



HAL
open science

Système d'information pour la production de connaissances. L'approche wiki sémantique

Eric Leclercq, Marinette Savonnet

► To cite this version:

Eric Leclercq, Marinette Savonnet. Système d'information pour la production de connaissances. L'approche wiki sémantique. *Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information*, 2012, 17 (3), pp.143-166. 10.3166/isi.17.3.143-166 . hal-00738682

HAL Id: hal-00738682

<https://u-bourgogne.hal.science/hal-00738682>

Submitted on 16 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Système d'Information pour la production de connaissances : l'approche wiki sémantique

Éric Leclercq* — Marinette Savonnet*

* Laboratoire LE2I - UMR CNRS 5158
Université de Bourgogne
9, Avenue Alain Savary
21078 Dijon, France
Prenom.Nom@u-bourgogne.fr

RÉSUMÉ. Dans cet article, nous montrons comment les wikis peuvent être des solutions pertinentes pour construire un corpus numérique dans des domaines caractérisés par une évolution rapide de la connaissance. Nous mettons en exergue les capacités d'adaptation des wikis, nous montrons que les extensions sémantiques développées permettent un meilleur contrôle de la qualité du contenu, améliorent la recherche par la possibilité d'écrire des requêtes complexes et supportent différents types d'utilisation. Nous commençons par une comparaison de la représentation de la connaissance et de son exploitation dans les applications. L'architecture du wiki sémantique, WikiBridge, ainsi que son modèle d'annotations et les mécanismes de contrôle de consistance sont présentés. Une section est consacrée à la validation de WikiBridge avec une application archéologique (le projet CARE) ainsi que la conception de l'ontologie qui représente la connaissance fondamentale nécessaire à cette application.

ABSTRACT. In this paper, we highlight how wikis can be relevant solutions for building applications in domains characterized by a rapid evolution of knowledge. We will point out the capabilities of semantic extension to provide better quality of content, to improve searching, to support complex queries and finally to carry out on different type of users. We detail the architecture of WikiBridge, a semantic wiki, its annotation model and the consistency checking mechanism. A specific section is dedicated to the description of the CARE project and ontology design which is the compulsory foundational knowledge for the application.

MOTS-CLÉS : wiki sémantique, ingénierie ontologique, connaissance évolutive

KEYWORDS: semantic wiki, ontological engineering, evolutive knowledge

1. Introduction

Les plates-formes collaboratives sont des outils essentiels pour les scientifiques. En effet, elles permettent une aide à la formalisation de connaissances et à la publication de résultats ou de jeux de données. Les éléments mis en ligne au travers de ces plates-formes doivent pouvoir être adaptés à différentes catégories d'utilisateurs (experts, visiteurs amateurs, novices, etc.). De plus, les plates-formes doivent pouvoir interagir avec des modules externes permettant par exemple la visualisation de données complexes ou l'analyse spatiale ou statistique.

Linster dans (Linster, 1992) montre que l'interaction entre les experts du domaine, les spécialistes en ingénierie de représentation des connaissances et les outils crée la connaissance. Il a mis en évidence que le processus pour élaborer un système à base de connaissances est un processus itératif de construction de modèles qui comprend : 1) un processus de discussion entre les spécialistes en ingénierie de représentation des connaissances et les experts du domaine et 2) la construction d'un modèle conceptuel c'est-à-dire d'un cadre général et abstrait. À l'issue de ce processus, différents types de modèles sont construits : des modèles conceptuels descriptifs dont l'objectif est de représenter la connaissance du domaine et des modèles exécutables dont l'objectif est d'implémenter l'application.

Cependant, pour atteindre de tels résultats, il est important de respecter des standards partagés par la communauté scientifique. Des standards bien établis existent dans de nombreux domaines comme CIDOC-CRM¹ pour le patrimoine culturel, OBO² pour les domaines biologique et biomédical.

La suite de l'article est organisée selon 5 sections. La section 2 présente une discussion sur la représentation de la connaissance entre les notions de modèles conceptuels, d'ontologie et de modèles exécutables. La section 3 décrit l'architecture de WikiBridge, la section 4 présente notre modèle d'annotation et la section 5 décrit l'application archéologique qui nous a permis de valider les concepts mis en œuvre dans WikiBridge. Finalement la section 6 conclut cet article.

2. Représentation de la connaissance

Dans cette section, nous proposons, dans un premier temps, d'aborder la représentation de la connaissance selon l'angle des modèles du Génie Logiciel puis selon l'angle des modèles de la représentation des connaissances (les ontologies). Nous comparons ensuite la conception d'une application collaborative centrée base de données et avec celle s'appuyant sur un wiki.

1. <http://www.cidoc-crm.org/>

2. <http://www.obofoundry.org/>

2.1. Langages de modélisation

La modélisation peut se faire dans différents espaces technologiques (Kurtev *et al.*, 2002), en utilisant par exemple une modélisation basée sur UML ou une modélisation basée sur un langage de description d'ontologie comme OWL. Les ontologies et les modèles UML sont similaires pour la représentation des connaissances d'un domaine dans la mesure où ils proposent tous les deux des modes de description à base de concepts et de relations entre ces concepts. Selon Spear (Spear, 2006), la description d'un domaine suppose un choix précis dans les limites des descriptions. Ces limites peuvent être appréhendées selon deux dimensions :

- la dimension horizontale ou pertinence a pour objectif de déterminer l'étendue de l'information qui sera incluse dans la description. Par exemple, si l'on représente le domaine de l'archéologie, la pertinence recouvre le choix d'inclure ou non des éléments aussi variés que les sources documentaires, la géologie, les techniques de construction ;

- la dimension verticale ou granularité a pour objectif de déterminer le niveau de détail de la représentation des connaissances. Par exemple, si l'on représente le domaine de l'archéologie, la granularité recouvre le choix de décrire un édifice de la structure des murs jusqu'aux éléments du décor, aux sols.

Dans un même modèle UML, il est difficile d'inclure une description générale de certains éléments et des détails sur d'autres (dimension verticale), sauf à prendre le risque de construire un modèle difficile à comprendre et à maintenir. En revanche, un modèle UML peut utiliser différentes sources pour représenter la connaissance, et ainsi ajuster l'étendue (dimension horizontale) de la connaissance qu'il recouvre. Au contraire, les langages de description d'ontologies offrent un grand degré de liberté pour gérer la granularité de la connaissance (dimension verticale). Ces deux approches de modélisation sont donc complémentaires :

- au niveau de la modélisation, la granularité est apportée par les ontologies et l'étendue du domaine est apportée par les modèles UML ;

- au niveau des capacités de raisonnement sous-jacentes : la classification des concepts et leurs propriétés peuvent être inférées dans les ontologies et vérifiées dans les modèles UML (via des contraintes OCL par exemple) ;

- au niveau de la logique formelle : l'hypothèse du monde ouvert des logiques de description utilisée dans le cadre du Web Sémantique en association avec les ontologies et l'hypothèse du monde clos associée aux modèles centrés données (exprimés par exemple par un diagramme de classe UML) et utilisés dans les SGBD. Dans (Motik *et al.*, 2007) les auteurs étudient plusieurs approches pour réconcilier ces deux hypothèses.

2.2. Modèles exécutables pour représenter l'application

Un modèle exécutable est la plus haute couche d'abstraction basée sur les langages d'implémentation (Jiang *et al.*, 2008). Par exemple, *executable UML* restreint le diagramme de classes en supprimant les associations de composition et d'agrégation, mais aussi les attributs multivalués pour permettre une traduction automatique vers le modèle relationnel.

La conception d'une application s'appuyant sur une base de données (modèle exécutable) impose : 1) de structurer les concepts sous la forme d'attributs simples et mono-valués, ainsi le schéma résultant de la normalisation et de l'adaptation au SGBD cible est en général éloigné du modèle conceptuel ; 2) de construire le modèle très tôt lors de la phase d'analyse en s'appuyant sur une connaissance du domaine à un instant donné et donc de figer la connaissance avec les éléments modélisés ; 3) d'effectuer une modélisation par un petit nombre d'experts du domaine. De plus, l'expérience a montré que des modèles conçus par deux équipes différentes pour deux bases de données différentes avec des champs d'application voisins, seront difficiles à intégrer *a posteriori* car chaque concepteur possède son propre style de modélisation.

