

FRAD

Modélisation des données d'autorité : les dernières nouvelles !

Le chantier entrepris par le groupe FRANAR¹ de l'IFLA, il y a tout juste dix ans, va bientôt être achevé avec l'accomplissement des trois tâches de son mandat :

- la définition des fonctionnalités requises des notices d'autorité,
- l'étude de faisabilité d'un numéro international normalisé pour les données d'autorité,
- la mise en liaison avec les autres communautés concernées par les données d'autorité.

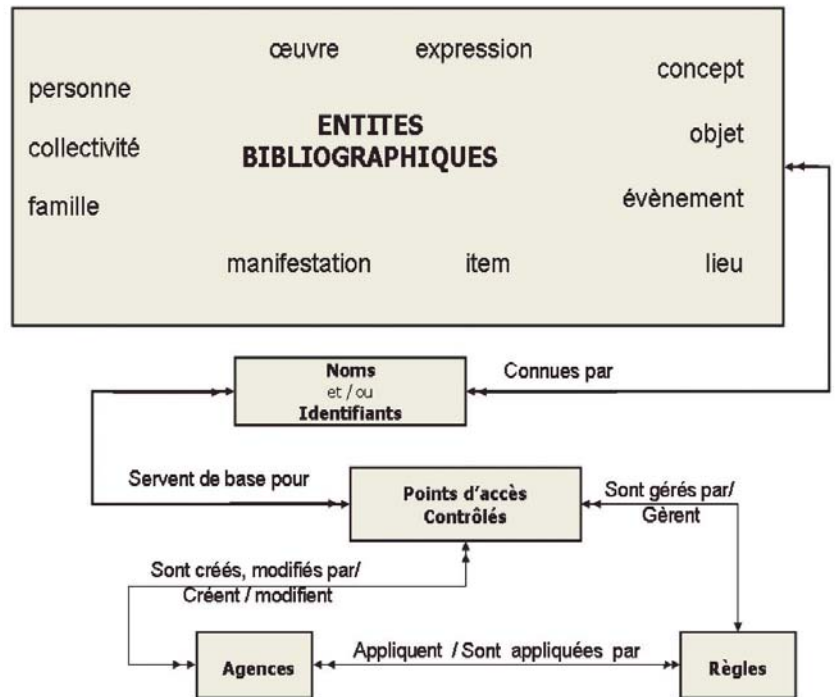
Le groupe a déjà rendu les résultats de l'étude de faisabilité d'un numéro international normalisé pour les données d'autorité dans un texte d'analyse et de positionnement publié sur le site de l'IFLA.² Il a aussi mis en place des échanges et des confrontations d'idées avec d'autres communautés, telles que les archivistes et les gestionnaires de droit. Cet article revient sur l'objectif principal du groupe FRANAR, à savoir la définition des fonctionnalités requises des notices d'autorités, qui s'accomplira avec la publication du modèle conceptuel **FRAD** (Functional Requirements of Authority Data), publication attendue d'un jour à l'autre³. Est ainsi couronnée une décennie d'efforts pour élaborer **un cadre permettant l'analyse des fonctionnalités requises pour supporter le contrôle d'autorité et pour partager des données d'autorité sur le plan international**. Au bilan des travaux : d'innombrables révisions successives du texte à travers lesquelles le nom du modèle s'est vu transformé de FRANAR à FRAR pour se stabiliser enfin en FRAD (où il n'est plus question de notices mais de données d'autorité), deux enquêtes internationales et... une forte attente vis-à-vis du modèle. Mais pourquoi toute cette attente ?

Les évolutions incessantes, notamment celles du web et des technologies sous-jacentes, avec les mouvements comme **le web sémantique**, misent sur la connexion de la connaissance et de l'intelligence accumulées dans les données. Dans ce contexte, les données d'autorité se révèlent comme le complément intelligent des catalogues de bibliothèques. De plus, elles sont utili-

sées dans d'autres applications en dehors des bibliothèques, notamment dans le domaine de la gestion des droits. Si le modèle FRBR a tant révolutionné la pensée catalographique, il lui manquait toutefois le volet « données d'autorité ». En effet, ce modèle n'analyse les concepts de l'univers bibliographique que du point de vue de la notice bibliographique et de l'accès à cette notice, à l'exclusion du volet « données d'autorité ». C'est chose faite avec le modèle FRAD.

