



**HAL**  
open science

Mise en œuvre d'une nouvelle structuration des données  
pour valoriser le fonds documentaire de l'IPE  
UNESCO : l'expérience FRBR

Aurélie Bergonnier Signoles

► To cite this version:

Aurélie Bergonnier Signoles. Mise en œuvre d'une nouvelle structuration des données pour valoriser le fonds documentaire de l'IPE UNESCO : l'expérience FRBR. domain\_shs.info.docu. 2011. mem\_00679338

HAL Id: mem\_00679338

[https://memic.ccsd.cnrs.fr/mem\\_00679338](https://memic.ccsd.cnrs.fr/mem_00679338)

Submitted on 15 Mar 2012

**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.

CONSERVATOIRE NATIONAL DES ARTS ET MÉTIERS  
Ecole Management et Société - Département CITS

INTD

MÉMOIRE pour obtenir le  
Titre professionnel « Chef de projet en ingénierie documentaire » INTD  
niveau I

Présenté et soutenu par  
*Aurélie Bergonnier Signoles*

le 24 novembre 2011

Mise en œuvre d'une nouvelle structuration des  
données pour valoriser le fonds documentaire de l'IPE  
UNESCO :  
l'expérience FRBR

Jury  
Mme Nadia RAIS  
Mme Asunción VALDERRAMA

Promotion XLI

# Remerciements

Mme Nadia Raïs, Professeur associé au CNAM, pour son enseignement, ses enseignements et sa patiente maïeutique.

Mme Asunción Valderrama, responsable de Centre de Documentation de l'Institut international de planification de l'éducation de l'UNESCO pour son accueil chaleureux et son dynamisme communicatif mais aussi son "intuition FRBR" et son soutien dans ce projet au long cours.

Toute l'équipe du Centre de Documentation de l'IPE pour son accueil et son soutien bienveillant, et surtout Corinne Bitoun pour son implication, en particulier dans l'équipée fantastique que fut la partie orientée objet de l'expérience.

Les experts, en particulier M. Patrick LeBœuf et Mme Anila Angjeli de la BnF pour leurs conseils et leurs orientations, mais aussi Mme Françoise Leresche (BnF), M. Patrick Verrier (Cité de la Musique) et Mme Catherine Di Sciullo (Gaîté Lyrique).

L'équipe pédagogique de l'INTD CNAM pour la qualité de la formation dispensée et l'équipe administrative pour son soutien constant.

Tous les collègues de la promotion 2009-2011 en souvenir de nos échanges fructueux et des moments intenses vécus au CNAM.

# Notice

BERGONNIER SIGNOLES Aurélie. Mise en œuvre d'une nouvelle structuration des données pour valoriser le fonds documentaire de l'IIPÉ UNESCO: l'expérience FRBR. INTD-CNAM, 2011, 152 p. Mémoire en ingénierie documentaire.

Dans le contexte d'une organisation internationale, l'Institut international de planification de l'éducation, une réflexion est engagée sur une nouvelle structuration de l'information. L'objectif est de rendre interopérables des bases de données hétérogènes et cloisonnées. L'expérience menée s'appuie d'abord sur le modèle conceptuel FRBR entité relation développé pour les bibliothèques puis sur un formalisme orienté objet afin de mieux prendre en compte les documents d'activité.

FRBR, métadonnée, modélisation, base de données, structuration de données, interopérabilité, catalogue

