

Métadonnées pour les solutions **BI et ANALYTIQUES**



RÉSUMÉ GÉNÉRAL :

Alors que les données sont des informations potentielles, les métadonnées constituent un ensemble de renseignements sur ces informations potentielles. Il existe plusieurs catégories distinctes de métadonnées fournissant différents types d'informations (métadonnées descriptives, administratives, structurelles, de langages balisés, d'utilisation), chacune comportant de nombreuses sous-catégories. Un schéma de métadonnées, aussi nommé ensemble d'éléments ou dictionnaire de données, fournit un cadre structuré pour les métadonnées. Dans le cadre de la Business Intelligence (BI), les métadonnées sont essentielles à l'extraction, la transformation et le chargement des données dans un entrepôt, et à l'analyse des données une fois stockées. Ces métadonnées peuvent être exploitées pour assurer la gouvernance, la gestion des risques et la conformité (GRC), l'automatisation des lançements, la refacturation, la simplification des mises à niveau et migrations, l'organisation du contenu à des fins de monitoring, ou encore à l'analyse l'adoption des outils BI. La plupart des organisations tire parti de solutions tierces de gestion de métadonnées, telles que 360Suite, qui centralisent et traitent les métadonnées et les convertissent en données métier. Ce qui distingue 360Suite d'un entrepôt de données c'est sa capacité à extraire uniquement les données pertinentes (métadonnées en lien avec la BI) vers un entrepôt de données fournissant aux utilisateurs autorisés un accès sécurisé, simple, rapide et fiable aux métadonnées pour répondre à leurs questions concernant la Business Intelligence. Au fur et à mesure de l'augmentation exponentielle du volume de données, il devient de plus en plus difficile de découvrir et comprendre le potentiel dans les informations potentielles. Les services de métadonnées représentent une nouvelle source d'informations à laquelle d'autres services ont la possibilité de se connecter. Ceci génère à son tour de nouvelles métadonnées et crée un plus grand besoin de métadonnées sur les métadonnées elles-mêmes (BI on BI) pour que l'information soit disponible aux outils de machine learning, d'intelligence artificielle et d'analyse commerciale.

QU'EST-CE QU'UNE MÉTADONNÉE ?	4
CATÉGORIES DE MÉTADONNÉES	5
Métadonnées descriptives	5
Métadonnées administratives	5
Sous-catégorie : métadonnées techniques	5
Sous-catégories : métadonnées de provenance	5
Sous-catégorie : métadonnées de conservation	5
Sous-catégorie : métadonnées de sécurité	6
Métadonnées structurelles	6
Langages balisés	6
Métadonnées d'utilisation	6
Sous-catégorie : Sortie de données	6
Sous-catégorie : Paradonnées	6
SCHÉMAS DE MÉTADONNÉES	7
Les métadonnées dans l'analyse commerciale (Business Analytics)	8
L'importance des métadonnées BI	9
Gouvernance, Gestion du risque et Conformité	9
Automatisation	9
Refacturation	10
Mises à niveau et migrations	10
Monitoring	11
Adoption	11
A propos de l'absence d'utilisation	12
Collecte de métadonnées BI	12
Services de métadonnées tiers	13
CONCLUSION	14

QU'EST-CE QU'UNE MÉTADONNÉE ?

Nous sommes entourés de métadonnées sans nous en rendre compte la plupart du temps. Mais qu'est-ce qu'une métadonnée ? Le terme «méta» (du grec : «au-delà») a été adopté dans l'usage populaire sous le sens «à propos de la chose elle-même». Par exemple, la métafiction est un style littéraire dans lequel on fait allusion au caractère fictionnel ou littéraire de l'œuvre. En d'autres termes, il s'agit de «fiction à propos de la fiction». De la même manière, un métalangage est un ensemble de mots et symboles évoquant un langage ou une langue, ou encore «un langage traitant du langage». Selon cette logique, les métadonnées sont des «données sur les données», ce qui constitue de fait la définition populaire. Mais quelle est la véritable signification ?

