

N°5 | BAROMÈTRE DES DIRECTIONS DATA 2024

PLUS DE 50 DIRIGEANTS
S'EXPRIMENT SUR LEUR
TRAJECTOIRE
DE MATURITÉ DATA ET IA

EXECUTIVE SUMMARY

Comment introduire cette 5ème édition de notre baromètre des Directions data par un autre sujet que l'IA générative ? Le tsunami a déferlé à une vitesse inédite pour une technologie qui est rapidement devenue l'une des priorités des Directions data.

Notre étude qualitative auprès d'une cinquantaine de dirigeants et responsables data de grandes et moyennes entreprises françaises confirme que le lancement de ChatGPT, il y a 1 an, a sorti l'IA d'une phase temporaire de désillusion au sein de plusieurs entreprises. En devenant accessible au grand public via une interface intuitive, la technologie s'est invitée à l'agenda des directions générales.

Un changement de paradigme qui suscite des attentes fortes et qui a démultiplié les sollicitations auprès des directions data et DSI : 80% des sondés estiment que la data gagne en priorité au sein de leur entreprise avec l'essor de l'IA générative.

Sur le sujet, les priorités des dirigeants interrogés se concentrent autour de 4 points :

- **Lancer rapidement de nouvelles expérimentations** pour répondre aux attentes et se forger de premières convictions sur la valeur des usages et les choix technologiques. A l'été 2023, 25% des sondés avaient déjà initié au moins une expérimentation et plus des 2/3 avaient lancé des groupes de travail stratégique sur le sujet. Des chiffres à considérer sensiblement à la hausse aujourd'hui. Si les usages potentiels sont nombreux, la majorité des expérimentations constatées tournent autour des « usages RAG » (Retrieval Augmented Generation), c'est-à-dire la capacité à récupérer des données spécifiques à un contexte dans une base de données externe, et à les mettre à la disposition d'un LLM lorsqu'on lui demande de générer une réponse. La prochaine vague devrait s'appuyer sur le déploiement d'applications bureautiques comme Microsoft Copilot mettant la technologie à disposition de toute l'entreprise.
- **Former les utilisateurs.** Ce déploiement à grand échelle va nécessiter de traiter l'interaction avec les IA génératives comme une compétence à maîtriser largement, sous peine de s'exposer à un usage sous-optimal ou à l'inverse excessif de la technologie.
- **Evaluer les risques et en dégager un cadre global.** Les craintes des Dirigeants data se portent notamment sur la fiabilité de modèles, connus pour halluciner, et dont la performance est difficile à mesurer avec tous les risques (opérations, financiers, réputation) que cela induit. Ce contexte, ainsi que la publication de l'AI Act Européen, remet l'IA de confiance au centre des préoccupations. Construire des IA génératives de confiance est un défi identifié par la majorité des entreprises, et des premières bonnes pratiques émergent.
- **Adapter son modèle opérationnel aux spécificités de l'IA générative.** Au-delà des POC et expérimentations qui sont menés actuellement, les sondés partagent plusieurs interrogations sur les éléments à intégrer pour pouvoir passer en 2024 de premiers usages en production. Coté technologie, si l'heure est aux tests, l'évolution des plateformes et outils MLOps commence à peine chez les acteurs les plus matures.

L'IA générative fait irruption dans un moment où la majorité des Directions data a globalement atteint un palier de maturité : en moyenne, 60% des usages IA passent désormais en production et 80% des entreprises ont inscrit la data dans leur plan stratégique.

Cependant, les Directions data peinent toujours à casser leur plafond de verre actuel et à avoir un impact fort à l'échelle de l'entreprise et de son modèle d'affaire. Bien que l'hybridation des modèles organisationnels entre équipe centrale et relais locaux dans les métiers soit effective, de nombreuses Directions data sont encore essentiellement dans une position de DOper de la data. Cela provoque à la fois un goulot d'étranglement autour des usages et des plateformes, et des difficultés à sortir d'usages « tactiques » et non stratégiques pour l'entreprise.

Dans ce contexte, les Dirigeants data identifient 3 défis pour augmenter leur impact à l'échelle de l'entreprise et casser ce plafond de verre :

1 – S'inscrire dans les enjeux stratégiques de l'entreprise. Pour réussir, les sondés identifient plusieurs facteurs clés de succès :

- Répondre à l'attente du top management et porter une vision stratégique sur l'apport de la data dans la transformation de l'entreprise et de son modèle d'affaire ;
- Rééquilibrer leur offre de service pour mieux se positionner comme *enabler* auprès des Directions métiers en demande d'autonomie et de capacités robustes, mais aussi auprès des DSI et des Directions digitales pour s'intégrer nativement dans leurs grands programmes ;
- Porter des usages transverses et emblématiques de la donnée permettant de se positionner en tierce partie et d'aligner les enjeux des différentes Directions métiers ;
- Co-sponsoriser des projets ambitieux Business/Data pour inscrire la data dans le modèle d'affaire de l'entreprise à travers des produits robustes et matures inscrits dans les processus métiers.

2 – Piloter la direction data par la valeur. Si l'estimation du ROI en amont des projets est devenue courante, à des fins de priorisation et de sponsorship, la valeur réellement dégagée par les solutions data et IA est encore largement méconnue. Seuls 16% des CDO sont responsabilisés sur des objectifs de création de valeur chiffrés. La hype autour de l'IA générative et de ses promesses de gains de productivité va accélérer la nécessité de mesurer réellement la valeur des investissements data.

Les engagements communs de création de valeur par le métier et la data, la Direction financière comme partenaire, et la formalisation d'une méthode partagée de valorisation sont autant de facteurs clés de succès partagés par les Dirigeants data.



3 – Démultiplier l'impact de la data à l'échelle avec le Data Mesh. Pour accélérer la mise en production de projets emblématiques, les data office doivent créer un écosystème fertile en mettant à disposition les capacités nécessaires aux différentes parties prenantes. Près de 80% des entreprises commencent à déployer les principes du Data Mesh pour atteindre cet objectif. Mais les « projets Data Mesh » émergent encore trop souvent de façon non coordonnée et sans trajectoire commune avec les parties prenantes. Pourtant, plus qu'une transformation technologique, le Data Mesh est avant tout un chamboulement organisationnel visant à décentraliser le data management. La majorité des CDO cherchent à délivrer des résultats rapidement pour maintenir la dynamique et ne pas rencontrer les mêmes écueils que sur la gouvernance des données. Là aussi, quelques bonnes pratiques émergent, s'éloignant du cadre théorique jugé souvent inatteignable pour cibler des victoires accessibles autour des data domains et des data products.

Relever ces 3 défis fait d'autant plus de sens que la capacité des organisations data à tirer pleinement partie des opportunités offertes par l'IA générative va être particulièrement scrutée. Pour de nombreuses entreprises, l'objectif est de faire de 2024 l'année du passage en production de ces nouveaux usages et du déploiement d'outils dans toute l'entreprise, embarquant nativement cette technologie. Cet objectif va constituer un « stress test » pour la maturité des organisations data et leur capacité à embarquer l'entreprise dans une transformation durable.

Jonathan CASSAIGNE,
Directeur des Expertises de Quantmetry

Présentation du baromètre

08

01

L'IA générative au cœur des préoccupations des entreprises

12

L'IA générative replace l'IA au cœur des préoccupations des entreprises	14
ZOOM - Qu'est-ce que l'IA générative ?	18
Les 4 enjeux des Chief Data Officers pour rester dans la course	22
La mise en production des IA génératives soulève de nouveaux défis	28
SUCCESS STORY – Leader dans le secteur de la parfumerie	32
SUCCESS STORY – Leader dans le secteur des services financiers	34
ZOOM - AI Act & IA de confiance : une accélération pour mitiger les risques réglementaires & opérationnels	36

02

Les organisations se transforment pour la data et cherchent désormais à se transformer par la data

44

Les entreprises maîtrisent l'industrialisation de la data et de l'IA	46
La data et l'IA sont adoptées par tous dans les entreprises	50
ZOOM - La data au service des enjeux RSE & développement durable	52
Les entreprises cherchent désormais à se transformer par la data	56

03

Trois grands défis à relever en 2024 pour adresser l'ensemble des leviers de génération de valeur par la data

64

DÉFI 1 - S'inscrire dans les enjeux stratégiques de l'entreprise	66
SUCCESS STORY – Leader dans le secteur de l'ingénierie	70
DÉFI 2 - Piloter les Directions data par la valeur	76
DÉFI 3 – Démultiplier l'impact à l'échelle avec le Data Mesh	82

Conclusions et perspectives

92

Qui sommes-nous ?

102

SOMMAIRE

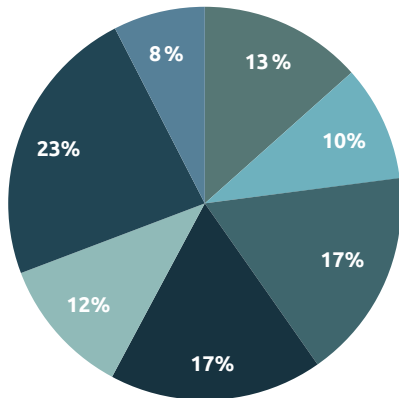
PRÉSENTATION DU BAROMÈTRE

Un panel de 53 répondants couvrant principalement 7 secteurs d'activités

Nous avons interrogé 53 Dirigeants et Responsables data de grandes, moyennes entreprises et start-up. Au cours d'entretiens, nous avons évoqué leurs enjeux, leurs organisations, leurs priorités, leurs réussites et leurs difficultés. L'objectif de ces échanges était de :

- Analyser les progrès réalisés depuis notre Baromètre 2022
- Faire ressortir les principaux enjeux partagés par la majorité des directions data
- Comprendre comment les directions data abordent ces enjeux

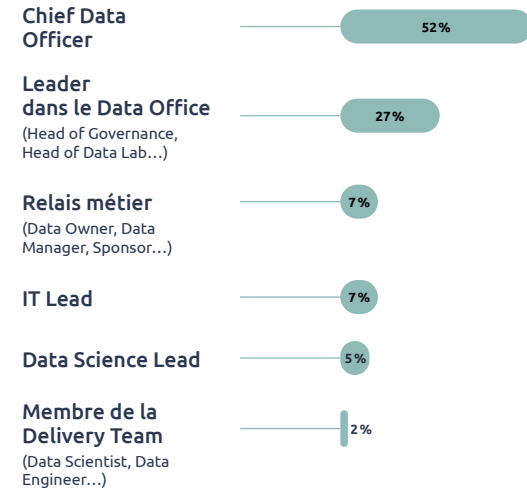
Répartition des entreprises par secteurs d'activité



- Assurance
- Banque
- Autre
- Biens de consommations/retail/distribution
- Energie / Utilités
- Industrie
- Service public

Postes occupés

% de répondants



- AG2R, April, CCR RE, Groupama, Matmut, SCOR
- Arkea, La Banque Postale, BPCE, BNP Paribas, Crédit Mutuel, LCL
- GSF, Orange, Santeclair, Sodexo, SNCF
- Chanel, Christian Dior Couture, Danone, Décathlon, ETAM, Eureden, L'Oréal, LSDH, Manutan, Terrena, Ynsect
- EDF R&D, Engie, GRTGaz, Orano, Schneider, Technip, Veolia Eau France
- Arcelor Mittal, EGIS, Eramet, MBDA, Michelin, Modulaire Group, Plastic Omnium, Safran, Saint Gobain, setec
- Ministère de la Culture, Ministère de l'Intérieur et des Outre-mer, Pôle Emploi, Région Ile-de-France

- Partie 1 -

L'IA GÉNÉRATIVE AU CŒUR DES PRÉOCCUPATIONS DES ENTREPRISES

L'IA générative replace l'IA au cœur des préoccupations des entreprises

Un retour en puissance de l'IA

Le lancement de ChatGPT fin 2022 et sa courbe d'adoption plus rapide qu'aucune autre application numérique avant elle, a ouvert la voie au boom de l'IA générative. Un retour en force de l'IA que même les observateurs les plus avertis n'avaient pas anticipé. L'IA se trouve pour la première fois à la portée de tous, accessible en ligne via une interface intuitive. Un changement de paradigme qui a modifié en profondeur les enjeux des directions data. Dissuadées par un ROI incertain et des difficultés à passer en production, les entreprises avaient, en 2022, diminué leurs investissements dans l'IA au profit de projets data moins spectaculaires (BI, Data Quality...). Mais l'IA générative, avec ses promesses de gain de productivité conséquent dans de nombreuses industries et ses premiers usages prometteurs, a remis l'IA au cœur des préoccupations des entreprises et de leurs dirigeants.

3 capacités techniques permettant d'envisager de nombreux usages

L'IA générative apporte de la valeur via 3 capacités clés

1 - Interfaces en langage naturel

L'IA générative peut interpréter une consigne exprimée par un utilisateur en langage naturel, par écrit ou à l'oral. Elle facilite ainsi grandement l'interface avec les outils, applications et services utilisés par les employés, les clients ou les partenaires de l'entreprise.

2 - Génération de contenu

Suivant un modèle et/ou des instructions données par l'utilisateur, l'IA générative produit des contenus variés sous la forme de code, d'image, de vidéo ou de texte par exemple. Le contenu généré s'appuie sur les milliers d'exemples connus par l'IA, mais il reste inédit car (a) il est produit spécifiquement pour répondre à la requête utilisateur (b) sa création est souvent déterminée de façon partiellement aléatoire.

3 - Recherche & synthèse documentaire

Sur la base d'une requête utilisateur, l'IA générative peut identifier, au sein d'un corpus, les documents contenant des informations pertinentes, puis s'appuyer sur leur contenu pour formuler une réponse informée.

Ces trois composants fondamentaux peuvent être déployés dans des solutions en réponse à de multiples besoins métiers couvrant toutes les fonctions de l'entreprise.

Les cas d'usage pré-identifiés aujourd'hui et les plus matures sont en grande majorité des sujets de RAG (Retrieval Augmented Generation) reposant sur les capacités 1 et 3.

Fonction	Exemples de cas d'usages	Recherche & synthèse documentaire	Génération de contenu	Interfaces langage naturel
R&D	Proposer des designs produit / formulations chimiques	●	●	●
Achat	Analyser les clauses contractuelles / propositions commerciales	●	●	●
Supply Chain	Générer des rapports d'inspection de flotte sur la base d'enregistrement audio terrain	●	●	●
Production	Analyser les causes racine d'incidents et proposer des pistes d'amélioration	●	●	●
Marketing & vente	Développer des scénarios de vente hyper-contextualisés	●	●	●
SAV	Proposer des réponses personnalisées aux clients via un chatbot	●	●	●
Finance	Automatiser la rédaction de rapports financiers	●	●	●
IT	Automatiser le cycle d'ingénierie logicielle, (spécifications, développement, support)	●	●	●
Data	Créer des données structurées sur la base d'images ou de texte	●	●	●
RH	Proposer des réponses personnalisées aux employés via un chatbot	●	●	●
Autres	Assister les employés dans leurs tâches du quotidien (emails, recherche de documents...)	●	●	●

● Capacité essentielle ● Capacité optionnelle

Leviers de valeur

- | | |
|--|--|
| <div style="background-color: #4682B4; color: white; padding: 5px; writing-mode: vertical-rl; transform: rotate(180deg); font-weight: bold;">EFFICACITÉ</div> <ul style="list-style-type: none"> • Travailler plus vite • Accélérer des processus métier • Se centrer sur des tâches à plus haute valeur • Abaisser le niveau de compétences nécessaires pour accomplir certaines tâches | <div style="background-color: #4682B4; color: white; padding: 5px; writing-mode: vertical-rl; transform: rotate(180deg); font-weight: bold;">QUALITÉ</div> <ul style="list-style-type: none"> • Être plus créatif • Mieux diffuser l'expérience / l'expertise au sein de l'entreprise • Passer plus de temps sur le perfectionnement que sur le brouillon |
|--|--|

Un mélange d'intérêt, de craintes et de scepticisme

Les entreprises entament à peine leur appropriation des IA génératives. Dans l'ensemble, elles cherchent à se forger de premières convictions sur la valeur, et à avancer avec prudence.

68%

des sondés ont déjà lancé des groupes de travail stratégiques ou opérationnels sur les IA génératives.

Parmi eux,

25%

avaient déjà initié au moins une expérimentation en juillet 2023

Les principales craintes se portent sur :

- la performance et la robustesse des modèles connus pour halluciner
- les impacts légaux et la propriété intellectuelle
- la dépendance envers les prestataires
- les coûts
- l'impact sur l'emploi et les compétences



Nous questionnons sur la manière d'utiliser ces modèles : réentraînement, coût, stockage et sur les problématiques de propriété intellectuelle. >>

CDO, secteur banque



Des attentes fortes envers les Directeurs data et DSI

La promesse d'une IA démocratisée a démultiplié les sollicitations auprès des Directions data et DSI.

80%

des sondés estiment que la data gagne en priorité au sein de leur entreprise avec l'essor de l'IA générative

D'un côté, cet intérêt représente l'opportunité d'obtenir un sponsorship et des budgets, y compris sur les fonctions support jusqu'alors moins actives sur la data et l'IA. De l'autre, les attentes COMEX et métiers vont mettre à l'épreuve le modèle opérationnel data mis en place au cours des dernières années. Les Directions data et DSI devront sans doute les faire évoluer autour d'interrogations stratégiques clés:

Quelle ambition ?

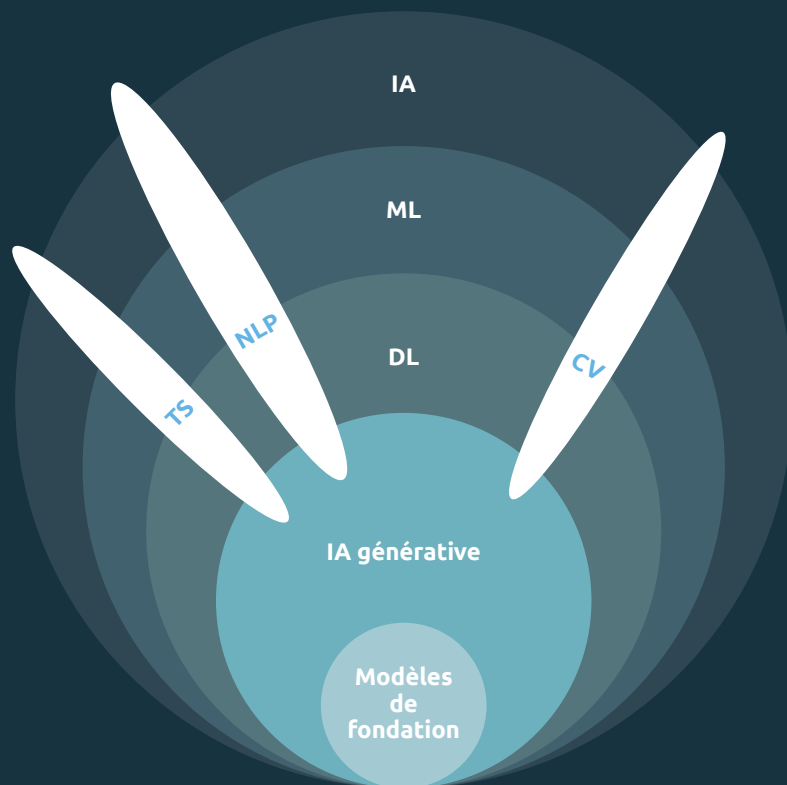
- Sur quels usages investir ?
- Faut-il prioriser l'IA générative au détriment des autres IA ?
- Comment mitiger les risques inhérents à cette technologie ?

