



Date : 17/06/2008

Web sémantique et interopérabilité des vocabulaires : une expérimentation dans le domaine des enluminures

Anila Angjeli

Bibliothèque nationale de France
anila.angjeli@bnf.fr

ANTOINE ISAAC

National Library of the Netherlands and Vrije Universiteit
Amsterdam
aisaac@few.vu.nl

Meeting: 129. Classification and Indexing

Simultaneous Interpretation: English, Arabic, Chinese, French, German, Russian and Spanish

WORLD LIBRARY AND INFORMATION CONGRESS: 74TH IFLA GENERAL CONFERENCE AND COUNCIL
10-14 August 2008, Québec, Canada
<http://www.ifla.org/IV/ifla74/index.htm>

L'expérimentation décrite a fait l'objet d'un rapport d'étude, interne à la Bibliothèque nationale de France, rédigé par Thierry Cloarec et Frédéric Martin, Département de la bibliothèque numérique à la Bibliothèque nationale de France en étroite collaboration avec Antoine Isaac et Lourens van der Meij, chercheurs à la Bibliothèque nationale des Pays Bas et à la Vrije Universiteit, Amsterdam.

Résumé :

Pendant la période 2006-2007 une collaboration entre la Bibliothèque nationale de France et la Bibliothèque royale des Pays-Bas a eu lieu autour du projet néerlandais STITCH. Ce projet, à travers des expériences concrètes, explore l'axe sémantique de la recherche. Comment effectuer des recherches sémantiques simultanées à travers plusieurs collections patrimoniales numériques ? Les métadonnées ayant trait à l'analyse du contenu sont souvent hétérogènes. Outre l'exploitation des méthodes de mise en correspondance manuelle des entités sémantiquement comparables, STITCH explore les techniques du web sémantique, notamment l'alignement d'ontologies.

Cette présentation porte sur l'expérience qui a été menée sur deux collections iconographiques numériques, Mandragore, base iconographique du département des Manuscrits de la Bibliothèque nationale de France et la Collection d'enluminures de la Bibliothèque royale des Pays-Bas.

Les contenus des deux collections sont comparables mais le traitement documentaire et les vocabulaires utilisés pour l'indexation de leurs contenus sont très différents. Les deux vocabulaires Mandragore et Iconclass sont certes contrôlés et hiérarchisés mais ils diffèrent

dans leur sémantique et dans leur structure. C'est précisément cette différence qui a intéressé le projet STITCH qui vise à étudier l'alignement automatique des deux vocabulaires.

L'expérience collaborative a démarré avec l'analyse fine de chacun des vocabulaires : les concepts et leur représentation, les propriétés lexicales des termes, les relations sémantiques, etc. L'équipe de chercheurs néerlandais a ensuite étudié et mis en place des mécanismes d'alignement des deux vocabulaires. Les modèles d'origine étant différents, une conversion dans un formalisme pivot (standard commun) était nécessaire pour engager ensuite les procédures d'alignement. C'est le RDF / SKOS qui a été utilisé. L'expérience a abouti à la réalisation d'un prototype qui permet une interrogation simultanée des deux bases à partir d'une même interface. Les descripteurs de chacun des deux vocabulaires servent de termes de requête sur l'ensemble des images appartenant indifféremment à l'une ou l'autre des deux collections.

Cette expérience n'est qu'une étape dans la recherche de solutions visant à faciliter la navigation entre les collections patrimoniales dont les métadonnées sont hétérogènes.

1. Contexte

Les collections patrimoniales affirment de plus en plus leur présence sur le World Wide Web.

Les institutions détentrices de ces collections ont modifié leurs pratiques intégrant progressivement les fonctionnalités et les technologies sous-jacentes du WWW, de l'espace vitrine à la représentation numérique des collections après la mise en ligne de leur référencement. Aujourd'hui de nouveaux défis sont à relever.

C'est l'heure de l'interconnexion et de l'interaction intelligente entre les différentes collections. L'espace de présence et d'action s'élargit et les collections ne sont plus considérées uniquement dans leur dimension institutionnelle, ancrées dans leurs lieux de conservation, mais comme parties d'un trésor européen, voire mondial, qu'il faut interconnecter. Les ambitions des institutions changent de dimension : d'une échelle institutionnelle à une échelle interinstitutionnelle, d'une échelle nationale à une échelle européenne.