Pour gérer une connaissance informelle, la structure orientée document des wikis est un avantage par rapport à une approche centrée base de données. En outre, les wikis sont des plates-formes qui fournissent des fonctionnalités collaboratives, ils sont simples à mettre en œuvre, ils supportent l'édition en ligne de documents. Les utilisateurs peuvent saisir et mettre en forme des articles, importer des ressources multimédia et lier les articles et les ressources au moyen de liens hypertexte. Les wikis proposent également un système de gestion de versions ainsi qu'un moteur de recherche textuelle. Ces caractéristiques contribuent largement à leur succès. Cependant, même si les wikis sont largement utilisés, le système de gestion d'articles et les liens ne sont pas suffisants pour modéliser finement la connaissance et garantir une cohérence sémantique des informations.

Les wikis sémantiques proposent des solutions afin de pouvoir décrire la sémantique. On distingue généralement deux approches dans la conception de wikis sémantiques. La première approche dite *wikis for ontologies* concerne le plus grand nombre de wikis. Cette approche considère les pages du wiki comme des concepts et les liens typés comme des propriétés. L'autre approche dite *ontologies for wikis* utilise une ontologie pré-existante importée dans le wiki pour la mise en place d'annotations. Cette approche fournit généralement des formulaires de choix multiple ou utilise l'auto-complétion pour baser les annotations sur l'ontologie. Ces moteurs de wiki sont le plus souvent destinés à des domaines spécifiques comme en atteste l'état de l'art des différents moteurs de wikis sémantiques réalisé par Meilender *et al.* (Meilender *et al.*, 2011).

En conclusion, les wikis sémantiques peuvent combiner le meilleur des deux mondes : la structure issue des bases de données et la flexibilité apportée par les wikis (figure1). La connaissance est ainsi représentée sous une forme semi-structurée gardant les avantages à la fois des wikis et des bases de données, la modélisation est

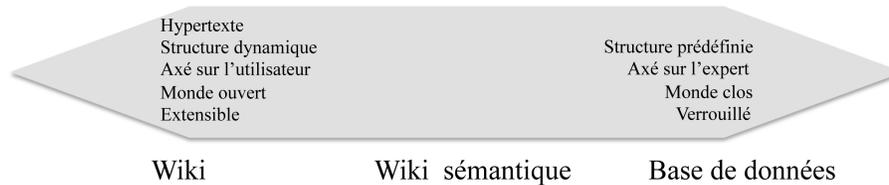


Figure 1. Wiki sémantique à la frontière des wikis et des bases de données relationnelles

effectuée selon un processus collaboratif, dynamique et évolutif. En effet, par rapport à une application s'appuyant sur une base de données, les wikis sémantiques ont les avantages suivants : 1) extension de la structure des documents via les formulaires (Haake *et al.*, 2005) et les annotations ; 2) utilisation de connaissances établies via les ontologies ; 3) support de la collaboration ; 4) émergence du modèle de document et de la sémantique à partir des usages. Cependant, la représentation de la connaissance et la vérification de la sémantique des annotations est un des enjeux majeurs des solutions basées sur les wikis sémantiques.

2.3. Aperçu de WikiBridge

Nous avons développé WikiBridge, un wiki de type *ontologies for wikis*, qui permet de travailler avec un contenu non formalisé c'est-à-dire des documents écrits en langage naturel, complété par des structures formelles c'est-à-dire un système d'annotation s'appuyant sur une ontologie. Un document contient trois structures de représentation (Poullet *et al.*, 1997) :

1) la structure physique définit la présentation du document. La présentation est dépendante de l'environnement, s'il s'agit d'une présentation "papier" le format de la feuille (A4, A3, etc.) ou son mode d'utilisation (portrait ou paysage) définira la présentation. Dans le cadre d'une présentation informatique, ce sont les fonctionnalités du système qui déterminent cet environnement comme les feuilles de style (CSS, XSLT) ;

2) la structure logique définit l'organisation (relations de composition, séquence) de l'information contenue dans le document. Cette organisation s'établit autour d'abstractions représentant les différentes parties du document : titres, chapitres, paragraphes, notes de bas de page, figures, etc. La structure logique est totalement indépendante de l'environnement alors que la structure physique en dépend ;

3) la structure sémantique représente le sens du document c'est-à-dire les informations qu'un utilisateur est susceptible de demander lors d'une recherche, en effet le besoin d'un utilisateur porte sur le contenu sémantique des documents recherchés, cette structure se superpose au document.

Nous avons mis en place ces trois structures à l'aide d'extensions que nous avons développées à partir du moteur de wiki MediaWiki diffusé sous une licence libre³. La sémantique y est décrite selon deux niveaux. Le niveau minimum proposé est une structure définie via des modèles de formulaires qui par analogie avec les schémas de bases de données permettent de décrire les grandes classes d'objets manipulés sans décrire précisément leurs propriétés. L'autre niveau est le mécanisme d'annotation qui travaille à un niveau de granularité plus fin en permettant d'annoter des portions de texte ou des ressources internes ou externes. Afin de fournir une sémantique précise, les annotations sont définies en utilisant les termes d'une ontologie (figure 2). Les fonctionnalités sémantiques ont été validées avec une application dans le domaine de l'archéologie au travers du projet CARE (*Corpus Architecturae Religiosae Europaeae - IV-X saec.* - ANR-07-CORP-011) (Chevalier *et al.*, 2008) présenté en section 5. Certaines caractéristiques de ce domaine viennent s'ajouter : 1) la complexité des données (hétérogènes, incomplètes, incertaines, inconsistantes, spatio-temporelles) ; 2) la barrière de la connaissance du domaine nécessaire pour utiliser l'outil ; 3) l'évolution de la connaissance et 4) les compétences des utilisateurs.

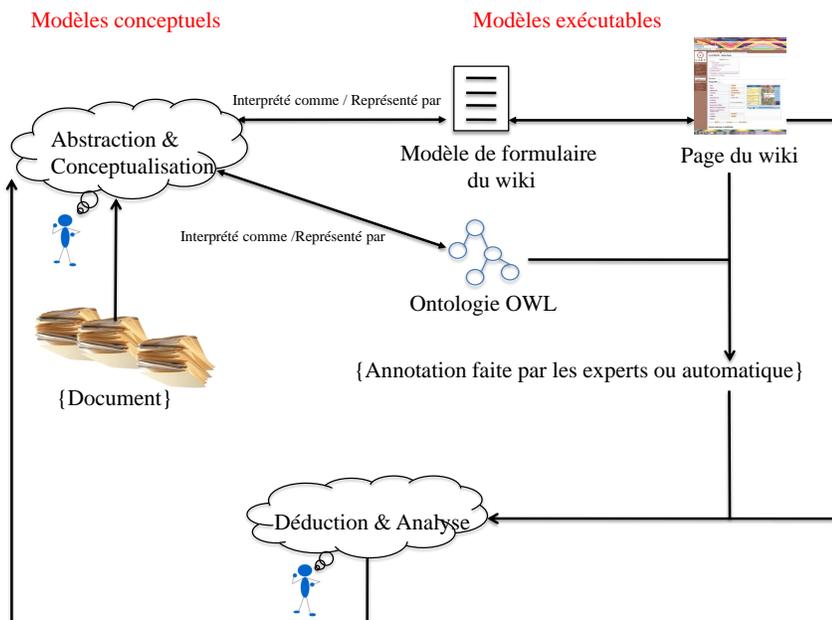


Figure 2. Interaction modèles conceptuels / modèles exécutables dans le cadre de WikiBridge

3. <http://www.mediawiki.org/>

3. Architecture de WikiBridge

Uren et al. (Uren *et al.*, 2006) ont identifié plusieurs exigences et passé en revue différents systèmes d'annotation dans leur état de l'art sur les annotations sémantiques. Dans la conception de WikiBridge, sept de leurs exigences ont été retenues : interface d'utilisation facile, conception collaborative du contenu, support de différentes catégories d'utilisateurs, support de données de format hétérogène, respect des standards du Web Sémantique, stockage des annotations, moteur de raisonnement.

La figure 3 présente les modules de WikiBridge et leur interactions. Le système de gestion des articles incluant la saisie ainsi que la mise en forme est implanté en utilisant MediaWiki que nous avons étendu avec plusieurs composants sémantiques : une interface d'acquisition de données basée sur des formulaires sémantiques permettant une annotation automatique, un moteur de requêtes SPARQL, un assistant pour la construction des annotations, un mécanisme de contrôle des annotations incluant la prise en compte du contexte et des contraintes sémantiques.

