

arbido

2013/4 Linked Open Data und Big Data

Prongué Nicolas, Deputy Director RERO+

Hügi Jasmin,

Les applications basées sur les LOD en bibliothèque: un tour d'horizon

Les Linked Open Data (LOD) sont devenus aujourd'hui un thème récurrent en bibliothèque, mais la plus-value qu'ils peuvent apporter est encore très difficile à percevoir. Quels résultats concrets ont déjà été atteints? Quelles sont, en bibliothèque, les applications LOD existantes à ce jour? Qu'en est-il de la situation en Suisse?

Tour d'horizon

De nombreux projets de mise à disposition de données bibliographiques en LOD sont en cours de réalisation ou ont abouti. En 2008 déjà, la Library of Congress (Etats-Unis) publiait le LCSH Library of Congress Subject Headings, fichier d'autorités matières en anglais, disponible sur <http://id.loc.gov/authorities/subjects.html> en RDF. Mais le phénomène touche en particulier les bibliothèques européennes. Par exemple, le réseau suédois Libris ¹ fut, la même année, le premier à publier l'intégralité de son catalogue en RDF.

D'autres institutions ont suivi. Plusieurs bibliothèques nationales (Allemagne, France, Grande-Bretagne, Hongrie, etc.) ainsi que des bibliothèques et réseaux universitaires, principalement allemands et français, ont publié certaines de leurs données.

Les données bibliographiques incluent également les fichiers d'autorités, qui sont particulièrement visés par les projets LOD. Ainsi, les fichiers d'autorités auteurs de plus de 30 bibliothèques ont été convertis en RDF et reliés (ils sont accessibles sur le site du VIAF [2](#)), tout comme les fichiers d'autorités matières RAMEAU en français, LCSH en anglais et GND en allemand [3](#). ISAAC, Antoine. RAMEAU subject headings as SKOS linked data. Department of Computer Sciences, University of Amsterdam (en ligne). 2012. [Consulté le 14 octobre 2013]. Disponible à l'adresse: www.cs.vu.nl/STITCH/rameau/. Du fait qu'ils représentent des concepts du monde réel, comme des personnes ou des lieux, ces fichiers ont un potentiel de liaison bien plus grand que les données décrivant des documents: ils peuvent être liés à des données qui ne sont pas d'ordre bibliographique.

Néanmoins, la plupart de ces initiatives, après avoir converti et lié les données, se contentent de les mettre à disposition. Mais au-delà de la simple exposition des données, la possibilité de les exploiter, au moyen d'applications, est encore rarement considérée par les bibliothèques. Très peu d'applications LOD existent à ce jour. Et pourtant, c'est là que réside le réel potentiel de plus value pour l'utilisateur.

Des données aux applications

Une application consiste en un logiciel traitant des données en vue de fournir un service à l'utilisateur. Dans cet article, nous avons donc considéré comme applications LOD en bibliothèque uniquement les services web destinés directement aux utilisateurs et apportant une plus-value basée sur les LOD. Ceci exclut notamment certaines interfaces dont l'utilisation nécessite des connaissances spécifiques, telles que le langage SPARQL par exemple.

Quelques bibliothèques disposent aujourd'hui de services web rendant accessibles les données bibliographiques en LOD. Néanmoins, la plus-value est bien souvent peu perceptible pour l'utilisateur direct, voire inexistante. De tels services sont entre autres proposés par le Sudoc [3](#) (France), Libris (Suède), le réseau HBZ [4](#) (Allemagne), la bibliothèque de l'Université de Winchester [5](#) (Angleterre) ou encore les bibliothèques nationales allemande [6](#), hongroise [7](#) et britannique [8](#).

Il existe néanmoins de réelles applications destinées directement à l'utilisateur lambda. En voici trois des plus abouties:

– Le Virtual International Authority File (VIAF). Maintenu par OCLC, ce portail fournit un accès centralisé aux notices d'autorités auteurs de différents réseaux et grandes bibliothèques, reliées entre elles.

– Le Centre Pompidou virtuel⁹. Le nouveau site du Centre Pompidou à Paris, inauguré en octobre 2012, propose une interface moderne donnant accès à toutes les ressources de ses diverses institutions. On y trouve notamment des notices de documents, des œuvres d’art numérisées, mais également toutes les ressources produites par le Centre lui-même: affiches, vidéos et enregistrements d’événements, dossiers pédagogiques, etc. Cependant, les données ne sont pas disponibles sous une licence ouverte, donc il ne s’agit pas d’une application strictement LOD.

– Data.bnf.fr. Cette application propose un accès unique à trois bases de données de la Bibliothèque nationale de France: le catalogue général, Gallica ainsi qu’Archives et manuscrits. Les données ont été FRBRisées pour être regroupées au sein de pages d’œuvres, d’auteurs et de sujets. De plus, des liens renvoient vers des notices d’autorités externes et des articles Wikipédia. Dans ses pages «ateliers», data.bnf.fr propose également de nouvelles manières d’exploiter les données, par exemple en les présentant sous forme de cartes et de frises chronologiques générées automatiquement.

