

La normalisation en bibliothèque à l'heure du web sémantique

Anila Angjeli

► **To cite this version:**

Anila Angjeli. La normalisation en bibliothèque à l'heure du web sémantique. Documentaliste - Sciences de l'Information, ADBS, 2011, 48 (4), http://www.cairn.info/article.php?ID_ARTICLE=DOCSI_484_0042&DocId=45005&Index=

HAL Id: hal-00875829

<https://hal-bnf.archives-ouvertes.fr/hal-00875829>

Submitted on 22 Oct 2013

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.

91

Repérer. Les standards du web sémantique concernent également le monde des bibliothèques, qui possède par ailleurs ses propres normes métier. Au-delà des transpositions et nécessaires adaptations, c'est la pensée normalisatrice qui s'en voit renouvelée.

92

Le web sémantique offre enfin aux bibliothèques une possibilité sans précédent de tirer bénéfice des données structurées de leurs catalogues mais aussi de sortir de leur isolement, dû à la conception de systèmes d'information en silos, aux technologies vieillissantes des formats Marc non adaptés aux technologies du Web. Ce dernier a besoin aujourd'hui d'une information typée, structurée, normée, qui tient compte de la nature des données. Relier ces dernières par le sens, tel est le postulat du web sémantique. Dans la grande partition du Web intelligent, si le W3C bâtit le socle normatif technique, les normes métier ont à jouer leur partie au niveau du « sémantique », car elles fixent le sens en traduisant des concepts métier.

Un existant – déjà complexe

93 Notre propos se limitera ici aux normes concernant les bibliothèques qui rendent possible la production, la diffusion, l'échange de l'information produite et l'accès à cette information. En premier lieu sont concernées les normes, règles et recommandations appliquées dans la réalisation des catalogues – celles relatives à la description bibliographique et aux dispositifs associés facilitant l'accès aux catalogues.

En font partie le bon vieux ISBD, les normes Afnor de catalogage¹ et autres codes, mais aussi les normes de description archivistique, principalement la norme ISAD(G).

Les normes jouent aussi un rôle pour les identifiants de produits éditoriaux (ISBN, ISSN, ISMN), de contenus (Isan, ISWC, ISTC) ou d'« agents » (Isni). N'oublions pas les normes génériques pour le codage des caractères, des écritures, des noms de langues et de pays, ou les normes de translittération. S'ajoutent à cet arsenal les standards des formats de structuration des données bibliographiques et d'autorité (Marc et Iso 2709, puis MarcXchange, et Dublin Core), et les protocoles d'échange afférents (Z39.50, SRU, OAI-PMH)

¹http://www.bnf.fr/fr/professionnels/normes_francaises/s.cat_normes_francaises.html

Mais où s'arrêter ? L'adaptation à un environnement en perpétuelle mutation amène les bibliothèques à s'impliquer dans des travaux de normalisation qui auparavant étaient hors de leur périmètre d'action : standards de métadonnées relatifs à la gestion des objets numériques, comme METS, d'encodage du contenu des documents comme la TEI, et la liste est longue. La création en 2010 du groupe d'incubation du W3C « Bibliothèques et web de données » en est la preuve la plus récente.

Dans ce contexte mouvant, le paysage des acteurs reste complexe. Des organismes de normalisation internationaux tiennent le palmarès, comme l'ISO - TC 46, où les professionnels participent *via* les organismes nationaux de normalisation tels que l'Afnor pour la France, Ansi et Niso pour les États Unis, BSI pour le Royaume Uni, ou DIN pour l'Allemagne, qui par ailleurs produisent des normes nationales. Néanmoins, de nombreux autres organismes, experts dans leur domaine, rallient les professionnels autour de travaux de normalisation spécifiques. Mentionnons ici des associations professionnelles comme l'Ifla, la SAA, l'ICA, des organisations interprofessionnelles comme la DCMI, mais également des institutions isolées comme la Library of Congress, ou des organismes dédiés à la réalisation d'un projet comme le JSC pour l'élaboration de RDA, etc. Ces travaux de normalisation aboutissent non seulement à des normes à proprement parler – statut réservé à ce qui est produit par des organismes de normalisation dédiés – mais aussi à des standards, codes, bonnes pratiques, recommandations, tous aussi valables dès lors qu'ils sont adoptés largement par les communautés intéressées.

97 Interopérabilité, partage, échange – principes historiquement fondamentaux du processus de normalisation - resurgissent sous un nouvel angle. Bon nombre de normes ont été élaborées bien avant le web sémantique et, de ce fait, hors de ses problématiques. Pour autant, leur valeur n'est pas remise en cause par les développements récents.

Transposer sur le web de données

98 Typiquement, des normes génériques existantes ont pu facilement être portées sur le web sémantique. Citons ici le cas des normes Iso 639-2 et -5 des codes de langues. La Library of Congress, autorité d'enregistrement Iso de ces normes, a attribué aux codes de celles-ci des URIs pérennes, permettant leur transposition sur le web des données.

99 Les standards traditionnels métier font ainsi alliance avec les standards du web sémantique. Des initiatives se multiplient pour exprimer ces standards en logique formelle ontologique et en RDF avec, en parallèle, la déclaration des espaces de nom, l'attribution

d'URI et la mise à disposition sur des registres ouverts, en vue de leur réutilisation aisée sur le web des données, par des machines ou des humains. À titre d'exemple, mentionnons l'expression en RDF et la déclaration des éléments de Dublin Core. Le mouvement se poursuit avec l'expression, sous forme d'ontologies Owl, d'autres standards existants comme Mads, ou le schéma EAC-CPF. Il se prolonge avec des projets de publication sur le web de données d'éléments des vocabulaires associés aux modèles et normes de l'Ifla, comme c'est le cas des modèles FRBR, FRAD et de l'ISBD.

