

Linked Open Data: quelles nouvelles compétences pour les professionnels de l'information?

Les LOD ne représentent pas la première thématique qui remet en question les compétences voire la formation des professionnels de l'information. Ces dernières années, la littérature scientifique a proposé des réflexions par rapport à des programmes d'études en lien avec les bibliothèques numériques, la «Digital Curation» (gestion des actifs numériques) ainsi qu'avec la gestion de données de recherche.

Des études concernant les nouvelles compétences des bibliothécaires

Les LOD ne représentent pas la première thématique qui remet en question les compétences voire la formation des professionnels de l'information. Ces dernières années, la littérature scientifique a proposé des réflexions par rapport à des programmes d'études en lien avec les bibliothèques numériques, la «Digital Curation» (gestion des actifs numériques) ainsi qu'avec la gestion de données de recherche.

Concernant les LOD en bibliothèque, encore très peu d'études ont été réalisées pour identifier les compétences qui devraient être enseignées aux futurs professionnels de l'information. Deux projets méritent d'être mentionnés dans ce contexte-là. D'une part, un curriculum concernant les LOD en général a été développé pour le projet EUCLID [1](#), sans considérer les institutions culturelles. Sur la base de ce curriculum, des matériels d'apprentissage sous forme de webinars, de tutoriels et d'un cahier d'accompagnement ont été élaborés et mis à disposition en ligne. Afin de définir les connaissances nécessaires pour travailler quotidiennement avec les LOD, des experts de différentes industries ont été interrogés. D'autre part, le projet Learning-Linked-Data [2](#) s'est focalisé spécifiquement sur les besoins des professionnels de l'information. Ce projet, qui s'est terminé en septembre 2012, avait pour but de faire un répertoire des sujets d'apprentissage afin de développer une plateforme soutenant les enseignants et les étudiants dans l'interprétation et la création des données.

Offres d'emploi et avis d'experts

Le Web des données et les LOD représentent un sujet très technique, avec lequel les frontières entre le métier d'informaticien et celui de professionnel de l'information s'estompent. L'étude qui est présentée dans cet article vise à identifier les compétences prioritaires à acquérir pour qu'un professionnel de l'information puisse travailler avec des LOD en bibliothèque. Pour ceci, des offres d'emploi trouvées en Suisse, en Allemagne, en Autriche et en France ainsi que des avis d'experts internationaux ont été analysés afin de dresser un premier constat. Seules les offres d'emploi exigeant un diplôme en science de l'information, en plus des compétences liées aux LOD, ont été retenues. Quant aux experts, ils devaient, pour être sélectionnés, avoir étudié les sciences de l'information ou avoir travaillé en tant que bibliothécaire, en plus de leur expérience en LOD. La méthodologie sera expliquée en détail dans un rapport qui paraîtra à la fin de cette année. Au final, huit offres d'emploi et douze réponses d'experts ont été prises en compte.

Huit compétences-clés

La nature technique des LOD avait une certaine influence sur les réponses. Ainsi, ces dernières mentionnent beaucoup de détails techniques, tels que des langages, des formats ou encore des logiciels. L'analyse du contenu des offres d'emploi et des réponses a permis d'identifier plus de 40 éléments de compétence différents. La figure 1 présente les compétences qui apparaissent dans le nombre de cas le plus élevé. Pour donner un exemple: si une compétence est mentionnée dans chacun des 20 cas, la compétence a une haute priorité. Si une compétence n'apparaît que dans un seul cas, sa priorité est basse (même si la compétence est mentionnée très souvent à l'intérieur du cas).

Ontologies et métadonnées

Dans cette première catégorie, il s'agit de connaître des schémas de métadonnées tels que Dublin Core, savoir évaluer et choisir un élément de métadonnées pour l'appliquer à ses propres données, savoir créer des vocabulaires contrôlés ainsi que des ontologies, et savoir effectuer un alignement de deux vocabulaires différents.

Développement web

Cet élément englobe des connaissances concernant le fonctionnement des protocoles web et des serveurs, ainsi que les principes du Web design. La maîtrise de ces connaissances est nécessaire afin de savoir créer et attribuer des URIs aux triplets, les publier en ligne ainsi que créer des interfaces web pour que les utilisateurs puissent interagir avec.