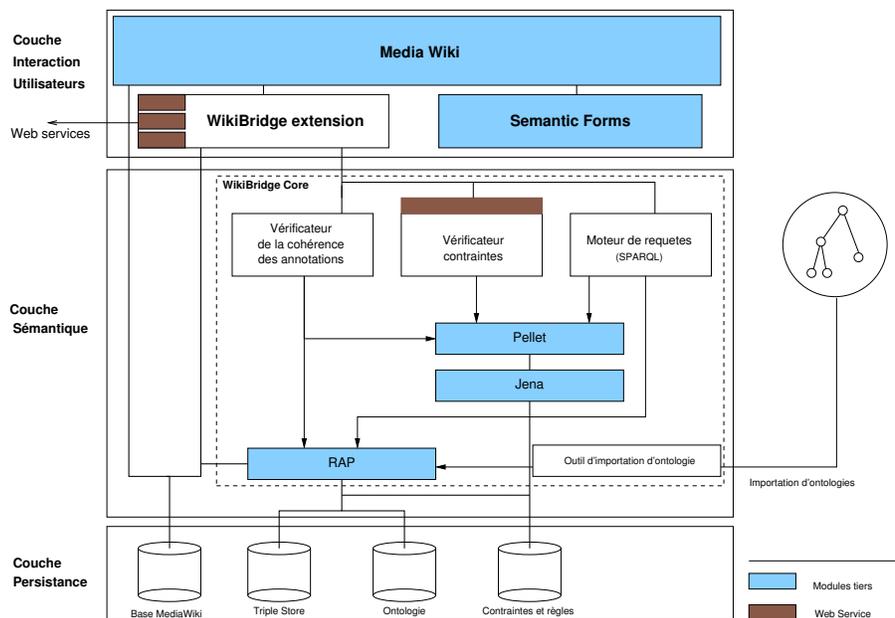


Figure 3. Architecture de WikiBridge

3.1. Couche d'interaction avec les utilisateurs

MediaWiki couvre majoritairement la couche d'interaction avec les utilisateurs en offrant un support aux structures physique et logique des documents.

Sources historiques et identification**Sources indirectes**

- G. Chevrier et M. Chauva, Chartes et documents de Saint-Brisigne de Dijon, t. 2 (990-1124), Dijon, Chartes 299 p. 194.
 Hubert J., 1952, L'architecture religieuse du haut Moyen Âge, Paris, p. 63, n°65. Gras P., Le séjour à Dijon des évêques de Langres du Ve au IXe siècle, ses conséquences, Dijon, p. 559.
 Veillard-Troiekouff M., 1976, Les monuments religieux de la Gaule d'après les œuvres de Grégoire de Tours, Paris, p. 116.

Sources archéologiques

- Albert X., 1935, « Fouilles de Saint-Philbert de Dijon », Mémoires de la Commission des Antiquités de la Côte-d'Or, t. 19, 1927-1932, Dijon, p. 103.
 Relaud J., 1994, Rapport de sauvetage de l'Église Saint-Philbert de Dijon, Dijon/SRA Bourgogne.

Sources Bibliographiques

- Foisnet P., 1982, « Saint-Philbert de Dijon et l'architecture romane en Bourgogne », Mémoires de la Commission des Antiquités de la Côte-d'Or, t. 6, Dijon, p. 1-32.
 Deshoulières M., 1929, « Église Saint-Philbert », Congrès archéologique de France, Dijon 1928, Paris, p. 96-109.
 Gras P., 1978, « Église Saint-Philbert », Mémoires de la Commission des Antiquités de la Côte-d'Or, t. 29, 1974-1975, Dijon, p. 30-33.
 Marlier J., 1978, « Saint-Philbert », Mémoires de la Commission des Antiquités de la Côte-d'Or, t. 30, Dijon, 1976-1977, p. 87-99.
 Provost M. et A., La carte archéologique de la Gaule, Côte-d'Or, 21/2, Paris, p. 300.

Conservation

L'édifice actuel est fermé au public et actuellement en travaux.

Titulature actuelle

Saint-Philbert

Titulature historique

Saint-Philbert

Diocèse actuel

Dijon

Diocèse historique

Langres

Contexte d'implantation

Description

Pays	FRANCE
Région	Bourgogne
Département	Côte-d'Or
Commune	Dijon
Insee	21231
Adresse/Lieu-dit	
Toponyme	
Propriétaire	Commune
Protection de l'édifice	Cl. M.H. site de 1982 et a
Références cartographiques	
Numéro parcellaire sur le Cadastre actuel	
Latitude	47.221015
Longitude	5.025240
Altitude	246 m

Figure 4. Rendu d'une fiche

3.1.1. Structure physique

MediaWiki propose une syntaxe et un moteur de rendu pour gérer la mise en page (texte en italique, gras, en couleur, réalisation de liste, de tableau, de chapitre, de section, etc.). Le wiki permet *un premier enrichissement* du document grâce : 1) aux liens internes entre des articles du wiki ou aux liens externes vers une page Web extérieure ; 2) aux supports multimédia comme des photographies, des plans, du son ou de la vidéo et 3) à des services Web. Dans le cadre du projet CARE, nous avons utilisé des services Web pour la géo-localisation comme l'extension Google Maps de MediaWiki⁴ et l'API Géoportail⁵.

3.1.2. Structure logique

Le concept de *wiki template* permet aux utilisateurs de déterminer la structure et l'apparence d'une page wiki (Haake *et al.*, 2005; Iorio *et al.*, 2006; Iorio *et al.*, 2008). Deux types de template ont été identifiés : 1) les templates fonctionnels correspondent à une page qui contient des espaces réservés complétés par des valeurs passées en paramètres, l'exemple type est l'infobox ; 2) les templates créationnels sont des pages utilisées comme point de départ pour la création de nouvelles pages avec la même structure. Un wiki utilisant des templates créationnels est appelé "*Lightly Constrained Wiki*".

Nous utilisons à la fois une infobox pour synthétiser les données (essentiellement géographiques) de l'édifice et des templates créationnels pour faciliter l'ajout de don-

4. http://www.mediawiki.org/wiki/Extension:Google_Maps

5. <http://depot.ign.fr/geoportail/api/doc/fr/developpeur/download.html>

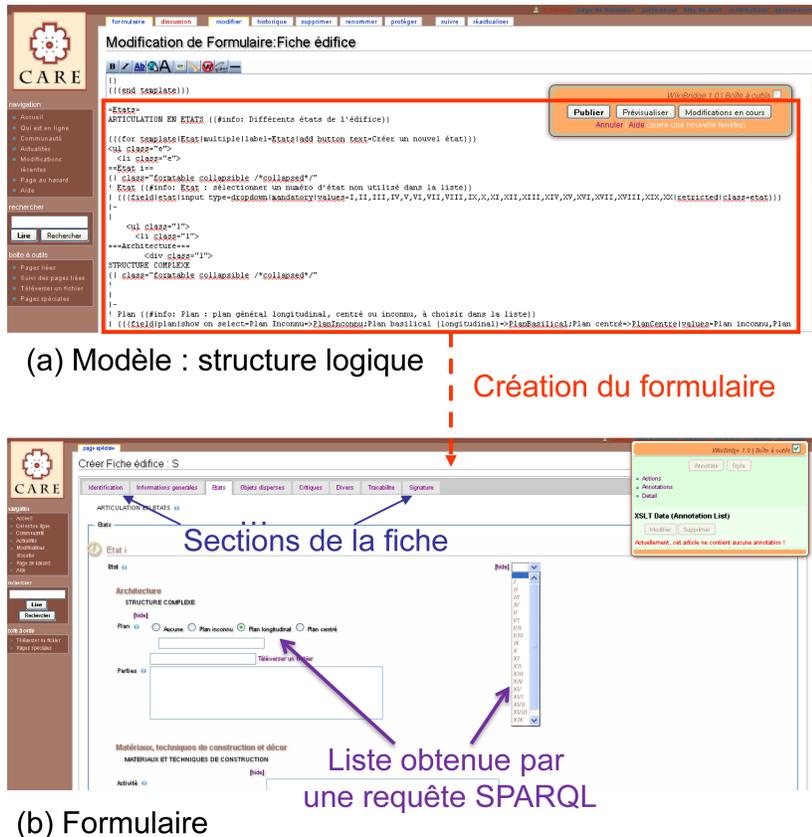


Figure 5. Processus de création d'un formulaire

nées (figure 4). Les templates créationnels sont pris en charge par l'extension Semantic Forms⁶ développée pour MediaWiki. La structure logique réalisée grâce à cette extension reflète l'organisation des documents. Les templates sont décrits en utilisant un langage spécifique, chaque partie du document est représentée par un template (figure 5.a), les templates peuvent être composés. Un formulaire de saisie est créé à la volée sur la base du template (figure 5.b). Les champs peuvent prendre la forme de boîtes de sélection, de cases à cocher ou de texte libre. Les boîtes de sélection et les cases à cocher sont construites à partir des individus de l'ontologie au moyen de requêtes SPARQL. Les modules correspondant à la couche d'interaction utilisateur sont représentés en haut de la figure 3.

