTRIALISATION DU MODÈLE

on déploie le modèle après s'être assuré qu'il sera exploitable en environnement de production. Cela nécessite de traiter plusieurs problématiques concernant les performances et les délais de traitement, de la capacité du modèle à bien fonctionner sur les données de production, et enfin le respect des critères environnementaux.

5 LA FORMATION CONTINUE

le modèle est constamment réentraîné sur des données rafraîchies et les nouvelles versions peuvent être automatiquement déployées lorsqu'elles apparaissent supérieures, à la fois sur le plan de la pertinence mathématique et de l'intérêt pratique.

2 LA CONSTRUCTION DU MODÈLE

lors de cette étape, essentiellement confiée aux data scientists, on ne se limite à la précision mathématique du modèle, mais on intègre d'emblée des indicateurs mesurant ses bénéfices pour le métier ainsi qu'un questionnement « by design » sur les aspects opérationnels, éthiques, réglementaires, etc.

4 LA SUPERVISION

une fois la solution en production, on met en place un suivi du modèle, de ses performances, de son utilisation, de la qualité des données, et on positionne des seuils d'alerte afin de détecter au plus tôt les éventuelles dérives.

« Contrairement à la plupart des solutions IT, le jour de la mise en production ne marque pas la fin du projet d'intelligence artificielle, mais seulement celle du premier cycle d'un processus d'apprentissage continu. »

L'EXEMPLE RÉUSSI D'UN ACTEUR MAJEUR DES TÉLÉCOMS

Un opérateur télécom a eu recours à cette démarche pour mettre en œuvre très rapidement une solution d'IA qui lui a permis de répondre aux injonctions de son autorité de régulation et de renforcer sa qualité de service pour ses clients. Pressée d'améliorer significativement la qualité des déploiements de fibre optique, l'entreprise s'est appuyée sur une information dont elle disposait déjà : les photos que prennent les installateurs avant et après le raccordement. En vérifiant certains points de contrôle sur ces images, une solution basée sur l'IA est capable d'identifier les branchements corrects et ceux qui nécessiteraient l'examen d'un technicien.

La démarche d'industrialisation a notamment permis de mettre en place un réentraînement continu du modèle afin de prévenir les éventuelles dérives de données (la qualité des photos évolue avec les nouvelles générations de smartphones, les conditions de prise de vue dépendent des saisons...) et d'usage (par exemple, s'il fallait prendre en compte des fils d'une couleur supplémentaire à la suite de l'arrivée d'un nouvel opérateur).

Permettant de sécuriser et d'accélérer les projets d'IA, la démarche d'industrialisation repose classiquement sur la mise en place d'un triptyque Outils/Processus/Organisation. Il faut cependant insister tout particulièrement sur l'importance d'associer étroitement les parties prenantes métier, IT et data. Un déficit de collaboration entre ces acteurs est en effet un fréquent facteur d'échec des projets, où l'on s'acharne parfois en vain sur des problèmes sans solution ou sur des solutions sans problème ! Mettre tout le monde autour de la table dès le départ, et à chacune des étapes, permet d'éviter de telles impasses et d'aborder les projets d'IA dans une perspective résolument tournée vers leur passage à l'échelle et la maximisation de leurs bénéfices.

À RETENIR

Seule une minorité de projets d'IA apporte les bénéfices attendus, la plupart butant sur l'étape délicate du passage à l'échelle.

Outre la multiplicité des contraintes à prendre en compte à la conception, un modèle d'IA doit être sans cesse suivi pour qu'il reste pertinent et performant.

Pour créer de la valeur, une approche industrialisée, orientée produit, couvrant l'ensemble cycle de vie des modèles et associant tous les acteurs est indispensable.

LA BUSINESS INTELLIGENCE À L'HEURE DU SELF-SERVICE



SYLVAIN
CORNILLE-AVOT

Managing Application
consultant

Stade ultime de la démocratisation de la data, le self-service se développe dans le domaine de la Business Intelligence. Les utilisateurs peuvent manipuler les données à leur guise tout en restant dans un cadre maîtrisé par l'IT.