Les technologies du web sémantique ouvrent de nouvelles perspectives pour réaliser ces ambitions et les collections patrimoniales deviennent ainsi les candidates idéales. Le but de ces technologies est de faire « travailler » les outils déjà créés par les professionnels de différents domaines, de libérer le sens contenu dans les métadonnées et de connecter

« l'intelligence » qui y est durablement capitalisé. Or typiquement les collections patrimoniales bénéficient de systèmes d'organisation de connaissances d'un haut niveau de professionnalisme et de spécialisation. Les métadonnées associées aux collections sont traditionnellement riches, fines et structurées et s'appuient sur des référentiels régulièrement mis à jour. Cette information, véritable valeur ajoutée, élaborée et accumulée dans le temps, pourra servir de moteur d'interconnexion des collections grâce aux techniques du Web sémantique.

Il n'en reste pas moins que les problèmes à résoudre sont de taille. La grande richesse et variété de ces collections, les différents contextes culturels de leur production, les approches souvent différentes adoptées pour leur traitement intellectuel et pour la production des métadonnées qui leur sont associées sont d'autant d'aspects qui rendent difficile l'interconnexion et l'interopérabilité des collections.

Cet article présente une expérimentation d'interconnexion des collections patrimoniales d'enluminures de la Bibliothèque Royale des Pays Bas et la Bibliothèque nationale de France. Dans la partie 2 nous plaçons cette expérience dans un cadre général de problèmes à résoudre pour l'interconnexion de collections avec des données hétérogènes. Les parties 3, 4, 5 et 6 traitent des spécificités du cas concret, de l'approche choisie et des choix, aussi bien intellectuels que techniques, jugés comme appropriés pour mener à bien l'expérimentation. Dans la partie 6 nous présentons le prototype réalisé pour l'expérimentation et nous concluons dans la partie 7 en mettant l'accent la réutilisation des résultats de l'expérimentation.

2. Problèmes à résoudre et état de l'art de la question

L'accès et l'interconnexion sémantique des collections patrimoniales est un des axes de recherche en cours. Les thèmes, sujets, concepts représentés dans les ressources sont comparables et constituent autant de points d'interconnexion pour rendre ces collections

interopérables. Mais ces collections sont généralement indexées, chacune avec ses langages spécifiques. Ces derniers constituent souvent des outils hétérogènes, conçus et élaborés selon des principes de représentation différents et avec des règles d'indexation qui leur sont propres. L'hétérogénéité réside autant dans le type d'outil et son mode d'élaboration et de représentation (thésaurus, système de classification, liste d'autorité) que dans les variations présentes dans la couverture sémantique de concepts comparables. Les environnements linguistiques hétérogènes se surajoutent également à cette complexité.

Comment permettre à un même public d'accéder à l'ensemble des documents de deux ou plusieurs collections indexées chacune avec un vocabulaire distinct en utilisant indifféremment les termes de l'un ou de l'autre de ces vocabulaires ? Ainsi donc nous souhaiterions utiliser les termes d'un langage de représentation spécifique pour chercher dans les collections qui, en effet, n'ont pas de lien direct avec ce langage ! Pour cela les termes en question doivent être alignés et les vocabulaires concernés doivent être rendus interopérables.

Deux orientations majeures se distinguent dans les travaux en cours. L'une a trait à l'établissement manuel des équivalences entre les termes des vocabulaires différents utilisés pour les collections visées à interconnecter, telle est la démarche adoptée par le projet MACS (Landry, 2007). L'autre, plus récente, utilise les techniques du web sémantique, partant du principe que les vocabulaires contrôlés sont de véritables systèmes d'organisation de connaissances (Knowledge Organisation Systems, KOS) et correspondent donc au type d'artefacts dont la conception constitue à la fois le but et le moyen de la vision du Web sémantique.

En d'autres termes, les techniques du Web Sémantique peuvent donner une nouvelle vie aux outils traditionnels que sont les vocabulaires contrôlés en les exploitant au moyen de

technologies standards développés pour les réseaux. Les vocabulaires contrôlés se voient ainsi utilisés dans des contextes élargis de mise en relation de ressources.

3. Une expérience concrète d'alignement de collections

Des chercheurs néerlandais étudiant ces techniques ont déjà réalisé des expériences dans le contexte des collections néerlandaises dans le cadre du programme de recherche CATCH¹ dont un des objectifs est d'étudier des solutions innovantes permettant l'interconnexion des collections patrimoniales. STITCH (SemanTic Interoperability To access Cultural Heritage)² est un des projets de ce programme visant précisément l'utilisation des technologies d'alignement d'ontologies (Shvaiko & Euzenat, 2005).

