



LIVRE BLANC

BUSINESS INTELLIGENCE

2024

GRANDES TENDANCES



DIGDASH

SOMMAIRE

Introduction	3
La gouvernance des données s'impose comme la priorité numéro 1	4
L'intelligence artificielle et le self-service BI convergent	7
L'IA générative se met au service du requêtage en langage naturel	8
Le data storytelling devient incontournable	9
L'intégration de l'open data	10
L'hybridation des architectures data se poursuit	11
Les alternatives aux data lakes gagnent du terrain	13
L'analytique embarquée est partout	14
Le déploiement des solutions BI est toujours plus rapide et agile	15
Conclusion	17

INTRODUCTION

Année après année, la Business Intelligence ne cesse d'évoluer sous l'effet des dernières avancées technologiques, mais aussi de l'émergence de nouveaux besoins au sein des entreprises.

2024 ne fait pas exception et de grandes forces transformatrices sont aujourd'hui à l'œuvre pour façonner le paysage de la BI, et ce dans de nombreux domaines, de la gouvernance à l'analyse des données, en passant par les architectures data.

Alors que les solutions de Business Intelligence jouent un rôle toujours plus important pour les entreprises, soucieuses d'extraire des informations pertinentes à partir de vastes ensembles de données, être à l'écoute de ces phénomènes émergents est une véritable nécessité.

Pour vous aider à y voir plus clair dans cet écosystème en pleine mutation, DigDash a sélectionné 9 grandes tendances qui transforment d'ores et déjà le monde de la BI.

LA GOUVERNANCE DES DONNÉES S'IMPOSE COMME LA PRIORITÉ NUMÉRO 1

Au cours des dernières années, le développement rapide de l'intelligence artificielle, du cloud hybride ou encore de l'internet des objets (IoT) a conduit à un véritable boom du Big Data.

Or, [face à des volumes de data toujours plus grands](#), et donc toujours plus complexes à gérer, **l'unification et la gouvernance des données devient une véritable priorité pour les entreprises**, qui doivent face à de nouveaux défis en matière de sécurité, ainsi qu'à des goulots d'étranglement nuisant à la prise de décision.

Cette véritable course à la data governance se décline en trois volets : les fabriques de données, le data lineage et les dictionnaires de données.

Les fabriques de données

Les solutions de **data fabric** (en français, fabrique de données) connaissent une popularité croissante au sein des entreprises, désireuses de réunir des systèmes de données disparates et d'y intégrer davantage de gouvernance, tout en renforçant la sécurité et la confidentialité.

En effet, la mise en place d'une fabrique de données vise à construire une architecture logique et unifiée. Ainsi, elle permet de **connecter l'ensemble des données d'une organisation**, mais aussi d'ajouter facilement de nouvelles sources de data à cet écosystème.

Concrètement, la data fabric automatise certaines tâches liées à la gestion des données, tout en éliminant les silos entre les différents systèmes. Cela se traduit par **des gains d'efficacité non négligeables**, mais aussi par **une amélioration globale de la qualité de la data**.

Par ailleurs, la fabrique de données est un gage de sécurité, car **elle facilite le masquage et le chiffrement des données les plus sensibles**. Elle permet aussi le déploiement de garde-fous supplémentaires au niveau des contrôles d'accès : autrement dit, chaque profil d'utilisateur accède uniquement aux données qui lui sont réservées.

Le data lineage

Le **data lineage** (que l'on pourrait traduire en français par « lignage des données ») peut être défini comme une cartographie permettant de connaître l'origine de n'importe quelle donnée, les différentes étapes qu'elle a suivies au cours du temps, ainsi que les raisons de son évolution.

Concrètement, le data lineage consiste à **documenter la source et la destination d'une donnée**, les différentes transformations qu'elle a subies **tout au long de son cycle de vie**, mais aussi la manière dont elle a transité.

En recueillant des informations relatives aux sources de données, l'entreprise peut résoudre plus facilement d'éventuelles erreurs, mais aussi simplifier les processus de migration ou de mise à jour d'un système.

De manière générale, **une meilleure traçabilité permet d'optimiser la gestion des données au quotidien et leur suivi opérationnel**. À tout moment, le service IT est en mesure de savoir qui a modifié une donnée et comment elle a été mise à jour.

Les dictionnaires de données

Un **dictionnaire de données** est un référentiel de [métadonnées](#) qui renseigne l'utilisateur sur le contexte d'une donnée et qui fournit les informations nécessaires pour l'interpréter.