Comment réussir ?

- Faut-il faire évoluer le socle technologique ?
- Quelle organisation et quelle gouvernance ?
- Faut-il recruter et former ?

ZOOM

Qu'est-ce que l'IA générative ?



- IA -

Intelligence Artificielle

Discipline informatique qui vise à créer des systèmes capables de réaliser des tâches qui nécessitent généralement une intelligence humaine.

- ML -

Machine Learning

Sous-domaine de l'IA qui se concentre sur le développement de techniques permettant aux ordinateurs d'apprendre à partir de données et d'améliorer leur performance sur des tâches spécifiques sans être explicitement programmés.

- DL -

Deep Learning

Sous-catégorie du Machine Learning qui utilise des algorithmes structurés en réseaux pour représenter les relations entre les données. Cette approche est plus coûteuse mais peut offrir de meilleures performances dans de nombreux contextes.

- NLP -

Natural Language Processing

Branche de l'IA qui se concentre sur la compréhension, la génération et la manipulation de textes et de discours par les ordinateurs.

- TS -

Time Series

Branche de l'IA qui se concentre sur les séquences de données chronologiques. Elle vise par exemple à comprendre et à prédire des tendances au fil du temps et à détecter des anomalies.

- CV -

Computer vision

Branche de l'IA qui permet aux ordinateurs de voir et comprendre en analysant des images et des vidéos.

IA générative

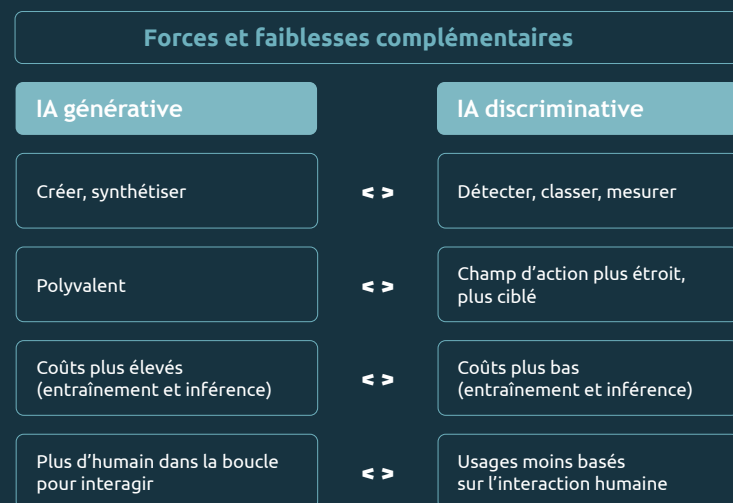
Théories et techniques permettant de générer de nouvelles données inexistantes auparavant (texte, image, son, etc.) en réponse à des instructions.

Modèles de fondation

Modèles d'intelligence artificielle entraînés de manière générique sur des données à grande échelle, et pouvant être adaptés à une large collection de tâches spécifiques.

Les IA génératives et discriminatives sont synergiques

L'IA générative ne doit pas être perçue comme la nouvelle génération d'IA venue remplacer les pratiques préalables. Les IA génératives et discriminatives vont continuer de prospérer côte à côte, et bien souvent coexister au sein d'une même solution.



Termes clés

Vecteur

Représentation mathématique de mots, d'images ou d'autres éléments permettant à un ordinateur d'évaluer la distance sémantique entre les concepts et d'effectuer des calculs du type *roi - homme + femme = reine*.

LLM - Large Language Model

Ces modèles servent de base pour les applications d'IA génératives fondées sur du texte. Ils sont des représentations statistiques d'une langue, construites sur la base de corpus de données pouvant atteindre plusieurs milliards de mots. Un LLM est conçu pour prédire le prochain mot le plus probable dans une séquence de type « le ciel est ... » (bleu, gris, etc).

Hallucination

Les LLM s'appuient sur leur connaissance des relations statistiques entre les concepts qui constituent une langue. Cela leur permet de construire des phrases souvent intelligibles et cohérentes. En revanche, la véracité des propos générés par les LLM n'est jamais garantie. Les LLM ont tendance à produire des « hallucinations » ou des contenus factuellement faux, qui ne sont pas toujours aisés à repérer sans connaissance préalable du sujet.

Capacités émergentes

La performance d'un LLM dépend de la taille, des thématiques et de la qualité de son jeu d'entraînement ainsi que du nombre de paramètres dont il dispose pour modéliser la langue. Cette performance n'est pas linéaire.

Au-delà de certains seuils, les modèles d'envergure produisent des capacités émergentes telles que la capacité à répondre à des questions, construire un raisonnement...

Fine tuning

Les LLM tels que GPT sont qualifiés de « fondationnels ». Ils servent de bases aux applications d'IA génératives (ex: ChatGPT) mais peuvent nécessiter un retraitement ou « fine tuning » pour obtenir des performances suffisantes en conditions d'utilisation réelles. Le fine tuning est un réentraînement qui nécessite de fournir un jeu de données d'exemples à reproduire par la machine.

Prompt engineering

En plus du fine tuning, la performance du modèle dépend beaucoup des données dont il dispose pour répondre à chaque requête. Ces données sont contenues dans le « prompt » ou l'instruction qu'il reçoit. Le prompt engineering est la discipline qui optimise la façon d'interroger les LLM au sein d'une application.

Il s'appuie sur de nombreux éléments comme :

- la question posée par l'utilisateur (prompt design)
- des exemples à suivre (few shots learning)
- des données de contexte issues de la base documentaire de l'entreprises (Retrieval Augmented Generation)
- l'utilisation de plusieurs LLM en séquence (Chaining)

Les 4 enjeux des Chief Data Officers pour rester dans la course

1. Lancer rapidement de premières expérimentations

Un enjeu de lucidité pour ne pas décevoir

Les CDO doivent aider le COMEX et les métiers à calibrer leurs attentes. Comme pour le Deep Learning il y a quelques années, l'essor de l'IA générative s'accompagne de fausses promesses et d'attentes exagérées.

Le POC fait son retour

A date, seules quelques fonctionnalités sont matures (cf. tableau ci-dessous). L'enjeu des entreprises sera donc d'éprouver rapidement les IA génératives sur leurs données et leurs cas d'usages pour en déterminer précisément la performance.

Un déploiement en 2 vagues

La première vague d'IA génératives proviendra des features intégrées aux applications bureautiques telles que Microsoft Copilot. Ces outils pourront couvrir une proportion importante des besoins courants des métiers mais atteindront leurs limites sur les cas d'usages plus spécifiques, nécessitant par exemple la maîtrise de concepts propres à l'entreprise ou l'accès à des bases documentaires internes.

Les entreprises devront rapidement identifier les usages nécessitant des développements dédiés, en veillant toutefois à anticiper l'évolution des fonctionnalités des éditeurs. Des expérimentations devront être initiées sans tarder pour rester dans la course.

Fonctionnalités de l'IA générative à ce jour

Texte/vocal	Code	Image / vidéo	Données structurées
Résumer	Compléter	Augmenter la résolution	Identifier des anomalies
Traduire	Traduire	Modifier / styliser	Générer des représentations graphiques
Interpréter / reformuler	Optimiser	Décrire / interpréter	Décrire / Interpréter
Répondre factuellement avec des sources	Documenter	Générer à partir de langage naturel	Compléter des valeurs manquantes
Résoudre des problèmes logiques / analytiques	Générer à partir de langage naturel	Passer de la 2D à la 3D	

● Mature ● Émergent ○ Embryonnaire

L'IA générative a fait remonter d'un cran les attentes du top management sur l'IA et la data. On doit maintenant accompagner les dirigeants sur le sujet et vite communiquer sur ce que l'on peut faire, et sur ce que l'on ne peut pas faire au risque sinon de faire des choses contre-productives. >>

Olivier MONNIER,
CDO, MATMUT

2. Évaluer les risques et en dégager un cadre global d'utilisation à affiner progressivement

Même si le potentiel est important, l'IA générative présente des risques et des limites dont les utilisateurs doivent prendre connaissance. Les entreprises doivent rapidement poser un cadre qui délimite pour les utilisateurs : (a) le champ des possibles et (b) les bonnes pratiques à adopter pour un usage responsable.

Cela doit se faire de façon itérative à mesure que la technologie et le cadre légal évoluent.

Limites

- **Hallucination**
Difficulté à garantir l'exactitude des résultats.
- **Toxicité**
Recours occasionnel à un langage inapproprié, grossier ou haineux.
- **Biais**
Reproduction des distributions statistique du jeu d'entraînement.
- **Mesure de performance**
Difficulté à automatiser la mesure de performance compte tenu de la subtilité du langage et de la nature aléatoire des contenus générés.
- **Coûts**
Coûts d'entraînement prohibitifs pour de nombreuses entreprises (plusieurs millions d'euros pour un modèle à l'état de l'art). Coûts d'inférences pouvant monter rapidement pour les applications fortement sollicitées et/ou nécessitant beaucoup de données en contexte.

Risques

- **Réputation**
Dysfonctionnement des applications en interface avec le public, les clients ou les partenaires commerciaux.
- **Juridique**
Zones grises autour de la propriété intellectuelle (données d'entraînement et contenu généré) et de l'utilisation éthique.
- **Fournisseur**
Dépendance possible aux fournisseurs d'IA générative et limite de souveraineté sur la gouvernance des modèles.
- **Opérationnel**
Prise de décision sur la base d'informations erronées.
- **Sécurité**
Cybercriminalité, fuites de données, deepfakes...

3. Former les utilisateurs pour éviter les déceptions

La manière d'interagir avec une IA générative impacte fortement sa performance. Modifier quelques mots dans les instructions et le résultat obtenu peut changer du tout au tout.

Exemple avec deux prompts différents

Le prompting standard

Input

Q : Roger a 5 balles de tennis. Il achète 2 autres boîtes de balles de tennis. Chaque boîte contient 3 balles de tennis. Combien de balles de tennis a-t-il maintenant ?

R : La réponse est 22.

Q : Le chef de la cafétéria avait 23 pommes. S'il en a utilisé 20 pour préparer le déjeuner et qu'il en a acheté 6 de plus, combien de pommes a-t-il ?

Model Output

R : La réponse est 27.



Le prompting Chain of Thought

Input

Q : Roger a 5 balles de tennis. Il achète 2 autres boîtes de balles de tennis. Chaque boîte contient 3 balles de tennis. Combien de balles de tennis a-t-il maintenant ?

R : Roger a commencé avec 5 balles. 2 boîtes de 3 balles de tennis chacune représentent 6 balles de tennis. $5 + 6 = 11$. La réponse est 11.

Q : Le chef de la cafétéria avait 23 pommes. S'il en a utilisé 20 pour préparer le déjeuner et en a acheté 6 de plus, combien de pommes a-t-il ?

Model Output

R : Le chef de la cafétéria avait 23 pommes à l'origine ; il en a utilisé 20 pour préparer le déjeuner. Il avait donc $23 - 20 = 3$. Il a acheté 6 pommes supplémentaires, il en a donc $3 + 6 = 9$. La réponse est 9.



Les entreprises devront donc traiter l'interaction avec les IA génératives comme une compétence à maîtriser largement par les employés. Si elles n'agissent pas rapidement, elles s'exposent à plusieurs risques.

- **Un usage sous-optimal** : les utilisateurs ne tirent pas pleinement avantage des capacités de l'IA
- **Un usage excessif** : les utilisateurs démultiplient les tentatives pour aboutir à un résultat satisfaisant, surconsommant ainsi les IA
- **Une perte d'intérêt** : les utilisateurs confondent leurs lacunes et celles de la technologie

Pour pallier à ces risques, les entreprises commencent à mettre en place des programmes de formation d'envergure à travers une "Prompt Academy".



Pour répondre aux forts enjeux et attentes liés aux LLM et à l'IA Générative, la réponse de Danone a été de monter un groupe de travail regroupant divers experts (Data Scientists, juristes, cyber...) afin de : Proposer des directives Groupe en matière de GenIA, Construire une offre de service interne simple et sécurisée, Mettre à disposition des collaborateurs un centre d'excellence sur le sujet. >>

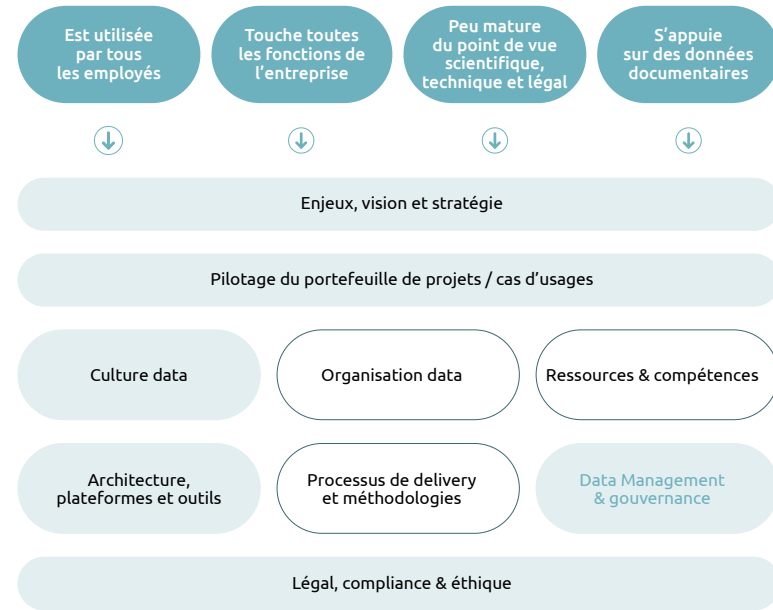
Anne-Laure Cebile,
Data Science Team Leader, DANONE

4. Adapter la stratégie et le modèle opérationnel data & IA

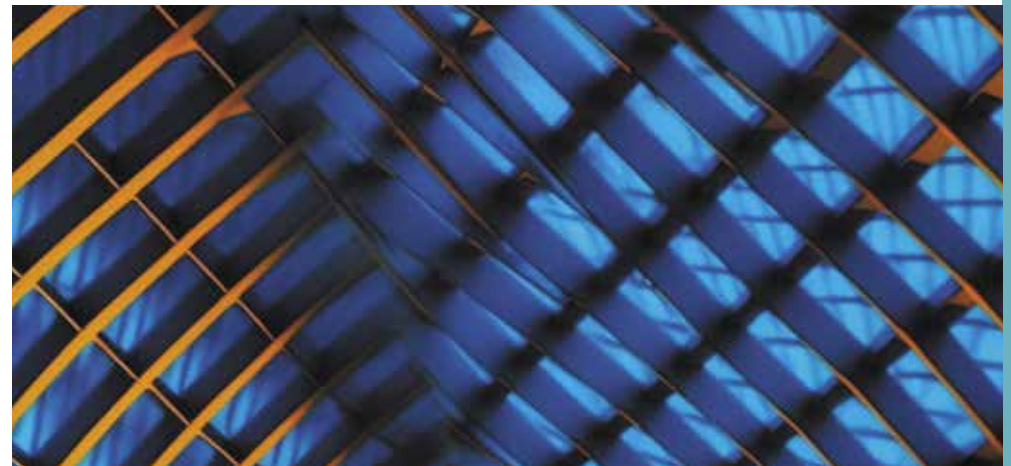
L'IA générative met à l'épreuve les stratégies et modèles opérationnels des Directions data et DSI. Elle présente en effet des spécificités qui peuvent la rendre complexe à maîtriser.



Vue d'ensemble des impacts du GenAI



Légende :
● Enjeux propres à l'IA générative ➔ Degré d'impact sur la stratégie et le modèle opérationnel data et AI
○ Impact modéré ● Impact fort



La mise en production des IA génératives soulève de nouveaux défis

Mesurer et optimiser la performance

Pour sortir les LLM du laboratoire, les Directions data vont devoir renforcer leur capacité à mesurer et à suivre les performances des modèles en condition d'usage.

Nos recommandations

A) Proposer un *Banc de Test* aux développeurs d'IA générative

La performance d'une IA générative dépend d'une multitude de facteurs. Au-delà du modèle, ce sont les choix de vector stores, les stratégies de prompt (chaînage, agents...), le design des prompts et même l'UX qui déterminent la performance de l'IA pour une tâche donnée.

Le Banc de Test permet aux développeurs d'élaborer un protocole pour comparer plusieurs configurations en parallèle. Aujourd'hui, le marché se développe pour proposer des solutions dédiées aux IA génératives via des acteurs comme Weight & Biases ou WhyLabs.

B) Intégrer des feedbacks utilisateurs à l'échelle

La mesure de performance est difficile à automatiser car le langage est une matière nuancée et ambiguë. Il est donc essentiel d'obtenir un feedback continu de la part des utilisateurs, autant en phase d'expérimentation que durant le run.

Ils auront pour rôle de remonter et qualifier les erreurs de l'IA afin que les développeurs :

- recommandent de nouvelles façons d'interroger le modèle
- optimisent le modèle via un fine tuning ou de nouvelles stratégies de prompt par exemple
- clarifient le domaine de validité du modèle : que sait-il faire ou ne sait-il pas faire ?

Éviter le « vendor lock-in » ?

Le marché des LLM se concentre progressivement autour d'une poignée de modèles à l'état de l'art détenus par des acteurs dominants comme OpenAI ou Anthropic. Cela rend les entreprises clientes vulnérables à :

• Obsolescence

La technologie progresse si vite qu'un modèle peut être supplanté par la concurrence en quelques mois.

• Décommissionnement

À mesure que de nouvelles versions sont créées, les modèles sur lesquels les applications sont fondées peuvent être décommissionnés.

• Changements tarifaires

Les fournisseurs stimulent actuellement la pénétration de marché en pratiquant la vente à perte. Il est possible que les tarifs augmentent à l'avenir.

• Éthique

L'entreprise qui consomme un modèle est tributaire des choix éthiques de son fournisseur. Par exemple, elle ne maîtrise pas la provenance des données d'entraînement, pouvant être soumises à des contraintes de propriété intellectuelle.

Nos recommandations

A) Construire des architectures agnostiques des modèles

Les architectures d'IA génératives doivent respecter un principe de découplage entre les modèles et les autres briques technologiques. Il doit être possible de substituer un ChatGPT par un autre sans apporter de modifications conséquentes au reste de l'architecture comprenant les pipelines de données, les vector stores, les agents ou les prompts préconçus.