Afin d'aborder le contexte multilingue, une expérimentation a été menée associant l'équipe de STITCH avec la Bibliothèque Nationale de France. Celle-ci a porté cette fois-ci sur deux collections dont le contenu est comparable :

- La base des manuscrits enluminés de la Bibliothèque royale des Pays-Bas³ - la plus riche collection iconographique médiévale des Pays-Bas, près de 11 000 miniatures, issues des manuscrits médiévaux enluminés de la Bibliothèque royale des Pays-Bas et du Musée Meermanno, numérisés et accessibles en ligne.
- La base iconographique Mandragore, du département des Manuscrits de la Bibliothèque nationale de France⁴, avec plus de 140.000 miniatures issues des dizaines de milliers de manuscrits des collections de la Bibliothèque nationale de

¹ CATCH : Continuous Access To Cultural Heritage <http://www.nwo.nl/catch>

² STITCH est financé par De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) – Organisation néerlandaise pour la recherche scientifique. Nous voudrions remercier l'institut RKD ainsi que Gerda Duijfjes-Vellekoop pour nous avoir permis d'accéder à Iconclass. Au sein de l'équipe STITCH, Frank van Harmelen, Henk Matthezing et Stefan Schlobach ont contribué à cette expérimentation par leur appui et nombreuses discussions. Pour en savoir plus sur STITCH voir : <http://www.cs.vu.nl/STITCH/>

³ Voir : <http://www.kb.nl/manuscripts/>

⁴ Voir : <http://mandragore.bnf.fr/html/accueil.html>

France dont les plus anciennes remontent à l’Egypte pharaonique et les plus récentes à l’époque contemporaine. En accroissement continu, plus de 50.000 des notices descriptives accompagnent l’image numérisée accessible en ligne.

L’expérience s’est déroulée pendant la deuxième moitié de l’année 2006 pour aboutir à la réalisation d’un démonstrateur en janvier 2007. Elle visait l’étude et l’alignement des vocabulaires Iconclass⁵ et Mandragore⁶ utilisés respectivement pour l’indexation de l’une et de l’autre collection. Cela a nécessité une analyse comparative préalable des modèles natifs des deux vocabulaires.

4. Analyse des vocabulaires

Le tableau ci-dessous met brièvement en regard les caractéristiques générales des deux vocabulaires Iconclass et Mandragore.

Iconclass	Mandragore
<p>Système de classification</p> <p>Élaboré dans les années 1970 par des chercheurs néerlandais pour la description et l’indexation des images - de portée internationale</p>	<p>Vocabulaire contrôlé</p> <p>Élaboré en interne pour les besoins d’indexation des enluminures du Département des manuscrits de la BnF.</p>
<p>Forme : Chaque terme est constitué</p> <ul style="list-style-type: none"> • d’un indice alphanumérique (notation) complexe • d’une définition (ou descripteur sous forme textuelle) • de références transversales (associatives) 	<p>Forme : Descripteurs sous forme textuelle un identifiant numérique est attribué aux descripteurs seulement pour la gestion interne</p> <ul style="list-style-type: none"> • des formes alternatives • des notes d’information

⁵ Voir : <http://www.iconclass.nl>

⁶ Le vocabulaire Mandragore est accessible à partir des écrans de recherche de la base iconographique Mandragore <http://mandragore.bnf.fr/html/accueil.html>, notamment à partir de la recherche par descripteur.