On ne peut définir les métadonnées sans assigner un sens aux données en premier lieu. Dans ce document, le terme «données» fera donc référence aux objets (ex. : faits) pouvant être traités pour en tirer des informations pertinentes. C'est à dire, les données représentent des informations potentielles. Jeffrey Pomerantz, chercheur en sciences de l'information et auteur de l'ouvrage *Metadata*, édité par MIT Press, définit les métadonnées comme «un relevé concernant un objet potentiellement informatif».¹ Pour plus de clarté, prenons l'exemple d'un appel téléphonique;

Dans la plupart des cas, la conversation orale constitue les données (informations potentielles si l'on considère que l'on comprend la langue utilisée et le sens de la conversation). Les métadonnées sont des informations sur ces données : heure et durée de l'appel, état final, numéro de l'appelant, numéros d'identification associés à l'échange (ensemble des équipements interconnectant les lignes téléphoniques), et route empruntée pour l'établissement et l'échange.

L'exemple de l'appel téléphonique illustre bien à quel point les métadonnées sont informatives. En fournissant des détails relatifs à un objet potentiellement informatif (dans le cas présent : un appel téléphonique), les métadonnées exposent le contenu de l'échange (les données) à un niveau surprenant. Par exemple, dans une étude de 2016 portant sur l'analyse des métadonnées téléphoniques de 800 volontaires, les chercheurs en sciences de l'information de l'université de Stanford ont pu corréliser des données privées, telles que des informations de santé, au niveau individuel, sur la seule base des métadonnées.² C'est donc pour cette raison que le grand public a découvert la portée des informations classifiées dévoilées par Edward Snowden, ainsi que les pratiques de l'agence Américaine NSA (National Security Agency), laquelle collectait en masse les métadonnées des citoyens et alliés. Ils ont réalisé que, ce qui était des métadonnées pour la plupart des gens (numéros de téléphone, durée des appels, relais, etc.), était en fait des données pour la NSA. Cette affaire a conduit le Congrès Américain à voter l'USA Freedom Act en 2015 pour interdire la collecte massive des métadonnées téléphoniques (en l'autorisant cependant pour des approches plus ciblées) par les services secrets américains.

Au-delà de cet exemple simplifié, les métadonnées sont bien plus complexes. Dans tous les cas, les métadonnées sont un type de données (informations potentielles). Ainsi, quelle est la différence entre données et métadonnées ? La limite peut être floue, mais celle-ci dépend du but de l'information. Est-elle destinée à fournir du contenu ou un contexte ? Une complexification supplémentaire relative aux métadonnées est leur nombre de catégories et de schémas distincts. Ajoutons à ceci le caractère abstrait des métadonnées, et il devient donc évident que lorsque deux personnes les évoquent, celles-ci ont de fortes chances (ou non) de parler de la même chose.

¹ Pomerantz, Jeffrey. *Metadata*. The MIT Press, 2015, 26.

² Mayer, Jonathan, Patrick Mutchler, and John C. Mitchell. "Evaluating the privacy properties of telephone metadata." *Proceedings of the National Academy of Sciences of the United States of America*. May 16, 2016.

<https://doi.org/10.1073/pnas.1508081113>

CATÉGORIES DE MÉTADONNÉES

La publication National Information Standards Organization (NISO), *Metadata: What Is Metadata, and What Is It For?*, fait référence à quatre catégories distinctes de métadonnées : descriptives, administratives, structurelles, et les langages balisés.³ Jeffrey Pomerantz y ajoute une cinquième catégorie : les métadonnées d'usage/utilisation.

Métadonnées descriptives

Les métadonnées descriptives sont les plus simples. Elles fournissent des informations descriptives des caractéristiques ou attributs d'une ressource donnée pour faciliter la «découverte de données» (découverte et compréhension de celles-ci). Par exemple, les catalogues de données s'appuient sur des métadonnées descriptives pour localiser les éléments individuels.

Métadonnées administratives

Les métadonnées administratives sont des ensembles d'informations relatives à la création d'une ressource donnée et à la gestion de celle-ci tout au long de son cycle de vie. Dans la mesure où de nombreux types de ressources existent, les métadonnées administratives représentent une catégorie conséquente comportant de nombreuses sous-catégories.

Sous-catégorie : métadonnées techniques

Les métadonnées techniques constituent un type de métadonnées administratives apportant des informations sur les caractéristiques d'une ressource. Ces métadonnées sont généralement automatiquement collectées par les programmes à la création ou à la modification des fichiers. Par exemple, les fichiers informatiques contiennent des métadonnées techniques liées au format, à la taille, à l'emplacement, etc.