# Table des matières

<b>REMERCIEMENTS</b> .....	<b>2</b>
<b>NOTICE</b> .....	<b>3</b>
<b>TABLE DES MATIERES</b> .....	<b>4</b>
<b>LISTE DES SIGLES</b> .....	<b>8</b>
<b>LISTE DES TABLEAUX</b> .....	<b>9</b>
<b>LISTE DES FIGURES</b> .....	<b>10</b>
<b>INTRODUCTION</b> .....	<b>11</b>
<b>PREMIERE PARTIE CLES DE COMPREHENSION</b> .....	<b>14</b>
<b>1 « L'information numérique et ses avatars » : document, ressource, données</b> .....	<b>15</b>
1.1 Document / objet documentaire / unité documentaire .....	15
1.1.1 Le document : une articulation information/ support .....	15
1.1.2 Évolution et complexification du document.....	16
1.1.2.1 Flux électroniques versus monographie ?.....	17
1.1.2.2 Les enjeux de la complexification du document pour les métiers de l'information documentation .....	17
1.2 Ressources .....	18
1.3 Données .....	18
1.4 Traiter l'information brute .....	19
<b>2 Structurer l'information : des mutations techniques, fonctionnelles et opérationnelles</b> .....	<b>20</b>
2.1 Les supports et leurs évolutions : catalogues et bases de données.....	20
2.1.1 Définitions: le catalogue et ses fonctions .....	20
2.1.2 Rappels historiques : les catalogues sur fiche.....	21
2.1.3 Support du catalogage, la base de données.....	22
2.2 Décrire et enregistrer : normes et formats.....	23
2.2.1 La normalisation de la description bibliographique .....	23
2.2.2 Enregistrer les notices : les formats documentaires informatiques.....	24
<b>3 Décrire et identifier pour trouver et obtenir sur le web</b> .....	<b>25</b>
3.1 Décrire, identifier sur le web : les métadonnées .....	25
3.1.1 Contexte et périmètre des métadonnées .....	25
3.1.2 Typologie des métadonnées .....	26
3.1.3 Comment écrire la métadonnée pour décrire la ressource .....	27
3.1.3.1 Utiliser des métadonnées généralistes.....	27
3.1.3.2 Utiliser des métadonnées par types de ressource .....	28

3.2 Trouver et obtenir sur le web : évolution des usagers et des usages .....	29
3.2.1 Le contexte de l'évolution du comportement des usagers .....	29
3.2.2 De nouvelles exigences .....	29
3.2.2.1 Les exigences des utilisateurs pour l'amélioration des catalogues .....	30
3.2.2.2 Les améliorations préconisées par les bibliothécaires.....	31
3.2.3 Les catalogues en évolution : d'un outil local à un outil global .....	31
3.2.3.1 Changer les pratiques : mutualiser, partager.....	32
3.2.3.2 Devenir visible sur le web.....	34
<b>4 Mettre les données dans le web .....</b>	<b>34</b>
4.1 L'évolution des langages .....	34
4.1.1 Les langages de description des formats de documents : HTML, XML .....	34
4.1.2 Ouvrir les métadonnées MARC au web avec le XML .....	35
4.2 La modernisation des règles de catalogage.....	36
4.3 Vers le Web de données par le web sémantique.....	37
4.3.1 L'évolution vers des données liées.....	37
4.3.2 Les enjeux du web sémantique.....	38
4.3.3 Comment identifier la donnée? .....	38
4.3.4 Les vocabulaires du web .....	39
<b>5 Un outil de médiation des données : le modèle conceptuel FRBR.....</b>	<b>41</b>
5.1 Présentation des FRBR .....	41
5.1.1 Définition .....	41
5.1.2 Un cadre théorique .....	42
5.1.3 Au centre du modèle : l'utilisateur .....	42
5.2 La structure du modèle .....	44
5.2.1 Définition des entités et articulation de l'arbre FRBR .....	44
5.2.1.1 Les entités structurelles du groupe 1 .....	45
5.2.1.2 Illustration.....	46
5.2.1.3 Approfondissement des entités du groupe 1 .....	47
5.2.1.4 Les entités de responsabilité et de sujet .....	49
5.2.2 Les relations entre entités .....	49
5.2.3 Les attributs .....	51
5.2.4 Les identifiants des entités.....	52
5.3 L'utilité du modèle FRBR <sub>ER</sub> .....	53
5.3.1 Le modèle entité relation, un support de réflexion théorique .....	53
5.3.2 Le modèle entité relation, une utilité pratique ? .....	54
5.3.2.1 Les conséquences pour l'utilisateur final.....	54
5.3.2.2 Les conséquences pour le gestionnaire du système d'information .....	55
5.4 L'élargissement : les modèles liés .....	57
5.4.1 FRAD et FRSAD.....	57
5.4.1.1 Les données d'autorité .....	57
5.4.1.2 Les autorités sujet .....	57
5.4.2 Vers un formalisme orienté objet .....	58
5.5 Du modèle théorique aux applications.....	59
5.5.1 Mettre en œuvre les FRBR dans les catalogues?.....	59
5.5.2 Des exemples de catalogues inspirés du modèle FRBR .....	61
5.5.2.1 Le catalogue bibliographique de la Cité de la Musique .....	61
5.5.2.2 D'autres expériences de catalogues inspirés du modèle FRBR .....	64
5.5.2.2.1 En France : le catalogue du Centre de ressources de la Gaîté lyrique.....	64
5.5.2.2.2 OCLC et WORLDCAT.....	64
5.5.2.3 Tableau de synthèse des expérimentations FRBR .....	65

