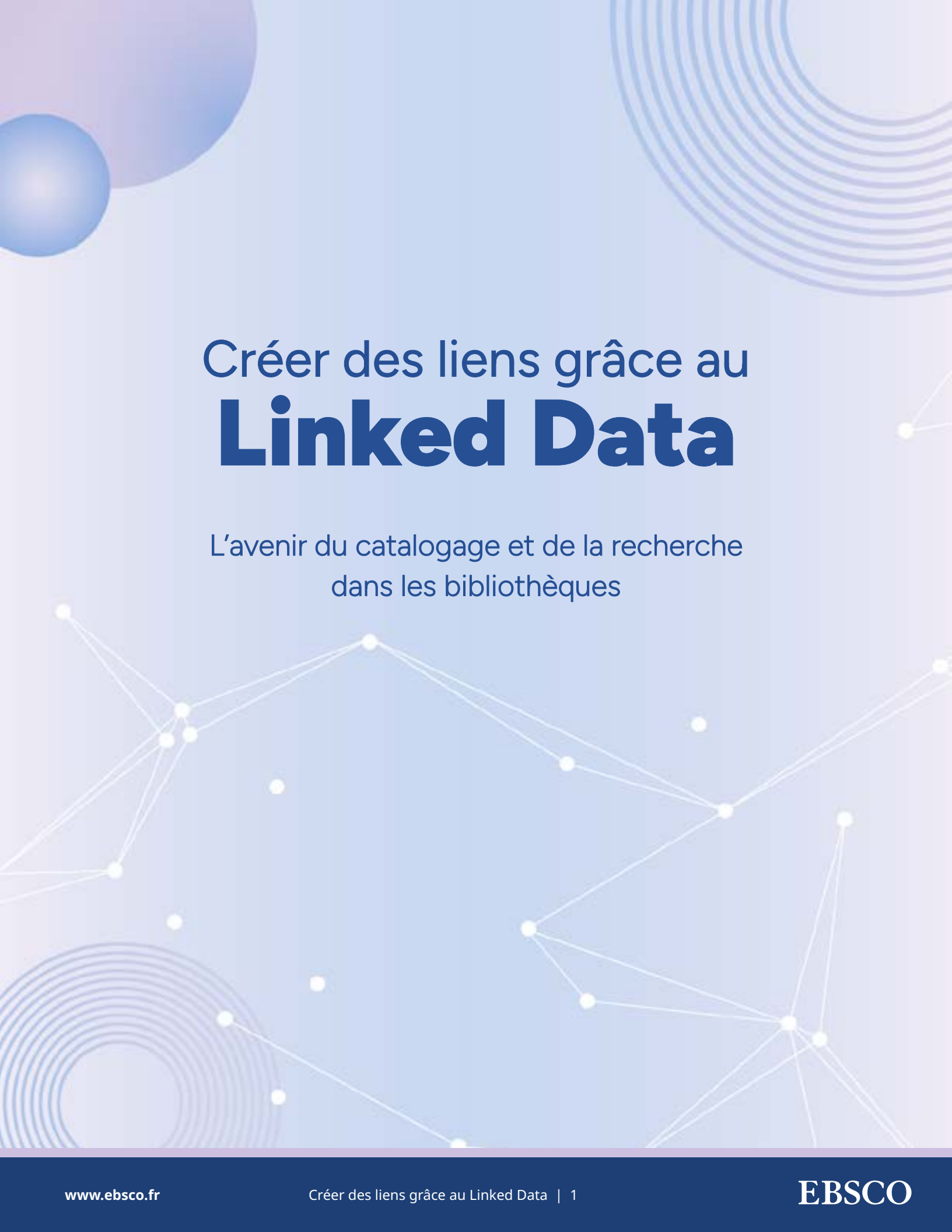




Créer des liens grâce au **Linked Data**

L'avenir du catalogage et de la recherche
dans les bibliothèques

MARC a un problème (ou plutôt trois)...

Développé dans les années 1960 comme format normalisé de catalogage pour les documents imprimés, MARC commence aujourd'hui à montrer ses limites. S'il a prospéré à l'ère des machines à écrire et des catalogues sur fiches, l'essor d'Internet et des technologies numériques a mis en évidence certaines de ses contraintes structurelles.