Sous-catégorie : métadonnées de provenance

Les métadonnées de provenance sont un type apportant des informations sur le contexte du cycle de vie dont les utilisateurs peuvent avoir besoin pour évaluer l'historique d'une ressource, y compris son authenticité et sa validité. Les enregistrements de métadonnées de provenance sont créés, la plupart du temps automatiquement, à chaque création ou modification de ressource, pour décrire les entités et processus impliqués dans sa production, fourniture, ou toute autre influence.⁴

Sous-catégorie : métadonnées de conservation

Les métadonnées de conservation sont un type administratif donnant des détails permettant la conservation des ressources numériques. Leur but est d'assurer l'existence continue des objets numériques, que ceux-ci restent utilisables, et que l'original soit bien distinct des versions dérivées. Par exemple : gestion du cycle de vie (ex. : les mises à niveau ou les migrations se basent sur les métadonnées de conservation pour identifier les formats d'objets, promouvoir la compatibilité et éviter les régressions).

Sous-catégorie : métadonnées de sécurité

Les métadonnées de sécurité appartiennent à la catégorie administrative et se rapportent aux droits de propriété et à la sécurité. Ces métadonnées conservent les informations relatives au droit d'auteur et à la propriété intellectuelle, y compris les droits d'accès, de numérisation, de collecte, ou d'octroi d'accès aux œuvres numériques.⁵ Dans le contexte de la BI, ces métadonnées comprennent les informations de permissions et de sécurité (ex. : qui a accès à quoi).

³ Riley, Jenn. "Understanding Metadata: What Is Metadata and What Is It For?" National Information Standards Organization, 2017.

⁴ "W3C Provenance Incubator Group Wiki," W3C.org, dernière modification le 14 septembre 2011, https://www.w3.org/2005/Incubator/prov/wiki/W3C_Provenance_Incubator_Group_Wiki.

⁵ Marcia Lei Zeng and Jian Qin, *Metadata* (New York: Neal-Schuman, 2008), 64.

Métadonnées structurelles

Les métadonnées structurelles fournissent des informations sur les caractéristiques et la structure des ressources, ainsi que sur les relations entre les objets. Ce type de métadonnées simplifie la navigation au sein du contenu et permet la création de nouvelles informations en rassemblant de plus petits éléments.⁶ Par exemple, il est possible de créer rapidement toutes sortes de documents par type (ex. : factures, bons de commande, rapports d'inventaire) en fonction de la structure. A noter : certains spécialistes en sciences de l'information considèrent les métadonnées structurelles comme un type de données administratives, tandis que d'autres les traitent comme une catégorie distincte.

Langages balisés

Un langage balisé est «un système (à l'image du HTML ou du SGML) de marquage ou balisage d'un document pour mettre en avant sa structure logique (ex. : paragraphes) et pour donner des instructions de disposition sur la page, tout particulièrement dans le cas des transmissions et affichages électroniques.»⁷ En insérant des balises au sein du contenu pour dénoter ses caractéristiques notables, les langages balisés mélangent les métadonnées et le contenu. Dans le contexte de la BI, les langages balisés permettent le marquage d'objets spécifiques (ex. : numéros de carte de crédit indiqués comme sensibles) et rendent les catalogues de données plus lisibles pour les humains et les machines.

Métadonnées d'utilisation

Les métadonnées d'utilisation fournissent des informations sur les actions effectuées par les utilisateurs sur des ressources données. Ce type d'information est traditionnellement considéré comme des données, et non des métadonnées. Hors, lorsque ces données sont considérées comme traitant d'un objet potentiellement informatif, celles-ci deviennent des métadonnées. Comme tout autre type de métadonnées, les métadonnées d'utilisation peuvent révéler une quantité incroyable de détails sur les individus et leurs réseaux. Cependant, contrairement aux autres types de métadonnées, celles-ci ne sont pas activement générées. Elles résultent davantage des conséquences d'autres processus. Dans son ouvrage intitulé Metadata, Jeffrey Pomerantz met en avant l'émergence de deux sous-catégories de métadonnées d'utilisation.

Sous-catégorie : Sortie de données

Les sorties de données décrivent les métadonnées d'utilisation résultant de certaines activités. Par exemple : l'ouverture d'un document génère des métadonnées constituant un enregistrement de cette activité.