## **DEUXIEME PARTIE APPLICATION DES FRBR AU CADRE D'UNE ORGANISATION INTERNATIONALE : L'EXPERIENCE MENEES SUR LES BASES INTERNES DE L'IIPE UNESCO..... 73**

<b>1 Présentation de l'IIPE.....</b>	<b>74</b>
1.1 L'institut international de planification de l'éducation.....	74
1.1.1 Une organisation internationale, l'IIPE.....	74
1.1.2 Les missions de l'IIPE.....	74
1.1.3 Ses spécificités.....	75
1.2 Le Centre de documentation de l'IIPE.....	76
1.2.1 Le personnel du Centre de documentation.....	76
1.2.2 Le profil des usagers.....	77
1.3 Les fonctions documentaires.....	78
1.4 Les documents du Centre de Documentation.....	80
1.5 Les bases de données de l'IIPE.....	81
1.5.1 Les bases internes.....	81
1.5.1.1 La base de signalement des activités du Governing Board (GB).....	81
1.5.1.2 La base des documents internes : INTEPI.....	82
1.5.2 Les bases en libre accès : la bibliothèque virtuelle.....	83
1.6 Outils et pratiques de catalogage.....	85
<b>2 Identification des besoins.....</b>	<b>87</b>
2.1 Les constats.....	87
2.2 Les objectifs.....	89
<b>3 Expérimentation des modèles entité relation et orientés objet : le parcours.....</b>	<b>91</b>
3.1 Scénario 1 : expérimentation du modèle FRBR.....	91
3.2 Scénario 2 : une démarche objet.....	97
3.2.1 Etape 1 : Modélisation des activités en CIDOC-CRM.....	97
3.2.2 Etape 2 : expérimentation avec FRBRoo.....	100

## **TROISIEME PARTIE LES PRECONISATIONS DANS LE CONTEXTE DE L'IIPE .....106**

<b>1 Les FRBR : un moyen de structurer l'information.....</b>	<b>107</b>
1.1 Retour sur expérience.....	107
1.2 L'avenir des FRBR à l'IIPE.....	108
1.2.1 Expérimenter le modèle FRBR.....	109
1.2.2 Cataloguer le fonds avec le code de catalogage RDA.....	110
1.2.3 Vers un autre système de gestion intégrant le modèle FRBR ?.....	111
<b>2 Mettre en place d'autres moyens de structurer l'information : le knowledge management.....</b>	<b>112</b>
2.1 La méthode.....	112
2.1.1 Typologie des documents IIPE reçus et enregistrés.....	113
2.1.2 Analyse sommaire du fonctionnement du circuit de l'information.....	115
2.2 Des solutions pour capitaliser les savoirs.....	115
<b>3 Attendre ...tout en se préparant au web de données.....</b>	<b>117</b>
3.1 Mettre en place des identifiants.....	117
3.1.1 Les initiatives à l'échelle internationale.....	117
3.1.2 A l'échelle de la France.....	119
3.2 Suivre les initiatives open data des institutions culturelles.....	119