Extension du modèle FRBR aux données d'autorité, le modèle FRAD est également élaboré selon la même méthode : celle **entité-relation**. Il reprend les mêmes objets informationnels que le modèle FRBR, les entités bibliographiques : **Œuvre, Expression, Manifestation, Item, Personne, Collectivité, Concept, Objet, Évènement, Lieu** en ajoutant l'entité **Famille** et s'intéresse à la nature intrinsèque de ces entités. Il analyse leurs propriétés et caractéristiques et en fait des **attributs** de ces entités. Il définit également les **relations** que ces entités peuvent entretenir dans l'univers bibliographique. Prenons par exemple l'entité **personne**. Le modèle FRBR traite des personnes uniquement dans le lien qu'elles peuvent avoir avec les ressources cataloguées et ne va pas plus loin que les caractéristiques d'un nom de personne en tant que point d'accès dans la description bibliographique. FRAD, en revanche, s'efforcera de répondre de manière explicite aux questions comme : Qu'est une personne en soi dans l'univers bibliographique ? Comment se fait-il qu'une même personne physique puisse

constituer deux entités bibliographiques (cas d'une même personne exerçant des activités différentes sous des noms ou pseudonymes distincts) ? Quels sont les **attributs** dont on a besoin et pourquoi en a-t-on besoin (par exemple les dates biographiques et autres qualificatifs servant à identifier celle-ci et à la différencier des éventuels homonymes) ? Quelles sont les relations qu'une personne peut entretenir avec les autres entités définies dans le modèle, comme celles d'une personne avec une collectivité en tant que membre de cette collectivité, avec une famille ou avec une autre personne, etc. ? **Par ailleurs FRAD modélise aussi les processus et les mécanismes de gestion des données d'autorité**. Ainsi, sont modélisés en tant qu'entités les acteurs qui créent et gèrent les données d'autorité – les **Agences** bibliographiques. Ce sont elles qui établissent les règles si importantes, notamment pour l'établissement des points d'accès. Les **Règles**, aussi bien que les **Points d'accès**, sont également modélisés comme des entités distinctes. Et enfin, chacune des entités bibliographiques (**œuvre, expression, manifestation, item, personne, collectivité, famille, concept, objet, évènement et lieu**) est connue par des noms et peut avoir des identifiants, comme c'est le cas de l'ISBN pour les manifestations, ou du numéro SIRET pour les collectivités. Les **Noms** et les **Identifiants** sont des éléments essentiels des systèmes d'information et le modèle FRAD en fait également des entités à part entière et les inscrit dans le schéma fonctionnel du modèle.



Contexte du projet

L'Institut de l'information scientifique et technique (INIST) a développé le portail TermSciences [www.termssciences.fr] en partant de deux constats :

1 la production terminologique des différents acteurs de la recherche est assez riche et diverse mais insuffisamment valorisée ;

2 la gestion des ressources terminologiques étant une tâche fastidieuse, une mise en commun de données et la fédération de compétences permettraient de mieux l'appréhender.

La réalisation du portail terminologique TermSciences vise donc à constituer un référentiel terminologique commun qui permet d'établir des passerelles entre les différents termes et variantes utilisés pour désigner une même notion tout en s'inscrivant dans une démarche de normalisation des données, d'interopérabilité des systèmes et de collaboration entre les spécialistes. Outre la valorisation et la mise en commun de la production terminologique, cet outil offre aux partenaires du projet la possibilité de gérer de manière collaborative le contenu ainsi produit et d'utiliser ce contenu pour l'indexation documentaire et/ou la recherche d'information dans les bases de données spécialisées ou sur le web.