Sous-catégorie : Paradonnées

Les paradonnées font référence aux informations d'audit résultant du déclenchement d'une action (ex. : check-out ou check-in de document) fournissant des informations générales (ex. : comment, où, pourquoi et par qui le document a été utilisé).

⁶ Michael Andrews, "Structural Metadata: Key to Structured Content," StoryNeedle.com, 11 octobre, 2017, <https://storyneedle.com/structural-metadata-key-to-structured-content/>.

⁷ Merriam-Webster, s.v. "markup language," accédé le 15 janvier 2019, <https://www.merriam-webster.com/dictionary/markup%20language>.

SCHÉMAS DE MÉTADONNÉES

Un schéma est un «cadre structuré ou un plan».⁸ Ainsi, un schéma de métadonnées est un cadre structuré pour les métadonnées. En d'autres termes, il s'agit d'un langage simple aux règles basiques indiquant ce qu'il est possible de communiquer des données. Un schéma unique peut être conçu pour inclure plusieurs catégories de métadonnées.

“Les schémas fournissent les structures formelles conçues pour identifier l'organisation des connaissances dans une discipline donnée et pour lier cette organisation aux informations relatives à cette discipline au travers de la création d'un système d'informations dédié à l'identification, la découverte et l'utilisation des informations propres à cette discipline.”

*Association for Library Services and Technical Collections
Committee on Cataloging: Description & Access
Task Force on Metadata Rapport final*

Si l'on considère le schéma comme une langue, alors les mots sont appelés éléments. Chaque schéma de métadonnées dispose de son propre ensemble d'éléments (ou sémantique), spécifique au type d'informations à générer. C'est pourquoi les schémas de données sont parfois nommés ensembles d'éléments ou dictionnaires de données.

Lors du développement ou de la sélection d'un schéma de métadonnées, la prise en compte de l'interopérabilité est essentielle. La NISO définit l'interopérabilité comme «la capacité pour plusieurs systèmes utilisant des différents matériels, plateformes, structures de données et interfaces, d'échanger des données en limitant au maximum la perte de contenu et de fonctionnalités». L'interopérabilité des schémas détermine la possibilité d'utilisation des métadonnées par les solutions BI d'Intelligence Artificielle (IA).

Liste non-exhaustive de schémas de métadonnées (par ordre alphabétique) :

- [Categories for the Descriptions of Works of Art \(CDWA\)](#): Schémas de métadonnées pour description des arts, architecture, ou autres œuvres culturelles
- [Creative Commons Rights Expression Language \(CC REL\)](#) : Schéma de métadonnées pour l'expression lisible à la machine des accords de licence de droit d'auteur et autres informations.
- [Dublin Core](#) : Schémas de métadonnées centrés sur les ressources en réseau
- [Exchangeable image file format \(Exif\)](#): Schéma de métadonnées fournissant une structure de marquage des métadonnées incluses dans un fichier image numérique
- [Metadata Encoding and Transmission Standard \(METS\)](#): Schéma pour l'encodage de métadonnées descriptives, administratives et structurelles relatives à des objets dans une bibliothèque numérique
- [Metadata Object Description Schema \(MODS\)](#): Schéma de métadonnées pour les ensembles d'éléments bibliographiques pour tous types d'usages, principalement dans les applications de bibliothèque
- [Preservation Metadata Implementation Strategies \(PREMIS\)](#): Schéma de métadonnées élaboré par la Bibliothèque du Congrès pour servir d'ensemble d'éléments principal pour la conservation d'objets numériques
- [VRA Core](#): Schéma de métadonnées utilisé pour décrire les arts visuels et culturels, comprenant des objets ou événements tels que des peintures, dessins, sculptures, ouvrages architecturaux, photographiques, lithographies, arts décoratifs, arts performatifs, ainsi que les images les documentant

⁸ Merriam-Webster, s.v. “schema,” accédé le 15 janvier 2019, <https://www.merriam-webster.com/dictionary/schema>

Les métadonnées dans l'analyse commerciale (Business Analytics)

L'analyse commerciale dépend des outils d'ETL pour l'Extraction des données depuis une source, la Transformation nécessaire et le Chargement dans un entrepôt de données (emplacement centralisant les données pour optimisation, analyse et reporting). Le processus d'ETL se fonde sur les métadonnées existantes pour trouver les données souhaitées, les organiser et les déplacer vers l'emplacement approprié. Le processus d'ETL génère également de nouvelles métadonnées liées à la transformation.