Cette structure logique, même si elle propose déjà de nombreuses fonctionnalités, ne peut pas représenter la connaissance et donc ne permet pas un accès facile à

6. http://www.mediawiki.org/wiki/Extension:Semantic_Forms

l'information. Nous avons besoin d'un moteur de requêtes SPARQL, et nous aboutissons à la proposition d'une structure en deux plans : le premier plan correspond à la structuration des concepts fondamentaux du domaine grâce au formulaire de saisie, le deuxième plan est une ontologie spécialisée pour l'application qui permet d'annoter le corpus et de faire émerger de nouvelles connaissances.

3.2. Couche sémantique

Afin de contrôler la qualité de données, durant le processus de saisie, nous proposons un ensemble de modules (boîtes blanches dans la figure 3) développés en utilisant des composants tiers.

Le mécanisme d'annotation permet d'annoter n'importe quel élément (portion de texte, image, lien, etc.) avec les concepts d'une ontologie. L'éditeur Protégé⁷ a été utilisé pour produire une ontologie OWL qui est ensuite importée dans WikiBridge.

L'assistant d'annotation permet aux utilisateurs de construire des annotations en sélectionnant les termes de l'ontologie dans des listes et en leur associant des propriétés et des valeurs. Le processus d'annotation étant sensible au contexte, les termes sont sélectionnés dans l'ontologie par rapport aux champs actifs du formulaire. Les annotations ainsi construites sont vérifiées par rapport à l'ontologie lors de l'enregistrement ou de la modification des articles.

Le processus de vérification de la cohérence des annotations comporte plusieurs composants spécifiques interagissant avec RAP⁸, Jena⁹ et Pellet¹⁰. Les contraintes sémantiques exprimées en logique du premier ordre sont vérifiées en utilisant Pellet et un service web interconnectant RAP et Jena. De plus, des règles peuvent être ajoutées pour interroger l'ontologie et les annotations afin de tester de nouveaux faits et ainsi produire de la connaissance.

3.3. Couche de persistance

La couche de persistance inclut quatre types de stockage : le contenu des articles, l'ontologie, les annotations sémantiques, les contraintes et les règles. Le stockage des articles est assuré par la base de données spécifique à MediaWiki. Chaque article est identifié par une URL et les annotations utilisent cette URL comme préfixe pour identifier les parties des articles ciblées (paragraphe, phrase, mot, etc.).

L'ontologie importée à partir d'un fichier OWL est stockée dans un schéma spécifique géré par RAP. L'ontologie peut également être interrogée en SPARQL et les

7. Protégé : <http://protege.stanford.edu/>

8. RDF API for PHP <http://www4.wiwiwiss.fu-berlin.de/bizer/rdfapi/>

9. <http://jena.sourceforge.net/>

10. <http://pellet.owldl.com/>

résultats inclus dans des pages du wiki. Cette technique est utilisée pour construire les listes de termes, de propriétés et de valeurs proposées par l'assistant d'annotation.

Les annotations sont stockées dans le triple store de RAP. Elles peuvent être interrogées au moyen de requêtes SPARQL incluses dans des articles (*in-line query*).

Les contraintes exprimées sur les termes de l'ontologie et utilisées par Pellet et Jena sont stockées dans le format natif des outils, c'est-à-dire sous forme textuelle. Un tag permet de dissocier les règles appliquées à l'ontologie des contraintes utilisées pour vérifier la cohérence des annotations.

3.4. Couche d'accès à l'information

Nous utilisons les annotations pour réaliser de la recherche d'information comme le propose Kiryakov et al. (Kiryakov *et al.*, 2004) : « *Semantic annotation . . . aiming to enable new information access methods and to extend the existing ones.* ». Pour répondre aux attentes d'utilisateurs ayant des besoins différents, nous offrons trois types de requêtes qui extraient des informations sans une connaissance *a priori* de son schéma :

1) navigation par facette (*faceted browsing*) partitionne l'espace d'information en utilisant des caractéristiques importantes des éléments d'information que l'on appelle facette. Chaque facette a de multiples valeurs et l'utilisateur sélectionne une valeur pour obtenir les éléments pertinents dans l'espace d'information. La théorie des facettes peut être directement associée à la navigation dans les données RDF : des éléments d'information sont des sujets RDF, les facettes sont des prédicats RDF et la restriction des valeurs sont des objets RDF. Dans WikiBridge, elle permet aux utilisateurs d'explorer, par filtration successive de la structure de l'ontologie, l'information disponible ;

2) recherche par formulaire permet de saisir des paramètres pour les requêtes portant sur des conjonctions de critères et identifiées lors de la phase d'analyse (figure 6) ;

3) vue agrégée affiche dans une factbox, pour chaque article, toutes les annotations qui lui sont liées.

Le moteur de requêtes SPARQL permet d'interroger la base d'annotations ou l'ontologie avec une connaissance du schéma. Les requêtes SPARQL peuvent être utilisées en ligne dans des articles (tableau 1) afin de produire des pages du wiki ou bien à partir de services web pour permettre par exemple, dans le cadre du projet CARE, une analyse spatio-temporelle des annotations relatives aux édifices. Le module d'analyse spatio-temporelle a été développé par l'équipe de géomatique de la MSH de Dijon¹¹.

11. Maison des Sciences de l'Homme : <http://msh-dijon.u-bourgogne.fr/pole-geomatique-et-cartographie.html>

CARE Interface de requêtes sémantiques

Bienvenue dans l'interface de requêtes sémantiques du corpus ANR CARE.

Type de recherche

Parcourir le triple store Requête personnalisée Utiliser un modèle de requête (connexion requise)

Then, choose a concept: Autel

Now, configure your annotation as wished.

Accès

*** Select a value ***

Dimensions

Épaisseur

Hauteur

Largeur

Longueur

Profondeur

Etat (S.V.P. : valeur comprise entre 4 et 11)

Forme

Autel caisse rectangulaire + colonnettes d'angles

Résultat de la requête exprimée ci-dessus

Id	Article
31	Avallon
27	NEVERS, cathédrale et baptistère
33	Saint Clément
35	Saint Aubin

Figure 6. Interface de requêtes

```
<wbt_ilsq type="db" model="http://care.u-bourgogne.fr/rdf/CAREO">
prefix ns: http://www.w3.org/1999/02/22-rdf-syntax-ns#
prefix careo http://care.u-bourgogne.fr/rdf/CAREO#
select ?s, ?longitude, ?latitude, ?pays, ?region, ?departement, ?batiment
where
{
  ?s    <aPourLongitude>    ?longitude ;
        <aPourLatitude>    ?latitude ;
        <aPourPays>        ?pays ;
        <aPourRegion>      ?region ;
        <aPourDepartement> ?departement ;
        <aPourBatiment>    ?batiment ;
        <aPourRegion>      'Bourgogne' .
}
</wbt_ilsq>
```

Tableau 1. Exemple de requête SPARQL

Ce module utilise OpenLayers¹², une capture d'écran du résultat est présentée en figure 7.

12. <http://openlayers.org/>

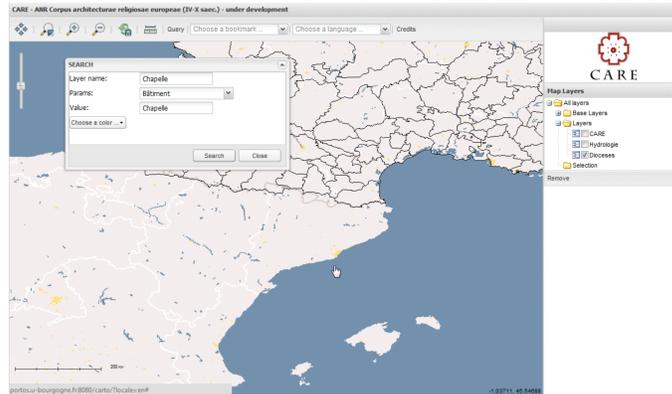


Figure 7. Capture d'écran de l'application de cartographie

4. Annotations

Dans cette section, nous détaillons le modèle d'annotation utilisé dans WikiBridge, puis nous proposons une analogie de la sémantique des annotations avec la sémantique des langages de programmation.