- ✗ MARC n'est pas **flexible**: il s'agit d'un format à champs fixes, ce qui signifie que chaque champ est associé à un usage précis. Cela complique l'intégration de nouveaux types d'informations ou l'évolution des modes d'organisation des données.
- ✗ MARC manque **d'adaptabilité**: bien que les catalogues en ligne puissent prendre en charge MARC, celui-ci n'est pas réellement conçu pour les environnements numériques actuels et s'intègre difficilement avec d'autres systèmes.
- ✗ MARC n'est pas **interopérable**: principalement pensé comme un format local, il rend le partage des données complexe entre différents systèmes, institutions ou services développés en dehors du secteur des bibliothèques.

Passer au Linked Data

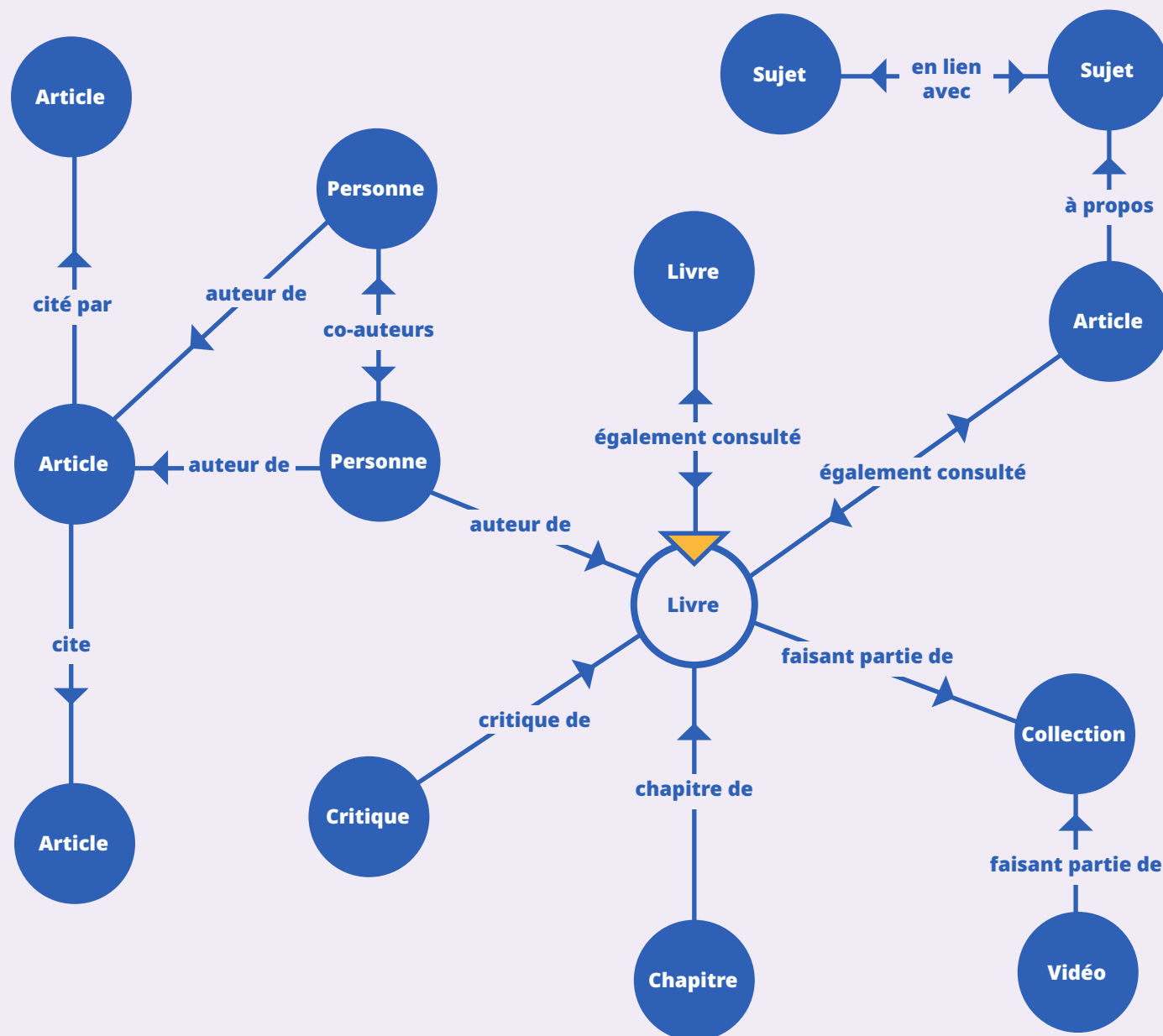
Les bibliothèques évoluent pour s'adapter aux nouvelles façons dont les usagers recherchent et utilisent l'information en ligne, et MARC n'a tout simplement pas été conçu pour le monde numérique. Le Linked Data offre une approche plus pertinente pour organiser et **partager les métadonnées des bibliothèques dans l'environnement actuel**.