B) Proposer un *Model Hub* aux développeurs d'IA génératives

Le Model Hub permet aux développeurs d'IA génératives d'accéder à une librairie étendue de modèles via APIs. Actuellement, le marché se structure pour offrir des solutions de Model Hub clés en main. Les principaux providers cloud développent leur offre sur ce marché. AWS Bedrock et Azure ML propose un accès vers une multitude de LLM propriétaires ou open source, hébergés sur leurs serveurs.

Votre entreprise est-elle « GenAI-Ready » pour saisir l'opportunité ?

AXES	ENJEUX CLÉS	CAPACITÉS FAVORABLES
#1 Enjeux, vision et stratégie	Acculturer le COMEX et proposer une stratégie sur le court et moyen terme	<ul style="list-style-type: none"> • Capacité à mobiliser les métiers sur toutes les fonctions de l'entreprise • Direction data proche du COMEX dans un rôle de conseiller stratégique
#2 Pilotage du portefeuille de projets	Canaliser l'afflux de nouveaux projets et orienter les choix make or buy	<ul style="list-style-type: none"> • Gouvernance et processus efficaces pour remonter les cas d'usages • Capacité de veille technologique en place
#3 Organisation & gouvernance	Poser un cadre pour une utilisation efficace et responsable tenant compte des limites et risques de l'IA générative	<ul style="list-style-type: none"> • Gouvernance mature intégrant métier, data, légal, DSI et cybersécurité • Capacité à mettre en place des task forces dédiées à l'IA générative
#4 Ressources et compétences	Evaluer les impacts sur les compétences des populations de développeurs et d'utilisateurs d'IA générative	<ul style="list-style-type: none"> • Existence d'une filière data structurée notamment sur les profils de Data Engineers cruciaux pour les cas d'usages d'IA générative • Existence d'une Data Academy et/ou d'un relai de formation équivalent • Relations de proximité entre équipe Data et DRH
#5 Culture data	Communiquer en continu au rythme des évolutions du marché et des technologies	<ul style="list-style-type: none"> • Relais en place dans les métiers pour faire de l'acculturation à l'échelle • Relations de proximité entre équipe Data et équipes communication / transformation
#6 Architecture, plateformes et outils	Faire évoluer les socles technologiques pour permettre le développement et l'exploitation d'IA génératives en production	<ul style="list-style-type: none"> • Partenariats avec les Cloud providers • Gouvernance du socle technologique mature
#7 Data management et gouvernance	Passer du Data Management au Data & Knowledge Management	<ul style="list-style-type: none"> • Centralisation au moins partielle des bases de données documentaires clés • Pratiques de Knowledge Management matures • Classification des données sensibles pour mitiger les risques de fuite de données
#8 Processus de delivery & méthodologies	Réaliser les premières expérimentations et en dégager des convictions sur les conditions de réussite en production et les spécificités des IA génératives en production	<ul style="list-style-type: none"> • Capacité d'expérimentation et de mise en production matures (stack MLOps) • Proximité avec les métiers utilisateurs pour mettre en place des boucles de feedbacks
#9 Legal, compliance & Ethics	Mitiger les risques sans brider l'organisation dans un contexte réglementaire encore flou	<ul style="list-style-type: none"> • Gouvernance en place entre les Directions data et juridique • Capacité de monitoring et de gouvernance des IA en anticipation de l'AI Act



SUCCESS STORY 1

Leader dans le secteur de la parfumerie

Problématique

Les collaborateurs au sein des métiers ont imaginé de nombreux usages de l'IA générative s'appuyant sur les LLM et les modèles de diffusion. L'enjeu majeur est de canaliser les initiatives et de profiter de cette impulsion pour identifier, prioriser et répondre aux besoins.

Pour faire cela correctement il est nécessaire de :

- Sensibiliser et démystifier les parties prenantes à l'IA générative
- Inspirer les collaborateurs au moyen d'exemples concrets
- Imaginer des cas d'utilisation percutants susceptibles d'apporter une forte valeur ajoutée à l'entreprise

Résultat

Une campagne de 40 ateliers de sensibilisation et d'idéation auprès des métiers a été menée. Les cas d'usages ont été classés en 3 catégories.

Catégorie 1

Les cas d'usages pouvant être lancés très facilement car ne nécessitant que la formation des utilisateurs et l'achat de licences.

Catégorie 2

Les cas d'usages à développer, dans un second temps, car soulevant des problèmes de confidentialité ou nécessitant de développer une interface ou une application complexe.

Catégorie 3

Les cas d'usages à ne lancer que lorsqu'un acteur externe proposera une solution car trop complexes à développer ou des cas d'usages que des acteurs majeurs sont en train de développer pour les intégrer à des outils (MS Teams, Gmail...).

Identifier les bons gisements de valeur >>

KPI

- **+ de 400 participants**
- **81 cas d'usages identifiés**
(20 Catégorie 1 / 40 Catégorie 2 / 9 Catégorie 3)
- **Des cas d'usages répartis** dans tous les départements de l'entreprise et souvent transverses.
- Le choix et le développement de **2 cas d'usage priorités**, à forte valeur ajoutée et facilitateur pour d'autres sujets proches.

Enseignements clés

- La priorisation des cas d'usages d'IA générative relève de critères différents des critères usuels de priorisation des cas d'usages IA. En effet, ces cas d'usages répondent souvent à des enjeux supplémentaires, comme la capacité à rayonner (valeur générée, économies de coûts...), en plus des critères de complexité encore peu maîtrisés et évoluant quotidiennement.
- La distinction entre les cas d'usages relevant des IA génératives et ceux pouvant être réalisés par d'autres techniques est nécessaire et permet de faire des économies de coût. Une partie des cas d'usage proposés par les métiers, car couverts par ChatGPT, peut l'être par des modèles moins coûteux avec une performance égale ou supérieure.
- Lors des campagnes d'idéation, les populations à prioriser sont celles faisant de la recherche d'information au sein de grande quantité d'informations textuelles et autres documents. Ces cas d'usages (nommés RAG) sont en effet à forte valeur et efficacement traités par l'IA générative et de nouveaux outils comme les bases de données vectorielles.



SUCCESS STORY 2

Entreprise dans le secteur des services financiers

Problématique

Un fond d'investissement gérant plus d'1Md d'assets souhaite accélérer significativement sa phase clé de due diligence. Une due diligence peut nécessiter d'analyser 10 000 à 20 000 documents pour vérifier des éléments très précis et valider ou non une décision d'achat de la société analysée.

Pour aider ses équipes, l'entreprise souhaitait :

- Construire un assistant intelligent avec des interactions de type humain en utilisant les capacités de GPT3.5/4
- Récupérer des informations à partir d'un large corpus de documents, les explorer et résumer les faits essentiels
- Fournir une interface de type ChatGPT pour interagir avec l'assistant et visualiser les documents récupérés

Résultat

Développement d'une application web aidant les équipes de due diligence à analyser un large corpus de documents (plus de 1 million de documents).

Cet analyste virtuel permet de répondre à plusieurs cas d'utilisations (recherche de documents en interne, résumé, extraction d'informations, recherche web...) au service de 3 catégories d'utilisateurs :

- L'analyste qui doit connaître et analyser l'ensemble des documents
- Le spécialiste qui va rechercher des informations très précises (sectorielle ou fonctionnelle) dans les documents
- Le partenaire qui souhaite avoir une réponse rapide à quelques questions spécifiques et préparer ses interactions avec la société analysée

Solution mise en production et utilisée sur chaque nouvelle due diligence.

Industrialiser rapidement une solution performante »

KPI

- **+ de 50 utilisateurs**
dont les partners de la société comme premiers ambassadeurs de la solution
- **Gain opérationnel**
significatif sur chaque due diligence

Enseignements clés

La performance de la solution ne dépend pas uniquement de la performance du LLM :

- L'interface utilisateur est clé dans l'adoption de la solution par les utilisateurs, notamment en personnalisant l'agent afin que l'utilisateur ait l'impression d'être dans une conversation. Pour cela, l'interface doit être intuitive, proposer des suggestions de prompt...
- L'intégration d'agents ayant chacun leur rôle, leur domaine de validité et leur spécialisation permet de proposer plus de capacités avec de meilleurs résultats au plus proche du besoin de l'utilisateur.
- L'utilisation d'outils performants tels que les bases de données vectorielles, les modèles d'embeddings multi-langues... est nécessaire pour garantir une solution optimale.

L'humain ne doit pas être délaissé afin d'assurer la performance et la pertinence de la solution. Il intervient dans l'ensemble du cycle de vie de la solution (entraînement, définition du domaine de validité, mise en place de contrôles et de jeux de test...). Ses feedbacks et jeux de tests permettent ainsi d'améliorer la solution.

ZOOM

AI Act & IA de confiance : une accélération pour mitiger les risques réglementaires & opérationnels

L'IA générative remet l'IA de confiance au centre des préoccupations

L'arrivée des IA génératives, et notamment des LLM

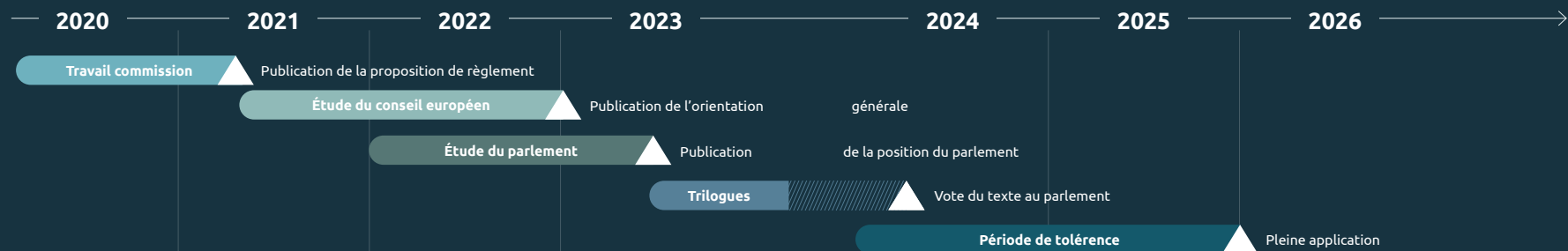
- Démultiplie les usages possibles de l'IA et incite donc les Dirigeants data à s'emparer des questions de confiance et de gouvernance des modèles
- Pose de nouvelles questions spécifiques liées à la nature des LLM comme le respect de la propriété intellectuelle ou encore la prémunition contre les hallucinations

En parallèle, l'IA Act se concrétise et s'enrichit pour couvrir les exigences relatives à l'IA générative

Le Parlement Européen propose une réglementation sur l'utilisation de l'IA. Par son positionnement sur les risques liés à l'usage et non à la technologie, l'AI Act a vocation à être un socle de référence résistant au contexte mouvant, centré sur la protection des droits fondamentaux.

Le vote du texte final et la période de tolérance avant les amendes effectives poussent les entreprises à suivre le calendrier de déploiement pour s'y préparer.

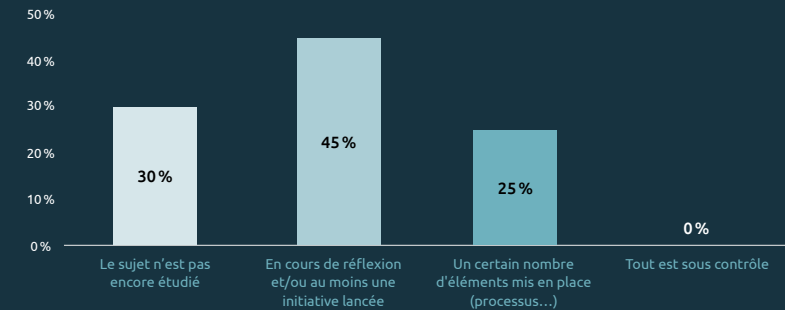
Calendrier de déploiement de l'IA Act



Les entreprises se préoccupent de l'AI Act mais restent majoritairement dans le flou sur la manière de s'y préparer

La réglementation RGPD a laissé des marques dans de nombreuses entreprises et pousse les organisations à prendre de l'avance par rapport à la réglementation à venir. Cependant, le périmètre du sujet n'est pas encore clairement compris et les actions d'état des lieux ou de remédiation pour mitiger les risques restent minoritaires.

Comparaison de l'appropriation de l'AI Act au sein des entreprises, % de répondants



ZOOM

AI Act & IA de confiance : une accélération pour mitiger les risques réglementaires & opérationnels

Pour répondre à ces enjeux, l'IA de confiance doit être déclinée sur 8 piliers

Ces thématiques sont celles assignées par l'IA Act, mais également celles qui sont des enjeux majeurs pour les organisations.



Bien que les LLM de confiance et l'audit de ces modèles constituent encore un domaine de recherche, nous avons d'ores et déjà identifié des principes directeurs spécifiques pour chaque pilier de l'IA de confiance.

	Points de faiblesse	Actions de conception	Actions d'audit
Responsabilité	<ul style="list-style-type: none"> Génération de contenu illégal ou toxique 	<ul style="list-style-type: none"> Acculturer et former les utilisateurs aux usages des LLM, notamment ChatGPT, pour éviter les utilisations sur des faiblesses connues des outils 	<ul style="list-style-type: none"> Caractériser les évasions possibles (jailbreaks)
Performance	<ul style="list-style-type: none"> Hallucinations Capacités ciblées 	<ul style="list-style-type: none"> Optimiser les prompts (prompt engineering) Ne pas avoir une approche «LLM tout nu» mais intégrer les modèles LLM dans un système plus large de briques techniques (couche métier ou modèle logique pour raisonner, chercher, calculer, optimiser...) afin d'utiliser la brique LLM pour sa capacité clé à date : le traitement du langage naturel 	<ul style="list-style-type: none"> Comptabiliser les hallucinations
Robustesse	<ul style="list-style-type: none"> Domaines d'utilisation spécifiques Non-reproductibilité 	<ul style="list-style-type: none"> Etablir et différencier les domaines d'utilisation de chaque LLM pour adapter la responsabilité et la performance sur des domaines de validité adaptés Réduire la température 	<ul style="list-style-type: none"> Quantifier l'incertitude des résultats
Qualité de données	<ul style="list-style-type: none"> Représentation arbitraire du monde 	<ul style="list-style-type: none"> Faire valider les sorties par l'humain pour analyser, fact checker et censurer les contenus générés 	<ul style="list-style-type: none"> Caractériser le jeu d'entraînement

.../...

.../...

	Points de faiblesse	Actions de conception	Actions d'audit
<p>Équité</p>	<p>Écarts de qualité de service</p>	<ul style="list-style-type: none"> Mettre en place un « pare feu » de contenu sur les inputs des utilisateurs pour s'assurer de ne pas adresser des thématiques non souhaitées en passant par exemple par un modèle de détection d'intention ou du prompt engineering. 	<p>Quantifier les écarts de performance par domaine</p>
<p>Explicabilité</p>	<p>Contenus générés sans citer aucune source</p>	<ul style="list-style-type: none"> Se brancher sur un moteur de recherche (RAG) Forcer une explicabilité by-design en mettant en place la citation de source avec une UX permettant aux utilisateurs de vérifier par eux-mêmes la fiabilité 	<p>Evaluer la pertinence des sources citées</p>
<p>Frugalité</p>	<p>Coûts d'utilisation non-maîtrisés</p>	<ul style="list-style-type: none"> Mettre en place une stratégie FinOps S'assurer que la solution est adaptée à l'usage : certains usages LLM peuvent être satisfaits avec des technologies moins consommatrices comme le NLP 	<p>Réaliser le bilan carbone du cas d'usage</p>
<p>Contrôle des dérives</p>	<p>Mises à jour permanentes des APIs</p>	<ul style="list-style-type: none"> Collecter les retours utilisateurs au fil de l'eau Mettre en place une stratégie de tests automatiques par domaine de validité pour détecter les faiblesses et monitorer l'évolution des performances et des retours utilisateurs 	<p>Faire des audits détaillés sur une base régulière</p>

Nous préconisons une démarche en 4 volets pour mitiger les risques efficacement et durablement.



Synthèse et analyse

Face à la rapidité, sans précédent, de l'émergence de l'IA générative, une technologie de masse, dont une partie des usages reposera sur des outils mis à la disposition du plus grand nombre, les Dirigeants data se trouvent confrontés à de nouveaux challenges. Tout d'abord, il leur faut réussir à répondre aux demandes de leur top management et lancer rapidement de nouveaux usages d'IA génératives. Mais si les opportunités potentielles générées par l'adoption de cette technologie sont immenses, l'environnement technologiques et les risques associés sont encore en phase de découverte.

Les Dirigeants data doivent alors définir une stratégie pour accompagner cette transformation :

- Expérimenter l'IA générative sur des premiers cas d'usages via des solutions logicielles (« buy ») et des développements internes (« make »)
- Gérer les risques et la propagation qui peut vite devenir incontrôlée sans une ligne directrice claire
- Préparer le passage de l'IA générative en production en :
 - adaptant le modèle opérationnel data pour prendre en compte les spécificités de l'IA générative en matière de compétences, de formation, de modèle de delivery...
 - faisant les bons choix technologiques en fonction de ses objectifs et de ses contraintes

Enfin, l'adoption de ces nouveaux usages ne sera pérenne que si les utilisateurs peuvent avoir confiance dans une technologie qui n'est nativement pas conçue pour être fiable. Le passage en production d'IA génératives de confiance, en capacité de s'intégrer dans le cadre défini par l'AI Act Européen, sera l'un des enjeux clé de l'année 2024.