Les métadonnées demeurent essentielles après le chargement des données dans l'entrepôt. Chaque table, colonne et ligne possède des métadonnées propres (ex. : format, sécurité, date de modification, etc.), rendant possible les activités d'analyse métier telles que l'intégration des données, la transformation, le traitement analytique en ligne (ou OLAP : Online Analytical Processing), et le data mining.⁹

Les outils de Business Intelligence (BI) ou logiciels décisionnels requêtent les entrepôts de données pour générer des rapports utiles à l'analyse métier. Ce processus génère à son tour ses propres métadonnées, parmi lesquelles :

- La connexion utilisée pour requêter l'entrepôt de données ;
- L'identité de la personne à l'origine de la requête ;
- Le(s) document(s) dans lesquels les données sont utilisées ;
- La sécurité associée aux données ;
- la date de création de l'instance/abonnement/publication.
- etc.

Chaque étape du processus (depuis l'ETL vers la base de données, puis vers l'analyse) dépend des métadonnées et chaque action génère de nouvelles métadonnées comprenant des données d'entrée et de sortie de métadonnées. Comme précédemment mentionné, les données d'entrée font référence aux métadonnées d'audit (métadonnées relatives à l'utilisation des données¹⁰ , et les sorties de métadonnées (ou sorties de données) se rapportent à la chaîne de données générée par les actions telles que les activités de recherche. Les deux sont essentielles à l'optimisation du potentiel des métadonnées en Business Intelligence.

⁹ Rahman, Nayem, Jessica Marz and Shameem Akhter. "An ETL Metadata Model for Data Warehousing." Journal of Computing and Information Technology 20, no. 2 (2012). 95-111.

<https://pdfs.semanticscholar.org/5013/bc35ad83319aaac456884f3e994e77ed2ce6.pdf>.

¹⁰ "What is Paradata," IGI-global.com, accédé le 15 janvier 2019,

<https://www.igi-global.com/dictionary/the-importance-of-being-honest/56385>.

L'importance des métadonnées BI

Dans le contexte de la Business Intelligence, ces métadonnées peuvent être exploitées pour répondre à de nombreux enjeux comme par exemple la gouvernance et la conformité des données (GRC), l'automatisation des lançements, la refacturation des coûts de fonctionnement, simplifier les mises à niveau et migrations, organiser le contenu à des fins de monitoring, ou encore pour mesurer l'adoption des outils BI par les utilisateurs finaux.

Voyons plus en détail comment l'exploitation des métadonnées peut permettre de relever ces enjeux.

Gouvernance, Gestion du risque et Conformité

Les métadonnées facilitent les tâches de GRC. Examinons le marquage, qui ajoute des métadonnées aux objets. Lorsque le marquage sert à indiquer les objets soumis aux réglementations (ex. : [SOX](#), [RGPD](#), etc.), les métadonnées facilitent la mise en conformité en permettant le suivi de l'utilisation du contenu sensible ou régulé. Les métadonnées simplifient également la [re-certification des comptes](#), la séparation des tâches, le maintien de la sécurité en détectant les faiblesses au niveau des permissions, l'identification des accès par utilisateur, et en déterminant si des individus profitent effectivement ou non de ces accès. Enfin, les métadonnées permettent le suivi de la création, modification, et suppression des objets et comptes, l'activité utilisateur, les adresses IP, et bien plus encore.

[Découvrez comment utiliser des tags et des flags dans SAP BusinessObjects.](#)

Automatisation

Les métadonnées sous formes de marqueurs (tags/flags) peuvent assister l'automatisation. Par exemple, les objets peuvent être marqués pour répondre à une planification ou à un événement déclencheur (ex. : [#apromouvoir](#)). L'automatisation garantit l'exécution des tâches selon un calendrier prédéfini (ex. : promotion chaque mercredi à 23h) et les métadonnées garantissent la prise en compte des objets appropriés par les tâches. En fonction de leur intérêt, les marqueurs peuvent être préservés ou automatiquement supprimés au cours du processus.