<p>Structure :</p> <p>Fortement hiérarchisé : 10 niveaux de profondeur. Chaque notation hérite de la sémantique des niveaux supérieurs.</p>	<p>Structure :</p> <p>Faiblement structuré : 2 niveaux de profondeur. Une liste de descripteurs à plat. Une hiérarchie à deux niveaux d'environ 150 éléments de classification (inspirés de la Dewey) rassemble les descripteurs selon des domaines généraux.</p>
<p>Langue :</p> <p>multilingue (anglais, allemand, français et italien, partiellement en finnois et norvégien)</p>	<p>Langue :</p> <p>français</p>
<p>Couverture sémantique :</p> <p>descripteurs désignant des objets, des personnes, des évènements et des idées abstraites susceptibles de constituer le sujet d'une image.</p>	<p>Couverture sémantique :</p> <p>descripteurs désignant des objets, des personnes, des évènements et des idées abstraites susceptibles de constituer le sujet d'une image.</p>
<p>Utilisation :</p> <p>selon le besoin, pour exprimer la signification :</p> <ul style="list-style-type: none"> • de scènes entières • d'éléments isolés contenus dans une image 	<p>Utilisation :</p> <p>selon le besoin, pour exprimer la signification :</p> <ul style="list-style-type: none"> • de scènes entières • d'éléments isolés contenus dans une image
<p>Objets informationnels que le vocabulaire vise à analyser :</p> <p>Peintures, dessins, photographies, etc.</p>	<p>Objets informationnels que le vocabulaire vise à analyser :</p> <p>Images contenues dans les manuscrits</p>
<p>Données volumétriques :</p> <p>28 000 descripteurs répartis en dix classes principales. Index alphabétique de 14 000 mots-clés utilisés pour localiser les indices dans le vocabulaire et les descripteurs textuels 40 000 références bibliographiques d'ouvrages ou d'articles d'intérêt iconographique</p>	<p>Données volumétriques :</p> <p>16.000 descripteurs</p>

L'expérience collaborative a été engagée dans un premier temps par l'analyse fine de chacun de ces langages de description: les concepts et leur représentation, les propriétés lexicales des termes, les relations sémantiques, la syntaxe qui lie les concepts, etc. L'équipe de chercheurs

néerlandais a ensuite étudié et mis en place des mécanismes d'établissement de correspondances entre les deux vocabulaires.

5. Web sémantique et alignement de vocabulaires

Le but de l'expérience était de tester les techniques du web sémantique sur des collections patrimoniales numérisées, notamment la technique d'alignement d'ontologies. Cette technique consiste à identifier les correspondances possibles entre les termes des vocabulaires considérés, pour ensuite les soumettre aux experts humains pour réajustement. L'alignement est de ce fait semi-automatique et permet d'économiser considérablement l'effort humain.⁷

Pour procéder à l'alignement des vocabulaires, le formalisme pivot RDF / SKOS⁸ (Simple Knowledge Organisation Systems) a été choisi. Modèle simple et standard de représentation des vocabulaires sur le Web, il permet de représenter en RDF :

- les concepts (Concept) ;
- la description du KOS en question (les cas échéant Iconclass et Mandragore) dans son intégrité, dans un SKOS Conceptscheme ;
- les propriétés lexicales des descripteurs (prefLabel, altLabel) y compris des variantes linguistiques ;
- les relations sémantiques entre les descripteurs (broader, related) ;
- les informations contenues dans les notes (scopeNote, definition).

L'analyse préalable des vocabulaires Iconclass et Mandragore a permis, tout d'abord, de définir les rattachements des éléments composants de chacun des deux vocabulaires aux éléments appropriés du modèle SKOS. Ainsi un terme préféré sera représenté au moyen de

⁷ Une publication de STITCH (Gendt, M. van et al., 2006) fait le point sur ces techniques d'alignement utilisées dans le projet.

⁸ <http://www.w3.org/2004/02/skos/>. Voir aussi (Miles & Bechhofer, 2008) et (Isaac & Summers, 2008) pour un aperçu de cette norme.

l'élément `skos:prefLabel`, un concept sera lié au concept hiérarchiquement supérieur au moyen d'un élément `skos:broader`, une note de définition a été formalisée par l'élément `skos:definition`, etc. Il faut toutefois noter que SKOS, dont le but est de permettre la portabilité des données dans le contexte du Web sémantique, est un modèle simple de représentation et que la conversion des données d'origine ne peut pas se faire sans pertes. De ce fait les modèles d'origine conservent toute leur valeur dans leur contexte initial. .

La connaissance des vocabulaires, acquise lors de l'étape de l'analyse, a permis de choisir la stratégie la plus adaptée pour élaborer les algorithmes d'alignement, en fonction des contraintes du projet.

La communauté du Web sémantique aborde le problème d'alignement d'ontologies par différentes approches, chacune s'appuyant sur un niveau différent d'information contenue dans les vocabulaires contrôlés considérés.

Les approches lexicales utilisent les différents types d'information linguistique contenue dans ces vocabulaires (termes préférés, rejetés, définitions, etc.).

Les approches structurelles utilisent l'architecture hiérarchique et associative qui lie entre eux les concepts d'un vocabulaire.