4.1. Modèle d'annotation de WikiBridge

Le système d'annotation mis en œuvre dans WikiBridge propose trois modes de fonctionnement : un mode d'annotation automatique à partir des formulaires, un mode d'annotation assisté où les annotations sont écrites par les utilisateurs à l'aide d'un assistant et un mode d'annotation manuel, dans l'article en utilisant la syntaxe du wiki, les annotations sont écrites par des utilisateurs expérimentés.

Ces annotations partagent le même modèle de base : $A = (s, p, o)$ où s est le sujet (la donnée annotée), p est le prédicat (le type de lien) et o est l'objet (la valeur de l'annotation). Les types de chacun des composants sont les suivants :

- s est une URI qui fait référence à un document (c'est-à-dire à un article dans le wiki ou à une partie d'un article) ou à un terme de l'ontologie ;
- p est une URI qui fait référence à un élément d'une ontologie (concept ou propriété) ;
- o est un littéral ou une URI qui fait référence à un individu d'une ontologie ou une URI d'un document ou null.

On distingue trois types d'annotation (tableau 2) :

- 1) une annotation simple est de la forme (s, p, o) telle que s et p sont non null. Si o est null et si p fait référence à un concept, l'annotation spécifie le type du sujet, cela

$$\begin{aligned}
\langle A \rangle &\rightarrow \langle A\text{-simple} \rangle \mid \langle A\text{-cplx} \rangle \mid \langle A\text{-recursive} \rangle \\
\langle A\text{-simple} \rangle &\rightarrow (\langle s \rangle, \langle p \rangle, \langle o \rangle) \\
\langle A\text{-cplx} \rangle &\rightarrow (\langle A\text{-simple} \rangle, \langle A\text{-liste} \rangle) \\
\langle A\text{-liste} \rangle &\rightarrow \langle A\text{-simple} \rangle \mid \langle A\text{-simple} \rangle, \langle A\text{-liste} \rangle \\
\langle A\text{-recursive} \rangle &\rightarrow (\langle s \rangle, \langle p \rangle, \langle o \rangle \langle A \rangle) \mid (\langle s \rangle, \langle p \rangle, \langle o \rangle \langle A \rangle \langle A\text{-rec-liste} \rangle) \\
\langle A\text{-rec-liste} \rangle &\rightarrow \langle A\text{-simple} \rangle \mid \langle A\text{-simple} \rangle \langle A\text{-rec-liste} \rangle \\
\langle s \rangle &\rightarrow \mathbf{URI} \mid \mathbf{URL} \\
\langle p \rangle &\rightarrow \mathbf{concept\ de\ l'ontologie} \mid \mathbf{propriété\ de\ l'ontologie} \\
\langle o \rangle &\rightarrow \mathbf{individu\ de\ l'ontologie} \mid \mathbf{littéral} \mid \mathbf{URI} \mid \mathbf{URL} \mid \mathbf{null}
\end{aligned}$$

Tableau 2. *Syntaxe abstraite de notre modèle d'annotation*

peut être assimilé à une contrainte de type restriction sur le domaine d'un attribut. Si o n'est pas null, o est soit un littéral soit un individu du concept spécifié par p. Cela peut être assimilé à une contrainte de base de données qui vérifie que la valeur d'un attribut est prise dans une liste énumérée ;

2) une annotation complexe ou annotation n-aire (notée A-cplx dans le tableau 2) permet de mettre en relation un même sujet avec une liste de couples (p, o), o pouvant être un littéral, un individu, ou une référence à un autre sujet. Tous les prédicats utilisés doivent être différents ;

3) une annotation récursive permet d'expliquer ou de préciser comment un prédicat et un objet sont liés, c'est-à-dire comment une propriété de l'ontologie est liée avec un individu. Elle a différents niveaux : une annotation de niveau i explique l'objet o de l'annotation de niveau i-1. Si toutes les annotations du niveau i sont une liste d'annotations alors toutes ces annotations partagent le même sujet o. Cette forme d'annotation est basée sur le modèle de valeur sémantique défini par Sciore et Rosenthal (Sciore *et al.*, 1994).

La figure 8 montre une annotation complexe et récursive qui associe deux prédicats (accès et dimension) à un sujet dont le type est spécifié par le prédicat installation liturgique. Le prédicat dimension comporte une liste de prédicats simples spécifiant la hauteur et la largeur. L'annotation engendrée par la grammaire est la suivante :

```

(SaintClément#Autel1, InstallationLiturgique, Autel
  (Autel, Accès, SaintClément#Escalier3)
  (Autel, Dimensions, 3D
    ((3D, Hauteur, null),
     (3D, Largeur, 0.90))
  )
)

```

Une annotation exprimée sur un sujet contribue à définir son contexte selon deux dimensions : 1) la dimension référentielle inclut la structure de l'ontologie et les règles portant sur les concepts utilisés ; 2) la dimension assertionnelle est l'ensemble de

toutes les annotations portant sur le même sujet. Une annotation est valide seulement si elle est consistante vis à vis des deux dimensions.

INSTALLATIONS LITURGIQUES

Etat	3	Installation liturgique
Délimitation		Autel
Autel	L' autel principal est un massif rectangulaire au centre de l'abside médiane ; les deux absides quadrangulaires maçonnés plus petit (0,75 x 0,50) et l'hémicycle.	Accès Select a value
		Dimensions Annotation n-aire
		Hauteur <input type="text"/>
		Largeur 0.90

Figure 8. Construction d'une annotation complexe et récursive avec l'aide de l'assistant

4.2. Sémantique des annotations : une analogie avec la sémantique des langages de programmation

L'assistant d'annotation offre une forme de contrôle des annotations permettant de vérifier deux types de contraintes en utilisant l'ontologie : 1) les domaines de valeurs des objets en utilisant les possibilités fournies par la ABox ; et 2) la consistance structurelle des prédicats en utilisant les possibilités fournies par la TBox (par exemple une cathédrale peut avoir une nef mais ne peut pas posséder un atrium). Cependant, toutes les contraintes ne peuvent pas être vérifiées avec la structure de l'ontologie et le mécanisme d'inférence des logiques de description et il est toujours possible d'écrire une annotation dans un document en utilisant la syntaxe du wiki. Afin de contrôler la cohérence des annotations c'est-à-dire à la fois leur structure mais aussi leur consistance dans un document (deux annotations ne peuvent pas être contradictoires) nous développons un ensemble de mécanismes basés sur les travaux dans la sémantique des langages de programmation. Ainsi, dans notre approche de wiki sémantique, l'ontologie et la syntaxe abstraite jouent le rôle de DSL (Domain Specific Language) puisqu'elles permettent d'exprimer quelles sont les annotations syntaxiquement correctes. L'ontologie permet un contrôle de cohérence structurelle des annotations selon trois niveaux :

- au niveau du typage : un sujet possède un type défini par des annotations de la forme (s, p, null) ;
- au niveau de la structure elle-même : il est impossible de créer une annotation complexe sur un sujet en utilisant des prédicats qui ne sont pas applicables à son type. Par exemple, dans la figure 8, le sujet "autel" est une installation liturgique et ce type d'installation a pour prédicat accès et dimension.

– au niveau des domaines de valeurs de la composante o de l’annotation, ces domaines sont contrôlés par les co-domaines des rôles et les individus de l’ontologie. Par exemple dans le cadre du projet CARE, l’annotation d’une installation liturgique exige que le matériau soit compatible avec la technique de construction, il est donc impossible de spécifier à la fois maçonnerie et en bois.

Cependant au delà de la cohérence structurelle des annotations, il existe des règles de cohérence plus globales dépendantes fortement du domaine. Par exemple dans le cadre du projet CARE il n’existe pas en France, pour les siècles concernés, d’église avec des murs en terre alors que ce type de construction peut exister en Irlande.

Afin de contrôler finement la sémantique des annotations nous proposons d’établir une analogie avec la sémantique des langages de programmation.

4.2.1. *Sémantique axiomatique*

Cette sémantique considère le processus d’annotation comme une transformation des propriétés attachées à un document. La cohérence du processus est donnée par la vérification de l’ensemble des propriétés par l’ensemble des annotations ajoutées sur le document à condition que l’ensemble des propriétés définies avant l’ajout d’annotations soit lui aussi valide.

La sémantique axiomatique du processus d’annotation se traduit par la vérification de contraintes ou de propriétés globales après annotation. La contrainte sur le fait qu’un édifice dédié à un Saint doit avoir été construit après la sanctification du personnage peut s’exprimer en logique par :

$$\text{isConsecrated}(\text{?b}, \text{?s}) \leftarrow \text{hasConstructionDate}(\text{?b}, \text{?d1}) \wedge \text{hasDateDead}(\text{?s}, \text{?d2}) \wedge \text{d1} \geq \text{d2}$$

4.2.2. *Sémantique dénotationnelle*

La sémantique dénotationnelle du processus d’annotation exprime la correspondance entre la structure de l’annotation et la sémantique du domaine au moyen de fonctions de l’ensemble des termes utilisés dans les documents vers les concepts de l’ontologie. Les règles de construction des annotations comme celles applicables sur les prédicats (exemples donnés en début de section) relèvent de la sémantique dénotationnelle.

