

LA GESTION ACTIVE DES MÉTADONNÉES, LEVIER INCONTOURNABLE DES ÉCOSYSTÈMES DE DONNÉES

Comment valoriser
les données,
face au défi des
nouveaux usages ?



Le partage des données représente **une opportunité de croissance énorme pour les entreprises**. Une gestion active des métadonnées est nécessaire afin d'accompagner ce changement de paradigme et **réussir l'accélération autour des écosystèmes collaboratifs de données**.

Un écosystème collaboratif de données se définit comme **un partenariat entre plusieurs institutions pour partager et gérer des données** afin de créer une nouvelle valeur qui n'aurait pas été possible dans le système précédent, cloisonné. Ce partage s'effectue dans le respect de toutes les réglementations et directives locales et protège la confidentialité et la vie privée, en particulier pour les données des consommateurs finaux ou des citoyens.

Pourtant, en 2021 seule une minorité d'organisations (14% sur l'ensemble des secteurs et pays analysés) sont engagées dans des modèles collaboratifs avec un partage étendu des données. La majorité des organisations participent à des initiatives impliquant un simple partage de données et de faibles niveaux de collaboration¹. **La gestion active des métadonnées fait partie des défis majeurs à relever afin de réussir un partage étendu des données.**

LES MÉTADONNÉES, UNE RESSOURCE CLÉ AU SERVICE DES USAGES

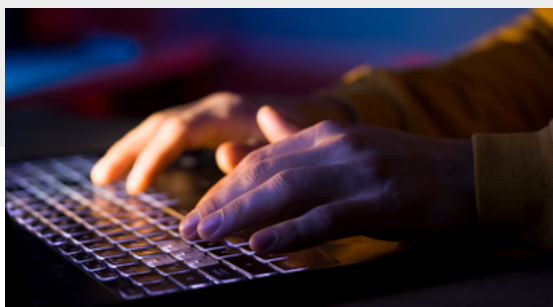
Une compréhension et gestion partielle des données ne favorise pas leur utilisation et partage, en particulier au sein des écosystèmes collaboratifs de données.

Les métadonnées permettent d'apporter du sens aux données et d'en assurer la gestion opérationnelle. Ainsi, les métadonnées peuvent être de nature descriptive ou opérationnelle, avec une capacité plus ou moins grande à évoluer au fil du temps.

NOUS DISTINGUONS DEUX GRANDES CATÉGORIES DE MÉTADONNÉES

MÉTADONNÉES DESCRIPTIVES

Elles permettent de décrire les caractéristiques associées aux données d'un point de vue métier et technique et d'en préciser la propriété. Ces métadonnées une fois produites restent souvent assez identiques tout au long du cycle de vie des données qu'elles décrivent.



MÉTADONNÉES OPÉRATIONNELLES

Elles permettent de préciser notamment les événements liés aux données ainsi que les usages associés. D'une manière intrinsèque, ces métadonnées seront amenées à être modifiées au cours du cycle de vie des données auxquelles elles sont liées.



Prenons l'exemple d'un écosystème de données mis en place pour répondre aux besoins de la taxonomie européenne des activités durables. Une entreprise X exerçant les deux activités économiques « 3.14 Fabrication de produits chimiques organiques de base » et « 4.3 Production d'électricité à partir d'énergie éolienne » se verra associer deux données « Activité économique ».

La définition de cette donnée « Activité économique » restera identique, indépendamment de l'activité en question et de l'entreprise pour laquelle elle est créée. Toutefois, le statut de certification de ces deux données sera probablement amené à être modifié à la suite de leur création (ex. données qui deviennent certifiées à la suite de leur identification dans le rapport extra-financier de l'entreprise X de l'année N-1).

1. Capgemini Research Institute, "Data Sharing Masters : Exploiter les écosystèmes de données pour un avantage concurrentiel clé" Septembre 2021

PASSER D'UNE LOGIQUE DE PRODUCTION À UNE LOGIQUE DE CONSOMMATION

L'accent est encore trop souvent mis sur le volume de métadonnées documentées, cette pratique peut s'avérer peu utile, voir superflu ou dispendieuse. **La valeur des métadonnées se mesure selon la valeur qu'elle apporte aux utilisateurs** et non selon la quantité documentée. De plus, les métadonnées sont souvent réparties à différents endroits en fonction notamment de leur nature ou de leur localisation au sein de la chaîne de valeur des données, ce qui freine leur usage.

Initialement incité par le besoin de répondre aux contraintes réglementaires (RGPD, BCBS 239, Solvabilité II, ...) et par la course à la démocratisation des données, les entreprises s'organisent et investissent pour documenter dans un glossaire métier et dans un dictionnaire technique un grand nombre de métadonnées descriptives, sans systématiquement identifier le consommateur qui en tirera une certaine valeur dans le cadre de ses activités. Les professionnels des données doivent ainsi chercher une aiguille dans une botte de foin afin d'accéder aux métadonnées dont ils ont besoin.

Par ailleurs, le volume des métadonnées opérationnelles produites est exponentiel du fait du volume croissant des données qui a été multiplié par 32 sur la période 2010-2020 (de 2 zettaoctets en 2010 à 64.2 zettaoctets en 2020²). Néanmoins, les métadonnées produites ne sont pas toujours facilement accessibles à ceux qui en ont le plus besoin.