3.2.1 Le projet data.bnf : diffuser les données de la BnF sur le Web .....	120
3.2.2 Décloisonner les institutions du savoir .....	122
3.3 D'autres pistes tous azimut .....	123
3.3.1 Collaborer avec la FAO.....	123
3.3.2 Envisager des solutions alternatives pour mettre les données sur le web.....	124
3.3.3 Mettre les données sur le web : la piste des éditeurs de logiciels .....	125
<b>CONCLUSION .....</b>	<b>127</b>
<b>BIBLIOGRAPHIE.....</b>	<b>129</b>
<b>ANNEXE 1 : FRBR QUIZ.....</b>	<b>148</b>
<b>ANNEXE 2: HYPOTHESE DE TRAVAIL : ESSAI DE FRBRISATION DES BASES INTERNES A L'AIDE DU MODELE CONCEPTUEL FRBR<sub>ER</sub> .....</b>	<b>151</b>



## Liste des sigles

**AACR** : Anglo-American Cataloguing Rules

**CCF** : Common Communication Format

**CIDOC** : Comité International pour la DOcumentation de l'ICOM (International Council of Museums)

**CRM** : Conceptual Reference Model

**FRAD** : Functional Requirements for Authority Data

**FRBR** : Functional Requirements for Bibliographic Records

**FRSAD** : Functional Requirements for Subject Authority Data

**FULBI** : Fédération des Utilisateurs de Logiciels pour Bibliothèque, Documentation et Information

**GB** : Governing Board

**IFLA** : International Federation of Library Associations and Institutions

**ISBD** : International Standard Bibliographic Description

**IIEPE** : Institut international de Planification de l'Éducation

**PIC** : Principes internationaux de catalogage

**PMT** : Plan à moyen terme

**RDA** : Resource Description and Access

# Liste des tableaux

Tableau 1 : Sélection de caractéristiques identifiant les entités structurelles du groupe 1

Tableau 2 : Document de synthèse sur la mise en œuvre des modèles conceptuels FRBR entité relation et orienté objet

Tableau 3 : Typologie des documents IIPÉ reçus et enregistrés au Centre de documentation

## Liste des figures

Figure 1 : Les entités FRBR du groupe 1, de l'abstrait au concret

Figure 2 : Exemple d'affichage avec le modèle FRBR

Figure 3 : Entités FRBR et principales relations verticales

Figure 4: Les ressources de l'IPE en 2010-2011

Figure 5 : L'organisation fonctionnelle du Centre de documentation de l'IPE

Figure 6 : Schéma simplifié de l'organisation des bases de données de l'IPE

Figure 7 : Représentation de l'interface de résultat "idéale" regroupant les ressources des trois bases

Figure 8 : Modélisation avec le CIDOC CRM

Figure 9 : De l'activité CRM aux œuvres FRBR<sub>oo</sub>

Figure 10 : Œuvre individuelle, agrégative et traductions

Figure 11 : Publication et identifiants

# Introduction

Ce mémoire professionnel est le résultat d'une expérience menée pour valoriser les données d'un système d'information.

Le contexte de l'étude est celui du fonds documentaire d'un institut de l'UNESCO spécialisé dans la planification de l'éducation : l'IIEP (Institut international de planification de l'éducation).

La responsable du Centre a eu l'idée d'expérimenter le modèle conceptuel FRBR de modélisation des données des bibliothèques pour améliorer l'accès aux données de ses bases internes (séparées entre base des activités et base des documents internes) et composées de ressources hétérogènes. Ce cloisonnement entre bases engendrait une perte d'information donc d'efficacité du système d'information documentaire.