La plus-value des applications

L’exploitation des LOD au moyen d’applications est encore très peu courante dans les bibliothèques. Pourtant, par rapport à un catalogue en ligne, cela peut apporter plusieurs avantages à l’utilisateur final, parmi lesquels:

La recherche fédérée

Dans data.bnf.fr par exemple, les diverses données, initialement encodées en MARC, Dublin Core et EAD, ont été converties dans un même format du modèle RDF **Bibliothèque Nationale de France. Web sémantique et modèle de données. data.bnf. fr (en ligne)**. [Consulté le 14 octobre 2013]. Disponible à l’adresse: <http://data.bnf.fr/semanticweb>. Grâce à cela, une seule requête permet d’interroger trois réservoirs de manière optimisée. La page d’un auteur fournit ainsi un accès unique à toutes les ressources liées à cet auteur, quels qu’en soient leurs réservoirs de base. Cela permet également, de pair avec la FRBRisation, d’éviter les doublons.

L’enrichissement des catalogues

Le principe de base des LOD réside dans la liaison des données. Les applications peuvent donc proposer des liens vers d'autres ressources, internes ou externes, pour favoriser la découverte et la sérendipité lors de la recherche. Le Centre Pompidou propose par exemple d'augmenter l'expérience de l'utilisateur par de nombreux liens internes, notamment vers des vidéos et des dossiers pédagogiques créés par ses institutions.

Les recherches multilingues

Les LOD sont prometteurs dans ce domaine. Par exemple, lors d'une requête dans The European Library¹⁰, le système propose des sujets connexes au terme recherché, en français, allemand et anglais, grâce à l'alignement des fichiers d'autorités matières RAMEAU, LCSH et GND en format RDF.

La visibilité des catalogues sur le Web

Un format compatible avec les standards du Web permet aux données d'apparaître dans les résultats des moteurs de recherche.

Et en Suisse?

Pour connaître l'état de la situation en Suisse, nous avons contacté les principaux acteurs susceptibles d'être intéressés par le développement des LOD, soit les grandes bibliothèques et les réseaux. Selon les résultats obtenus, un certain retard peut être constaté. En effet, seules quatre institutions sur les onze ayant répondu ont un projet LOD en cours de préparation: RERO, la Bibliothèque nationale, Swissbib et la bibliothèque du CERN. Cependant, rien n'a encore abouti à un résultat concret. RERO, par exemple, a déjà préparé le terrain pour le passage aux LOD, notamment en participant au VIAF, en intégrant RAMEAU, en encourageant la création de notices d'autorités auteurs au sein du réseau ainsi qu'en attribuant des URI à ses données.

Ce retard peut être dû à différents facteurs. D'une part, les institutions suisses sont très petites en comparaison à des géants tels que la Bibliothèque nationale de France, et cela représente un frein. D'autre part, les institutions suisses n'ont pas encore toutes déterminé à qui revient la responsabilité du développement des LOD. Est-ce la tâche des réseaux ou de chaque bibliothèque?

La Bibliothèque nationale doit-elle montrer l'exemple? Faut-il organiser un projet au niveau national? Cette problématique est essentiellement d'ordre stratégique.

Conclusion

Les LOD reposent sur une technologie très jeune et encore en pleine phase de recherche. Très peu d'applications ont déjà été créées par les bibliothèques directement pour leurs utilisateurs. La plus-value des LOD est encore difficile à percevoir: aucune «killer application», révolutionnant la recherche et l'exploitation des données, n'a pour l'instant été inventée. Les bibliothèques suisses, en retard dans ce domaine, devraient se fixer des objectifs et prendre position par rapport aux LOD. Si elles ne souhaitent ou ne peuvent pas encore investir dans un tel projet, elles peuvent du moins déjà soigner et enrichir leurs fichiers d'autorités, qui sont la clé de voûte des LOD en bibliothèque.

1 <http://libris.kb.se>.

2 Virtual International Authority File: <http://viaf.org/>.

3 www.sudoc.abes.fr.

4 Hochschulbibliothekszentrum (Rhénanie-du-Nord-Westphalie): www.lobid.org/.

5 <http://prism.talis.com/winchester-ac>.

6 <https://portal.dnb.de/>.

7 <http://nektar2.oszk.hu>.

8 <http://bnb.data.bl.uk>.

9 www.centrepompidou.fr/.

10 www.theeuropeanlibrary.org/tel4.



Nicolas Prongué

Nicolas Prongué is information specialist and deputy director at RERO+. He studied information science while working in several libraries, in Geneva and La Chaux-de-Fonds. After obtaining his Master's degree at the HEG Genève, he worked as a research assistant in this same institution, before being hired by RERO+ for its transition project.



Jasmin Hügi

spécialiste en information documentaire

Abstract

Deutsch

LOD beruht auf einer sehr jungen Technologie, die sich noch im Forschungsstadium befindet. Es gibt erst sehr wenige Anwendungssoftware, die von Bibliotheken für ihre Nutzer entwickelt wurde. Der Wert der LOD kann noch nicht beziffert werden: Es wurde bis jetzt keine eigentliche «Killer-Applikation» erfunden, die die Suche und die Verwaltung von Daten revolutioniert hätte. Die Schweizer Bibliotheken, die diesbezüglich der Entwicklung etwas hinterher hinken, müssen in Bezug auf LOD Stellung beziehen und sich Ziele setzen. Wenn die Bibliotheken sich noch nicht in solche Projekte einbringen können oder wollen, so können sie zumindest ihre Autoritätsdaten, die den Zugangsschlüssel zu LOD in Bibliotheken darstellen, anreichern und pflegen.