Il contient différents types d'informations, par exemple :

- Identifiant de la donnée
- Description
- Traduction (dans un contexte multilingue)
- Le type de donnée
- Les relations et les dépendances
- Les règles de calcul utilisées
- Les règles de validation de la qualité des données
- L'emplacement de la source
- Le responsable

Bon à savoir

*La solution de Business Intelligence DigDash dispose d'un **catalogue de données interne**, permettant de documenter les modèles de données et les graphiques. Il s'intègre d'ailleurs avec de puissantes solutions tierces comme **DataGalaxy**, un data catalog agile à 360 degrés, conçu pour explorer les données et améliorer leur gouvernance.*

L'INTELLIGENCE ARTIFICIELLE ET LE SELF-SERVICE BI CONVERGENT

L'intelligence artificielle est sans aucun doute l'une des principales forces transformatrices dans le paysage de la BI en 2024. Non seulement l'IA révolutionne la façon dont les organisations analysent et extraient de la valeur à partir des données, mais elle apporte aussi une rapidité et une précision inédites dans les processus de prise de décision.

À mesure que l'intelligence artificielle se perfectionne, son intégration à la Business Intelligence devient bien plus qu'une simple tendance : c'est un véritable impératif stratégique. Là où l'analyse de données traditionnelle peine à traiter et à interpréter de vastes ensembles pour en tirer une compréhension globale, **l'IA se distingue par sa capacité à découvrir des modèles cachés dans la data.**

Une spécificité d'autant plus appréciable qu'elle peut être mise au service de tous les utilisateurs, y compris les non-spécialistes. En effet, **la convergence entre l'intelligence artificielle et l'analyse en libre-service** est en passe de transformer la Business Intelligence.

Le self-service BI, qui permet aux utilisateurs d'interagir directement avec la data, est aujourd'hui un enjeu majeur pour améliorer l'autonomie des collaborateurs, tout en les aidant à accéder immédiatement aux informations les plus pertinentes.

Or, l'IA ne fait qu'accentuer la démocratisation de la data et l'accès à des informations à la demande. Des atouts qui se traduisent par un véritable avantage concurrentiel pour les entreprises capables de tirer parti de cette convergence entre intelligence artificielle et self-service.

L'IA GÉNÉRATIVE SE MET AU SERVICE DU REQUÊTAGE EN LANGAGE NATUREL

Pendant longtemps, les outils de Business Intelligence ont été confrontés à un problème de taille : certaines informations cruciales restaient inaccessibles pour les utilisateurs non experts, dissimulées derrière des interfaces complexes et des reportings indigestes.

Heureusement, les choses ont bien changé avec l'apparition du **traitement du langage naturel (NLP)** et des interfaces conversationnelles. Ces technologies permettent aujourd'hui à n'importe quel utilisateur d'interroger la data et de trouver des informations pertinentes, mais aussi de générer les data visualisations les plus adaptées pour chaque requête.

Toutefois, la Business Intelligence est sur le point de franchir un nouveau cap, en s'appuyant sur les dernières avancées dans le domaine de **l'intelligence artificielle générative**.

Une tendance qui modifie en profondeur la façon dont les organisations interagissent avec leurs données.

En effet, cette « BI générative » transforme la data en une véritable ressource, permettant de **fournir des informations personnalisées à chaque utilisateur**, dans une démarche à la fois prédictive et proactive.

Mieux encore, **elle convertit les données brutes en récits attrayants**, directement actionnables pour la prise de décision. Ainsi, chaque entreprise peut tirer le meilleur parti de sa data afin de répondre à ses défis et à ses objectifs propres.

LE DATA STORYTELLING DEVIENT INCONTOURNABLE

C'est une tendance qui se maintient d'année en année : le data storytelling joue un rôle de plus en plus important dans la Business Intelligence. Le cabinet Gartner prévoit même que d'ici à 2025, le storytelling sera la méthode d'analyse de données la plus courante et que [75 % des récits seront automatiquement générés via des techniques d'analytique augmentée](#).

Fournir des informations qualitatives, à la bonne personne et au bon moment, ne suffit plus : désormais, la BI va plus loin et transforme les données complexes en des récits captivants et compréhensibles.

Un enjeu majeur dans un monde ultra-connecté où les décisions doivent être prises dans des délais toujours plus courts et où la data est omniprésente au quotidien.

C'est là qu'intervient le data storytelling, ou l'art de raconter des histoires à partir des données. C'est un fait : les récits ont un véritable impact émotionnel sur les individus, font mieux ressortir les enjeux et incitent davantage à l'action. Ainsi, **les données sont plus compréhensibles, plus faciles à mémoriser, mais aussi plus exploitables par les décideurs.**

Toutefois, le storytelling ne se limite pas à l'ajout d'éléments narratifs dans la data. Idéalement, des tableaux de bord dynamiques et des récits doivent être introduits de manière fluide dans les outils utilisés au quotidien par les collaborateurs.

