

RAMEAU et SKOS

Parallèlement, RAMEAU était représenté au I^{er} congrès de l'Association internationale francophone des bibliothécaires et documentalistes (AIFBD) qui se tenait à Montréal, où une communication sur le fonctionnement du FNPR a été donnée.

Après la Pologne et la Roumanie dans les années précédentes, la coopération internationale autour de RAMEAU se développe actuellement avec les pays du Maghreb, Tunisie et Maroc principalement : ainsi, une mission de formation à RAMEAU a-elle été programmée en décembre dernier à Tunis, pendant une semaine.

Enfin, RAMEAU est partie prenante de deux projets européens très prometteurs : d'une part, en relation avec le multilinguisme, le programme MACS (accès multilingue par sujet) qui a bénéficié d'un coup d'accélérateur en 2008, avec son inscription officielle dans le programme européen TELplus et le lancement d'une phase de test sur le moteur de recherche prévue en 2009 ; d'autre part, en relation avec le web sémantique, la traduction à titre expérimental de RAMEAU dans le langage formel SKOS. De tels projets (et l'on pourrait en mentionner d'autres, cette fois en France) montrent qu'un langage comme RAMEAU, tout en poursuivant ses missions traditionnelles, voit s'ouvrir de nouvelles perspectives dans le cadre des évolutions techniques en cours, qui sont autant d'encouragements pour tous ceux qui participent à son enrichissement et à son évolution.

Michel Mingam

 Michel.MINGAM@bnf.fr

Centre national RAMEAU

 www.rameau.bnf.fr

Michel Mingam

 01 53 79 86 68  81 50

 BNF - Centre national RAMEAU

Quai François-Mauriac

75706 PARIS CEDEX 13

LCSH

Library of Congress Subject Headings

MACS

Multilingual Access to Subjects

MeSH

Medical Subject Headings

SKOS

Simple Knowledge Organisation System

Une expérimentation

Le langage RAMEAU, dans le cadre du projet européen TELplus et en collaboration avec la Bibliothèque nationale de France et le Centre national RAMEAU, a fait l'objet, à titre expérimental, d'une première conversion au format web sémantique RDF (Ressource Description Framework), suivant le modèle de représentation SKOS (Simple Knowledge Organization System). Nous présentons dans cet article quelques détails du travail réalisé, ainsi que les perspectives que celui-ci ouvre quant aux usages futurs du langage.

Une brève introduction au web sémantique

La notion de *web sémantique* [Berners-Lee] – ou plus récemment, celle de *web de données* – fait référence à un ensemble de technologies promues par le consortium W3C, permettant de représenter, d'échanger et d'exploiter des informations structurées sur le *World Wide Web*. Les pages web actuelles permettent aisément à un utilisateur humain d'accéder à l'information qu'elles contiennent, et de naviguer vers d'autres sites contenant des informations liées. Cependant, il est beaucoup plus difficile pour des machines – services web, moteurs de recherche – d'exploiter l'information publiée sur le web de façon fine, « intelligente », sans avoir à la réinterpréter via des outils d'analyse linguistique et/ou statistique.

L'objectif du web sémantique est de pallier les déficiences du web actuel en lui ajoutant un niveau supplémentaire, entièrement dévolu à la publication de (méta-)données aisément exploitables par des machines. Concrètement, l'approche du web sémantique est relativement simple mais ambitieuse : publier sur le web, selon un formalisme standard appelé RDF, des connaissances en tant que *descriptions de ressources web*. Pour ceci, tout élément dont on veut donner une description, comme une personne, un livre ou un sujet, se voit attribuer un *identifiant web*, une