Doyenne de l'usage des données, la Business Intelligence (BI) continue de bien se porter dans les entreprises, où les informations qu'elle apporte restent indispensables à la prise de décision et au pilotage de la performance. Son développement continu s'accompagne aujourd'hui d'une évolution majeure avec l'essor de solutions en libre-service qui offrent davantage d'autonomie aux métiers.

Le self-service reporting permet aux utilisateurs métier de créer eux-mêmes des rapports dont la mise en œuvre effective est ensuite assurée par l'IT. Cette nécessaire intervention entraîne des délais liés aux cycles incompressibles des projets et, bien souvent, au déficit de compétences métier des équipes IT. Pour s'affranchir de cette contrainte, les métiers sont parfois tentés de développer seuls leur propre solution, mais ce « shadow IT », en marge des principes directeurs de la DSI, est une source importante de risques (sécurité, conformité...) et d'inefficiences (silos, coûts, dette technique...).

Pour réconcilier l'aspiration des métiers à davantage d'agilité et de rapidité, et le nécessaire contrôle des

systèmes d'information, les entreprises investissent de plus en plus dans des solutions de self-service BI, adoucies par la DSI, qui permettent aux différents départements de l'organisation d'explorer, analyser et visualiser les données directement à la source et en toute autonomie.

Grâce au self-service BI, les métiers sont en mesure de construire et modifier eux-mêmes leurs tableaux de bord au rythme de leur activité : suivre les ventes d'un nouveau produit, créer un nouveau segment de clientèle, introduire un nouvel indicateur de performance... Plus proches de la donnée, les utilisateurs peuvent l'exploiter à leur guise, en tirer des connaissances nouvelles, et prendre plus vite des décisions plus pertinentes.

Toutefois, la mise en place d'un self-service BI ne se résume pas à la simple implémentation d'un outil car, en arrière-plan, les enjeux de sécurité, de conformité, et de maîtrise (des technologies, des coûts, des données...) demeurent et doivent impérativement être traités.

UNE NOUVELLE RÉPARTITION DES RÔLES ENTRE IT ET MÉTIERS

Le self-service BI ne signifie pas l'abandon par l'IT de toutes ses prérogatives au bénéfice des métiers, mais une répartition différente des rôles. L'IT reste en effet la garante de la plateforme. Elle en assure le déploiement, l'exploitation et la maintenance. Elle en réalise les évolutions techniques et fonctionnelles souhaitées par les utilisateurs. Elle met en place les règles et les contrôles de sécurité et de conformité. En phase de production, et particulièrement pour les solutions cloud, elle pilote l'optimisation dynamique des coûts et des infrastructures (FinOps), et surveille les éventuelles dérives. Elle assure par ailleurs la formation des utilisateurs à leurs nouveaux outils et les sensibilise à leurs possibilités comme à leurs limites. Elle joue aussi un rôle clé en amont dans la gouvernance des données et la gestion de leur qualité, gage de la fiabilité et de la valeur des résultats. Enfin, elle accompagne les métiers dans la mise en œuvre de leurs cas d'usage dans un rôle de support, d'expertise et de garde-fou.

Pour leur part, les métiers ont entièrement la main sur l'exploitation des données. Ils prennent donc en charge l'ensemble le cycle de l'analyse : identification des enjeux pour l'entreprise, traduction de ces enjeux en indicateurs de performance, construction de ces KPI à partir des données disponibles et mise en forme de la restitution et de la visualisation des résultats. Leur périmètre englobe donc à présent une importante composante technologique. Ceci suppose le développement de nouvelles compétences et la mise en place de nouveaux rôles afin d'être en mesure d'identifier, maîtriser et diffuser les nouveaux moyens d'analyse désormais à leur disposition, et ainsi pouvoir en exploiter tout le potentiel.