En termes simples, le Linked Data permet d'**identifier et de structurer des connexions dynamiques** entre les ressources. Au lieu de conserver les données enfermées dans des catalogues locaux, il aide les bibliothèques à diffuser plus largement leurs connaissances et à les relier à d'autres sources.

Compatible avec les sites web, les moteurs de recherche et d'autres systèmes, il **facilite la création et la découverte de liens** entre des éléments tels que les auteurs, les ouvrages ou les sujets, et permet de rendre ces connexions visibles sur le web.

Prenons un exemple :

Le schéma ci-dessous nous donne un aperçu des connexions rendues possibles par le Linked Data. Suivez ces liens et laissez-vous guider dans un parcours de découverte intuitif entre œuvres, auteurs, sujets et bien plus encore.





Que permet le Linked Data?

Un roman se déroulant dans le Paris du XIXe siècle peut être relié à des cartes historiques, des photographies ou encore à des œuvres contemporaines de la même époque. Ces connexions permettent au lecteur d'explorer le contexte avec davantage de profondeur et de mieux comprendre, par exemple, comment l'architecture de la ville ou le climat politique ont pu influencer un récit de fiction. Le lecteur peut également découvrir des auteurs ayant réellement vécu à Paris à cette période, comme Victor Hugo. Comment son expérience personnelle a-t-elle nourri son écriture ? Et comment celle-ci se compare-t-elle au roman par lequel nous avons commencé ? Cette mise en contexte transforme une simple expérience de lecture en un **parcours multidimensionnel**, à la croisée de la fiction et de la non-fiction.





Que permet le Linked Data?

Imaginons maintenant un chercheur travaillant sur les politiques liées au changement climatique. Il peut découvrir que deux scientifiques spécialisés dans les technologies d'énergies renouvelables ont coécrit des articles avec un troisième expert du stockage d'énergie par batterie, révélant ainsi une possible convergence entre disciplines. Grâce au Linked Data, il peut également suivre la manière dont un article fondateur a été cité dans différentes disciplines telles que l'économie, la sociologie ou l'urbanisme. Cet **ensemble de citations** met en lumière l'évolution des idées initiales, qu'elles aient été approfondies, remises en question ou élargies, et **aide le chercheur à repérer les tendances émergentes** ou les lacunes dans la littérature scientifique.



Pourquoi le Linked Data est-il essentiel ?



Améliorer la découverte et l'accès aux ressources

Le Linked Data enrichit l'expérience de recherche en **reliant les métadonnées des bibliothèques à des jeux de données internes et externes**. Il met en lumière les liens entre auteurs, œuvres, sujets et bien plus encore, permettant aux utilisateurs de découvrir des ressources connexes, au-delà des institutions, des formats et des langues.

Conçu pour le web, le Linked Data rend ces connexions non seulement lisibles par les machines, mais aussi accessibles en ligne. Les collections de votre bibliothèque peuvent ainsi apparaître dans des environnements de recherche élargis, non plus uniquement dans votre catalogue, mais sur l'ensemble du web. Cette visibilité accrue **facilite l'accès à vos ressources et favorise leur utilisation**.



Améliorer la qualité des métadonnées

Les formats de métadonnées traditionnels, comme MARC, sont souvent limités à un contexte local. À l'inverse, le Linked Data est **flexible, dynamique et cohérent à l'échelle internationale**. Il peut être réutilisé et adapté à différentes plateformes et contextes, tandis que les fichiers d'autorité (par exemple pour les auteurs ou les sujets) sont mises à jour en temps réel. Des identifiants partagés, tels que VIAF ou ORCID, garantissent une harmonisation des métadonnées entre institutions.