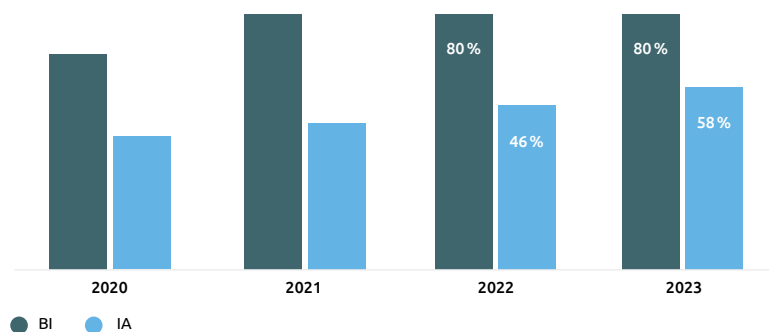
- Partie 2 -

LES ORGANISATIONS SE TRANSFORMENT POUR LA DATA ET CHERCHENT DÉSORMAIS À SE TRANSFORMER PAR LA DATA

Les entreprises maîtrisent l'industrialisation de la data et de l'IA

De plus en plus de cas d'usages en production

Évolution du taux d'industrialisation des solutions BI et IA



Presque **6 produits IA sur 10 industrialisés**, soit **25% de +** qu'en 2022

Nous industrialisons maintenant avec plus de facilité, mais le challenge devient le time-to-market. >>>

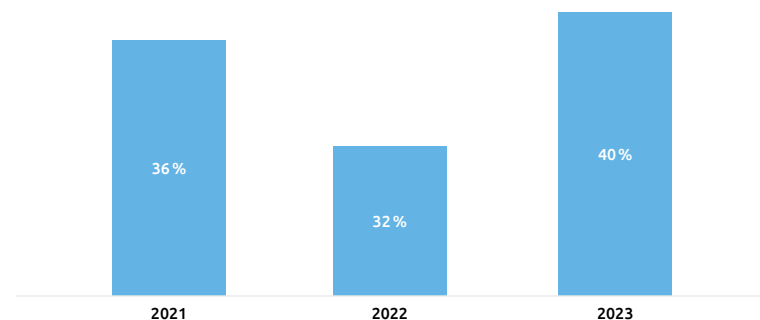
Head of AI Factory, secteur luxe

Nous avons un bon taux d'industrialisation grâce à une sélection et priorisation réaliste des cas d'usages que l'on instruit. >>>

Antoine Ly, Chief Data Science Officer, SCOR

L'IA gagne du terrain dans les portefeuilles data

Evolution de la part des projets IA (vs BI)



En 2023, l'IA représente **40%** des cas d'usages des portefeuilles de projets

Après une dépriorisation des projets IA en 2022, au profit de projets reposant sur des pratiques et technologies maîtrisées, **l'IA gagne du terrain dans les portefeuilles car :**

- l'arrivée de l'IA générative a redonné de l'intérêt pour l'IA
- les efforts d'acculturation portent leurs fruits et les métiers remontent désormais spontanément des cas d'usages IA
- le développement des équipes data au sein des métiers permet de développer plus de cas d'usages et de leur laisser l'ownership sur la réalisation
- l'aboutissement des projets favorisés ces dernières années passées (création de jeux de données, APIisation, mise sous contrôle de référentiels...) a permis d'adresser les projets IA plus facilement et sereinement

La relation avec la DSI comme facteur clé de succès

Des DSI mieux acculturées à la data

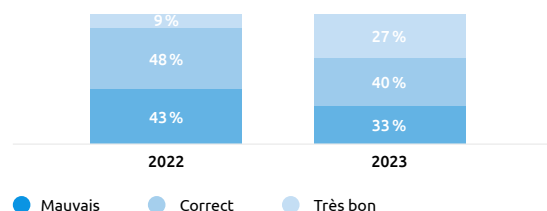
La collaboration entre DSI et Directions data s'améliore et les mandats respectifs se clarifient, fluidifiant fortement le passage des IA en production. **Sur les projets, les facteurs de succès principaux demeurent :**

- l'hybridation des compétences data et IT (Data Engineer, Machine Learning Engineer)
- une meilleure implication de la DSI de la phase de cadrage jusqu'au déploiement

Sur l'alignement stratégique entre DSI et data, les CDO constatent une amélioration en matière de :

- synchronisation des feuilles de route data et IT
- meilleure entente sur les besoins d'évolution des socles data

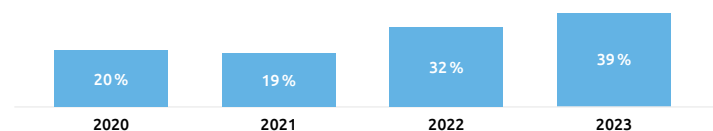
Évolution du niveau d'acculturation des DSI



Un retour de la data au sein des DSI

Les évolutions d'organigramme constatées depuis quelques années sont à la fois causes et résultantes du rapprochement entre data et DSI.

Évolution de la part des directions data rattachée aux DSI



En 2023, **39%** des Directions data **sont rattachées aux DSI**. Un chiffre en augmentation constante depuis 3 ans.

Et d'autres facteurs clés de succès identifiés

- La **capacité à impliquer** et responsabiliser les métiers grâce aux démarches d'acculturation
- Les **bons choix technologiques** selon des écosystèmes SI cohérents (Move-to-Cloud, Suites Editeurs...)
- La maîtrise des **approches MVP & Agile** pour apporter de la valeur à court terme
- La **productisation des cas d'usages IA** de mieux en mieux maîtrisée : intégration dans les processus opérationnels (produits IA/digitaux) et intégration dans l'écosystème SI
- La constitution d'**équipes produits complémentaires** avec des rôles bien identifiés entre Data Scientist, Data Engineer, Machine Learning Engineering & Ops et Product Owner/Manager
- Une **culture data** chez les métiers qui permet désormais d'être réaliste sur la promesse de valeur des cas d'usages

Ainsi que des nouvelles difficultés

Le time-to-market

de mise en production des produits à cause des travaux préalables de data management (clarification des règles de gestion, consolidations et nettoyages de référentiels, ingestion sur la Data Platform...).

La gestion de l'évolutivité

de l'environnement business et SI impactant à la fois les données et les produits en production, notamment la conservation et documentation de la connaissance.

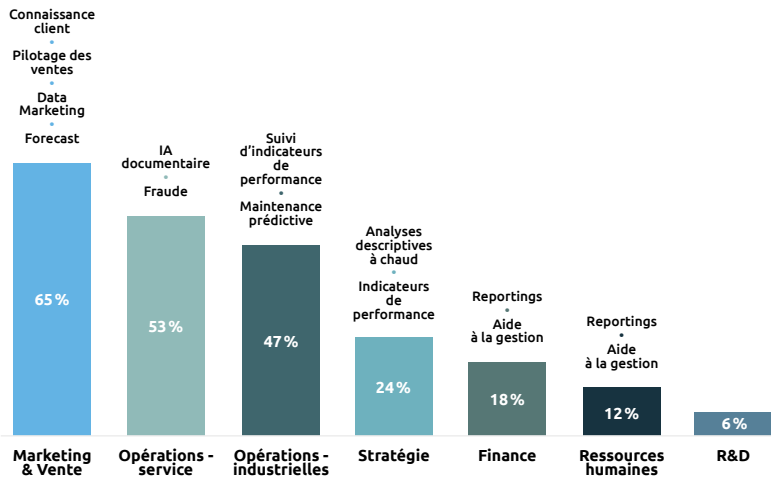
La data & l'IA sont adoptées par tous dans les entreprises

Des cas d'usages couvrant toutes les fonctions de l'entreprise

Comme en 2022, les fonctions Marketing & Vente et Opérations bénéficient fortement des cas d'usages en production. La majorité des cas d'usages pour les fonctions Marketing et Vente sont au service de l'amélioration de la relation client et les cas d'usages pour les opérations sont destinés à l'amélioration des processus.

Nous observons également une forte hausse de cas d'usages en production au bénéfice de la stratégie, les dirigeants s'étant emparés de la BI comme élément d'aide à la décision et au pilotage grâce à des dispositifs BI fiables, et de l'IA comme enjeu stratégique.

Fonctions de l'entreprise bénéficiant du plus de cas d'usages en production et principaux « grands » cas d'usages répandus



En 2023, le **Marketing & Vente** fait partie du **TOP 3** des Directions métiers bénéficiant du plus de cas d'usages data pour **65% des entreprises**

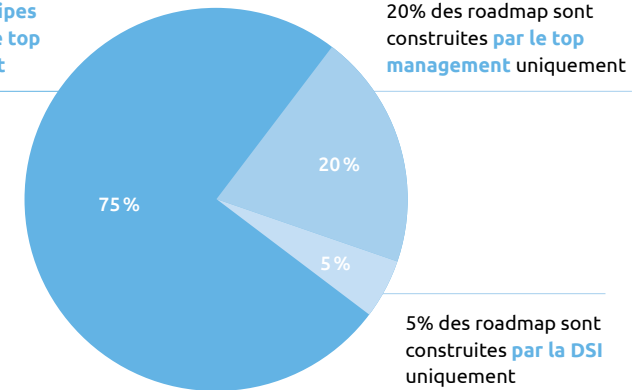
Les métiers proposent de plus en plus de cas d'usages d'eux-mêmes. Ce phénomène s'accélère d'ailleurs avec l'essor de l'IA générative. >>

Chief Data Officer, secteur banque

Des roadmaps de cas d'usages construites en collaboration avec les équipes métiers

La data étant désormais attendue sur la réalisation de cas d'usages par l'ensemble des parties prenantes de la transformation, les feuilles de routes sont construites en collaboration avec les principaux partenaires projets :

75% des roadmap sont construites avec les équipes métiers et le top management



La data bien identifiée comme enjeu clé de la transformation de l'entreprise

Les capacités data sont reconnues comme un asset clé et s'inscrivent de plus en plus dans les plans stratégiques des entreprises, souvent comme pilier de la transformation digitale. En ce sens, les activités data font désormais majoritairement l'objet d'une planification cohérente à travers une feuille de route annuelle.

77% des Directions data ont formalisé une roadmap stratégique

80% des entreprises ont inscrit la data dans leur plan stratégique

ZOOM

La data au service des enjeux RSE et du développement durable

La data doit maintenant faire face aux défis RSE & développement durable

Dans l'ensemble des domaines métiers adressés par les Directions data, les sujets de développement durable se multiplient et semblent s'inscrire comme une toile de fond dans les nouveaux défis à relever pour la fonction data.

50%
des Directions data sont mobilisées pour répondre à des enjeux de développement durable

« La dynamique autour de la data a été accélérée par des besoins RSE, notamment par la volonté de rendre notre reporting plus robuste. »

Laurent MISSET,
Directeur Data & Analytics,
CHANEL

Une intervention data pour répondre aux enjeux réglementaires RSE

Les équipes data doivent s'organiser pour répondre aux enjeux des équipes métiers et enrichir leur offre de service pour répondre à la fois aux challenges liés à la mesure et au reporting réglementaire, mais aussi être au rendez-vous pour porter des cas d'usages de réduction d'impact environnemental, au cœur des activités métiers.



40%

Des organisations data contribuent au reporting réglementaire

La mesure et l'élaboration des reportings RSE reste l'une des principales contributions des équipes data au service des directions RSE. La principale difficulté rencontrée réside dans l'acquisition de données granulaires, pour passer d'un reporting ponctuel à partir d'ordres de grandeur à un pilotage dynamique industrialisé répondant aux obligations réglementaires (notamment CSRD).

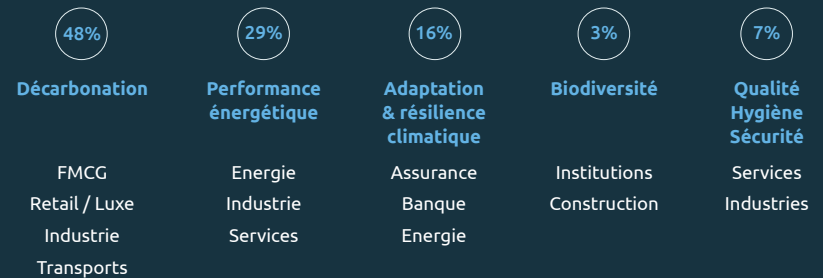
« Une des difficultés est de trouver la bonne data, factueliser une situation et avoir une mesure cohérente. »

Lilian PLACET, Head of Data, GSF

Une intervention data pour répondre aux enjeux opérationnels

Les autres cas d'usages adressés ciblent principalement des enjeux de réduction d'empreinte carbone, d'amélioration de performance énergétique ou encore de résilience climatique selon les secteurs.

Part des organisations adressant chaque thématique du développement durable, et principaux secteurs s'étant emparés du sujet



48% des Directions data adressent des sujets de décarbonation, et **29%** adressent des sujets de performance énergétique

ZOOM

Des premières initiatives apparaissent comme des facteurs clés de succès

Pour accélérer la démarche « Data for Sustainability », afin de bâtir des Fondations data robustes à l'échelle de l'entreprise, et pouvoir ainsi construire des cas d'usages « actionnables », des bonnes pratiques se détachent dans les organisations matures.

1 Structuration d'un Data Domain Sustainability dédié

« Nous avons créé un Data Domain Sustainability comme socle pour nos cas d'usages, notamment 80% des rapports extra-financiers du groupe. »

Charles-Antoine ROBELIN,
Head of Data Platform, DECATHLON

2 Capture de données chaudes/dynamiques pour des analyses et actions réalistes

« Nous nous constituons des socles de données et avons structuré un data product carbone scope 1,2 et 3 pour ainsi mesurer avec des données « chaudes/physiques » et ne plus convertir des euros en carbone. »

Alexis TRICHET,
Directeur de la Stratégie, Data & Connaissance client, ORANGE

3 Rétroaction dans les processus opérationnels et inscription dynamique dans les prises de décisions

« Nos modèles d'IA nous permettent de mieux utiliser l'énergie, en accélérant ou en ralentissant nos usines pour réduire la quantité de gaz brûlé, faute de consommation. »

Chief Data Officer, secteur industrie

De premiers grands cas d'usages génèrent d'ores et déjà de la valeur

3 exemples de cas d'usages en production, intégrés aux processus cœur métier

RETAIL INTÉGRATION D'INDICATEURS CARBONE DANS LES RÉFÉRENTIELS

Intégration des indicateurs carbone dans les référentiels produits ou fournisseurs, pour une prise en compte de l'impact environnemental dans la gestion des stocks et approvisionnements au quotidien

INDUSTRIE LOGISTIQUE DÉCARBONÉE

Optimisation des opérations, des flux et des chaînes d'approvisionnement, et gestion intelligente des entrepôts

ASSURANCE PRÉVISION DE RISQUES CLIMATIQUES

Utilisation des bases de données Climat historiques et des modèles prédictifs, afin de prévoir le niveau de probabilité d'occurrence d'un risque climatique, et quantifier ses impacts humains, matériels ou financiers

« L'objectif : transformer la base Achat en CO2 pour piloter par les émissions. »

Chief Data Officer, secteur industrie

Des grands challenges à adresser

Les principaux challenges à adresser par les organisations data pour être au rendez-vous des enjeux de développement durable

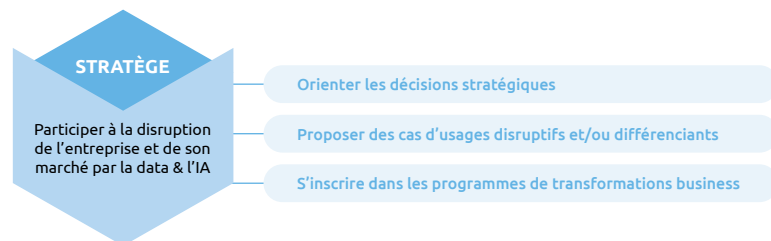
- **S'approprier les frameworks** méthodologiques spécifiques (équivalences carbones, ACV...) pour comprendre les données et être en capacité de les modéliser selon les besoins de l'entreprise
- **S'approprier les bases de données** de références externes (ADEME...) et les mettre à disposition pour accélérer les reportings et permettre la réalisation des cas d'usages
- **Poser une gouvernance** des données RSE sur l'ensemble du cycle de vie de la donnée
- **S'organiser en filière** pour mesurer et agir de manière cohérente selon des standards data partagés entre partenaires / sous-traitants / fournisseurs afin de faciliter la mise en place de cas d'usages data & IA impactants
- **Accoster les roadmaps RSE & data** pour identifier les principaux leviers de réduction d'impact tout en mettant sous contrôle le périmètre de données associées pour pouvoir mesurer, agir et monitorer grâce à la data & l'IA
- **Limiter l'impact environnemental** des usages de la data et de l'IA, notamment énergétique

Les entreprises cherchent désormais à se transformer par la data

Générer de la valeur à tous les niveaux de l'entreprise

Pour créer de la valeur par la data, les attentes et besoins de l'entreprise évoluent et poussent les Data Offices à diversifier leurs offres de services pour jouer à la fois un rôle DOer, d'enabler et de stratégie.

Les CDO attendus comme stratèges pour transformer en profondeur



Les CDO sont de plus en plus mobilisés dans leur fonction de conseiller stratégique pour actionner au mieux les opportunités générées par la data & l'IA et impacter le modèle même de l'entreprise.

« Nous collaborons directement avec les CEO et le management fonctionnel des différentes enseignes pour réfléchir aux grands cas d'usages stratégiques au-delà de l'instruction de quick wins opérationnels apporteurs de valeur immédiate. »

Sophie GALLAY, Directrice Data Groupe, ETAM

« Les ambitions portées par la data s'inscrivent dans la stratégie globale de l'entreprise. L'un de nos objectifs clé par exemple est de favoriser et soutenir l'innovation à travers le Groupe. »

Chief Data Officer, secteur banque

« L'ambition globale est de soutenir le business qui a un programme d'amélioration de l'EBITDA dont nous sommes contributeurs. »

Fabien DENEUVILLE, Head of Data Science, Modulaire Group



Les modèles d'affaires encore peu impactés

En capacité de délivrer les bons cas d'usages au service des métiers et d'apporter des gains d'efficacité opérationnelle de manière tactique, les Directions data & IA doivent à présent démontrer leur capacité à générer des avantages stratégiques et des nouveaux revenus pour impacter le modèle d'affaires de l'entreprise.

80% des entreprises cherchent désormais à créer de la valeur en agissant à la fois sur leurs coûts et leurs revenus

Création de valeur sur la top line



Création de valeur sur la bottom line

- Chiffre d'affaires
- Parts de marchés,
- Volume de ventes
- Panier moyen
- Coûts
- Marge
- Productivité

« Avant, la data, c'était assez opérationnel (usages internes), alors que maintenant, on recherche de la croissance : nous sommes désormais un réel support de la proposition de valeur, avec des services et produits data intégrés à notre catalogue d'offres. »

Group Chief Data & Digital Officer, secteur services

Le potentiel de valeur encore mal estimé à l'échelle de l'entreprise

Un ROI rarement mesuré en production

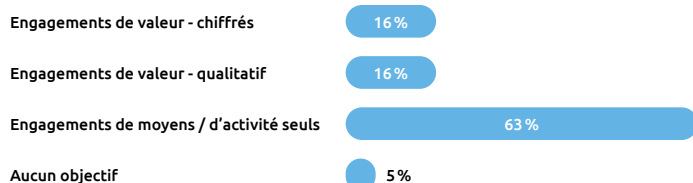
L'estimation du ROI en amont des projets est devenue courante à des fins de priorisation et de sponsorship. En revanche, la valeur réellement dégagée par les solutions data et IA est encore largement méconnue.

Un manque de pilotage par la valeur

Les 2/3 des Directions data pilotent leur activité exclusivement sur la base d'objectifs de moyens ou d'activités. Le tiers restant s'engage partiellement sur des métriques métier, mais sans engagement chiffré.

Il est évident qu'une part importante de la valeur dégagée reste diffuse et qualitative : satisfaction client / collaborateur, image, mitigation des risques... Néanmoins, l'estimation des euros économisés ou gagnés grâce à la data reste un prérequis pour assoir sa légitimité et convaincre les métiers.

Type d'objectifs suivis par les Directions data en 2023



Seuls **16%** des CDO sont responsabilisés sur des objectifs de création de valeur chiffrés

Les Directions data peinent à démultiplier leur impact à l'échelle

Les Directions data peinent à développer des cas d'usages répliquables à l'échelle de l'entreprise : impact sur plusieurs Directions métiers, déploiement à des échelles géographiques importantes (plusieurs usines, plusieurs pays...). Les cas d'usages développés jusque-là sont majoritairement tactiques et répondent à des besoins spécifiques, peu transverses.

On observe plusieurs difficultés :

- disparité des process métiers
- disparité des systèmes IT
- plateformes peu matures
- une gouvernance des données ne permettant pas de passer à l'échelle de manière uniforme / homogène.