100 Mais ne banalisons pas le propos, car l'esprit du web sémantique et ce qui gravite autour infiltrent la pensée normalisatrice, avec pour corollaire un changement de posture dans les travaux. La matière à normaliser est pensée différemment. Dès le processus d'élaboration des normes et standards, corrélations et synergies entre les travaux de normalisation et d'autres initiatives sont recherchées. Les travaux adoptent une dynamique itérative, celle inhérente au mouvement du web sémantique.

Élaborer en vue du web sémantique

101 Une illustration exemplaire est l'élaboration du nouveau code international de catalogage RDA, dont l'objectif est la modernisation de la production des données catalographiques et la structure des catalogues. Dès l'origine, il est fondé sur un modèle conceptuel métier (FRBR) et vise à concevoir l'information bibliographique comme un réseau d'informations interconnectées. Ses éléments et valeurs sont, au fur et à mesure, déclarés et publiés dans un registre de métadonnées, exprimés en RDF et identifiés par des URI.

102 Mais, au-delà du code lui-même, c'est le processus d'élaboration qui illustre cet esprit. S'y entrecroisent des collaborations avec d'autres initiatives dans un souci d'interopérabilité et de synergie. Ainsi le groupe RDA/ONIX, qui travaille à faciliter le transfert et l'utilisation de la description des ressources dans les deux communautés des bibliothèques et des éditeurs, a rendu son premier travail : un cadre ontologique distinguant enfin le contenu du contenant. Parallèlement, d'autres initiatives continuent dans ce sens et d'autres se mettent en place, comme l'étude des passerelles entre DCMI et RDA, l'harmonisation entre RDA et l'ISBD, etc. En France, au sein de la commission CG46 Documentation de l'Afnor, un groupe technique et un groupe stratégique étudient l'opportunité d'adoption du code et ses implications au niveau national²

²Ressource Description and Access (RDA) : en France http://www.bnf.professionnels/rda/s.rda_en_france.html

Dernier-né de la famille des identifiants ISO, l'Isni est un identifiant unique, international, pour toute partie impliquée dans les maillons de la chaîne, allant de la création des contenus intellectuels et artistiques à la gestion des droits, qu'il s'agisse d'une personne ou d'une collectivité. Bien que les travaux de normalisation aient démarré en dehors de la problématique du web sémantique, ils ont eu pour toile de fond l'économie du numérique et les besoins et usages qu'elle génère. Émanant d'une collaboration entre acteurs des secteurs public et privé, Isni se positionne en identifiant passerelle. Il est destiné à relier entre eux des systèmes d'information, à la fois ceux contenant des données sensibles et ceux destinés à une diffusion large de données. La connexion se fera *via* la « couche » publique des données, susceptible d'être exposée sur le Web. Les identifiants Isni et les métadonnées publiques qui les accompagnent (notamment des URI vers des réservoirs contenant des informations d'autorité sur l'identité en question) seront exposés sur le web des données en RDF, de même que les liens entre Isni et Viaf.

104 Un identifiant unique pour les « identités publiques » des parties, ne tenant compte ni de leur appartenance nationale, ni de leur secteur d'activité, de portée mondiale et dont le succès réside dans l'adoption large par tous les secteurs concernés ! Avec ces caractéristiques, Isni s'inscrit pleinement dans la logique d'interopérabilité par l'identification des objets informationnels à mettre en relation. D'autres collaborations sont en cours : l'Orcid pour les acteurs du domaine scientifique et académique et le Niso I² pour les institutions.

105 On le voit : le mouvement du web sémantique ne fait que confirmer les motivations et principes fondamentaux de la normalisation : consensus des parties prenantes, optimisation des processus de travail, retombées économiques, interopérabilité. La tension permanente entre normalisation, au sens de réduction des variétés, et besoin d'expression des diversités et des spécificités s'atténue par la mise en avant du *sens*.

Les sigles en jeu

Ansi : American National Standards Institute | **BSI** : British Standards institution | **DCMI** : Dublin Core Metadata Initiative | **Din** : DeutschesInstitutfürNormunge.V. | **EAC-CPF** : Encoded Archival Context – Corporate bodies, Persons and Families | **FRAD** : Functional Requirements for Authority Data | **FRBR** : Fonctionnal Requirements for Bibliographic Records | **ICA** : Conseil international des archives | **Isad(G)** : International Standard Archival Description-General | **Isan** : International Standard Audiovisual Number | **ISBD** : International standard bibliographic description | **ISMN** : International Standard Music Number | **Isni** : International Standard Name Identifier | **ISTC** : International Standard Text Code | **ISWC** : International Standard musical Work Code | **JSC – RDA** : Joint Steering Committee for Development of Ressource Description

and Access | **Mads** : Metadata Authority Description Schema in RDF | **Marc** : Machine-Read Cataloging | **Mets** : Metadata Encoding & Transmission Standard | **Niso** : National Information Standards Organization | **OAI-PMH** : Open Archives Initiative – Protocols for Metadata Harvesting | **Onix** : Online Information Exchange | **Orcid** : Open Researcher & Contributor ID | **SAA** : Society of American Archivists | **SRU** : Search/Retrieval via URL | **TEI** : Text Encoding Initiative | **Viaf** : Virtual International Authority File •