Triplets

En troisième position arrivent les compétences en lien avec les triplets, c'est à dire la connaissance des spécifications RDF émises par le W3C, qui représentent la base pour la création de triplets. Il est également important de savoir développer et modéliser un graphe qui peut être appliqué aux données. Puisque les triplets peuvent être exprimés à travers différentes syntaxes (RDF/XML, Notation 3, etc.), il est utile de comprendre ces dernières ou de connaître des outils qui peuvent les produire.

Web sémantique et Linked Data

Il s'agit de connaître les principes émis par le W3C et par Tim Berners-Lee, ainsi que leurs objectifs. Cette catégorie inclut également la compréhension basique de l'inférence ou du raisonnement, par exemple.

Standards de catalogage

Les compétences concernant les standards de catalogage sont essentielles dans un projet LOD, car pour transformer des données en triplets, il faut bien connaître les données de base. Il s'agit donc de comprendre les règles de saisie des données (telles que les AACR2), le format d'enregistrement (tel que MARC), ainsi que le fonctionnement du SIGB utilisé (tel que Virtua).

Programmation

Certains experts interrogés ont évoqué la connaissance basique d'au moins un langage de programmation (Java, Ruby, PERL, etc.) comme compétence nécessaire. Cela servirait avant tout de moyen de communication avec les informaticiens. Ces connaissances peuvent également être utiles pour la programmation de codes simples qui permettent la conversion en masse de données en LOD.

Technologies du Web sémantique

Les données organisées en triplets nécessitent une base de données spécifique (un «triple store»), un langage d'interrogation spécifique (SPARQL) et une interface d'interrogation spécifique (SPARQL Endpoint). La maîtrise de ces technologies est fortement souhaitée pour travailler avec les LOD.

XML

Enfin, dernier élément mais non des moindres, des connaissances de XML et des langages annexes comme XSLT ont été mentionnées dans les cas analysés. XML est à la base de nombreux standards du Web sémantique, et il est donc indispensable de comprendre et de savoir exploiter des données structurées en XML.

Un monde changeant

Le travail avec des LOD en bibliothèque nécessite visiblement des compétences très spécifiques et au premier abord, cela peut sembler intimidant. Néanmoins, il est important de noter que les compétences déjà existantes parmi les professionnels de l'information se retrouvent tout à fait dans les éléments cités dans le diagramme. Ainsi, la première compétence touche aux ontologies et aux métadonnées, qui sont utilisées pour la description des ressources, et ceci a toujours fait partie des connaissances de base des professionnels de l'information. Concernant les standards de catalogage, il est inutile d'évoquer que cela appartient à la bibliothéconomie. En outre, les formations de bachelor en information documentaire en Suisse contiennent des modules d'enseignement de développement web et de XML. Par rapport à la programmation, la frontière entre informaticien et professionnel de l'information se montre fluctuante, et l'attribution de cette compétence à l'un des deux métiers représente un sujet de discussion.

Ce qui manque réellement en ce moment, ce sont les compétences strictement liées au Web sémantique (principes, standards et technologies). Selon nous, l'enseignement de ces compétences devrait être inclus dans le curriculum des professionnels de l'information. Une introduction obligatoire à la thématique ainsi que des cours à choix devraient être dispensés au niveau bachelor, alors que les étudiants au niveau master devraient atteindre une compréhension globale afin d'être capables de prendre des décisions stratégiques. Des cours de formation continue sont par ailleurs à mettre en place pour que les professionnels sur le terrain puissent eux aussi se former aux LOD. En effet, le Web sémantique est bien plus qu'une tendance, et le monde des bibliothèques ne peut se permettre d'attendre.

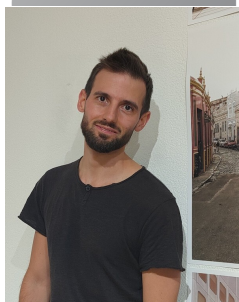
1 www.euclid-project.eu.

2 <http://lld.ischool.uw.edu/wp/>.



Jasmin Hügi

spécialiste en information documentaire



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.

Abstract

Deutsch

Die Linked Open Data (LOD) verbreiten sich progressiv in der kulturellen Domäne. Eine wachsende Anzahl an Projekten entstehen in Institutionen und LOD-Applikationen halten Einzug in Bibliotheken. Dieser Einzug stellt gewisse Kompetenzen der I&D Fachleute infrage, stärkt andere und erfordert neue Kenntnisse und neues Know-how. Dieser Artikel präsentiert die Resultate einer Studie, die zum Ziel hat, diese neuen Kompetenzen, die sich I&D-Fachleute angesichts LOD aneignen müssen, zu bestimmen.