

## Interfaçage d'une ontologie et d'une BDD relationnelle pour la saisie, la corrélation et la visualisation de données de conservation du patrimoine bâti

<b>Opportunités</b>	<ul style="list-style-type: none"> <li>• Développement de compétences en modélisation d'ontologie</li> <li>• Interfaçage de briques technologique dédiée à la gestion de données et des connaissances</li> <li>• Découverte du domaine de la conservation du patrimoine au regard du numérique.</li> </ul>
<b>Lieu de stage</b>	En présentiel au Laboratoire MAP équipe CRAI, à l'ENSA Nancy, France & en distanciel
<b>Encadrants</b>	Tommy MESSAOUDI & Gilles Halin
<b>Indemnités</b>	Gratification minimum réglementaire (3,90 €/heure)
<b>Date</b>	2 jours par semaine en présentiel et le reste en distanciel entre avril et Juillet 2021

### Compétences attendues

**A** Compétences en déploiement de nouvelles bases de données relationnelles, compréhension de la structure de BDD relationnelles existantes, modélisation des connaissances, déploiement d'outil de visualisation de données graphe, mis au point d'un système de requête SQL, SPARQL, et Cypher.

+ Français ou anglais courant

**B** Installation et exploitation de librairie dédiées à la conception d'interface faisant dialoguer les briques permettant de :

- Mettre à jour une ontologie
- Saisir, consulter et extraire de nouvelle données dans la base de données courantes,
- Visualiser des graphes.

### Contexte

Ce travail entre dans le cadre du Projet LINKS (Proposition d'une ontologie de domaine générique dédiée au suivi de la conservation d'églises rupestres peintes). Dans le domaine de la conservation du patrimoine bâti, l'élaboration de système de gestion d'une quantité importante de données hétérogènes, complémentaires et distinctes est devenue un enjeu incontournable. Destinées aux acteurs du patrimoine culturel elles ont pour objectifs de produire et de partager des données et des connaissances spatialisées sur des objets du patrimoine à des fins de restauration et conservation. Le partage de ces données (provenant de bases de données relationnelles distinctes) et connaissances nécessite une structuration informatique ad-hoc formelle facilitant leur exploitation en termes d'acquisition, de recherche et de croisement. Pour cela, un travail conséquent, en phase de finalisation, est de réaliser une ontologie générique décrivant de tels aspects. Le cas d'étude exploité est l'église Qorqor Maryam pour laquelle une description conceptuelle a été formalisée : ses peintures, son état de conservation et sa sauvegarde numérique. L'objectif aujourd'hui est d'exploiter cette ontologie de domaine comme structure pour les données analytiques de description

de tels édifices et de créer un système d'information manipulant ces données instanciant cette ontologie. Bien que les données de description se situent actuellement dans des bases de données de types relationnelles distinctes, l'usage de système d'information structuré par une ontologie permettrait potentiellement de non seulement rapatrier, mais également d'uniformiser ces données au sein d'une référence conceptuelle commune, et par conséquent, mais également de les consulter, de les corréliser et de les visualiser. Des expérimentations ont d'ores et déjà été opérées pour éprouver cette possibilité à l'aide d'une suite de logiciel fournissant des librairies opérationnelles en vue d'être déployé dans des environnements web.

## Objectifs

L'ambition ici est donc de faire la preuve de concept (PoC, Proof of Concept ou étude de faisabilité) d'un prototype appliqué à une portion de l'ontologie générique élaborée avec la prise en charge des données de description distinctes.

## Tâches

- Prise de connaissances de l'aspect technique du **projet LINK** et des enjeux.
- Compréhension approfondie des notions clés : **OBDA** (Ontology-Based Data Access), **SPARQL**, bases de données **graphes**, **Triplestore**.
- Réalisation d'un modèle de processus **BPMN** pour l'élaboration du prototype,
- Réalisation du **tutoriel** pour expérimenter **ONTOP** (Figure 1)
- **Formalisation** d'une portion du modèle conceptuel LINK dans Protégé (Fusion CRM-Core et des extensions utiles importables directement depuis le site)
- Réalisation d'une **base de données relationnelle** simulant une réalité de stockage distant des données d'analyses
- Élaboration des **mapping SQL/SPARQL**
- Mise en place de **requêtes SPARQL** simples inférant sur les données
- **Export** des résultats des requêtes,
- Mis en place du système de **visualisation Neo4J**.
- Réaliser l'ensemble dans un **prototype web** (Figure 2) permettant à la fois : la saisie de nouvelles données à travers des formulaires simples et leurs visualisations sous forme de tableurs et de graphes.