Les chercheurs peuvent ainsi **s'appuyer sur des métadonnées fiables, à jour et interopérables**, ce qui réduit les doublons et améliore la qualité des notices bibliographiques.

Pourquoi le Linked Data est-il essentiel ?

✓ Interopérabilité entre systèmes

Le Linked Data permet une intégration fluide entre bibliothèques, archives, musées et autres institutions du patrimoine culturel. Il relie différents vocabulaires et ontologies issus de collections variées, allant des manuscrits et jeux de données aux enregistrements audio et aux expositions.

Cette interopérabilité offre aux utilisateurs la possibilité d'**explorer des ressources diversifiées depuis leur plateforme documentaire habituelle**. Pour les chercheurs, cela signifie moins de temps passé à naviguer entre les systèmes et davantage de temps consacré à l'analyse et à l'exploitation des contenus.

✓ Soutien aux initiatives de science ouverte

Le Linked Data s'inscrit dans les principes de la recherche et de la science ouverte en favorisant la transparence dans la structuration et le partage des connaissances. Il encourage l'accès aux métadonnées et aux ressources au-delà des institutions, et renforce la collaboration entre bibliothèques grâce au partage des données.

Les bibliothèques qui participent à des écosystèmes de Linked Data contribuent à **un environnement collectif de connaissance**. Elles bénéficient des métadonnées produites par d'autres tout en enrichissant à leur tour cet ensemble commun. Cette dynamique favorise l'innovation à l'échelle du réseau, qu'il s'agisse de systèmes de recommandation plus performants, d'un meilleur suivi des citations ou d'outils de découverte plus inclusifs.

D'un point de vue technique, comment fonctionne le Linked Data ?

Vocabulaires et structure

Le Linked Data s'appuie sur les standards du Web sémantique pour rendre les données issues de différentes sources compréhensibles, réutilisables et interopérables. Ces vocabulaires permettent de décrire les ressources et leurs relations de façon cohérente :

- ✓ **BIBFRAME** (Bibliographic Framework): Développé par la Library of Congress pour remplacer MARC, il permet de relier et de partager les données bibliographiques entre systèmes.
- ✓ **Bulde**: Vocabulaire sous licence ouverte conçu par EBSCO pour relier et étendre le Linked Data entre communautés incluant le catalogage, la recherche académique et les archives.
- ✓ **Dublin Core**: Ensemble simple et largement utilisé d'éléments de métadonnées pour décrire des ressources web telles que des documents, des images et des jeux de données.
- ✓ **Schema.org**: Vocabulaire collaboratif utilisé sur le web, notamment par Google et d'autres moteurs de recherche, pour décrire personnes, organisations, œuvres et contenus créatifs de manière structurée et lisible par les machines.
- ✓ **SKOS** (Simple Knowledge Organization System): Utilisé pour représenter des vocabulaires contrôlés tels que des thésaurus, des taxonomies et des systèmes de classification.

Tous ces vocabulaires reposent sur le RDF (Resource Description Framework), qui exprime les données sous forme de triplets (sujet, prédicat, objet), permettant d'établir des liens entre des concepts issus de différents jeux de données et plateformes.

D'un point de vue technique, comment fonctionne le Linked Data ?

Gouvernance des données

Le Linked Data contribue à renforcer la fiabilité et la qualité des données grâce à des dispositifs tels que SHACL (Shapes Constraint Language), une norme du W3C utilisée pour valider des données RDF au regard d'un ensemble de contraintes, appelées « shapes ». Il joue un rôle essentiel dans la gouvernance des données en permettant de :

- ✓ Garantir que les données respectent les structures attendues (par exemple, une personne doit avoir un nom et une date de naissance)
- ✓ Détecter les erreurs dès les premières étapes des flux de traitement des données
- ✓ Faire respecter les règles et la cohérence sémantique

En appliquant la norme SHACL, les organisations peuvent automatiser les contrôles qualité, réduire les silos de données et maintenir un haut niveau d'intégrité des données au sein de systèmes distribués.

D'un point de vue technique, comment fonctionne le Linked Data ?