Les approches statistiques, quant à elles, étendent le champ d'exploration au-delà des vocabulaires eux-mêmes. Elles puisent dans les métadonnées des documents indexés au moyen de ces vocabulaires ; le nombre d'occurrences de tel ou tel concept dans les documents indexés est utilisé comme critère de pertinence pour tester la validité de l'alignement avec un concept candidat dans un autre vocabulaire.

D'autres approches encore vont chercher l'aide d'outils externes pour pallier les défaillances des vocabulaires considérés (des dictionnaires externes pour les problèmes lexicaux, des ontologies externes pour ce qui est de la structure, etc.).

6. Stratégies retenues

Pour les besoins de la présente expérience, parmi les techniques ci-dessus mentionnées, il a été jugé préférable d'utiliser la technique lexicale en testant et combinant plusieurs stratégies. Les architectures des deux vocabulaires présentant des différences très considérables (grand écart de structuration, principe d'héritage opérant dans Iconclass mais totalement absent dans Mandragore, etc.), l'approche structurelle a été estimée inadaptée. La mise en œuvre des autres approches nécessitait plus de moyens et de temps que ne disposait l'équipe de recherche.

Ci-dessous quelques indications sur les stratégies retenues dans l'application de l'approche lexicale pour l'établissement des alignements. Des équivalences ont été établies :

- entre les formes préférées des concepts exprimés avec des formes lexicales identiques, par exemple entre la forme préférée Grange dans Iconclass et sa correspondante parfaite Grange dans Iconclass où elle a également le statut d'une forme préférée ;
- entre une forme préférée dans Iconclass et sa correspondante lexicale se trouvant en forme rejetée dans Mandragore, par exemple entre le terme Enterrement, en forme préférée dans Iconclass et le terme Inhumation, en forme rejetée dans Mandragore.
- entre un terme qui ne représente qu'une partie d'une forme préférée dans un vocabulaire et un terme constituant une forme préférée dans l'autre, par exemple
- entre Hercule est découvert par Junon et Minerve, celle-ci le met au sein de Junon (Iconclass) et le terme junon (Mandragore)

- entre zoologie (généralités) (Mandragore) et le terme zoologie (Iconclass) qui a une portée sémantique plus large que le terme de Mandragore ;
- entre une définition ou une note dans Mandragore et une forme préférée dans Iconclass. Dans Mandragore les définitions donnent des indications permettant de rattacher un concept à d'autres plus généraux. Par exemple, un lien d'équivalence a été établi entre le terme Manche, présent dans une définition dans Mandragore, et le terme mer, en forme préférée dans Iconclass.

Certaines de ces stratégies provoquent du bruit documentaire, elles n'ont pourtant pas été négligées car elles permettent d'obtenir des réponses, un réajustement pouvant se faire à posteriori en combinant les résultats de différentes stratégies.

D'autres problèmes sont restés non résolus. Les modèles originaux des vocabulaires étant complexes et non standards, certaines informations, importantes pour le processus d'alignement, n'ont pu être converties en SKOS et n'ont pu être exploitées pour les alignements.

7. Accès intégré aux collections multilingues : le démonstrateur

L'étape suivante a été la construction d'un démonstrateur⁹ permettant d'accéder simultanément aux collections d'enluminures de la KB et de la BnF soit à partir des termes du vocabulaire Iconclass soit de ceux de Mandragore.

Le démonstrateur n'étant qu'un prototype, il ne tourne que sur un échantillon des collections d'origine, 2170 notices de la collection des enluminures de la BnF et 3987 notices de celle de la KB.

⁹ Le démonstrateur est accessible à l'adresse : http://www.cs.vu.nl/STITCH/BNF_KB_demo.html

Le démonstrateur s'inspire de l'interface à facettes *Flamenco*¹⁰, développée par l'Université de Californie à Berkeley. Elle permet d'affiner les critères de recherche et de les combiner entre eux. La recherche est totalement guidée. Les termes indexés s'affichent dans leur environnement hiérarchique sous forme de listes arborescentes. L'utilisateur lance une requête en cliquant sur un des termes. Il a la possibilité de déployer la hiérarchie et de restreindre ou élargir sa recherche. Les résultats sont recalculés instantanément.

Il est à noter que les alignements sont faits sur la base des termes en français dans les deux vocabulaires. Iconclass est un vocabulaire multilingue, aussi la conséquence indirecte de l'alignement des deux vocabulaires est l'opportunité pour Mandragore de bénéficier de ce multilinguisme. Ainsi les requêtes lancées à partir des termes d'Iconclass en anglais ou en allemand portent également sur les documents de Mandragore, bien que ces derniers soient indexés en français exclusivement.