4.2.3. *Sémantique opérationnelle*

La sémantique opérationnelle des annotations assimile le processus à un changement d’état vu non pas sous l’angle des propriétés comme dans la sémantique axiomatique mais sous la forme des états accessibles à partir d’un état donné.

Par exemple dans le projet CARE, si nous utilisons trois concepts essentiels pour modéliser l’évolution des édifices : l’usage religieux, les entités spatiales, le temps. L’édifice est représenté par un ensemble d’annotations qui le situe dans un espace à trois dimensions. Soit \mathcal{U} l’ensemble des usages, \mathcal{E} l’ensemble des entités spatiales

et \mathcal{T} l'ensemble des entités temporelles. Un édifice a est sous-ensemble du produit cartésien des trois ensembles $\mathcal{U} \times \mathcal{E} \times \mathcal{T}$, soit $a = \{(u, e, t), u \in \mathcal{U}, e \in \mathcal{E}, t \in \mathcal{T}\}$. Les évolutions valides sont déterminées par les changements de valeur dans une des composantes de l'un des triplets.

Cette sémantique permet : 1) de restituer tous les états possibles ; 2) d'analyser et observer tous les changements d'état possibles c'est-à-dire de procéder à une différence entre deux états ; 3) de caractériser les évolutions c'est-à-dire le processus de changement d'état. Le principe d'indépendance entre l'usage religieux, l'espace et le temps permet d'observer les facteurs influant sur le changement, de restreindre l'étude à des produits deux à deux des trois ensembles et d'estimer ainsi le rôle ou la prépondérance de l'un par rapport à l'autre.

4.2.4. Principes de raisonnement sur les annotations

Parmi les différentes sémantiques associées au processus d'annotation, nous pouvons déduire les modes de raisonnement à mettre en œuvre pour valider les annotations. Dans le cadre des annotations globales relatives à la sémantique axiomatique un mode de raisonnement avec l'hypothèse du monde clos est nécessaire. Pour la validité structurelle des annotations relatives à la sémantique dénotationnelle c'est l'hypothèse du monde ouvert qui sera activée. Pour les transitions d'état des édifices il est probable que l'hypothèse du monde clos soit nécessaire.

5. Terrain d'expérimentation : le projet CARE

L'objectif du projet international CARE (*Corpus Architecturae Religiosae Europaeae - IV-X saec.*) est la constitution d'un corpus des monuments chrétiens antérieurs à l'an Mil (<http://care.u-bourgogne.fr>). Il s'agit de recenser tous les édifices religieux et de décrire leurs évolutions entre le IV^e et le début du XI^e siècle dans plusieurs pays (Italie, Espagne, Croatie, etc.). Le projet a commencé en France en 2008, après avoir été accepté par l'ANR, pour une durée de quatre ans. Le corpus français, d'environ 2 700 édifices, met l'accent sur les VII^e-VIII^e siècles et sur les décennies précédant ou suivant l'an Mil très riches en monuments. Le travail sur le corpus d'informations textuelles et graphiques concernant chaque édifice se décompose en deux tâches principales : la collecte des données grâce à la rédaction des fiches de dépouillement puis leur analyse et interprétation. Les données de terrain recueillies par les archéologues, complétées aussi bien par des sources littéraires, d'archives et épigraphiques font l'objet d'un dépouillement systématique. Après dépouillement, chaque édifice ou groupe d'édifices est décrit par une fiche comportant différentes sections (par exemple, des informations générales sur l'édifice, l'évolution des éléments architecturaux décrite par la notion d'état, les considérations critiques sur les états et la chronologie, etc.).

D'un point de vue organisationnel, le projet CARE prend la forme d'un réseau d'experts assurant l'alimentation du corpus et collaborant à son exploitation au moyen

de travaux de recherche. Le caractère novateur de la démarche est apprécié au regard des caractéristiques suivantes :

- exhaustivité : à terme, la quasi totalité des sources disponibles aura été dépouillée ;

- interdisciplinarité : le programme de dépouillement et d'interprétation des données rassemble des archéologues, des historiens, des historiens de l'art, des dessinateurs topographes. Il est conçu comme un processus collaboratif permettant de traiter un édifice dans son ensemble, en identifiant et en intégrant toutes les relations entre les différents éléments manipulés et donc de comprendre un objet complexe dans sa totalité. L'objectif est de synthétiser et de relier chaque savoir disciplinaire et de le replacer dans un cadre systémique plus large ;

- transdisciplinarité : cette coopération se double d'un partenariat entre la recherche archéologique et la recherche informatique. La création d'une plate-forme informatique fournit aux informaticiens un terrain d'expérimentation des concepts qu'ils élaborent dans les thématiques d'ingénierie des ontologies et du Web Sémantique ;

- innovation : la mise en place d'une plate-forme informatique vise à moderniser les méthodes de travail en archéologie en libérant l'archéologue de routines fastidieuses lorsqu'elles sont effectuées manuellement. Il permet également de limiter les risques de biais introduits par une sélection lacunaire des données ;

- valorisation de la recherche : le projet fait l'objet d'une intégration sur la grille ADONIS¹³ et des services web sont en cours de développement pour une utilisation avec le système ISIDORE¹⁴. WikiBridge est conçu de manière à pouvoir être transposé pour répondre aux besoins d'autres domaines.

5.1. Compréhension du domaine

Nous avons appliqué les directives de Linster au projet CARE de façon à initier une connaissance fondamentale à partir des documents. Notre approche demandant une représentation spatio-temporelle d'un édifice, quatre aspects doivent être pris en compte. Premièrement, il faut décrire les éléments constitutifs de l'édifice par rapport à leur position dans l'espace. Deuxièmement, il est nécessaire de décrire le niveau physique d'un édifice par des mesures, des matériaux. Troisièmement, il faut ajouter une dimension temporelle pour prendre en compte les différentes évolutions. Finalement, l'aspect documentaire est à considérer car la reconstitution de tous les états s'appuie sur les sources disponibles. La figure 9 schématise cette représentation spatio-temporelle particulière. Ensuite, pour chacun des concepts nous associons un vocabulaire approprié. Dans la suite, l'ontologie d'application que nous avons construite est présentée.

13. <http://www.tge-adonis.fr/>

14. <http://www.rechercheisidore.fr/>

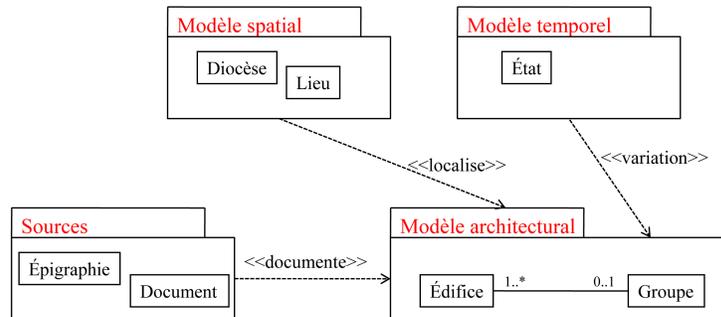


Figure 9. Représentation spatio-temporelle

5.2. Ontologie développée

Il existe un grand nombre de vocabulaires contrôlés pour décrire et indexer des objets du patrimoine culturel¹⁵. Le CIDOC (Comité International pour la Documentation) soutenu par l'ICOM (International Council of Museums) a pour objectif d'améliorer la gestion des collections, des archives et des produits scientifiques ou administratifs liés au patrimoine artistique et culturel. L'idée de base est de mettre en place une ontologie de domaine pour décrire des œuvres d'art, des vestiges archéologiques mais aussi des monuments. L'ontologie CIDOC-CRM (Conceptual Reference Model) a été élaborée depuis 1994 et publiée en 2006 par l'ISO (<http://www.cidoc-crm.org>). Parce que CIDOC-CRM est une référence dans le domaine, et notamment parce qu'elle propose la notion d'*événement* qui décrit à la fois l'environnement de l'objet au cours du temps et l'énoncé de tout ce qui a pu lui arriver, nous l'avons utilisé pour servir de base à l'ontologie CARE. Notre base de travail est constituée d'extraits de langage naturel donnant des descriptions des différents éléments d'un édifice. Pour aboutir à l'ontologie, notre démarche a comporté quatre étapes :