L'objectif est d'améliorer l'accessibilité de l'information, mission fondamentale d'un Centre de documentation. Il s'agit pour cela de mieux structurer les informations produites, collectées par le Centre et stockées sur les bases internes. Le but est de relier les activités de l'IIEP aux produits qu'elles génèrent (documents de la base interne) mais aussi aux publications de l'IIEP et aux productions de ses chercheurs (base bibliographique) afin de permettre à l'utilisateur d'accéder à partir d'une plate forme unique à des informations interconnectées. Cette expérience s'est déroulée suite à la réflexion engagée sur la mise en place d'un nouveau logiciel de traitement documentaire (ré informatisation du fonds documentaire).

Nous essaierons d'évaluer dans quelle mesure la mise en place d'une nouvelle structuration de l'information basée sur le modèle conceptuel FRBR peut valoriser un fonds documentaire. Pour cela, nous engagerons une réflexion sur l'accès à l'information, et en amont sur la façon d'organiser l'information afin d'améliorer les services offerts. Il nous faudra aussi dépasser cette problématique d'accès à l'information et d'amélioration de l'interface pour réfléchir à une échelle plus fine sur une nouvelle structuration des données.

Plus globalement, ce mémoire se situe dans le contexte des mutations engagées dans le secteur de l'information et de la documentation.

D'une part on assiste à une multiplication des canaux d'information, dans le contexte concurrentiel des informations diffusées par le web, mais aussi à des exigences et des pratiques nouvelles des usagers.

D'autre part, l'information diffusée change de nature et de volume (augmentation des ressources électroniques et de leur accessibilité et profusion de données / métadonnées). Enfin, avec le numérique, les frontières traditionnellement définies tant au niveau des institutions patrimoniales que des catégories d'objets documentaires s'estompent. De nouvelles problématiques émergent, rendant nécessaires l'identification des objets numériques et l'interopérabilité des systèmes d'information.

L'étude aborde en premier lieu les bases théoriques nécessaires à la compréhension du système d'information de l'IIEP et de l'expérience FRBR : définition des éléments constitutifs d'une information et leur évolution dans un contexte numérique, mais aussi présentation des mutations techniques, fonctionnelles et opérationnelles engagées dans les milieux de l'information documentaire. Ces éléments didactiques correspondent à la découverte personnelle d'un « tout nouveau monde ». Ils sont cependant nécessaires pour comprendre et "aligner" des vocabulaires dont la signification varie selon les métiers. Nous détaillerons ensuite le modèle FRBR, sa structure, son intérêt et l'illustrerons avec des exemples de catalogues l'ayant mis en œuvre. Ce modèle ayant d'abord et surtout été conçu pour les bibliothèques, nous insisterons tout particulièrement

sur la structure des données prévues dans les catalogues bibliographiques et leurs adaptations possibles dans l'optique du web de données.

La deuxième partie présente le contexte de l'expérience : un fonds documentaire dont les activités et les ressources sont tournées vers l'international avec des problématiques spécifiques (linguistiques notamment). L'originalité de ce mémoire est de présenter une expérience se basant sur un modèle conceptuel récent (1998 pour FRBR dans sa version entité relation et 2008 pour le formalisme orienté objet) pour lequel peu d'études ont été jusqu'alors menées. L'intérêt est aussi de l'appliquer in vivo à des documents hétérogènes qui ne correspondent pas à un fonds de bibliothèque classique. La mise en application du modèle FRBR dans le contexte d'un système d'information reposant sur des activités sera détaillée, avec le parcours suivi et la méthode mise en œuvre.

Nous ferons enfin un point sur l'applicabilité du modèle à l'IIPE et préconiserons les pistes à suivre pour structurer autrement -non seulement les données- mais plus globalement l'information produite.

Première partie  
Clés de compréhension

Avant d'aborder les outils de structuration et de signalement de l'information, il convient d'abord de clarifier la terminologie employée (document, ressources, données) dont le sens a évolué avec l'environnement numérique.

En effet, les éléments de compréhension que nous allons aborder dans ce chapitre seront ensuite transposés dans un cadre autre que bibliographique.

# 1 « L'information numérique et ses avatars » : document, ressource, données

Malgré leur différence d'échelle, un document, une ressource, une donnée, ont comme point commun d'être des informations primaires.