« Le self-service BI ne signifie pas l'abandon par l'IT de toutes ses prérogatives au bénéfice des métiers, mais une répartition différente des rôles. »

COMMUNICATION ET AGILITÉ, LES MAÎTRES-MOTS DE LA COLLABORATION

Initiée au niveau de la gouvernance, la collaboration étroite entre l'IT et les métiers doit se prolonger sur le terrain à l'occasion de la mise en œuvre des cas d'usage. Si les métiers sont désormais à la manœuvre, l'apport de l'IT reste essentiel pour accélérer et sécuriser les projets, et l'approche agile est idéale pour articuler cette complémentarité dans un environnement extrêmement changeant. Une bonne communication n'est seulement nécessaire entre l'IT et les métiers, mais aussi entre les métiers eux-mêmes. Bien souvent, chacun dispose d'une pièce du puzzle que les anciens silos empêchaient de réunir. Avec un self-service BI transverse, les différents services peuvent plus aisément mettre leurs connaissances en commun pour mieux analyser les situations et apporter une réponse optimale. La communication dans un cadre agile devient un vecteur de progrès pour l'entreprise.

En matière de Business Intelligence, self-service n'est donc pas synonyme de laisser-aller, bien au contraire ! En contrepartie de plus d'autonomie dans la gestion et la production de leurs indicateurs et de leurs tableaux de bord, les métiers ont à endosser des responsabilités accrues dans le cadre redéfini de leur coopération avec l'IT. Si cette relation fonctionne de manière fluide et efficace, alors les bénéfices seront indéniables : les décisions seront plus rapides et plus pertinentes, l'information sera mieux partagée et mieux exploitée, et les coûts comme les risques associés à la plateforme seront réduits et maîtrisés. Pour les métiers, ce sera aussi un vecteur important pour diffuser une culture de la donnée et une maturité sur les questions techniques qui pourra, par exemple, trouver son prolongement dans la mise en place d'un Data Mesh.

LA GOUVERNANCE, PIÈCE ESSENTIELLE DE LA RÉUSSITE

Dans le cadre d'un self-service BI, IT et métiers ont donc des attributions très complémentaires et le succès de la plateforme tiendra à leur capacité à jouer leurs rôles en harmonie. C'est pourquoi il est essentiel de mettre en place une gouvernance globale afin d'encadrer et coordonner ce travail en commun autour de la plateforme. C'est à ce niveau qu'il conviendra de délimiter précisément les champs d'intervention et de responsabilité de chacun, de choisir les outils, de définir les règles et les paramètres généraux (de sécurité, de qualité de service...) et d'établir une feuille de route partagée. On mettra également en place un comité chargé de valider les cas d'usage envisagés au regard de critères à la fois techniques et métiers.



À RETENIR

Grâce au self-service BI, les utilisateurs métiers peuvent **manipuler les données en toute autonomie**, gagnant ainsi en agilité et en rapidité.

La responsabilité de la plateforme se partage entre les métiers, davantage impliqués, et l'IT, qui garde la main sur les grands enjeux techniques.

Encadrée par une gouvernance partagée, la collaboration entre IT et métier autour du self-service BI doit être **agile et reposer sur une communication fluide**.

POUR EN SAVOIR PLUS

www.capgemini.com/fr-fr/services/data-intelligence-artificielle/

À PROPOS DE CAPGEMINI

Capgemini est un leader mondial, responsable et multiculturel, regroupant 360 000 personnes dans plus de 50 pays. Partenaire stratégique des entreprises pour la transformation de leurs activités en tirant profit de toute la puissance de la technologie, le Groupe est guidé au quotidien par sa raison d'être : libérer les énergies humaines par la technologie pour un avenir inclusif et durable. Fort de 55 ans d'expérience et d'une grande expertise des différents secteurs d'activité, Capgemini est reconnu par ses clients pour répondre à l'ensemble de leurs besoins, de la stratégie et du design jusqu'au management des opérations, en tirant parti des innovations dans les domaines en perpétuelle évolution du cloud, de la data, de l'Intelligence Artificielle, de la connectivité, des logiciels, de l'ingénierie digitale et des plateformes. Le Groupe a réalisé un chiffre d'affaires de 22 milliards d'euros en 2022.

Get the Future You Want* | www.capgemini.com

**Capgemini, le futur que vous voulez.*

Les informations contenues dans ce document sont la propriété exclusive de Capgemini. Copyright © 2023 Capgemini. Tous droits réservés.