1) classification des éléments architecturaux d'un édifice : éléments maçonnés (chapiteau, fût, etc.), charpentes, sols, etc. ;

2) définition des relations entre les éléments. Goulette (Goulette, 1999) a énuméré les différentes relations entre éléments : relation partie-tout, relations spatiales. Nous avons dans notre analyse pris en compte les relations qu'un élément d'architecture peut avoir avec une technique de construction, un élément stylistique, etc. Les relations spatiales concernent l'orientation, la délimitation, la distance et le positionnement. Vieu (Vieu, 1997) distingue trois types d'orientation : a) l'orientation absolue qui fait référence à un système de coordonnées externes comme les directions cardinales ; b) l'orientation intrinsèque pour laquelle le repère est lié à l'objet de référence ;

15. http://www.pro.rcip-chin.gc.ca/normes-standards/vocabulaire_vocabulaires-vocabulary_vocabulary-fra.jsp

c) l'orientation contextuelle pour laquelle le repère est lié à une entité différente de l'objet de référence. Pour le projet CARE, l'orientation contextuelle est donnée par l'orientation de l'édifice. En effet, depuis les origines et jusqu'au XV^e siècle, dans tous les pays chrétiens, une église est orientée vers l'est. La délimitation permet de définir les frontières entre deux éléments avec les notions d'intérieur et d'extérieur. La distance représente la notion de proximité/éloignement entre deux éléments. Le positionnement relatif permet de préciser la position d'un élément ou d'un attribut d'un élément, l'axe par exemple, relativement à un autre élément ou à un de ses attributs. Le deuxième élément devient alors le référentiel architectural qui permet de préciser la position du premier par le biais de préposition de localisation. Cette localisation se fait principalement d'une manière qualitative par le biais de relations topologiques. Nous utilisons les travaux de Egenhofer et Herring (Egenhofer *et al.*, 1991) définissant un ensemble minimal de huit relations décrivant les relations entre deux régions.

3) intégration du vocabulaire religieux en utilisant soit la représentation symbolique des éléments soit leur décomposition logique en différents espaces. Ainsi par exemple une caisse en bois représente un cercueil, une table un autel, la nef est l'espace où est rassemblé les fidèles, le chœur est l'espace où se trouve l'autel et où se déroule les liturgies.

4) prise en compte de la dimension temporelle. En archéologie, le temps est construit à partir d'indices spatiaux, stylistiques ou naturels, les indices sont ensuite croisés. Leur validité est perpétuellement remise en question par l'émergence de nouvelles techniques. CIDOC-CRM offre une branche spécifique pour les concepts liés au temps. Le concept *Entité temporelle E2* regroupe des notions telles que celles de tranche temporelle, période, événement etc. Comme nous étudions des édifices allant du IV^e au X^e siècle, nous travaillons donc sur des intervalles (siècle, début, fin et tiers de siècle). Le concept *Tranche chronologique E52*, qui est une plage temporelle ayant un début, une fin et une durée sans autre connotation sémantique, a été spécialisé pour gérer ces intervalles. Le projet CARE a pour objectif de suivre les états (création, modification, destruction) d'un édifice et de ses éléments. Le concept d'activité est pour nous primordial car un état est caractérisé par une activité dans l'ontologie CIDOC-CRM, nous utilisons donc cette sous-branche de CIDOC-CRM.

Cette méthodologie nous a permis de dégager un ensemble de concepts connectés à des termes architecturaux organisés par des relations méronomiques (décomposition morphologique) et représentationnelles (dimensions, matériaux, relations spatiales) (voir figure 10). La haute technicité des descriptions fournies (nom et organisation des parties d'un édifice, techniques de construction, éléments de décoration, etc.) a rendu possible la réalisation de l'ontologie. Notre spécialisation, développée avec Protégé, comporte 124 classes et 715 individus.

Dans la suite, nous présentons des approches similaires dans le domaine du patrimoine culturel.

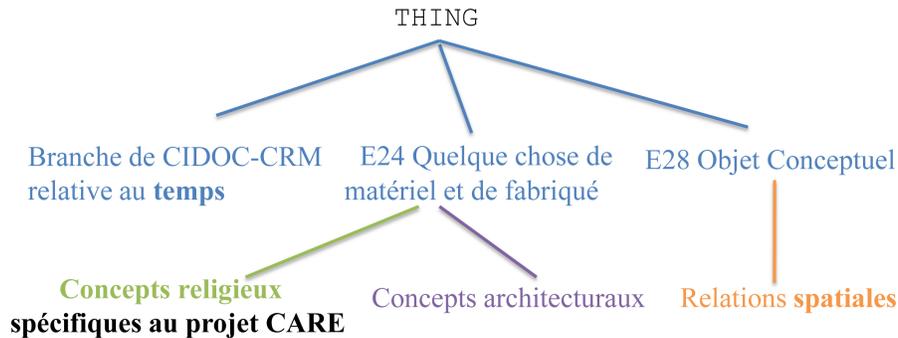


Figure 10. Grandes branches de l'ontologie CARE

5.3. Utilisation des wikis sémantiques dans le domaine du patrimoine culturel

Le projet 3C2MA (« Climat, Catastrophes naturelles et Crises sanitaires des Mondes péri-méditerranéens dans l'Antiquité et au Moyen-Âge ») a pour objet la collecte d'informations historiques concernant le climat, les événements tectoniques et les crises sanitaires ayant touché les pays riverains de la Méditerranée (<http://www.3c2ma.com>). Analysées et interprétées par des historiens associés à des archéologues, des géographes, des médecins et des vétérinaires épidémiologistes, les informations recueillies au sein de textes d'origines multilingues et polygraphes (arabe, hébreux, latin et grec) alimentent une base de ressources termino-ontologiques. Cette base est développée à partir du wiki sémantique SweetWiki (Buffa *et al.*, 2008).

Le projet « *German Handbuch der Architektur* » a pour objectif de construire un wiki au moyen de la numérisation d'un volume (506 pages) du même nom (Witte *et al.*, 2010). Les auteurs cherchent à représenter deux sous-domaines par des ontologies : le domaine de la gestion des documents (phrase, nom, numéro de page, etc.) et le domaine de l'architecture (mur, matériau de construction etc.). Le traitement automatique des langages permet de connecter des concepts architecturaux avec un document spécifique, par exemple, les phrases qui mentionnent des éléments de construction utilisant un matériau donné. Une version publique est disponible à <http://durm.semanticsoftware.info/wiki>. Dans la même perspective, Plantec *et al.* (Plantec *et al.*, 2009) utilisent le traitement automatique des langages pour transformer les pages d'un wiki en pages d'un wiki sémantique dans le domaine des collections scientifiques d'un musée. Les auteurs utilisent Semantic MediaWiki (Krötzsch *et al.*, 2006) et CIDOC-CRM.

Le projet HermesWiki (Reutelshoefer *et al.*, 2010) est un plugin du wiki sémantique KnowWE (Reutelshoefer *et al.*, 2008). L'objectif est de fournir un aperçu concis et fiable de l'histoire grecque ancienne à des étudiants. Une ontologie pour

le domaine historique a été élaborée. Une version publique est disponible à <http://hermeswiki.informatik.uni-wuerzburg.de>.

Dans la même idée, NavEditOW est un cadre pour construire des sites web basés sur des ontologies. Il a été testé dans deux projets : 1) des notices bibliographiques dans un portail sur la Préhistoire et la Protohistoire dans la région italienne (Bonomi *et al.*, 2006), et 2) le projet SilkRoDE dont l'objectif est de recueillir, structurer et diffuser toutes les connaissances sur le patrimoine culturel de l'Asie centrale dans des domaines tels que l'archéologie, la géographie ou l'histoire. NavEditOW a intégré un moteur de wiki pour le rendu des documents stockés (Bonomi *et al.*, 2008). MANTIC est un portail web sur les informations archéologiques de la ville de Milan (Mantegari *et al.*, 2010). Il intègre différentes sources de données et son schéma global est basé sur CIDOC-CRM. Porphyry (<http://www.porphiry.org>) propose un outil de mise en ligne de corpus, un système hypermédia pour la manipulation et l'annotation des documents par les experts. Deux domaines ont permis de valider les concepts : l'histoire sociale du monde chinois du XX^e siècle et des photographies de vases anciens.