Pour pouvoir faire atterrir des projets emblématiques, les Data Offices doivent recentrer leurs efforts sur la mise en œuvre d'un écosystème « fertile » en mettant à disposition les capacités nécessaires (Data Platform, cadre de gouvernance des données et des outils, offre self-service, modèle organisationnel déployé...).

« Nous priorisons le passage à l'échelle des cas d'usages actuels par rapport au développement des nouveaux « Proofs Of Concept ». Cela nous permet de démontrer l'apport de valeur aux métiers et donc de mieux les engager sur les projets de transformation en cours. »

Hind MECHBAL,
CIO, CCR RE

Les Data Offices attendus comme «enabler» pour transformer largement et rapidement

Volontairement dans un rôle de DOer pour insuffler une dynamique en phase d'amorçage, les CDO cherchent désormais à se positionner comme «enablers» pour multiplier la capacité à industrialiser et couvrir l'ensemble des opportunités de création de valeur par la data.

Pour cela, les Data Offices semblent vouloir dédier leur force de frappe en tant que DOer à la réalisation de cas d'usages stratégiques pour le groupe et déléguer la réalisation de cas d'usages tactiques aux équipes métiers afin de se concentrer sur leur rôle d'enabler.

« La Direction data n'a pas vocation à prendre toute la chaîne de valeur data, notre valeur doit être demain sur la partie high level et la mise en capacité. »

Directeur Data Groupe, secteur luxe

Les attentes vis-à-vis du Data Office comme «enabler» proviennent à la fois des équipes métiers...

- 1 Les programmes d'acculturation lancés ces dernières années et le niveau de culture data des populations métiers leur permettent aujourd'hui d'identifier des usages, d'exprimer des besoins et de solliciter des moyens pour les réaliser eux-mêmes
- 2 Certaines technologies no code/low code sont considérés comme accélérateurs de la décentralisation (principalement PowerBI et Dataiku)
- 3 Les métiers matures veulent pouvoir expérimenter de manière autonome sans s'inscrire dans des démarches projets formelles

...et de la DSI

- 1 Des attentes grandissantes de la DSI vis-à-vis de la data comme pilier dans les programmes SI (CRM, ERP, E-Commerce...)
- 2 Des produits digitaux nécessitant un apport d'expertise data pour bâtir des architectures robustes et un écosystème digital cohérent

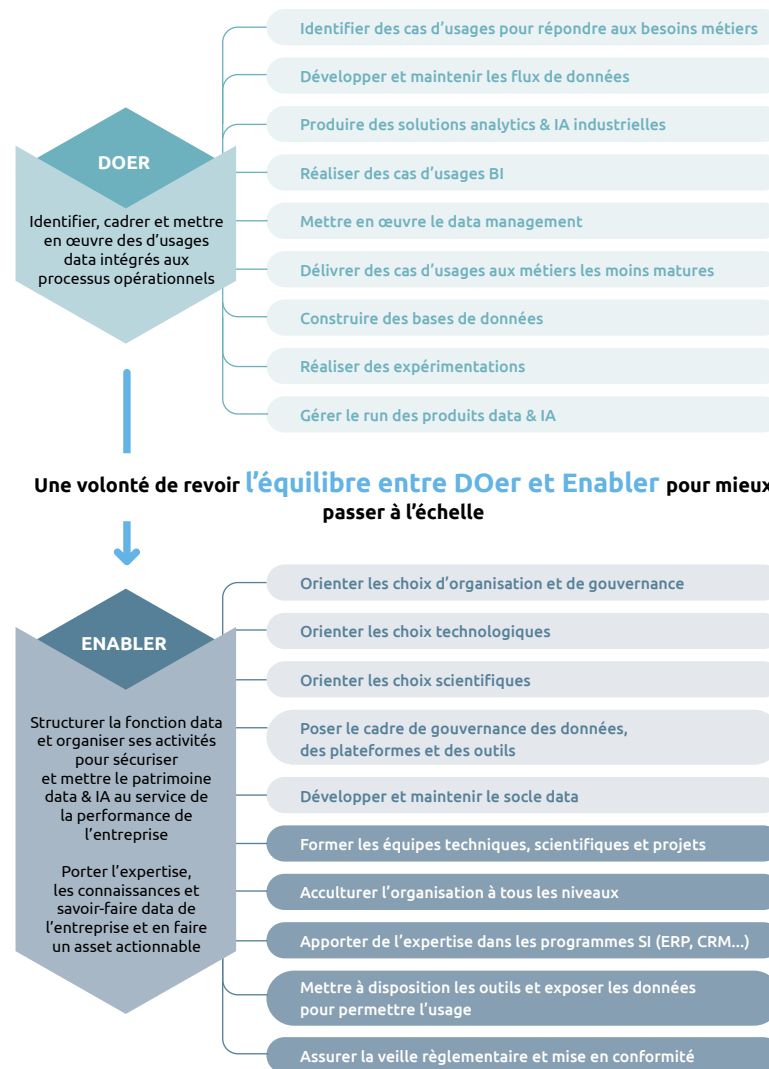
La majorité des sondés estiment être principalement orientée « livraison de produits finis » et


+ de 60% des directions


veulent tendre vers une offre de service « Mise à disposition de moyens »


Un nécessaire rééquilibrage de l'offre de service des Directions data


Pour passer à l'échelle, les Data Offices les plus matures doivent progressivement mettre la réalisation de cas d'usages « tactiques » aux mains des équipes métiers, et s'appuyer sur l'hybridation et le Data Mesh pour ne pas être goulot d'étranglement face à l'augmentation du volume de cas d'usages à adresser. Ils se placent alors dans un rôle d'enabler.




 Nous avons mis en place une offre de service data pour tous : mise à disposition d'outils pour faciliter la réalisation de l'exploration des données à l'industrialisation. >>
CDO, acteur des services


 Aujourd'hui, nous travaillons surtout des produits finis; nous devons poser la gouvernance du self-service pour mettre à disposition des données et répondre à l'engouement métier autour de PowerBI. >>
*Lilian PLACET,
DSI Adjoint CIO Office, GSF*


 Nous avons réussi à répandre l'utilisation de Dataiku, couvrant aujourd'hui 80% de nos projets Data Science. >>
Chief Data Officer, secteur banque


 Nous avons la volonté en 2023 de nous rapprocher du 50/50 entre produits finis et mise à disposition d'outils. >>
*Laurent BEUGNET,
Chief Data Officer, Orano*

Synthèse et analyse

Les Directions data délivrent désormais efficacement de la valeur de manière tactique grâce à une maîtrise de l'industrialisation des cas d'usages au service de l'ensemble des métiers de l'entreprise. Cependant, bien que des usages soient maintenant en production et que le calcul d'un ROI prévisionnel par projet se soit systématisé, l'apport de la data à l'échelle des enjeux stratégiques de l'entreprise est encore difficilement estimé et chiffré.

Cette génération de valeur par les cas d'usages semble se heurter à deux difficultés :

- la position de goulot d'étranglement du data office dans son rôle de DOer
- le passage à l'échelle des cas d'usages

En parallèle, la culture data des métiers leur permet maintenant d'être autonome dans l'identification proactive de cas d'usages et parfois dans leur réalisation. Cependant, ces équipes attendent maintenant les Directions data dans leur rôle d'ENABLER pour structurer un écosystème fertile (outils, données, compétences et cadre de gouvernance) et favoriser l'innovation au service des métiers.

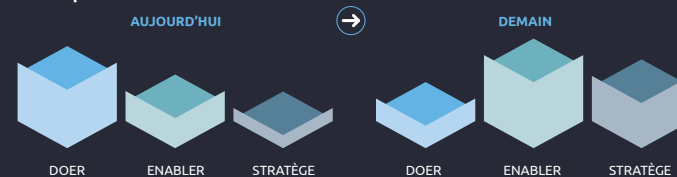
D'autre part, les Directions data sont désormais attendues par leur COMEX en tant que STRATEGES pour réussir à impacter les modèles d'affaires.

Les Data Offices sont donc attendus dans un triple rôle pour adresser la création de valeur à tous les niveaux de l'entreprise :

- la data comme STRATEGES, auprès des dirigeants et du top management en demande de vision technologique / business
- la data comme ENABLER, auprès des Directions métiers en demande d'autonomie et de capacités robustes
- la data comme DOer, pour accélérer l'entreprise et la transformation auprès des métiers peu matures et des fonctions transverses

Les organisations cherchent aujourd'hui à rééquilibrer leur offre de service pour mieux répartir le rôle de DOer et mieux répondre aux attentes dans leur rôle d'enabler et de stratège. Les Data offices devront pour cela s'appuyer sur un modèle opérationnel adapté et une gouvernance des plateformes, données et outils mature et cohérente.

Transformation du rôle des Directions data pour répondre aux enjeux de l'entreprise



- Partie 3 -

TROIS GRANDS DÉFIS À RELEVER EN 2024 POUR ADRESSER L'ENSEMBLE DES LEVIERS DE GÉNÉRATION DE VALEUR PAR LA DATA

Défi #1

S'inscrire dans les enjeux stratégiques de l'entreprise

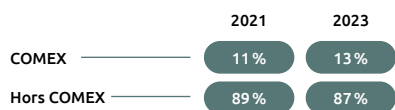
Une fonction encore trop éloignée du Comex ?

Les initiatives data & IA transforment les usages, l'organisation et souvent même les relations clients / fournisseurs / employés, et en ce sens, les Chiefs Data Officers (CDO) occupent un rôle de plus en plus stratégique au sein de l'entreprise.

Pourtant, depuis que le rôle de CDO s'est développé, son positionnement hiérarchique par rapport au COMEX a peu évolué.

Ainsi, moins de 15% des CDO sont directement rattachés au COMEX. Cela concerne des entreprises de toutes tailles et maturité, partageant toutefois un point commun : une volonté d'accélérer sur la data pour en faire un axe de différenciation ou rattraper un retard perçu par rapport au marché.

Part des Directions data rattachées au COMEX



En 2023, seulement **13%** des CDO sont directement rattachés au COMEX

La majorité des Directions se situent en N-2 par rapport au COMEX, mais de nombreux CDO et professionnels de la data militent pour un rattachement plus direct à la Direction, avec pour ambition de gagner en visibilité et en influence sur les enjeux stratégiques de l'entreprise.

Bien que ce rapprochement puisse être bénéfique, il n'est pas strictement nécessaire. En effet, les CDO les plus matures développent déjà leur posture de leader et leur capacité d'influence sans rattachement formel au COMEX. **Pour réussir, ils s'appuient sur 3 principes.**

- 1 Placer la data et l'IA dans l'agenda des décideurs : passer de l'acculturation à l'action
- 2 S'inscrire dans les grands programmes de transformation SI et digitaux en tierce partie transverse
- 3 S'inscrire dans les transformations métiers : bâtir des partenariats forts et continus avec les Directions métiers pour co-construire des trajectoires communes

1

Placer la data et l'IA dans l'agenda des décideurs : passer de l'acculturation à l'action

Depuis 2020, les CDO les plus matures ont généralement accordé une place importante dans leur feuille de route à l'acculturation data & IA des dirigeants et du top management. Les programmes initiés ont dans l'ensemble porté leurs fruits et permettent à la data de figurer à présent dans les plans stratégiques des entreprises avec des budgets alloués.

Toutefois, l'écosystème data et l'IA évolue rapidement, mû par d'importants progrès :

- technologiques (exemple: IA générative)
- réglementaires (exemple: AI Act Européen)
- socio-économiques (exemple: impact de l'IA sur l'emploi et l'organisation du travail)

De plus en plus, les décideurs se tournent alors vers la Direction data pour décrypter ces enjeux. La capacité d'actions des CDO varie en fonction de leur maturité et de leur champ d'actions. Les plus matures saisissent l'occasion de façon proactive, pour sortir de leur pré-carré naturel et animer des groupes de réflexions interdisciplinaires afin de proposer des stratégies à l'échelle de l'entreprise via une offre de services structurée. Ils jouent alors un rôle de conseiller stratégique et plus seulement d'évangéliste. D'autres, souvent moins matures et plus centrés sur un rôle de DOer, ont un positionnement plus attentiste.

Pour répondre à ces attentes de conseil, les Directions data s'emparent également de nouveaux mandats :

- **Proposer une interprétation stratégique des tendances technologiques**
S'assurer d'être en capacité de « digérer » les disruptions technologiques comme l'IA générative en expérimentant rapidement pour se bâtir des convictions sur le potentiel pour l'entreprise et les opportunités à saisir ou non
- **Représenter l'entreprise dans l'écosystème data et dans la filière**
Faire rayonner les initiatives data de l'entreprise au sein du secteur et des acteurs tech, pour bénéficier de partenariats technologiques, rester à l'état de l'art, être au cœur des réflexions sectorielles et ainsi être leader sur son marché dans la transformation de l'entreprise

2

S'inscrire dans les grands programmes de transformation SI et digitaux en tierce partie transverse

Les Directions data semblent de plus en plus attendues dans l'apport d'expertise aux programmes transverses. Pour répondre à ce besoin, elles se mettent en capacité de s'inscrire dans les programmes SI et participer à la cohérence globale du SI d'entreprise (règles métiers, modèles de données, décisions d'architecture...) pour en faire un « terrain fertile » aux cas d'usages à venir.

La valeur d'une donnée dépend en grande partie de sa capacité à circuler sans entrave à travers l'entreprise. Dans une banque par exemple, les données produites au sein de la division paiement présentent un intérêt majeur pour le marketing, en matière de connaissance client.

Or, les données restent encore largement silotées by design, franchissant difficilement les barrières organisationnelles. En cause, un manque de leadership reconnu et de ressources au service de la transformation data transverse, qui poussent les métiers à prioriser les initiatives data dont ils sont directement bénéficiaires sur leur périmètre.

Au lieu d'opérer au sein des divisions organisationnelles, les CDOs les plus matures interviennent en tierce partie. Ils alignent les intérêts des métiers producteurs et consommateurs de données, sécurisent le sponsorship, et lorsque nécessaire, appellent les métiers à mutualiser leurs ressources sur des projets communs.

Cela passe par le lancement de chantiers « fondations » pour amorcer la pompe: identifier des terrains de jeu transverses à fort potentiel (domaine de données cœur métier par exemple) pour préparer le patrimoine de données en tierce partie afin de le rendre exploitable par les métiers et encourager la réalisation de cas d'usages.

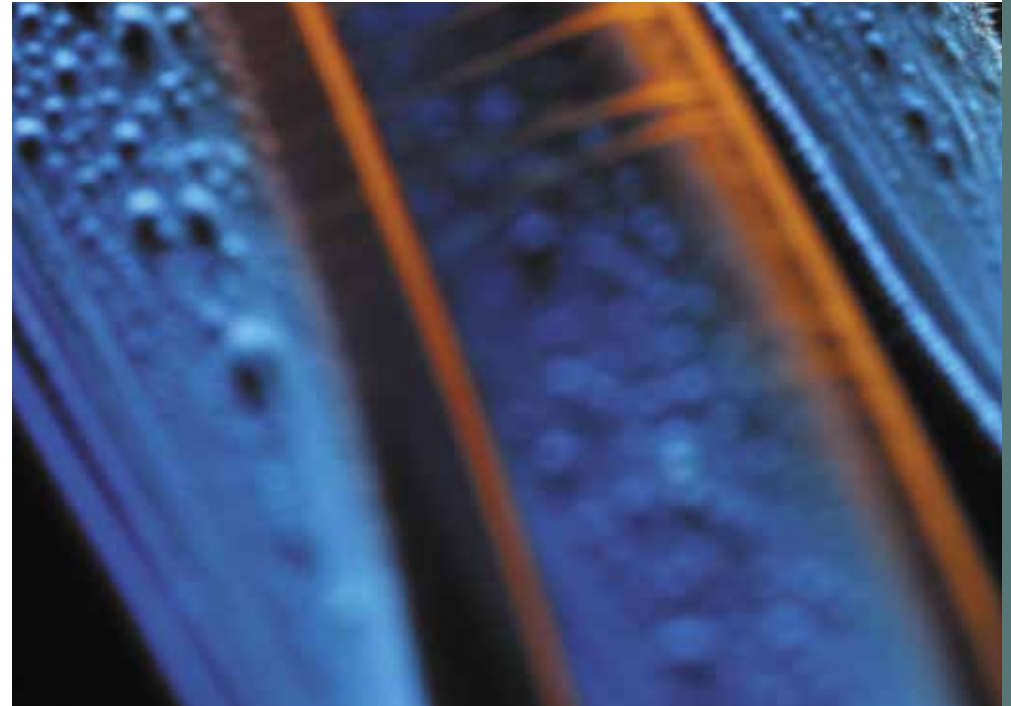
3

S'inscrire dans les transformations métiers : bâtir des partenariats forts et continus avec les Directions métiers pour co-construire des trajectoires communes

Les Directions data les plus matures se retrouvent désormais en position de co-sponsoriser des projets ambitieux business/data pour inscrire la data dans le modèle d'affaires de l'entreprise à travers des produits robustes et matures, construits sur la durée et inscrits dans les processus métiers (ou dans les catalogues de produits, offres de services...).

Cela donne alors lieu à des produits matures intégrant, par exemple, les catalogues de services proposés aux clients comme différenciateur marché.

Cette collaboration de proximité passe par un accompagnement du métier dans les réflexions business de valorisation de la donnée (business model d'un produit data, monétisation de données...) et des binômes CDO/Directeurs alignés.





SUCCESS STORY

Leader dans le secteur de l'ingénierie

Problématique

Porté par un important plan de transformation de l'entreprise visant à devenir un leader mondial dans les secteurs de l'ingénierie et de la mobilité, le groupe souhaite avancer rapidement sur sa transformation data pour en faire un fort accélérateur de croissance pour le groupe.

Une importante évolution de l'offre de valeur data était donc attendue pour répondre favorablement aux enjeux opérationnels et stratégiques de l'entreprise.

Démarche

Dans une optique d'offre de valeur holistique, l'organisation centrale en charge de la data a défini et mis en œuvre un plan stratégique complet selon un horizon temporel à 2/3 ans à partir de :

- l'ambition du groupe, pour définir des objectifs chiffrés
- les besoins de transformation du groupe pour utiliser la data comme socle dans les programmes SI & digitaux (CRM, ERP, BI...)
- les besoins opérationnels des équipes métiers pour identifier les bons cas d'usages (IA, Optimisation...).

Ce plan stratégique a ensuite été opérationnalisé rapidement sur l'ensemble des dimensions data : Data Platform, cas d'usages, déploiement de la fonction data, data foundations... grâce à un sponsoring COMEX fort et aux bonnes synergies avec les autres programmes de transformation du groupe (IT & Business).

Le modèle opérationnel cible a été défini en prenant en compte les spécificités métiers comme la maturité hétérogène des business lines, la technicité des données à adresser ou encore la demande d'autonomie des équipes terrain dans la manipulation de données. C'est notamment ce qui a appuyé le choix d'un déploiement rapide du Data Mesh technique et organisationnel.