L'INTÉGRATION DE L'OPEN DATA

Enrichir vos données grâce à l'open source sera bientôt possible dans les outils de BI.

Qu'est-ce que l'open data ?

De nombreux **services publics**, instituts de recherche, universités et entreprises du monde entier ouvrent **l'accès à des données**, qui deviennent **publiques et disponibles gratuitement.**

Cette pratique est largement **démocratisée**, mais ces données ne sont que très **peu exploitées** à l'heure actuelle, et leur usage est pour l'instant **limité à un petit public d'experts**.

De plus, il faut être en mesure de **fiabiliser et d'uniformiser** ces données avant traitement, ce qui complique encore davantage leur exploitation.

En **2024**, les **outils de BI** s'emparent de ces données afin d'en **démocratiser l'accès et l'usage** au plus grand nombre.

Par exemple, **DigDash** travaille actuellement au développement d'une **interface conversationnelle** qui, grâce à l'IA, permettra aux utilisateurs de **poser des questions en langage naturel** à l'outil. Bien sûr, cette fonctionnalité, seule, existe déjà. L'innovation réside dans sa capacité à se connecter à des **données ouvertes** (open data) et à jour, afin de les **recroiser avec les données existantes** de l'entreprise, toujours de manière simple et non technique.

L'objectif : **élargir les capacités d'analyse**, en améliorer la pertinence et proposer de nouveaux axes de réflexion pour accompagner la prise de décision.

LE CLOUD FAIT PLACE AUX STRATÉGIES HYBRIDES

Dans le domaine de la Business Intelligence comme ailleurs, la volonté de migrer vers le cloud est palpable au sein des organisations. Néanmoins, cette démarche ne fait pas toujours l'unanimité, car elle soulève des questions essentielles :

- Quel fournisseur cloud choisir ?
- Faut-il opter pour un cloud public, privé ou hybride ?
- Quelles données et applications déployer sur cette plateforme ?
- Quelles sont les garanties en matière de sécurité et de confidentialité des données ? Quid de la souveraineté numérique ?

Des interrogations légitimes, alors que le partage des données dans le cloud est au cœur de nombreux débats éthiques, mais aussi législatifs. D'ailleurs, il semblerait que **le passage au 100 % cloud n'est pas encore d'actualité**.

En effet, en 2024, l'hybridation des architectures data se poursuit, [à mi-chemin entre le cloud et les infrastructures on-premise](#). En effet, ce modèle hybride présente plusieurs avantages indéniables pour les entreprises.

D'abord, **sa flexibilité de déploiement et sa scalabilité permettent de répondre de manière adéquate aux besoins des métiers**. De plus, il est très adapté à certains types de projets, comme des migrations planifiées sur plusieurs années.

Il permet aussi de **capitaliser sur les lourds investissements réalisés en interne** par certaines organisations, comme la création de data lakes. Enfin, dans certains cas, une architecture hybride est même incontournable pour garantir la conformité réglementaire de l'entreprise.

En contrepartie, l'hybridation soulève de nouveaux problèmes en matière de performance, de gouvernance et de maîtrise des coûts. Les organisations, qui utilisent de plus en plus de solutions SaaS, tendent à stocker des données dans le cloud, puis à les transférer et à les synchroniser, que ce soit vers une infrastructure on-premise ou vers un autre hébergement cloud.

Or, la multiplication de ces flux d'entrée et de sortie du cloud pourrait engendrer d'importants surcoûts.

LES ALTERNATIVES AUX DATA LAKES GAGNENT DU TERRAIN

Les [data lakes](#), ces emplacements de stockage contenant des données structurées, semi-structurées ou non structurées, sans aucune organisation ou hiérarchie particulière, ont le vent en poupe depuis plusieurs années.

En effet, ils permettent de conserver tout le potentiel des données d'origine (qui ne sont structurées qu'au dernier moment), mais aussi d'exploiter la data pour de nombreux cas d'usage.

Cependant, l'engouement pour les lacs de données devrait ralentir en 2024, au profit d'alternatives moins complexes. À commencer par **l'urbanisation de la data**, qui consiste à dé-siloter les données à l'aide de différents pipelines d'alimentation. Cette approche présente plusieurs avantages, puisqu'elle **limite les risques de déduplication des données**, tout en évitant certains processus de synchronisation laborieux.

Par ailleurs, les entreprises réalisent qu'il est difficile, voire impossible, de centraliser toute la data en un seul et même endroit. Elles se tournent donc de plus en plus vers le **data mesh**, une architecture décentralisée qui organise les données par domaine d'activité spécifique : par exemple, le marketing, les ventes ou le service client.