6. Conclusion

Dans cet article, nous avons présenté WikiBridge, un wiki sémantique pour les applications collaboratives dans les domaines de la connaissance scientifique. Ce type d'application repose sur des ontologies de domaine. Cependant l'intervention de multiples utilisateurs autour de la plate-forme nécessite de contrôler la cohérence des annotations définies à partir des ontologies. Des vérifications sont mises en place pour trouver les annotations inconsistantes. Ce processus permet d'établir itérativement une connaissance très précise du domaine. De plus, grâce à des requêtes et à l'analyse des données, de nouveaux concepts peuvent apparaître et de nouvelles contraintes peuvent être exprimées. En résultat, les ontologies peuvent être modifiées. Cet axe constitue une perspective de développement de notre wiki sémantique.

Notre expérience d'utilisation de WikiBridge dans le cadre de l'ANR CARE montre des possibilités intéressantes pour la communauté scientifique et ouvre des perspectives de recherche sur les aspects communautaires : 1) la possibilité donnée aux scientifiques de partager et construire de façon collaborative la connaissance ; 2) la possibilité de mettre en place un système de recommandation de termes de l'ontologie en fonction du contexte de l'annotation. Nous envisageons d'étendre ce mécanisme au profil de l'utilisateur (archéologue, médiéviste, historien d'art, etc.) pour lui proposer les termes utilisés par sa communauté. De plus, la composante recherche collaborative peut être mise en place au moyen des requêtes in-line qui constituent elles-mêmes des ressources qui peuvent être annotées.

7. Bibliographie

- Bonomi A., Mantegari G., Vizzari G., « A Framework for Ontological Description of Archaeological Scientific Publications », *Semantic Web Applications and Perspectives (SWAP)*, 2006.
- Bonomi A., Mosca A., Palmonari M., Vizzari G., « Integrating a Wiki in an Ontology Driven Web Site : Approach, Architecture and Application in the Archaeological Domain », *Proceedings of the Third Workshop on Semantic Wikis - The Wiki Way of Semantics (SemWiki)*, 2008.
- Buffa M., Gandon F. L., Erétéo G., Sander P., Faron C., « SweetWiki : A semantic wiki », *Journal of Web Semantics*, vol. 6, n° 1, p. 84-97, 2008.
- Chevalier P., Sapin C., ANR Corpus architecturae religiosae europeae [CARE], saec. IV-X, Technical report, Centre d'études médiévales d'Auxerre, 2008.
- Egenhofer M., Herring J., Categorizing Binary Topological Relations Between Regions, Lines and Points in Geographic Databases, Technical report, National Center for Geographic Information and Analysis, CA, 1991.
- Goulette J.-P., « Sémantique formelle de l'espace. Une application au raisonnement spatial qualitatif en architecture », *Intellectica*, vol. 29, n° 2, p. 9-34, février, 1999.
- Haake A., Lukosch S., Schümmer T., « Wiki-templates : adding structure support to wikis on demand », *Int. Sym. Wikis*, p. 41-51, 2005.
- Iorio A. D., Vitali F., Zacchiroli S., « Wiki Content Templating », *Proceedings of the 17th International Conference on World Wide Web (WWW)*, p. 615-624, 2008.
- Iorio A. D., Zacchiroli S., « Constrained Wiki : an Oxymoron ? », *Int. Sym. Wikis*, p. 89-98, 2006.
- Jiang K., Zhang L., Miyake S., « Using OCL in Executable UML », *ECEASST*, 2008.
- Kiryakov A., Popov B., Terziev I., Manov D., Ognyanoff D., « Semantic annotation, indexing, and retrieval », *Web Semantics : Science, Services and Agents on the World Wide Web*, vol. 2, n° 1, p. 49 - 79, 2004.
- Krötzsch M., Vrandečić D., Völkel M., « Semantic MediaWiki », *International Semantic Web Conference*, p. 935-942, 2006.
- Kurtev I., Bézin J., Aksit M., « Technological Spaces : an Initial Appraisal », *International Symposium on Distributed Objects and Applications (DOA)*, 2002.
- Linster M., « Viewing Knowledge Engineering as a Symbiosis of *Modeling to Make Sense* and *Modeling to Implement Systems* », *6th European Knowledge Acquisition Workshop (GWA)*, p. 87-99, 1992.
- Mantegari G., Palmonari M., Vizzari G., « Rapid Prototyping a Semantic Web Application for Cultural Heritage : The Case of MANTIC », *Proceedings of the 7th Extended Semantic Web Conference (ESWC)*, p. 406-410, 2010.
- Meilender T., Jay N., Lieber J., Palomares F., « Les moteurs de wikis sémantiques : un état de l'art », *Extraction et Gestion de Connaissances (EGC)*, p. 575-580, 2011.
- Motik B., Horrocks I., Sattler U., « Bridging the gap between OWL and relational databases », *Proceedings of the 16th International Conference on World Wide Web (WWW)*, p. 807-816, 2007.

- Plantec A., Ribaud V., Varma V., « Building a Semantic Virtual Museum : from Wiki to Semantic Wiki using Named Entity Recognition », *Companion to the 24th Annual ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*, p. 769-770, 2009.
- Pouillet L., Pinon J.-M., Calabretto S., « Semantic Structuring Of Documents », *3rd Basque International Workshop on Information Technology (BIWIT)*, p. 118-124, 1997.
- Reutelschöfer J., Baumeister J., Puppe F., « Ad-Hoc Knowledge Engineering with Semantic Knowledge Wikis », *SemWiki*, 2008.
- Reutelschöfer J., Lemmerich F., Baumeister J., Wintjes J., Haas L., « Taking OWL to Athens - Semantic Web technology takes Ancient Greek history to students. », *Proceedings of the 7th Extended Semantic Web Conference (ESWC)*, p. 333-347, 2010.
- Sciore E., Siegel M., Rosenthal A., « Using Semantic Values to Facilitate Interoperability among Heterogeneous Information Systems », *ACM Trans. Database Syst.*, vol. 19, n° 2, p. 254-290, 1994.
- Spear A. D., *Ontology for the Twenty First Century : An Introduction with Recommendations*, Technical report, INFOMIS, Sarrbrück, Germany, 2006.
- Uren V., Cimiano P., Iria J., Handschuh S., Vargas-Vera M., Motta E., Ciravegna F., « Semantic Annotation for Knowledge Management : Requirements and a Survey of the State of the Art », *Web Semantics : Science, Services and Agents on the World Wide Web*, vol. 4, n° 1, p. 14-28, 2006.
- Vieu L., « Spatial Representation and Reasoning in Artificial Intelligence », *Spatial and Temporal Reasoning*, p. 5-41, 1997.
- Witte R., Krestel R., Kappler T., Lockemann P. C., « Converting a Historical Architecture Encyclopedia into a Semantic Knowledge Base », *IEEE Intelligent Systems*, vol. 25, n° 1, p. 58-67, 2010.

ANNEXE POUR LE SERVICE FABRICATION
A FOURNIR PAR LES AUTEURS AVEC UN EXEMPLAIRE PAPIER
DE LEUR ARTICLE ET LE COPYRIGHT SIGNE PAR COURRIER
LE FICHER PDF CORRESPONDANT SERA ENVOYE PAR E-MAIL

1. ARTICLE POUR LA REVUE :

ISI. Volume X – n° Y/2012

2. AUTEURS :

Éric Leclercq — Marinette Savonnet**

3. TITRE DE L'ARTICLE :

*Système d'Information pour la production de connaissances : l'approche
wiki sémantique*

4. TITRE ABRÉGÉ POUR LE HAUT DE PAGE MOINS DE 40 SIGNES :

Wikibridge : un wiki sémantique

5. DATE DE CETTE VERSION :

5 octobre 2011

6. COORDONNÉES DES AUTEURS :

- adresse postale :
 - * Laboratoire LE2I - UMR CNRS 5158
 - Université de Bourgogne
 - 9, Avenue Alain Savary
 - 21078 Dijon, France
 - Prenom.Nom@u-bourgogne.fr
- téléphone : 00 00 00 00 00
- télécopie : 00 00 00 00 00
- e-mail : corine.cauvet@lisis.org

7. LOGICIEL UTILISÉ POUR LA PRÉPARATION DE CET ARTICLE :

\LaTeX , avec le fichier de style `article-hermes.cls`,
version 1.23 du 17/11/2005.

8. FORMULAIRE DE COPYRIGHT :

Retourner le formulaire de copyright signé par les auteurs, téléchargé sur :
<http://www.revuesonline.com>

SERVICE ÉDITORIAL – HERMES-LAVOISIER
14 rue de Provigny, F-94236 Cachan cedex
Tél. : 01-47-40-67-67
E-mail : revues@lavoisier.fr
Serveur web : <http://www.revuesonline.com>