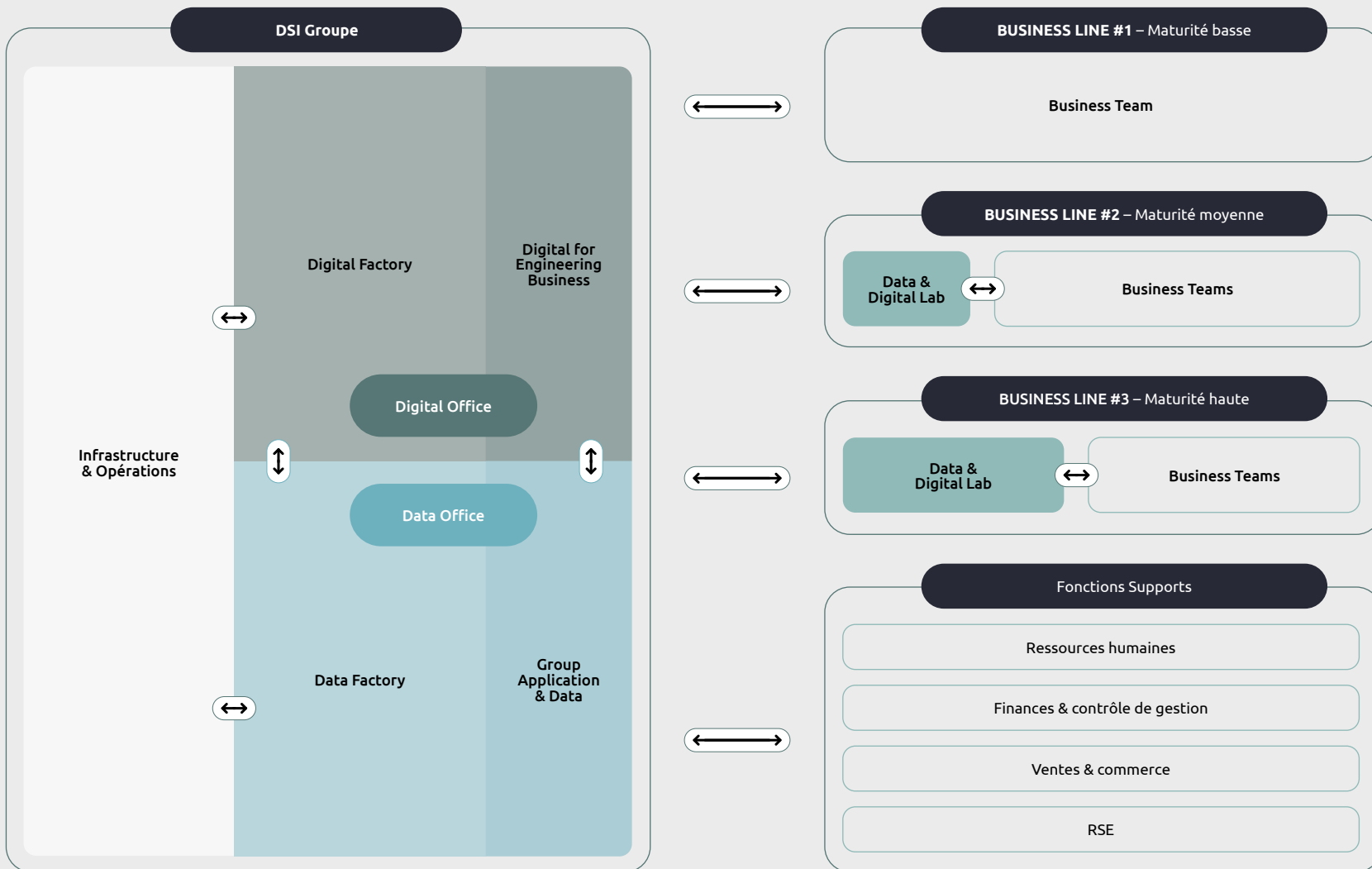
Résultats

- Une vision claire, structurée et partagée de la stratégie data du Groupe déclinée pour les business lines
- 5 dimensions majeures adressées en parallèle pour structurer et conduire la transformation data sur 2 ans
- Un modèle organisationnel et un cadre de gouvernance permettant de :
 - structurer rapidement les activités et assurer une montée en puissance des métiers
 - adapter le niveau d'hybridation et l'offre de service de la data factory et du data office selon les besoins métiers, les domaines concernés et les types de projets : delivery de produits fini, mise à disposition de données, apport d'expertise...
- Un nouveau processus de valorisation des cas d'usages data/IA objectivant et validant les intérêts économiques, métiers et environnementaux des cas d'usages en collaboration avec les Directions transformation et financière
- Une nouvelle Data Platform Cloud désignée selon les principes du Data Mesh, à même de supporter la transformation data et technologique du Groupe en tant que tierce partie fournissant un écosystème propice à la génération de cas d'usages transverses et locaux : délivrer les cas d'usages transverses et outiller la transformation digitale des métiers
- De premiers cas d'usages data/IA sont valorisés, cadrés et développés par la Data Factory et par des entités data locales dès les premiers mois de transformation pour les fonctions cœur métier et fonctions support

Exemple de schéma d'organisation de la fonction data dans l'ingénierie

SUCCESS STORY

Une entité digitale proche des équipes cœur métier pour transformer les activités en profondeur en mettant la puissance des produits digitaux et data au service du métier d'ingénieur



Un modèle organisationnel qui s'adapte à la maturité de chaque business line avec un mandat de la DSI qui évolue en conséquence : d'abord D'Oper puis Enabler une fois le métier autonome

Une collaboration DSI & fonctions supports pour accompagner efficacement les programmes transverses en s'appuyant sur la force de frappe Data Factory

ORGANISATION & ARCHITECTURE DATA MESH

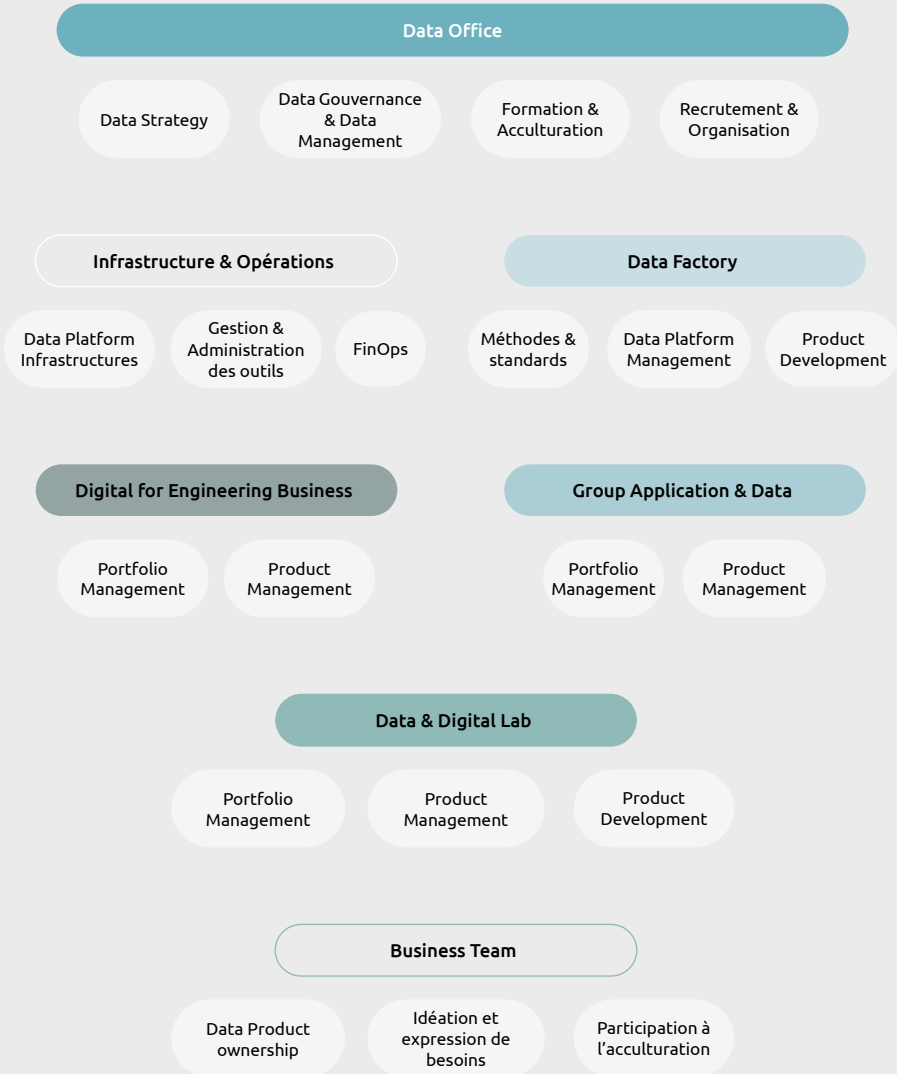
- Data as a Product
- Data Contracts
- Data Product Owners



Interfaces définies selon le cadre de gouvernance de la fonction data commun et partagé : mandats, rôles, processus et comitologie

Exemple de schéma d'organisation de la fonction data dans l'ingénierie

Principales activités data par entité



Défi #2

Piloter les Directions data par la valeur

Des engagements de valeur partagés par le métier et la data

Les entreprises matures sont plus nombreuses à demander des engagements business communs aux CDO et aux lignes métiers pour renforcer l'alignement de leurs feuilles de route.

Cette approche fonctionne généralement sur les principes suivants :

- chaque ligne métier s'engage sur un cash-in annuel data en €
- les projections sont fixées en commun avec un représentant métier rattaché à la Direction data, lui-même objectif
- pour atteindre l'objectif, chaque représentant aide sa ligne métier à :
 - établir une feuille de route
 - piloter la réalisation des usages en lien avec la direction data et la DSI

La Direction financière comme partenaire

Pour occuper pleinement leur rôle de stratège, les CDO matures ont mis en place des systèmes de mesure en partenariat avec la Direction financière¹.

Celle-ci est principalement impliquée à 3 niveaux :

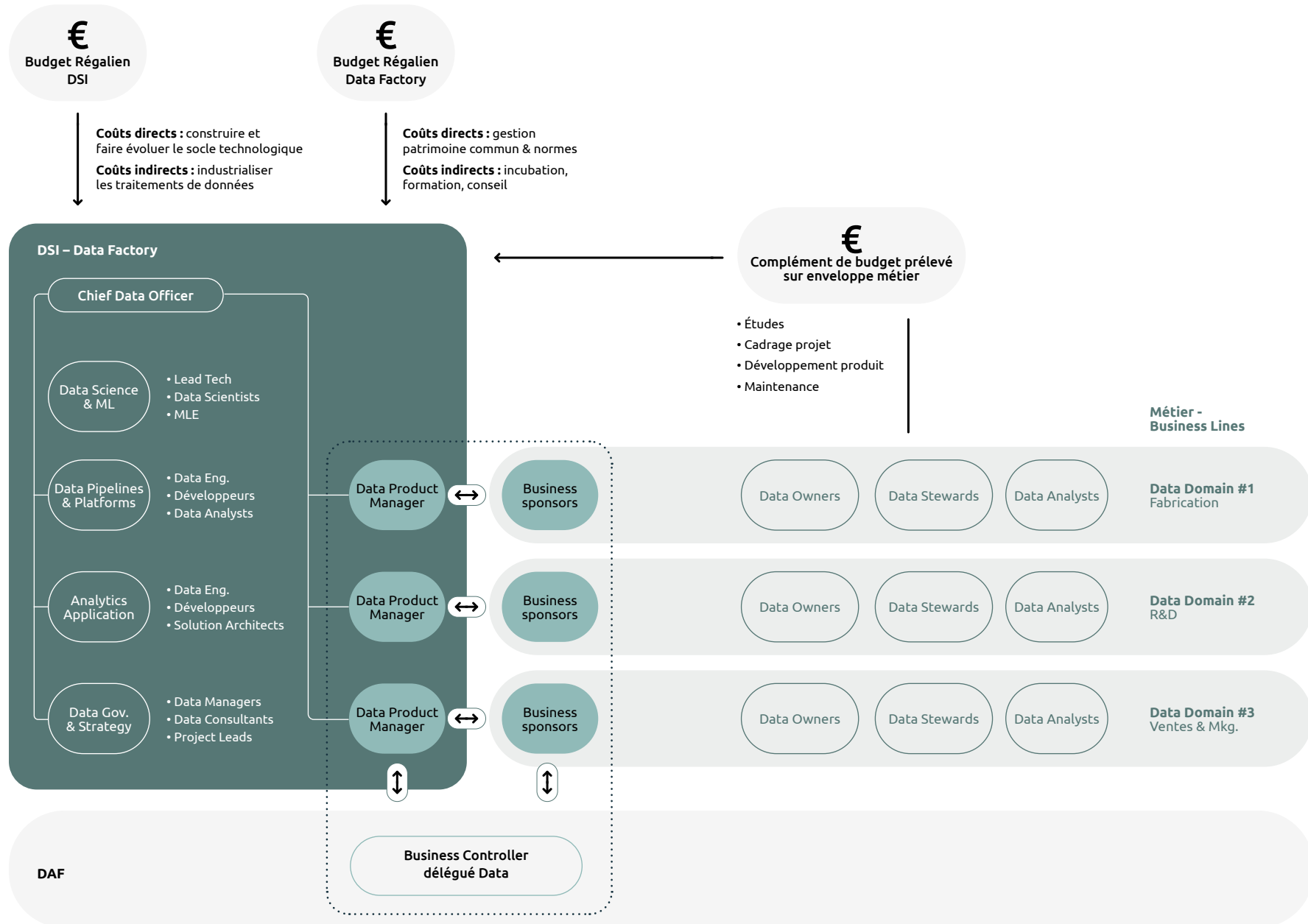
- définir la méthodologie de chiffrage
- contrôler les hypothèses pour chaque cas d'usages
- valider les résultats

C'est une façon, pour les métiers, de sécuriser leurs objectifs business, et pour la data, de légitimer son offre de valeur et ses demandes d'investissement ou de ressources.

(1) Dans certains cas, la Direction de la transformation peut jouer ce rôle de partenaire.



Exemple d'une organisation data orientée ROI chez un acteur industriel (Organigramme simplifié)



Une méthode de valorisation à formaliser

Formaliser le pilotage de bout-en-bout

La plupart des entreprises ont mis en place un processus de pilotage par la valeur en amont du lancement des projets data, pour prioriser les opportunités et cibler les cas d'usages prometteurs. En revanche, seules les entreprises les plus matures pilotent la valeur dégagée durant le développement et l'exploitation des solutions data.

Cet exercice nécessite une revue régulière du portefeuille et s'appuie de plus en plus sur des Dashboards dédiés, permettant aux équipes data, DSI et métiers de centraliser les informations relatives au niveau d'utilisation, à la disponibilité et la performance des solutions data.

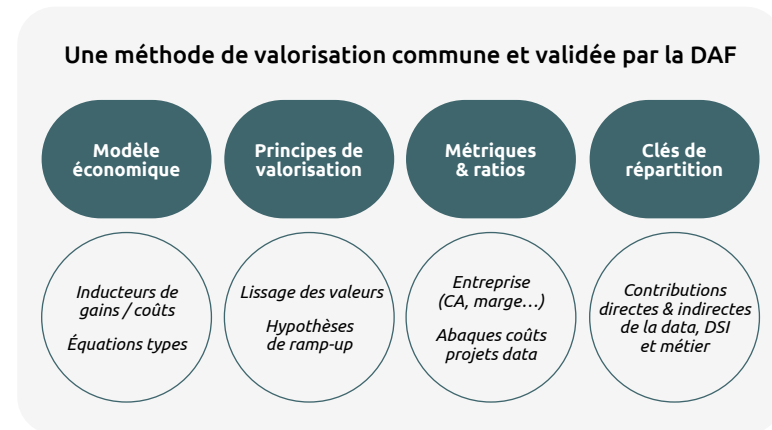
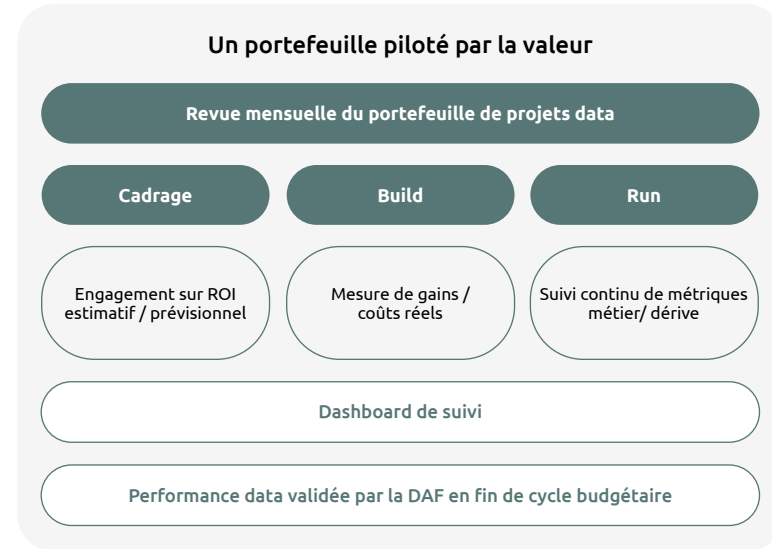
Formaliser le chiffrage des cas d'usages

Les entreprises matures s'appuient sur des méthodes de valorisation articulées autour de 4 éléments.

- 1 **Modèle économique** : quelle métrique utiliser pour la valorisation (ex: ROI, Cash-in, score...)? Comment estimer chaque métrique sur la base de valeurs connues (ex: nombre d'heures économisées x coût horaire moyen x nombre d'utilisateurs)?
- 2 **Principe de valorisation** : quelles sont les règles de calcul et hypothèses admises pour simplifier/robustifier l'estimation (ex : courbes d'adoption type, taux de croissances annuels moyens...)
- 3 **Métriques & ratios** : quels sont chiffres clés ou les ordres de grandeurs sur lesquels s'appuyer par défaut (ex: chiffre d'affaires, salaires moyens, coûts d'hébergement...)?
- 4 **Clé de répartition (optionnel)** : comment répartir la valeur générée entre Data, métier et DSI dans le cadre d'une comptabilité intra-entreprise formelle?

Ces méthodes permettent de simplifier le travail d'estimation, de fiabiliser les comparaisons entre projets et de générer un meilleur consensus autour des estimations.

Exemple d'une méthode de valorisation à l'état de l'art d'un acteur industriel



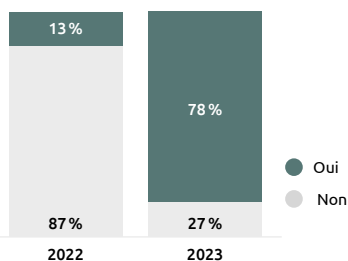
Défi #3

Démultiplier l'impact de la data à l'échelle avec le Data Mesh

Un modèle qui continue de se populariser

La part des entreprises qui s'appuie sur des concepts Data Mesh pour transformer l'organisation data a bondi de 13% à 78% en l'espace d'un an. Encore en phase exploratoire, la plupart des entreprises déploient le modèle sur un nombre restreint de data domains, et testent seulement les premiers Data Products en condition de run. Cependant, la majorité des sondés sont enthousiastes et souhaitent persévérer dans la direction choisie pour 2024.