Les métadonnées peuvent également servir à déclencher des actions automatiques plus complexes telles que le nettoyage de la plateforme BI afin d'éviter l'augmentation incontrôlée des volumes de données ou d'objets. Dans ce cas, les métadonnées ont deux rôles. Premièrement elles fournissent une vue détaillée de l'utilisation BI. Deuxièmement, sous forme de marqueurs (tags/flags), elles guident les procédures automatiques. Le marquage automatique des objets pour archivage (ex. : [#archive](#)) est possible lorsqu'un seuil de déclenchement donné est atteint pour ces objets (ex. : non-utilisé depuis X jours). Comme dans l'exemple précédent, l'automatisation assure l'exécution des tâches en fonction de la planification définie, et les métadonnées garantissent l'application de celles-ci sur les objets appropriés.

Enfin, les métadonnées permettent le reporting automatisé de la sécurité, de la re-certification des comptes, du contenu marqué, etc.

Découvrez en davantage sur [l'automatisation SAP BusinessObjects](#).

Refacturation

Dans ce document, la refacturation fait référence à la répartition des coûts de fonctionnement du Système d'Information (SI) aux différents départements métier qui en ont l'usage. Les organisations peuvent déterminer le pourcentage d'utilisation des ressources SI par les départements afin de mesurer les retours sur investissement et de planifier les budgets. Les outils BI (faisant partie intégrante des ressources SI) sont particulièrement difficiles à refacturer en raison de leur complexité. C'est précisément là où les métadonnées trouvent leur utilité. Elles permettent l'analyse de l'utilisation BI, par exemple, pour déterminer que 5% des ressources d'analyse métier ont été associées à un document spécifique, et que 70% de l'activité sur ce document proviennent du département X. De la même manière, les métadonnées permettent de déterminer que le département X a consommé 20% de la mémoire totale associée aux outils BI.

Quelle que soit l'approche analytique, le but de l'analyse des métadonnées pour la refacturation est la compréhension du coût total de propriété par groupe d'utilisateurs ou unité organisationnelle. Ceci simplifie la gestion des technologies métier et la planification des futurs investissements BI.

Découvrez en davantage sur la [refacturation SAP BusinessObjects](#).

Mises à niveau et migrations

Les métadonnées apportent une compréhension de l'utilisation de la BI comprenant :

- Les documents utilisés sur une période donnée
- La dernière utilisation en date d'un document
- Les documents les plus fréquemment utilisés
- Les documents dupliqués
- Les objets sensibles
- etc

Alors que les données obtenues via des questionnaires utilisateur sont subjectives, les métadonnées restent objectives. L'analyse des métadonnées pour mieux comprendre l'utilisation permet l'archivage des contenus superflus, réduisant de fait le temps et les coûts nécessaires aux mises à niveau et aux migrations.

Comme mentionné précédemment, les métadonnées sous formes de marqueurs (tags/flags) peuvent assister l'automatisation. Lors de la préparation d'une mise à niveau ou d'une migration, le contenu non-utilisé peut être marqué pour archivage automatique et le contenu activement utilisé pour être promu grâce aux marqueurs appropriés.

Suite à ces opérations, les métadonnées peuvent également simplifier les tests de régression. Bien que les tests de régression se concentrent majoritairement sur les données, le contrôle des métadonnées peut mener à l'identification de modifications de la structure et de la sécurité des documents.

Monitoring

Les métadonnées autorisent le filtrage du contenu BI en fonction des descriptions (ex. : caractère sensible, sécurité, activité, etc.) afin d'en effectuer le suivi. Ceci s'avère particulièrement utile dans les scénarios suivants :

- Activités de BI externalisées. Les métadonnées relatives aux actions sur les documents (ex. : création, suppression, modification, etc.), durée des sessions par utilisateur, activités de check-in/check-out, etc. permettent aux organisations de générer des rapports contenant les métriques nécessaires au bon suivi de la BI externalisée.
- Monitoring des accords sur les niveaux de service (service-level agreements ou «SLA») par les organisations sur plusieurs solutions BI (ex. : SAP BusinessObjects, Tableau, Power BI, etc.). Les métadonnées relatives à l'activité de la plateforme, aux performances serveur, aux planifications (réussies/en échec) etc, permettent aux responsables décisionnels et métier de comparer les outils BI et d'évaluer leur niveau de conformité avec les SLA.