Identifiants uniques

Au cœur du Linked Data se trouve l'utilisation d'URI (Uniform Resource Identifiers) comme identifiants uniques globaux des entités. Ces identifiants permettent :

- ✓ À tout utilisateur d'accéder aux données associées à un URI via un navigateur ou une API
- ✓ La création de liens sur le web ouvert afin d'établir des connexions entre les données
- ✓ La construction de graphes de connaissances dans lesquels chaque nœud représente un concept ou une ressource identifié de manière unique

Par exemple, un URI comme <http://id.loc.gov/authorities/names/n79107741> renvoie vers des données structurées sur Beethoven, notamment les différentes formes de son nom, ses dates de naissance et de décès, sa profession, des correspondances exactes avec d'autres référentiels, les sources, etc. Ces données associées peuvent être réutilisées et reliées à d'autres jeux de données.

À propos des graphes de connaissances...

L'**EBSCO Scholarly Graph** (ESG) est un ensemble de données structuré et interconnecté qui décrit des entités du monde académique telles que des personnes, des publications, des institutions et des thématiques, en s'appuyant sur le Linked Data. Il alimente certaines fonctionnalités des solutions EBSCO qui permettent aux utilisateurs d'**explorer les relations** entre auteurs, d'**identifier l'impact d'une recherche** et de **naviguer plus intuitivement dans les domaines académiques**. Construit à partir de l'agrégation de métadonnées issues de sources fiables comme OpenAlex et Crossref, le Scholarly Graph enrichit ces données par des relations contextuelles et des identifiants normalisés. EBSCO calcule également des informations complémentaires afin d'étendre le graphe.

Cela permet des expériences de découverte plus avancées, comme des pages auteurs dynamiques présentant leurs publications, leurs co-auteurs et les thématiques associées.

[En savoir plus](#)

L'IA au service du Linked Data

EBSCO voit dans l'IA un levier pour enrichir l'expérience de recherche, à condition qu'elle soit utilisée de manière responsable. Combinée à des technologies avancées et à l'expertise humaine, l'IA contribue à améliorer l'ergonomie, à garantir la fiabilité et à maintenir des standards élevés.

Dans cette démarche, nous utilisons l'IA pour renforcer nos capacités en Linked Data. Par exemple, notre Universal Subject Index (USI) s'appuie depuis longtemps sur le Linked Data pour établir des correspondances entre synonymes issus de différents référentiels d'autorité, aidant ainsi les utilisateurs à trouver des ressources pertinentes même sans maîtriser la terminologie exacte. Nous avons mis en place une recherche en langage naturel assistée par l'IA, qui fonctionne en complément de l'USI afin de mieux interpréter les requêtes des utilisateurs et de fournir des résultats encore plus précis.

Par ailleurs, notre fonctionnalité de suggestions de recherche utilise l'IA pour analyser une requête et proposer des pistes alternatives ou connexes, guidant souvent les chercheurs vers une exploration plus approfondie d'un sous-thème spécifique.

L'IA responsable est au cœur de notre engagement en faveur de l'intégrité et de l'innovation dans la recherche.

En savoir plus

Pourquoi le Linked Data est-il important pour les bibliothèques ?

	Sans Linked Data	Avec Linked Data
Découverte et accès	Les chercheurs travaillent plus dur pour trouver du contenu lié à leurs recherches.	Les métadonnées sont reliées entre jeux de données pour mettre en lumière les liens entre auteurs, œuvres et sujets, et améliorer la recherche et la découverte.
Qualité des métadonnées	Les données sont cloisonnées et peuvent nécessiter des mises à jour manuelles.	Des données flexibles et réutilisables, associées à un contrôle d'autorité dynamique, produisent des notices plus cohérentes et actualisées.
Interopérabilité	Les ressources de la bibliothèque ne sont pas reliées à d'autres ressources.	Des passerelles entre vocabulaires et ontologies hétérogènes unifient l'accès aux ressources des bibliothèques, archives et musées.
Science ouverte	Les systèmes fermés restent isolés de l'écosystème global.	Les contributions aux écosystèmes mondiaux de données ouvertes s'inscrivent dans des valeurs de transparence et de collaboration.

Comment le Linked Data est-il intégré aux produits et services d'EBSCO ?

[Cliquez ici](#)