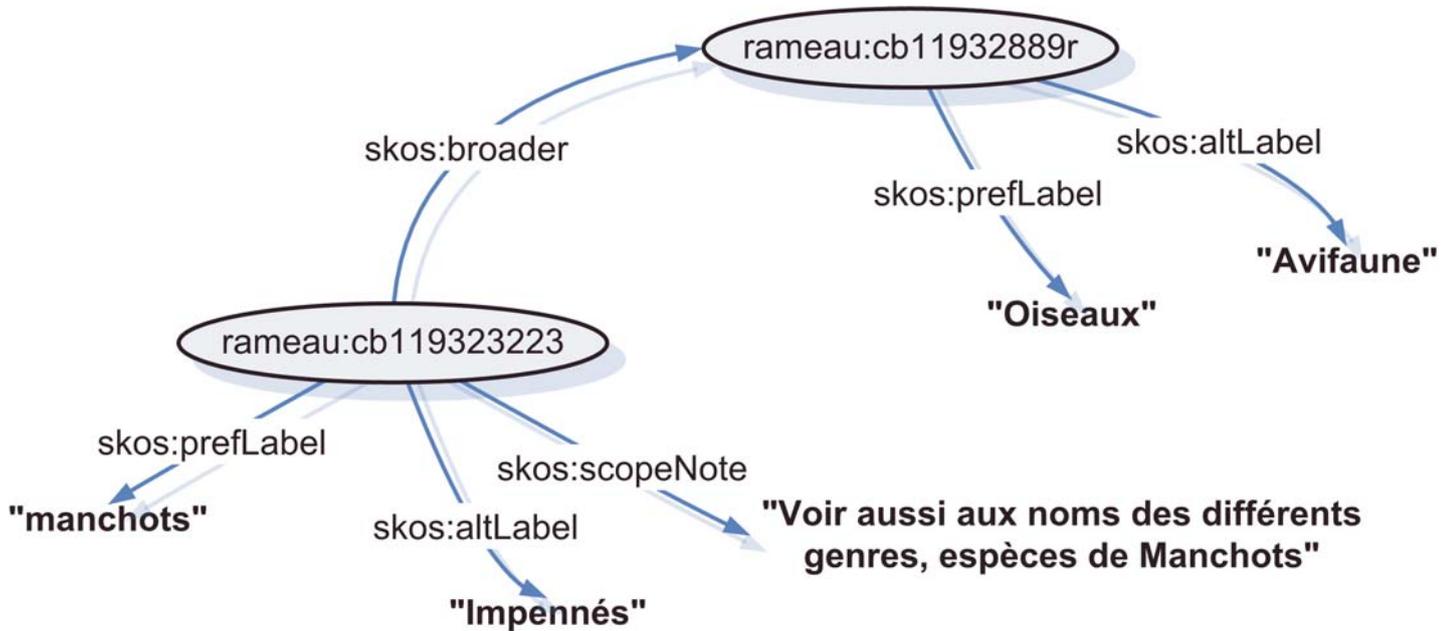
URI : par exemple, *rameau:manchot*. Les URIs sont ensuite liées entre elles par des assertions ternaires, des *triplets*, formalisant le lien qui existe entre elles. Par exemple, (*rameau:manchots*, *a-pour-terme-générique*, *rameau:oiseaux*). De proche en proche, ces assertions simples permettent de constituer un réseau de connaissances complexes. Ce réseau, tout comme celui des sites web « classiques », est accessible via le protocole internet HTTP, soit par le biais des entrepôts interrogeables en utilisant le langage de requêtes SPARQL, soit par celui de serveurs web qui renvoient directement les informations dont ils disposent sur des ressources web de leur domaine – comme un serveur web standard renvoie une page HTML lorsqu'un navigateur l'interroge avec une adresse donnée.

SKOS

RDF ne suffit pas à lui tout seul à représenter des informations réellement partageables et exploitables : tout comme pour XML, il faut lui adjoindre des schémas – on parle d'*ontologies* – qui définissent les briques de base des descriptions, comme le lien *a-pour-terme-générique* de notre exemple. SKOS fournit une ontologie dédiée aux vocabulaires contrôlés et structurés (*knowledge organization systems*) tels que les thésaurus, classifications, taxonomies, etc. [SKOS] Il repose sur un modèle qui se veut à la fois simple et compatible avec une majorité d'approches existantes. À ce titre, il est inspiré par des formats ou des guides comme la norme ISO 2788:1986 pour les thésaurus, mais son objectif n'est pas de remplacer ceux-ci dans leur contexte d'utilisation initial. Il s'agit de faciliter la publication, l'échange et l'interconnexion de ces vocabulaires dans le contexte nouveau du web sémantique.