Implémentation de concepts Data Mesh au sein des entreprises sondées



« Data Mesh, nous sommes en plein dedans, notamment pour accélérer notre vélocité sur les fondamentaux data (plateforme & data gouvernance) via la définition des data domains et la création de notre Data Platform Office. »

Antoine LY,
Chief Data Science Officer, SCOR

« Nous avons un objectif de 6 data products d'ici à la fin de l'année. Nous en avons déjà un en place sur l'algorithme de recommandations personnalisées par exemple. En top down, nous avons plus de difficultés à identifier les data products susceptibles de répondre de façon directe à des douleurs exprimées par le top management, même si de très bonnes pistes commencent à émerger. »

Brice MIRANDA,
Data, AI & Automation Deputy, ORANGE

« La logique Data as a Product est inscrite dans les objectifs cibles. »

Chief Data Officer, Secteur Banque

« Certains choix d'architecture s'appuient sur les concepts du Data Mesh, notamment en matière d'accès à la donnée. »

Loïc BRIENT,
Chief Data Scientist, BPCE

« On veut tendre vers une organisation Data Mesh à partir de l'année prochaine pour décentraliser la gestion de l'infrastructure et ainsi alléger les processus de notre datalake. »

Head of AI Factory, secteur luxe

Une réponse aux enjeux des Directions data et DSI

Le Data Mesh adresse des enjeux partagés par la plupart des Directeurs data :

- équipes centralisées sous dimensionnées par rapport à des besoins métiers croissant
- pratiques data hétérogènes à travers l'organisation
- règles de gouvernance difficiles à faire respecter
- données silotées peu partagées

En réponse à ces enjeux, le Data Mesh propose une vision claire :

- autonomiser au maximum les métiers en leur donnant le mandat, les incitations et les moyens nécessaires
- favoriser l'échange de données entre pairs sans passer par l'intermédiation systématique du Data Office

OBJECTIFS DU DATA MESH

- 1 Réduire les goulots d'étranglement sur les équipes centralisées
- 2 Conserver l'expertise métier au plus près de la donnée
- 3 Faciliter et étendre l'usage des données en dehors des silos
- 4 Clarifier les responsabilités de chacun autour de la data

Autonomiser les métiers

Relais métiers renforcés et responsabilisés
Données gouvernées par leur producteur
Gouvernance centralisée réduite au minimum
Stack modulaire en self-service

Favoriser le Peer 2 Peer

Producteurs incités à partager leurs données
SLA / QLA
Données cataloguées
API first





Nous avons prévu, dès le départ, du Data Mesh car les filiales sont très indépendantes, avec des P&L et des statuts juridiques indépendants. Donc on peut dire aujourd'hui que chaque société a son datalake et il y a des passerelles qui sont créées. >>

Chief data Officer, secteur industrie



Depuis l'an dernier, nous souhaitons aller vers une logique Data Mesh, pour améliorer l'accès à la donnée. >>

CDO, secteur luxe

Une transformation profonde de l'organisation data

Le Data Mesh est avant tout une transformation organisationnelle visant à décentraliser le data management (self-service, autogestion des données par le métier...), tout en offrant la cohérence nécessaire pour que les données circulent sans entrave à travers l'organisation (standards, socles communs, interopérabilité...).

D'un point de vue technologique, le Data Mesh ne repose pas sur des composants d'architecture particulièrement innovants (ex: Data catalogues, ETLs, Datalakes...).

C'est l'alignement entre les choix technologiques et la transformation organisationnelle qui représente la réelle innovation.

4 principes fondamentaux*

1- Domain Ownership

- Déléguer la gestion des données aux métiers les plus proches du contexte d'utilisation
- Promouvoir les échanges de pair à pair sans intermédiation

2 - Data as a Product

- Responsabiliser des équipes pluridisciplinaires sur le cycle de vie d'une donnée
- Promouvoir une utilisation fiable et autonome notamment via des APIs, et des SLAs

3 - Federated Computational Governance

- Equilibrer l'autonomie des Data Domains et l'imposition de normes collectives (interopérabilité, sécurité...)
- Favoriser l'implémentation des normes directement dans « le code » des Data Products

4 - Self-Service Infrastructure as a platform

- Permettre aux utilisateurs de construire leur stack technique sur un socle commun garantissant l'interopérabilité
- Offrir des capacités centralisées utiles à tous (ex : catalogue de données)

* Source : Zhamak Dehghani, Analyse Quantmetry

4 leçons apprises sur le terrain

1

Piloter la transformation de manière unifiée

Trop souvent, des projets Data Mesh émergent indépendamment dans l'entreprise de façon non coordonnée. Or, le principal apport du Data Mesh est justement d'offrir une vision, un vocabulaire et une trajectoire commune aux multiples parties prenantes de l'entreprise.

Assurez-vous d'avoir

- un programme officiel et visible au niveau COMEX
- un leader et une gouvernance dédiée

2

Prioriser l'organisation avant la technologie

Le Data Mesh est avant tout une transformation organisationnelle. Les arbitrages technologiques doivent succéder aux choix d'organisation et de gouvernance.

Assurez-vous d'avoir

répondu aux questions clés avant d'engager des transformations technologiques ou de développer des produits data :

- Comment se redéfinissent les responsabilités entre data, DSI et métier ?
- Qui est en charge de construire et d'opérer les produits data ?
- Quelles compétences et ressources pour rendre cela possible ?

Transformer l'organisation

Création de nouvelles structures (ex : Data Domains)

Nouvelles façons de travailler (ex: mode produit vs projet)

Nouveaux rôles (Data Product Owner)

Changement de culture / incitations (partage vs silotage des données)

Transformer l'architecture

Modularisation de la stack

Interopérabilité des composants

Harmonisation des modèles sémantiques

Implémentation d'une couche d'observabilité

Une transformation progressive et ciblée

Les organisations relèvent ces défis de façon progressive en insistant sur les éléments du Data Mesh qui leur paraissent prioritaires.

Exemples de stratégies Data Mesh d'entreprises issues de secteurs variés

	Acteur Bancaire	Acteur de l'automobile	Acteur de la grande distribution	Acteur du transport	Acteur de la santé	Acteur du divertissement
Domain Ownership	<ul style="list-style-type: none"> Chaque objet métier est assigné à un Data Domain Les Data Domain Owners sont en place avec un budget dédié 	<ul style="list-style-type: none"> Equipes de delivery de la Data Factory structurées selon les domaines fonctionnels (vs découpage technique) 	<ul style="list-style-type: none"> Le corporate et les régions sont alignés sur les Data Domains 	<ul style="list-style-type: none"> Structuration des Data Domains avec les Business Units Identification des principaux objets business 	<ul style="list-style-type: none"> 1 Data Domain par fonction métier Ownership métier sur chaque domaine Alignement sur les principes de fonctionnement entre Data Squads et Digital Products 	<ul style="list-style-type: none"> 1 Data Domain par fonction métier Ownership métier sur chaque domaine Alignement sur les principes de fonctionnement entre Data Squads et Digital Products
Federated Computational Governance	<ul style="list-style-type: none"> Gouvernance inter-domaines Standards d'interopérabilité 		<ul style="list-style-type: none"> Standardisation des organisations locales et de la gouvernance en région La France comme colonne vertébrale pour produire des cas d'usages répliquables 	<ul style="list-style-type: none"> Accords et alignements inter-métiers sur la Data Policy et la Sharing Strategy Définition de processus communs 	<ul style="list-style-type: none"> Le modèle opérationnel est aligné entre les entités 	<ul style="list-style-type: none"> Le modèle opérationnel est aligné entre les entités
Data As a Product	<ul style="list-style-type: none"> Framework de Data Contract Templates de Data product Plateforme API 	<ul style="list-style-type: none"> Standardisations des formats d'exposition selon les cas d'usage 		<ul style="list-style-type: none"> Accords et alignements sur les standards d'exposition de la donnée (Data Product) 	<ul style="list-style-type: none"> Approche Produit Standardisation des Data Products 	<ul style="list-style-type: none"> Définition de 3 types de Data Products : Data Analytics, Data Science (API) and Dashboard products
Self-Service Infrastructure as a Platform			<ul style="list-style-type: none"> Normalisation des interfaces avec la Data Plateforme Groupe Choix imposé aux stakeholders 	<ul style="list-style-type: none"> Une offre de service unique à l'échelle du groupe (Snowflake) 	<ul style="list-style-type: none"> Data Products exposés à travers une solution unique 	<ul style="list-style-type: none"> Data Products exposés à travers une solution unique (Snowflake) avec des espaces dédiés par Data Domain et 3 couches de stockage (Raw, Silver, Optimized)

- High implementation
- Medium implementation
- No implementation

3

Rester pragmatique

Le Data Mesh offre un cadre théorique utile, mais ses principes se heurtent parfois à la réalité terrain. Les entreprises qui réussissent suivent un parcours de transformation qui leur est propre.

Assurez-vous d'avoir

- une approche pragmatique, itérative et adaptée à votre maturité et vos priorités

4

Délivrer des résultats rapidement

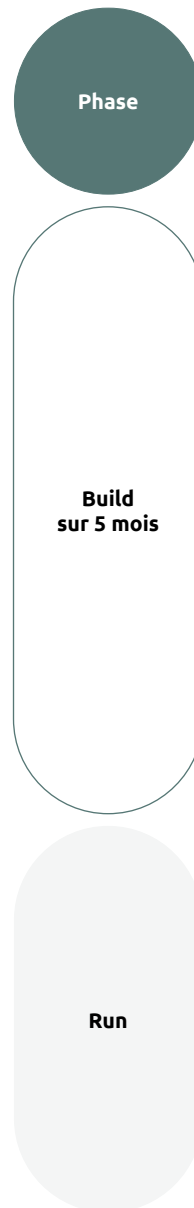
Le Data Mesh est une transformation profonde et progressive. Pour autant, il faut cibler ses efforts sur des victoires accessibles afin de maintenir la dynamique.

Assurez-vous d'avoir

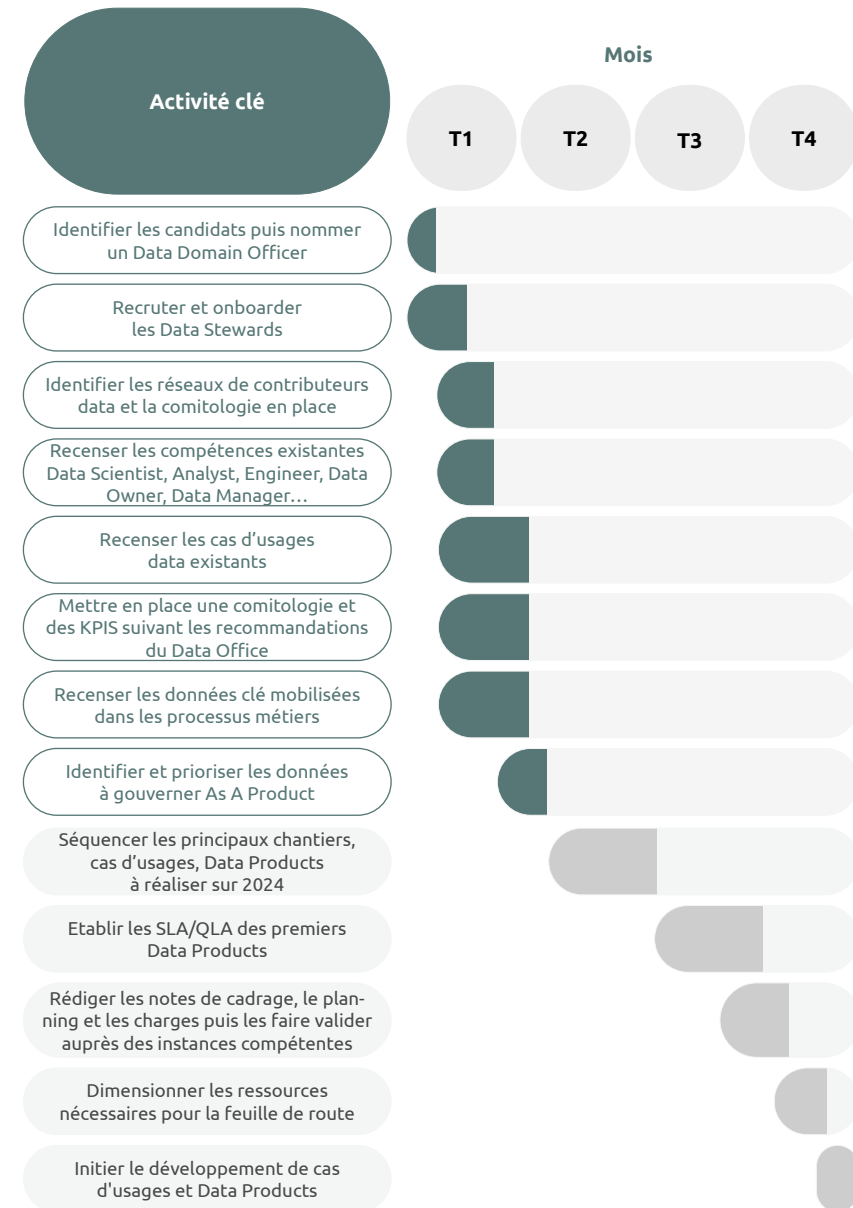
- fractionner le déploiement en phases : lancer au plus 2 à 4 Data Domains à la fois
- bien choisir ses Data Domains : privilégier ceux qui s'engagent dans une relation client/fournisseur de données avec un bon niveau de sponsorship
- aboutir vite sur des Data Products : viser le développement d'un Data Product en 6 mois ou moins

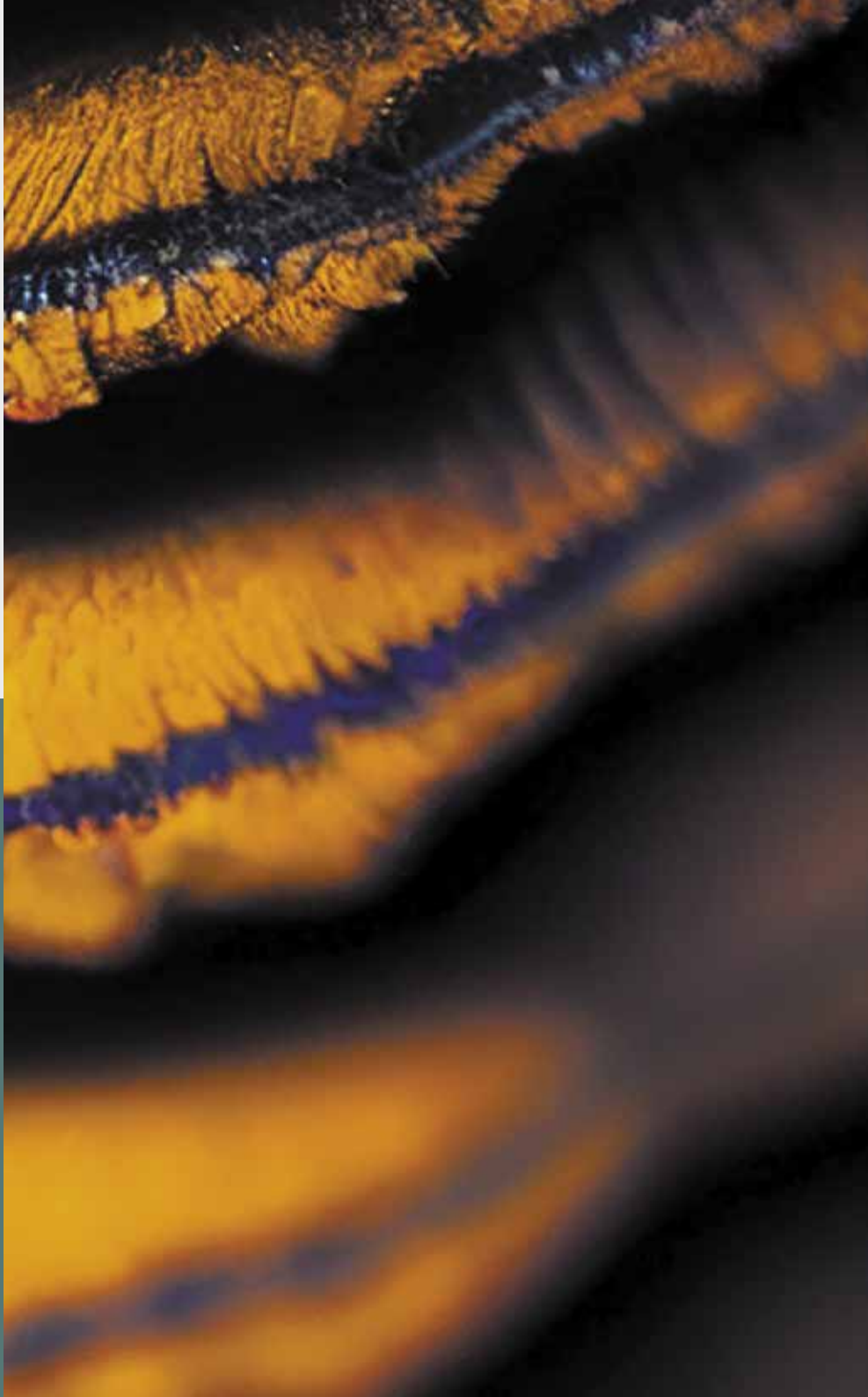
« Nous essayons d'impulser une logique Data Product. On ne dit pas Data Mesh car ça ne parle que très peu aux instances de décision pour l'instant, ce n'est pas encore assez vulgarisé. »

Brice MIRANDA,
Data, AI & Automation Deputy, ORANGE



Exemple de feuille de route type pour le déploiement d'un Data Domain





Synthèse et analyse

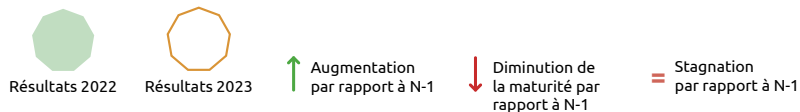
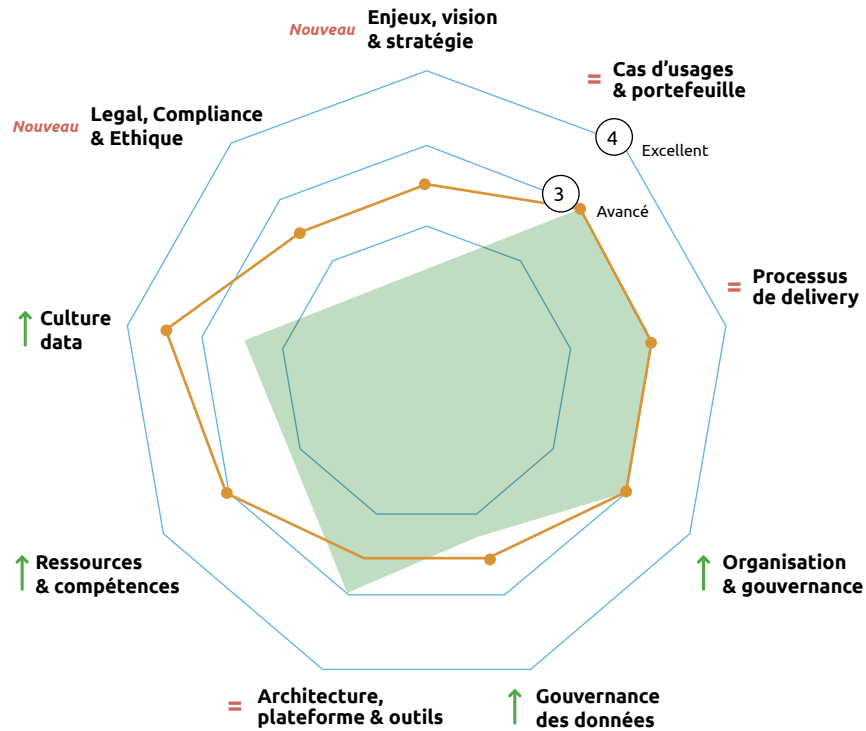
L'engouement autour de l'IA générative a remis la data dans l'agenda des COMEX et c'est ainsi une magnifique opportunité à saisir pour les CDO afin de relever les défis de la transformation de l'organisation par la data :

- ce coup de projecteur doit permettre plus que jamais aux CDO de s'inscrire dans les enjeux stratégiques de l'entreprise aux côtés de la DSI et des Directions métier
- les promesses de création de valeur et de gain de productivité de l'IA générative vont également provoquer des attentes plus élevées de mesure concrète de cette valeur. L'occasion d'aligner métiers et contrôle de gestion dans un pilotage par la valeur de la data
- enfin, le déploiement en production et à l'échelle d'IA génératives, comme celui de tous les usages data, reposera notamment sur des outils mis à la disposition du plus grand nombre... si, et seulement si, les données et les connaissances qui alimentent ces outils sont disponibles et pilotées ! Les grands principes du Data Mesh répondent à cet enjeu et une large majorité d'entreprises avance dans ce sens. Mais les initiatives restent très hétérogènes et souffrent d'un manque de pilotage globale au niveau de l'entreprise. Cela remet en avant la nécessité au préalable de bien aligner l'ensemble des parties prenantes, data, IT et métiers vers un objectif et des moyens communs

Pour éviter l'écueil du gouffre des désillusions dans lequel l'IA « traditionnelle » était tombée, l'IA générative et ses promesses de transformation profonde des entreprises devra s'accompagner d'un investissement stratégique et humain sur ces piliers de la transformation par la data.