Base de données terminologique

La base terminologique du portail TermSciences résulte de la fusion de plusieurs ressources terminologiques produites par différents organismes de recherche :

- les vocabulaires d'indexation de l'INIST-CNRS ;
- le thésaurus MeSH bilingue de la National Library of Medicine, traduit par l'Inserm avec la participation de l'INIST ;
- le thésaurus de la banque de données en santé publique (BDSP) ;
- des dictionnaires et lexiques produits par l'INRA ;
- le thésaurus de linguistique (thésaulangue) de l'ATILF-CNRS (Analyse et traitement informatique de la langue française) ;

Mais quelles opérations peut-on effectuer avec les données d'autorité ? Si à l'origine, le contrôle d'autorité est né d'un objectif – celui d'identifier et de contrôler la forme des points d'accès aux notices bibliographiques d'un catalogue – le souci permanent d'identifier personnes, collectivités, œuvres, etc. a conduit à un enrichissement progressif des fichiers d'autorité avec des informations supplémentaires. Ces fichiers sont devenus, ainsi, de véritables réservoirs d'informations aptes à être réutilisés pour de multiples applications, même en dehors du contexte des catalogues des bibliothèques. FRAD prend note de ce constat et définit les tâches que les différents utilisateurs, bibliothécaires ou autres chercheurs d'information, peuvent effectuer en utilisant les données d'autorité. Ces tâches sont : **Trouver, Identifier, Contextualiser et Justifier**. Le modèle s'efforce également de répondre à la question : quelles relations et quels attributs sont nécessaires pour effectuer telle ou telle tâche ? Par exemple, la précision du type de relation entre deux personnes, tel que la relation de *collaboration* ou la relation *parent/enfant*, sert à effectuer la tâche contextualisation, à savoir à clarifier le contexte de vie de ces personnes ainsi que leurs liens avec les contenus et documents catalogués.

Enfin quel est l'impact de FRAD ?

Les évolutions en cours dans le domaine du traitement bibliographique n'ont pas attendu la publication du modèle pour en intégrer ses concepts. Il est déjà devenu la référence en matière de données d'autorité dans l'univers bibliographique. Ainsi, les nouveaux Principes internationaux de catalogage⁴ s'appuient sur FRAD pour tout ce qui concerne les données d'autorité. De même les RDA⁵ en cours d'élaboration se veulent respecter FRAD.

Et ensuite ? Notons que pour compléter l'univers « FR.. » un autre modèle est en cours d'élaboration par un groupe de travail de l'IFLA, le groupe FRSAR (Functional Requirements for Subject Authority Records)⁶. Vous l'avez compris, il s'agit de la modélisation des données d'autorité relatives à l'indexation matière et de l'indexation matière elle-même.

Rendez-vous donc sur le site de l'IFLA, dans un premier temps, pour l'accès au texte final du modèle FRAD et, espérons le bientôt, pour des nouvelles sur le modèle FRSAR.



A. Angjeli

 anila.angjeli@bnf.fr

BNF

Direction des services et des réseaux
Département Information
bibliographique et numérique (IBN)

Anila Angjeli

 01 53 79 53 95  85 86

1 FRANAR (Working group on Functional Requirements and Numbering of Authority Records) est le nom du groupe de travail de l'IFLA <<http://www.ifla.org/VII/d4/wg-franar.htm>> ; FRAD est le nom du modèle conceptuel élaboré par ce groupe (IFLA : International Federation of Library Associations and Institutions).

2 A review of the feasibility of an International Standard Authority Data Number (ISADN) / prepared for the IFLA Working Group on FRANAR By Barbara B. Tillett ; Edited by Glenn E. Patton, 1 July 2008. <<http://www.ifla.org/VII/d4/franar-numbering-paper.pdf>>

3 En attendant la publication de la version finale du modèle FRAD, c'est la version soumise à l'enquête internationale en 2007 qui est accessible sur le site de l'IFLA <<http://www.ifla.org/VII/d4/franar-conceptual-model-2ndreview.pdf>>.

4 La dernière version de la déclaration des *Principes internationaux de catalogage*, soumise à l'enquête internationale est accessible en ligne <http://www.ifla.org/VII/s13/icc/imeicc-statement-of-principles-2008_french.pdf>

Le texte final a été validé par l'IFLA et est en cours de publication.

5 Resource Description and Access <<http://www.collectionscanada.gc.ca/jsc/rda.html>>

6 FRSAR : groupe de travail de l'IFLA <<http://www.ifla.org/VII/s29/wqfrsar.h>