Mise en correspondance d'Intermarc et SKOS

Les informations lexicales et conceptuelles de RAMEAU sont représentées dans des notices au format Intermarc [Intermarc].



Une fois les entités à décrire de RAMEAU munies d'URIs, produire une description en SKOS de celles-ci implique donc de mettre en correspondance (*mapping*) le modèle d'Intermarc et celui de SKOS, puis de mettre en œuvre un programme de conversion qui réalise cette correspondance. Nous signalons ici les aspects les plus importants de cette mise en correspondance :

- chaque notice (regroupant une vedette matière et les formes rejetées) donne naissance à une ressource de type `skos:Concept` ;
- les données terminologiques (zones 166 et 466 du format Intermarc) sont représentées au moyen des propriétés `skos:prefLabel` et `skos:altLabel` ;
- les liens sémantiques (zones 302 et 301) sont représentés par les propriétés `skos:related` et `skos:broader`, qui pointent vers les concepts correspondants à d'autres notices ;
- les notes, comme la zone 300, sont représentées au moyen de propriétés de documentation dédiées, comme `skos:scopeNote`.

Il est à noter que la conversion n'est pas complète, en particulier concernant les constructions syntaxiques de RAMEAU. SKOS en tant que tel ne permet pas de représenter les notions liées à la coordination de vedettes : tête de vedettes, subdivisions, vedettes construites... Si des usages web sémantique non encore présentés au moment de la création de SKOS et de cette expérience de conversion requièrent d'échanger de telles informations, il faudra alors réutiliser (ou créer) une extension appropriée. Il faut aussi remarquer que notre expérimentation ne propose pas de forme définitive pour les

identifiants web des ressources RAMEAU. Il appartiendra aux instances représentatives du réseau national RAMEAU de décider d'une convention de « nommage » des concepts du langage.

Perspectives

Le travail présenté ici est une étape importante pour mieux exploiter et partager les données de RAMEAU dans le cadre du web sémantique. Ceci peut se révéler très intéressant pour les bibliothèques utilisant RAMEAU et pour lesquelles les problèmes d'interopérabilité et d'ouverture des données deviennent de plus en plus importants. Une telle conversion permet en effet de relier plus facilement et plus explicitement RAMEAU aux objets de son contexte institutionnel et conceptuel. Par exemple, lorsque la Library of Congress aura publié le vocabulaire LCSH en SKOS sur son propre domaine [LCSH], officialisant les expérimentations précédentes [Summers], il sera possible de représenter d'une manière explicite, immédiatement exploitable, les liens entre RAMEAU et LCSH. De façon plus générale, nous espérons qu'une large mise à disposition (sous réserve d'acceptation par les instances représentatives du réseau national) de RAMEAU en SKOS, à savoir selon un format ouvert et compatible avec le web, permettra le développement d'applications innovatrices, dans la communauté RAMEAU actuelle et au-delà.

Antoine Isaac
✉ aisaac@few.vu.nl
Thierry Bouchet
✉ thierry.bouchet@bnf.fr

Références internet
[TELplus]

- <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/> [SKOS]
- <http://www.w3.org/2004/02/skos/> [RAMEAU]
- <http://rameau.bnf.fr/> [Intermarc]
- http://www.bnf.fr/pages/infopro/produits/p-b-aidprod_inter.htm [LCSH]
- <http://id.loc.gov/>

Articles
[Berners-Lee]

- Tim Berners-Lee, James Hendler, Ora Lassila
The Semantic Web
Scientific American 284(5):34-43, 2001
Traduction française
<http://www.urfist.cict.fr/archive/lettres/lettre28/lettre28-22.html> [Summers]
- Ed Summers
Antoine Isaac, Clay Redding, Dan Krech
LCSH, SKOS and Linked Data
Proceedings of the International Conference on Dublin Core and Metadata Applications, (DC 2008)
<http://arxiv.org/abs/0805.2855>

Antoine Isaac

Vrije Universiteit Amsterdam
✉ <http://www.few.vu.nl/>

✉ Vrije Universiteit Amsterdam
De Boelelaan 1081a, 1081HV
AMSTERDAM

Thierry Bouchet

Bibliothèque nationale de France
✉ www.bnf.fr