CONCLUSIONS ET PERSPECTIVES

La maturité data des entreprises à l'épreuve de l'IA générative



Ce qui est nouveau cette année dans la construction de notre grille d'analyse

La grille d'évaluation des maturités passe de 7 à 9 axes pour mieux représenter les dimensions data et introduire des sujets prenant de l'envergure au sein des fonctions data :

- Séparation de l'axe « Enjeux & cas d'usages » en 2 axes distincts « Enjeux, vision & stratégie » d'une part et « Portefeuille de cas d'usages » d'autre part afin d'aborder spécifiquement l'inscription de la data dans la stratégie d'entreprise puis la déclinaison dans le portefeuille de cas d'usages et les impacts sur sa composition. La grande majorité des Directions data s'est fixée des ambitions, une vision des objectifs, une stratégie et a construit une feuille de route. Cela a été favorisé par l'inscription grandissante de la data comme sujet clé dans les stratégies d'entreprise.
- Création d'une dimension d'étude « Légal, compliance & éthique » pour préciser la maturité des entreprises face aux contraintes réglementaires actuelles et futures (RGPD, AI Act) ainsi qu'aux enjeux éthiques soulevés par l'IA et l'IA générative (GenAI). La réglementation RGPD est plutôt maîtrisée sur les cas critiques, et les entreprises se préparent désormais à l'arrivée de l'AI Act afin de ne plus revivre l'adaptation chaotique au RGPD.

Les piliers des organisations data à l'épreuve de l'IA générative

- En 2023, plus de 1000 outils d'IA générative ont été lancés, provoquant un nouveau far-west technologique. Cela vient réinterroger les choix réalisés ces dernières années par les entreprises. **Pour retrouver le même niveau de maturité que sur l'IA/ML, les organisations devront mettre en place des plateformes qui prennent en compte les nouveaux besoins de l'IA générative.** La volonté de créer des plates-formes d'IA générative avec une vraie composante FinOps devrait s'intensifier en 2024.
- Le **processus de delivery** stagne cette année. Si pour les entreprises les plus matures, le déploiement de l'agilité, la mise en place de rôles dédiés (PO, PPO...) et la systématisation des méthodologies de cadrage permettent désormais de passer en production plus facilement, l'IA générative vient bousculer légèrement les pratiques. L'intégration du LLMOps au sein des pratiques « Ops » est attendu pour 2024.
- La gestion du **portefeuille de cas d'usages stagne également car**, malgré la mise en place de gouvernance de portefeuille permettant de réunir les parties prenantes autour des prises de décisions, les entreprises peinent encore à mesurer le ROI effectif de leurs cas d'usages. De plus le foisonnement des initiatives IA génératives partout dans l'entreprise n'est pas encore bien maîtrisé par l'ensemble des acteurs.

Des fondamentaux qui continuent de progresser ou se relancent

- **Le concept de Data Mesh a mis sous les projecteurs le besoin d'hybrider les organisations data** entre des ressources centrales et des relais locaux. Petit à petit, chaque entreprise progresse pour avancer vers le niveau d'hybridation adapté à ses enjeux. Cependant, cela ne se reflète pas dans la matrice car le franchissement d'un cap de maturité a amené à une redéfinition des standards à la hausse. Les modèles opérationnels permettent de gouverner la fonction data en répartissant les mandats entre central/local et entre data/DSI/métier. Le défi des entreprises sera de proposer une offre de service et de se munir de ressources pour satisfaire à la fois les entités métiers et locales les plus matures comme celles ayant besoin de plus d'accompagnement.
- **La gouvernance des données** progresse (enfin!) sous l'effet de la mise en application de concepts désormais maîtrisés (data catalogue, data lineage...) et du développement du Data Mesh, notamment de l'appropriation du principe de « Data product » pour répondre aux demandes métiers : autonomie et accès à des données fiables. **De nombreux acteurs restent néanmoins encore insatisfaits des résultats tangibles. Enfin, l'impact de la diffusion des LLM nécessitant de passer du data management au knowledge management n'est pas encore mesuré ni pris en compte.**
- **Les dimensions ressources, compétences et culture continuent de gagner en maturité** logiquement sous l'effet de l'expérience, des programmes d'acculturation et de formation. Les **compétences/rôles/postes** data s'inscrivent de plus en plus dans les référentiels et dispositifs RH afin d'être reconnus et structurés. L'acculturation permet désormais aux métiers d'identifier des cas d'usages et l'IA générative en devenant grand public apporte aussi aux populations métiers une culture générale data/IA.



Grille d'évaluation des maturités data

	Enjeux, vision & stratégie	Cas d'usages & portefeuille	Ressources & Compétences	Architecture, plateforme & outils	Organisation & Gouvernance	Processus de delivery et méthodologies	Data Gouvernance et Data Management	Culture Data	Legal, Compliance & Ethics
Excellent	La data & l'IA sont inscrits comme des sujets clés de la stratégie d'entreprise - La roadmap de la fonction data est mise en œuvre et actualisée continuellement pour aligner les moyens et l'ambition	La roadmap est continuellement alimentée avec des cas d'usages portés par les métiers et un ROI effectif est suivi	Les ressources sont disponibles, les RH autonomes et matures pour faire croître le dispositif	Les plateformes intègrent des outils de data management, d'éthique et de gestion du cycle de vie des modèles - Des outils avancés sont mis à disposition des utilisateurs techniques et métiers pour répondre à des usages multiples (no code,..) - Appropriation du Data Mesh (technique) «anti monolithique» + Self Service	Le modèle opérationnel est éprouvé sur l'ensemble du périmètre, et le cadre de gouvernance permet une bonne articulation des stakeholders de la fonction data : IT, Métier, Digital, Data. L'ensemble des ressources nécessaires est recruté et/ou formé	L'industrialisation est la norme et se passe de façon fluide grâce à des méthodologies déployées sur les phases de cadrage, pilotage (agilité), industrialisation (CI/CD, DevOps) et exploitation	La totalité du patrimoine de données est cartographiée et documentée, la data est gérée comme un produit dont le métier s'est approprié la responsabilité fonctionnelle - Les data users/producers appliquent les standards sous la supervision des responsables Data Gouv	La data est un axe stratégique majeur. COMEX, métier et IT soutiennent véritablement la transformation data via des engagements et budgets concrets	RGPD déployé, framework d'audit et remédiation IA de confiance sur l'ensemble des cas d'usages
Avancé	L'ambition est définie à travers des objectifs chiffrés et les moyens à aligner sont identifiés et planifiés	Les cas d'usages sont cartographiés et priorités dans une roadmap. Le ROI est un critère de sélection	Les effectifs nécessaires sont en place. Des écarts de compétences restent à combler par de la formation / prestation et des rôles/positions restent à inscrire dans la fonction RH (valorisation des compétences, parcours de formation,..)	Les plateformes couvrent la totalité du patrimoine de données et sont rationalisées pour une bonne capitalisation du Dev. et un passage à l'échelle fluide des I.A. - Les outils avancés adaptés sont mis à disposition des utilisateurs techniques.	Les fondamentaux du modèle opérationnel sont posés (rôles, responsabilités, funding). Ils sont effectifs sur un périmètre partiel. - Les instances permettent de prendre des décisions aux bons niveaux de manière fluide; une politique RH dédiée permet le recrutement et la formation des principaux rôles data	Les processus de cadrage des produits et de pilotage agile des projets permettent à certains usages de passer en production avec difficulté	Une entité promeut des standards qui s'appliquent encore petit à petit dans l'organisation et des piliers sont en cours de mise en place (data catalog, data quality, data lineage...)	L'organisation reconnaît l'importance de la data mais l'enthousiasme peine à se concrétiser dans les faits - Les métiers sont suffisamment acculturés pour identifier des cas d'usages pertinents et porter le product ownership	La norme RGPD est mise en œuvre sur l'ensemble du patrimoine de données, des objectifs et une feuille de route IA de confiance existent
Emergent	Une feuille de route partielle fournie une visibilité sur les projets et chantiers en cours et à venir à court et moyen terme	Des cas d'usages sont lancés de façon éparse dans l'organisation avec peu de pré-qualification	Les ressources data permettent de couvrir partiellement les besoins techniques, scientifiques et d'assurer un pilotage	Des plateformes sont déployées. Elles disposent d'environnements de développement et de production - Des outils sont administrés au niveau groupe	Une ou des entités de type Datalab, Datafab, Data Office existent. Le modèle d'organisation est encore expérimental	Les cas d'usages data en sont encore à un stade expérimental sans processus d'industrialisation	Le patrimoine de données est partiellement cartographié mais n'est pas sous-contrôle (rôles, documentation, qualité) - Les questions d'accessibilité, de sécurité/compliance et de qualité sont gérées de façon ad hoc selon les usages	Les sujets data sont connus mais considérés comme contraignants, secondaires ou hors de portée	Les procédures RGPD existent mais ne sont pas appliquées à l'ensemble du patrimoine et la dimension éthique reste non considérée
Embryonnaire	Pas d'ambition formalisée ni de transformation data en tant que telle	Peu ou pas de besoins/opportunités identifiés sur la data	Pas ou peu de ressources présentes	La donnée est silotée dans les systèmes applicatifs ou dispersée dans des documents - Les outils sont éparpillés et «actionnés» par les utilisateurs sans rationalisation/administration au niveau groupe	Les sujets data sont portés de manière désordonnée	Les cas d'usages data ne dépassent pas le stade d'idéation ou de cadrage	Le patrimoine de données est peu/pas connu (cartographié) La donnée est silotée par système. L'accessibilité, la sécurité/compliance et la qualité ne sont pas pilotées -	Les opportunités, enjeux et besoins propres à la data sont globalement méconnus	Peu ou pas de préoccupations pour les aspects réglementaires et éthiques de l'utilisation des données - L'ensemble de mon patrimoine n'est pas couvert par les procédures RGPD

Cases pleines : niveau moyen du marché français en 2023.

Pendant des années, l'enjeu des Directions data fut de réussir à industrialiser des cas d'usages data & IA puis à fluidifier cette industrialisation. Ces deux dernières années, le déploiement des pratiques MLOps, les investissements dans les infrastructures technologiques et la montée en maturité progressive des organisations et des compétences ont permis l'atteinte d'un seuil optimal pour les entreprises ayant investi dans leur transformation data depuis plusieurs années.

L'irruption soudaine du tsunami IA générative va constituer un « stress test » pour ces organisations data dans ces prochains mois. En effet, passée la vague actuelle de POC et expérimentations en tout genre, 2024 devra être l'année du passage en production de ces nouveaux usages et du déploiement d'outils dans toute l'entreprise embarquant nativement cette technologie.

En ayant remis la data sur le devant de la scène, la capacité des CDO à piloter ces enjeux sera d'autant plus attendue par le top management.

Les organisations qui ont déjà réussi à fluidifier le passage en production des cas d'usages IA pourront s'appuyer sur leur fondamentaux tout en adaptant leur modèle opérationnel et leur stack technologique aux nouveaux enjeux de l'IA générative.

Mais pour devenir un moteur de la transformation de l'entreprise, l'enjeu du CDO est également de sortir de son positionnement trop souvent unique de DO-eur pour adopter une offre de service d'Enabler auprès des métiers et de la DSI, ainsi qu'une posture de stratège auprès du top management.

Pour cela, il faudra relever 3 défis que l'IA générative va naturellement accélérer : s'inscrire dans les enjeux stratégiques des entreprises, piloter ces usages par la valeur et organiser le bon niveau d'hybridation de son organisation entre central et local via les grands principes du Data Mesh.

Enfin, l'IA générative ne sera pas le seul blockbuster à l'agenda des directions data en 2024. La publication de l'AI Act Européen va lancer véritablement la course de mise en conformité vis-à-vis de cette nouvelle réglementation. Cela imposera d'anticiper les nécessaires plans de remédiation et de conduite du changement, mais ce sera aussi l'opportunité, pour les plus ambitieux, de développer une plus grande confiance dans leurs IA et ainsi d'accélérer leur adoption au sein de l'entreprise et de la Société.



QUI SOMMES-NOUS ?

L'équipe à l'initiative du baromètre



Jonathan CASSAIGNE

jonathan.cassaigne@capgemini.com

Jonathan est Associé et Directeur des expertises de Quantmetry. Il anime notre pôle Data Strategy et accompagne nos clients dans la définition de leur ambition de transformation par la data et la mise en place de leur modèle opérationnel. Il dispose de plus de 15 ans d'expérience dans le conseil en stratégie digitale et data.



Matthieu VANNIER

matthieu.vannier@capgemini.com

Matthieu est Manager de l'expertise Data Strategy et accompagne les dirigeants dans la compréhension de leurs enjeux business vis-à-vis de la data et de l'intelligence artificielle puis les aide à «transformer l'essai» en élaborant les bons programmes de transformation sur l'ensemble des dimensions data : technologiques, organisationnelles et humaines.

Avec la collaboration de l'équipe commerciale Quantmetry dans la préparation et la conduite des entretiens, ainsi que de Claire Monjoin pour l'édition du présent rapport.



Clément LAVERRIERE

clement.laverriere@capgemini.com

Clément est Consultant Senior au sein de la practice Data Strategy. Il accompagne nos clients dans la définition de leurs stratégies data & IA, leur gouvernance et leur organisation. Il intervient également dans leur mise en œuvre opérationnelle.



Vlad FLAMIND

vlad.flamind@capgemini.com

Fort de 10 ans d'expérience dans le conseil en stratégie et organisation, Vlad accompagne nos clients dans la construction de leur modèle opérationnel et dans la mise en œuvre concrète de programmes de Data Management et Data Gouvernance. Il intervient aussi sur le cadrage et le pilotage d'usages data et IA au service des métiers.



Pierre DEMEULEMEESTER

pierre.demeulemeester@capgemini.com

Pierre est vice-président chez Capgemini Invent et dirige la pratique Capgemini Invent Data. Il a mené de nombreux projets de définition de vision stratégique, de conception de modèle opérationnel, d'évaluation de cas d'usage dans le domaine de la gestion des données dans tous les secteurs : industrie, services publics, santé, services, etc. Il a plus de 13 ans d'expérience dans le conseil en stratégie Data.



Baptiste SMAGGHE

baptiste.smaghe@capgemini.com

Baptiste est Senior manager dans l'entité Data Stratégie de Capgemini Invent. Il a mené de nombreux projets de stratégie Data, AI et Gen-AI et piloté des portefeuilles de cas d'usages IA. Il a construit un benchmark Data & IA auprès d'entreprises du CAC40 afin d'évaluer la maturité des organisations et identifier les meilleures pratiques sur le marché.

Merci aux entreprises et à leur dirigeant data pour la confiance qu'ils nous ont accordée, ainsi que pour la qualité de nos échanges

AG2R La Mondiale

April

ArcelorMittal France

Arkea

BNP PARIBAS SA

BPCE

CCR

Chanel

Danone

Décathlon

Dior

EDF R&D

EGIS

ENGIE

Eramet

Eureden

Groupama

Groupe ETAM

GRTgaz

GSF

L'Oreal

La Banque Postale

LCL

LSDH

Manutan

MATMUT

MBDA

Michelin

Ministère de l'Intérieur

Ministère de la Culture

Modulaire Group (ex Algeco)

Orange

Orano Corp

Plastic Omnium

Pôle Emploi

Région Ile-de-France

SAFRAN ANALYTICS

Saint Gobain

Santeclair

Schneider

SCOR

Sodexo

Setec

Shift Technology

SNCF Matériel

Technip Energies

Terrena

Veolia

Ynsect

Quantmetry, part of Capgemini Invent, est un cabinet de conseil spécialisé dans l'intelligence artificielle de pointe. Il est animé par la volonté d'offrir une gouvernance des données de très haut standard et des solutions d'intelligence artificielle à l'état de l'art.

Lauréat du consortium "Confiance.ai", Quantmetry et son laboratoire de R&D contribuent activement au développement de l'IA de confiance en France.

Il s'appuie sur des méthodes et des outils capables de mesurer l'incertitude liée à l'IA, de manière maîtrisée et transparente.

Au sein de Capgemini Invent, Quantmetry est le cœur de réacteur pour la transformation numérique et le développement de produits et services innovants à fort impact, reposant sur l'IA de confiance.

Capgemini  invent

 Quantmetry
Part of Capgemini Invent