Adoption

Les métadonnées permettent l'évaluation de l'adoption des outils BI. Le niveau d'adoption BI représente une métrique essentielle pour les responsables BI pour l'optimisation des ressources et pour l'évaluation des points suivants :

- Pourcentage de documents maintenus activement utilisés, et par qui.
- Quels outils BI attribués et activement utilisés, et par qui. (Ce point est particulièrement important dans la mesure où la plupart des organisations investissent dans plusieurs outils BI)

Au-delà de la mesure du taux d'adoption, les métadonnées aident également les responsables BI à mieux appréhender les connexions, identifier les sources, déterminer quand les documents sont utilisés uniquement pour télécharger des données (devenant source de donnée pour d'autres outils d'analyse), localiser les données non-gérées, et bien plus encore.

A propos de l'absence d'utilisation

Le contraire de l'adoption c'est l'absence d'utilisation. L'absence d'utilisation n'est pas mise en évidence par les métadonnées, mais par l'absence de métadonnées. Celle-ci représente ce qu'il reste après identification des usages à l'aide des métadonnées. La connaissance de ce qui n'est pas utilisé est tout aussi importante pour le responsable BI et le métier que la connaissance de ce qui est utilisé.

Par exemple, la comparaison entre des tendances d'usage et l'absence d'utilisation apportent des informations sur le reporting BI et le cycle de vie de la plateforme. Cela guide les décisions liées à la formation, aux investissements, au nettoyage, etc. C'est pourquoi l'absence d'utilisation est un concept clé de l'analyse métier.

- GRC : Dans les cas de non-utilisation, il est possible d'identifier les utilisateurs n'ouvrant jamais de session, ainsi que les données sensibles non-exploitées.
- Automatisation : L'absence d'utilisation peut déclencher des actions ou un reporting automatiques.
- Mise à niveau et migration : L'absence d'utilisation permet la détection d'objets et documents pouvant être archivés avant une mise à niveau ou une migration.
- Adoption : L'absence d'utilisation est l'opposé de l'adoption et identifie les objets et documents non-utilisés ou les utilisateurs inactifs.

A noter : l'utilisation et l'absence d'utilisation sont tridimensionnelles, c'est à dire basées sur l'existence (date de création de l'objet, actions sur celui-ci), ou le manque de métadonnées auxquelles s'ajoute une dimension temporelle.

En raison de l'augmentation des volumes de métadonnées, des limites à la quantité et à la durée de conservation de celles-ci sont mises en place. Les possibilités d'analyse du taux d'utilisation disparaissent avec la perte des métadonnées. C'est pour cela que le suivi continu de l'utilisation est important, de même que l'adoption de solutions d'exploitation des métadonnées telles que 360Suite pour la capture et la rétention des métadonnées d'audit.

Collecte de métadonnées BI

Les solutions de gestion des métadonnées comprennent des registres, référentiels et services de développement et production.¹¹ Les services de métadonnées collectent les informations depuis des outils d'ETL, de base de données, et outils BI, puis les stockent dans un entrepôt de données pour les organiser à l'aide d'une couche sémantique.

Les entrepôts de données apportent plusieurs avantages. Premièrement, ils simplifient l'accès aux données. Deuxièmement, ils réduisent l'impact sur les systèmes de production actifs en chargeant les données pendant les heures creuses, le plus souvent en mode delta. Troisièmement, ils permettent aux responsables BI d'effectuer des analyses d'impact (pour mieux comprendre la portée des modifications d'objets sur les documents). Quatrièmement, ils leur permettent de suivre le cycle de vie des objets depuis l'ETL jusqu'au reporting.

¹¹ Zeng, Metadata , 212.

Services de métadonnées tiers

La plupart des organisations tire parti de solutions tierces de services de métadonnées, telles que 360Suite, qui centralisent et traitent les métadonnées et les convertissent en données métier. De manière intéressante, des métadonnées d'adoption sont à leur tour générées par les services leur faisant appel.

Les catalogues de données, parfois nommés «glossaires métier interrogeables», constituent le plus souvent la partie émergée des solutions de services de métadonnées. Les catalogues de données étant en libre service, il est donc possible pour les utilisateurs du SI et métier d'effectuer des recherches de données, retracer l'origine de celles-ci, ajouter des commentaires aux objets et marquer le contenu.

Les solutions de gestion des métadonnées, comme 360Suite, aident à retrouver les informations, marquer le contenu, documenter, et même déclencher des actions grâce aux métadonnées. Ce qui distingue 360Suite d'un entrepôt de données est le fait que cette solution n'extrait que les données les plus pertinentes (métadonnées en rapport avec les outils BI) vers un magasin de données. Cette approche ciblée simplifie le contrôle d'accès et donne aux utilisateurs autorisés des moyens simples, rapides, fiables et sûrs d'obtenir les métadonnées répondant à leurs questions autour de la BI.