Passer d'une gestion passive à une gestion active des métadonnées signifie replacer le consommateur et l'usage au centre de la production et de la gestion des métadonnées. Dans le cadre des écosystèmes collaboratifs de données, il est d'une part particulièrement essentielle de s'interroger initialement sur les consommateurs potentiels des différentes métadonnées ainsi que leurs besoins spécifiques, et d'autre part de rendre le système scalable et évolutif dès sa conception.

2. Statista, "Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2024", consulté en Octobre 2022



AUTOMATISER LA MAÎTRISE DES MÉTADONNÉES DE BOUT EN BOUT

La gestion active des métadonnées peut être facilitée par l'outillage, qui permet un certain niveau d'automatisation et une meilleure vue d'ensemble. Parmi les fonctionnalités automatisables, certaines sont devenues indispensables à une bonne gestion des métadonnées.

CRÉATION ET MISE À JOUR DES MÉTADONNÉES

RENSEIGNER

Selon les données à documenter, l'outil déterminera les métadonnées essentielles qui serviront un grand nombre d'utilisateurs et qui devront systématiquement être renseignées lors de la création d'une donnée dans l'écosystème collaboratif.

COLLECTER

Certaines métadonnées descriptives seront automatiquement collectées via une exploration des systèmes connectés à l'outil. Pour les métadonnées jugées moins essentielles, le processus de documentation sera inversé. Dès qu'un utilisateur aura besoin d'une métadonnée manquante, il pourra initier une demande en explicitant son besoin. L'expert de cette donnée pourra alors compléter l'information qui servira à présent un besoin clairement identifié.

METTRE À JOUR

Concernant les métadonnées opérationnelles, elles seront mises à jour automatiquement sur la base de certains événements détectés par l'outil. Ces informations seront alors jugées dignes de confiance et pourront alimenter par exemple un tableau de pilotage des données certifiées ou des utilisateurs ayant accès aux données.

PARTAGE ET CONSULTATION DES MÉTADONNÉES

GÉRER

Les utilisateurs internes ou externes souhaitant accéder aux métadonnées sont à catégoriser selon leur rôle, l'expert qualité de données aura un besoin qui divergera du besoin d'un architecte des données, d'un Data Scientist ou encore d'un analyste métier. Ainsi, l'outil permettra de gérer les habilitations et de personnaliser l'affichage des métadonnées en fonction du profil de l'utilisateur et de ses besoins précis. Ce dernier pourra aller à l'essentiel grâce à un tableau de bord allégé et à un moteur de recherche personnalisé.

PRENDRE PART AUX ÉCOSYSTÈMES DE DONNÉES

La gestion active des métadonnées permettra aux organisations de s'impliquer davantage dans des écosystèmes de données collaboratifs. Les organisations impliquées bénéficieront ainsi d'un **modèle financier plus avantageux** par rapport aux organisations moins impliquées. **Les gains potentiels sont estimés à plus de 10% annuellement**, en ajoutant à cela un revenu incrémental, des coûts plus faibles et une productivité plus élevée³.

Au-delà des aspects d'outillage et d'automatisation, prendre le tournant de la gestion active des métadonnées demande un **fort changement culturel et organisationnel** pouvant être porté par une stratégie des données ou un projet de transformation par exemple.

Outre le partage des données, d'autres changements de paradigme (ex. Data Mesh) demanderont une gestion active des métadonnées différente, et ceci afin de prendre en compte les spécificités liées aux nouveaux usages.

3. Capgemini Research Institute, "Data Sharing Masters : Exploiter les écosystèmes de données pour un avantage concurrentiel clé" Septembre 2021

AUTEURS

Arnaud Rolland, [Directeur Data](#)

Nissim Naccache, [Manager Data](#)

Cassandra Sillère, [Consultante Data](#)

À PROPOS DE CAPGEMINI INVENT

Capgemini Invent est la marque d'innovation digitale, de design et de transformation du groupe Capgemini, qui permet aux dirigeants de façonner l'avenir de leurs entreprises. Etablie dans plus de 36 bureaux et 37 studios de création dans le monde, elle comprend une équipe de plus de 10 000 collaborateurs composée d'experts en stratégie, de data scientists, de concepteurs de produits et d'expériences, d'experts en marques et en technologie qui développent de nouveaux services digitaux, produits, expériences et modèles d'affaire pour une croissance durable.

Capgemini Invent fait partie du groupe Capgemini, un leader mondial, responsable et multiculturel, regroupant 340 000 personnes dans plus de 50 pays. Partenaire stratégique des entreprises pour la transformation de leurs activités en tirant profit de toute la puissance de la technologie, le Groupe est guidé au quotidien par sa raison d'être : libérer les énergies humaines par la technologie pour un avenir inclusif et durable. Fort de 55 ans d'expérience et d'une grande expertise des différents secteurs d'activité, Capgemini est reconnu par ses clients pour répondre à l'ensemble de leurs besoins, de la stratégie et du design jusqu'au management des opérations, en tirant parti des innovations dans les domaines en perpétuelle évolution du cloud, de la data, de l'Intelligence Artificielle, de la connectivité, des logiciels, de l'ingénierie digitale et des plateformes. Le Groupe a réalisé un chiffre d'affaires de 18 milliards d'euros en 2021.

*Get The Future You Want**

Plus d'informations sur www.capgemini.com/invent

** Capgemini, le futur que vous voulez*