Le démonstrateur propose deux types d'accès aux collections : un accès simple, par le biais des termes de l'un ou de l'autre vocabulaire, l'autre accès permettant de combiner les termes des deux vocabulaires. Ces deux types d'accès se déclinent en plusieurs facettes :

- une à partir des termes du vocabulaire Mandragore
- une à partir des termes d'Iconclass (au choix : en français, en anglais ou en allemand)
- une combinant le vocabulaire de Mandragore et d'Iconclass
- une pour restreindre l'affichage à une des deux collections.

¹⁰ L'interface à facettes *Flamenco* est présentée sur le site de Berkeley à l'adresse : <http://flamenco.berkeley.edu/index.html>

Les réponses de la requête se présentent sur l'écran sous forme d'un ensemble de vignettes donnant accès directement aux enluminures numérisées et aux notices descriptives de ces enluminures.

8. Conclusion

Cette expérience n'est qu'une base de réflexion dans la recherche de solutions visant à faciliter la navigation entre les collections patrimoniales dont les métadonnées sont hétérogènes. En tant que telle, elle constitue un des multiples « use cases » étudié par le groupe de travail « Semantic web deployment » du W3C (Isaac, Phipps & Rubin, 2007). L'objectif ici est d'étudier des solutions en vue de créer un continuum entre les collections patrimoniales européennes.

Les informations recueillies lors de cette expérience, cumulées avec des informations et enseignements issus d'autres expériences, serviront à d'autres projets et réalisations et conduiront à des améliorations substantielles de ces techniques. Pour illustrer ces propos signalons l'axe « Amélioration de l'accès aux collections » du programme TELplus¹¹. Cet axe vise, entre autres, l'exploration des techniques d'alignement automatique des vocabulaires à l'échelle de tout un ensemble de collections européennes.

Références

Balikova, M. (2005) Multilingual Subject Access to Catalogues of National Libraries (MSAC) Czech Republic's collaboration with Slovakia, Slovenia, Croatia, Macedonia, Lithuania and Latvia. Paper presented at the 71th IFLA General Conference and Council "Libraries - A voyage of discovery", August 14th - 18th 2005, Oslo, Norway. Accessible à l'adresse : <http://www.ifla.org/IV/ifla71/papers/044e-Balikova.pdf>

¹¹Portail du projet TELplus : <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/>

Gendt, M. van, Isaac, A., Meij, L. van der, Schlobach, S. (2006) Semantic Web Techniques for Multiple Views on Heterogeneous Collections: a Case Study. In: Julio Gonzalo et al. (Eds.). Research and advanced technology for digital libraries: proceedings of the 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006), Alicante, Spain, September 17-22 2006. Berlin, Heidelberg: Springer. (Lecture Notes in Computer Science, 4172), 426-437. Accessible à l'adresse :

<http://www.few.vu.nl/~aisaac/papers/STITCH-ECDL06.pdf>

Isaac, Antoine, Phipps, Jon and Rubin, Daniel (Editors) (2007). SKOS Use Cases. W3C working draft, 17 May 2007. Dernière version accessible à l'adresse:

<http://www.w3.org/TR/skos-ucr/>.

Isaac, Antoine and Summers, Ed (Editors) (2008). SKOS simple knowledge organization system primer. W3C working draft, 21 February 2008. Dernière version accessible à l'adresse: <http://www.w3.org/TR/skos-primer/>.

Landry, P. (2007) Multilingual Access to Subjects Access: mise à jour du projet. Présentation à la BnF le 19 janvier 2007.

Accessible à l'adresse : <http://rameau.bnf.fr/informations/pdf/MACS-bnf-2007.pdf>

Miles, Alistair and Bechhofer, Sean (Editors) (2008). SKOS simple knowledge organization system reference. W3C working draft, 25 January 2008. Dernière version accessible à l'adresse:

<http://www.w3.org/TR/skos-reference/>.

Sémantique et interopérabilité. Journée d'étude BnF / AFNOR CG46 Référentiels, données d'autorité, thésaurus, ontologies, taxonomies... Pour en savoir plus ! Bibliothèque Nationale de France, Paris, 28 mars 2008. Accessible à l'adresse :

<http://www.bnf.fr/pages/infopro/journeespro/pdf/AFNOR2008/Isaac.pdf>

Shvaiko, P and Euzenat, J (2005) Ontology Matching. D-Lib Magazine, In Brief, 11(12). December 2005