Le data mesh répond également à plusieurs problématiques récentes, liées à l'émergence des microservices, de l'architecture basée sur les événements (EDA) ou encore de l'approche CI/CD (« Continuous Integration/Continuous Delivery »). De nouvelles pratiques qui supposent **un accès libre aux jeux de données**, sans modèle de data rigide ni technologies contraignantes.

Au final, **le data mesh permet aux utilisateurs de gagner en autonomie**, que ce soit pour la production ou la consommation des données. Ceci dit, nous sommes encore bien loin de la fin des data lakes, qui conservent une importance indéniable pour les entreprises. En revanche, une tendance croissante en 2024 concerne la mise en place de lacs de données on-premise, au détriment du 100 % cloud.

L'ANALYTIQUE EMBARQUÉE EST PARTOUT

Tandis que les outils de Business Intelligence deviennent de plus en plus robustes et performants, leurs capacités d'intégration se perfectionnent également. L'essor de [l'analytique embarquée](#) illustre parfaitement ce phénomène.

Grâce à l'embedded analytics, qui permet de **greffer des capacités d'analyse de données dans un large éventail d'applications** utilisées quotidiennement par les entreprises, la data se démocratise et se veut plus accessible que jamais. Elle permet en effet de rationaliser le processus décisionnel dans l'ensemble de l'organisation, quels que soient le métier et le niveau hiérarchique. Les utilisateurs bénéficient d'un accès fluide à des informations pertinentes, fournies en temps réel.

Ainsi, le développement de l'embedded analytics est non seulement une tendance forte, mais aussi **un élément clé pour optimiser les processus liés à la Business Intelligence dans leur globalité**. En plus d'accélérer la prise de décision, l'analyse intégrée favorise l'efficacité et la flexibilité des organisations, car les données font partie intégrante de leurs flux de travail quotidiens.

LE DÉPLOIEMENT DES SOLUTIONS BI EST TOUJOURS PLUS RAPIDE ET AGILE

En matière de déploiement et d'infrastructure data, les approches traditionnelles sont souvent synonyme de lenteur et de rigidité, nuisant à la bonne exploitation des données. Cependant, un changement de paradigme est en cours, poussé par le développement des **pratiques d'intégration et de déploiement continus (CI/CD)**.

En bref, l'approche CI/CD garantit une automatisation et une surveillance constantes tout au long du cycle de vie des applications : intégration, test, distribution, déploiement...

Les **pipelines CI/CD**, utilisés pour améliorer la distribution de logiciels en s'appuyant notamment sur la méthode DevOps, s'imposent donc naturellement dans le domaine de la Business Intelligence, avec des promesses alléchantes. Parmi elles, **une meilleure gestion des flux de données et des modèles d'apprentissage automatiques** (machine learning), mais aussi un important gain de productivité.

En effet, **cette approche libère l'agilité des équipes chargées des données et de l'analyse** grâce au développement itératif. Concrètement, le CI/CD permet une expérimentation rapide, un développement accéléré des modèles ou encore une correction efficace des bugs. Finies les lenteurs, place à la flexibilité ! De quoi aider les organisations à s'adapter aux exigences de leur marché, en constante évolution, ce qui leur confère un avantage concurrentiel indéniable.

Par ailleurs, l'explosion des volumes de données traités nécessite une utilisation plus efficace des ressources. Justement, **les pratiques CI/CD facilitent le déploiement et la scalabilité des outils**, permettant aux entreprises de traiter de grands ensembles de données, d'effectuer des analyses complexes et de fournir des informations à grande échelle.

Pour ce faire, il est possible de s'appuyer sur différentes technologies, notamment :

- **Kubernetes** : une plate-forme conçue pour automatiser le déploiement, la montée en charge et la mise en œuvre de conteneurs d'application sur des grappes de serveurs.
- **Ansible** : un logiciel de gestion des configurations qui automatise le déploiement des applications et la livraison continue des mises à jour. Sa conception permet d'ailleurs d'optimiser la consommation des ressources des serveurs.

CONCLUSION

En 2024, la Business Intelligence poursuit son évolution, portée par des technologies de plus en plus matures (à l'instar de l'intelligence artificielle), ainsi que des usages et des besoins naissants.

Les fonctionnalités telles que la BI en libre-service, le requête en langage naturel ou le data storytelling offrent aujourd'hui des possibilités inédites, tandis que l'amélioration de la data governance reste une priorité absolue.

Mais la Business Intelligence se transforme aussi sur le plan du déploiement et de l'infrastructure, allant vers toujours plus d'agilité, de scalabilité et de vitesse.



DIGDASH