## Équipements de travail ---- Environnement de développement

- Au vu de la situation sanitaire actuelle, il sera nécessaire que les stagiaires aient à disposition leur propre machine pour des questions de praticité.
- L'environnement de développement pour la phase de test sera composé de plusieurs logiciels : **uWamp** (bases de données relationnelles portable), **Protégé** (formalisation de l'ontologie), **ONTOP** (mappings) et **Neo4J** ou **VOWL** (visualisation). L'environnement de développement pour l'interfaçage des différentes briques sera composé, pour certains, des mêmes librairies et pour d'autres, proposées par les éditeurs mentionnés ci-dessus. Le prototype pourra être fait dans **Bootstrap** (négociable). Le langage de programmation pourra être du java (négociable).

# Annexes

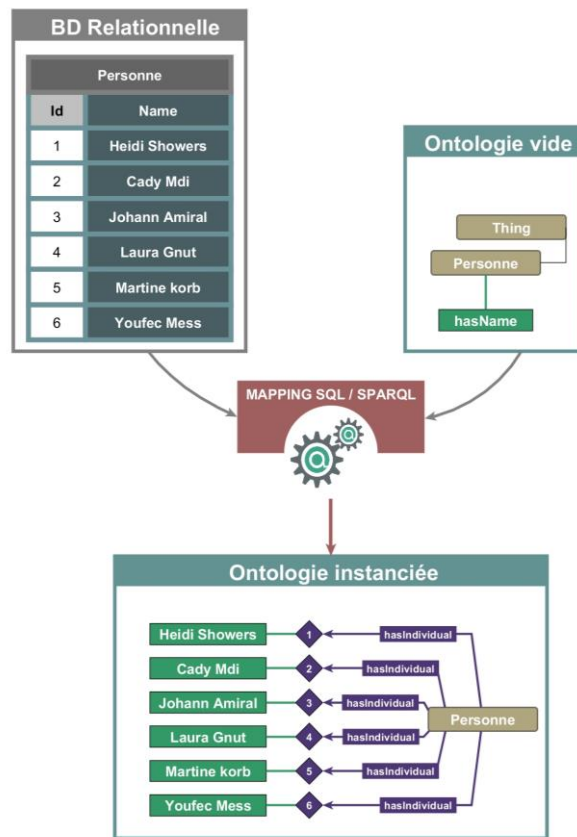


Figure 1: Principe d'ONTOP.

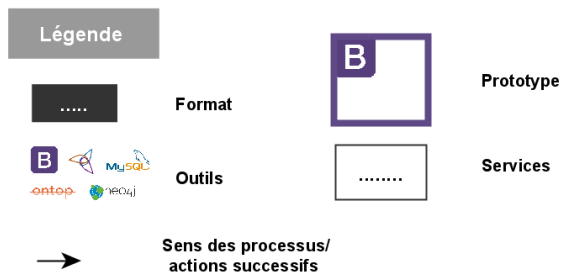
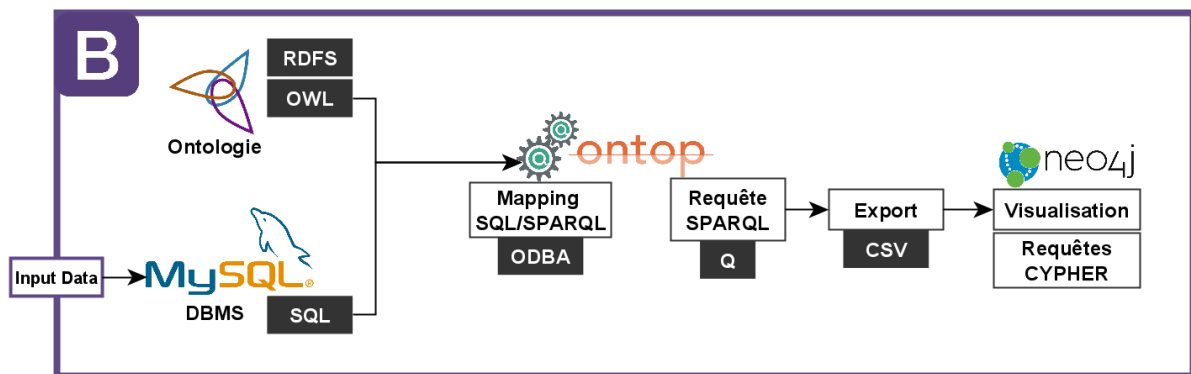


Figure 2: Plan d'implémentation du prototype LINK.