360Suite inclut également des modèles de rapports en libre service construits autour de requêtes simples et complexes, assistant l'interrogation des données et rendant les métadonnées utilisables par le SI (analystes et informaticiens) et par les utilisateurs métier. L'interrogation des métadonnées a bien des avantages, dont les possibilités suivantes :

- Comparaison de photos de métadonnées au fil du temps pour une meilleure compréhension du cycle de vie BI ;
- Mise en évidence de l'utilisation et de l'absence d'utilisation pour une meilleure prise de décision ;
- Mesure et contrôle de la qualité des données grâce aux tests de régression (comparaison des métadonnées historiques, y compris les informations sur la source) ;
- Nettoyage des métadonnées pour limiter les volumes dans le magasin de données ;
- Affichage de l'origine des données ;
- Exécution d'analyses d'impact.

Enfin, 360Suite crée des données qui peuvent être utilisées comme sources pour l'automatisation et l'analyse plus approfondies, pour ajouter des traitements d'Intelligence Artificielle (IA) ex. : IBM Watson). De cette manière, des informations de contexte sur les données sont mises à disposition, permettant à l'IA de tirer parti des métadonnées.

CONCLUSION

Que les données soient ou non le nouvel or noir, il n'y a en tout cas aucun doute quant à leur grande valeur. Au fur et à mesure de l'augmentation exponentielle du volume de données (structurées, semi-structurées, ou non structurées), il devient de plus en plus difficile de découvrir et comprendre le potentiel dans les informations éventuelles. La seule solution de gestion effective des données en masse reste l'utilisation des métadonnées sous forme de couche sémantique pouvant être optimisée pour l'IA. L'exploitation des métadonnées et de l'absence d'utilisation en tant que couche sémantique contribue aux analyses descriptives, prédictives et prescriptives. Les métadonnées apportent une plus grande flexibilité et une meilleure compréhension. Dans les faits, les données sans métadonnées sont virtuellement inexploitable.

L'interopérabilité est la clé du machine learning. Avec les nombres grandissants de sources BI et de connexions, l'interopérabilité des métadonnées est en passe d'imposer un important défi. En conséquence, les services de métadonnées gagneront en importance. Ils représentent une nouvelle source d'informations à laquelle d'autres services ont la possibilité de se connecter. Ceci génère de nouvelles métadonnées et crée un plus grand besoin de métadonnées sur les métadonnées elles-mêmes (BI on BI) pour que l'information soit disponible aux outils de machine learning, d'Intelligence Artificielle et d'analyse commerciale.

L'importance des métadonnées pour la BI et l'analyse métier n'est donc pas à négliger. Ces éléments apportent aux responsables BI et aux informaticiens les détails nécessaires à toute prise de décision informée, dont : la compréhension de l'adoption utilisateur, la comparaison des taux d'utilisation et d'absence d'utilisation, la gestion du risque, l'optimisation des plateformes, l'optimisation de l'attribution des ressources. Les métadonnées sont le pilier commun au service des objectifs SI et métier : l'expérience utilisateur optimale au moindre coût.

360Suite est une suite de **solutions agiles de gouvernance** dédiée pour SAP BusinessObjects, développée par Wiiisdom.

Chez Wiiisdom, nous vous accompagnons pour faire de votre **patrimoine Analytics**, un lieu fiable qui permette de maximiser l'utilisation de vos données et de prendre de **meilleures décisions** au quotidien.

360Suite est un ensemble de solutions permettant d'assurer la qualité, la fiabilité, la performance et l'efficacité de SAP BusinessObjects au travers de méthodologies de **tests**, d'audit, de monitoring, de catalogage et de planifications de tâches. 360Suite s'adresse aux grandes organisations souhaitant se prémunir des **risques** liées aux données, d'**automatiser** leurs opérations, et représente une solution de choix pour tout type de projet de migration.

[DEMANDER UNE DÉMO](#)

Auteurs: Bruno Masek  
Kristen Champagne Gray 

Métadonnées pour les solutions **BI et ANALYTIQUES**



 360 Suite



Visiter www.wiisdom.com/fr/360suite