Afin lever les ambiguïtés, étudions les éléments qui les différencient.

## 1.1 Document / objet documentaire / unité documentaire

L'objectif d'un système d'information documentaire est d'assurer la gestion et la communication des informations utiles contenues dans des documents. « Cette gestion implique le repérage et la validation des sources et ressources d'information, la collecte, le traitement, le stockage, la diffusion, ainsi que les moyens de classement et de récupération des documents porteurs d'information en utilisant des outils classiques ou informatisés »<sup>1</sup>.

### 1.1.1 Le document : une articulation information/ support

Un document est un « ensemble formé par un support et une information » (ISO TC-46), un « support porteur d'informations »<sup>2</sup>. Cette définition est large : elle recouvre une information de différente nature (texte mais aussi son, image fixe, image animée ou multimédia) accessible sous un ou différents supports (papier, numérique, multimédia). Le document traditionnel est inhérent à son aspect matériel, comme « lié » de manière permanente et unique à un support physique. [7]

Anne-Marie Chabin approfondit la définition du document en introduisant la notion d'activité et de cycle de vie. Au cours de son cycle de vie, le document passe par trois statuts qui correspondent à trois intervalles : [2] [3]

1- un statut de production : l'information est élaborée, mise en forme, validée puis diffusée ;  
2- un statut de trace : l'information validée par son auteur (et figée) présente un intérêt de preuve, de gestion ou de mémoire ;

3- un statut de source de connaissance : on arrive au stade de l'information.

---

<sup>1</sup> BOULOGNE Arlette [et al.], INTD- er. Vocabulaire de la documentation. Paris, ADBS éditions, 2004. 338 p. Sciences et techniques de l'information. ISSN 1762-8288. ISBN 2-84365-071-2

<sup>2</sup> CACALY Serge, LE COADIC Yves- François, POMART Paul- Dominique, SUTTER Eric. Dictionnaire de l'information. 3e éd. Paris, Armand Colin, 2008. 295 p. ISBN 978-2-200-35132-8



L'objet soumis à un traitement documentaire pour devenir source de connaissance peut être catégorisé en fonction de :

- la forme des données enregistrées : texte écrit, image, son, données...
- la nature des données : statistiques, références bibliographiques
- du support d'enregistrement : papier, audio-visuel (magnétique, argentique), numérique
- la forme éditoriale, liée au mode de production des documents : livre, périodique, compte-rendu, rapport, thèse, brevet, carte, annuaire, répertoire...
- du contenu documentaire : document primaire (document originel/source), secondaire (références bibliographiques, métadonnées), tertiaire (synthèses, bilans...)
- du champ disciplinaire [...]
- la source : document interne (compte-rendu, rapport d'activités, étude), administratif (courrier, mél), document externe (littérature éditée, périodiques)...
- du statut : confidentiel (accessible en lecture seule ou aussi en écriture) ou du moment de son cycle de vie (en relecture, validé...). [6]

Tout objet peut devenir un document s'il est considéré du point de vue des informations dont il est porteur : un ticket de métro, un vase, une bouteille identifiée par son étiquette.

Ainsi, les activités d'une institution, ensemble d'actions décomposables en tâches recouvrant différentes opérations, génèrent des documents. Ces documents d'activité sont les produits d'une chaîne de fabrication précédée d'une décision administrative. Ils sont modélisables à l'aide de processus, « ensemble d'activités corrélées ou interactives qui transforme des éléments d'entrée en éléments de sortie »<sup>3</sup>.

Avec le numérique, le document n'obéit plus à la définition classique imbriquant un support et une information : il évolue et se complexifie. [8]

### 1.1.2 Évolution et complexification du document

Le document numérique présente trois caractéristiques : [5]

- sa granularité c'est-à-dire sa capacité à être séquencé, coupé, égrainé en unités informationnelles porteuses de sens
  - sa plasticité. Cela signifie la capacité du document à changer de forme, « à recomposer des unités informationnelles selon des usages, des besoins, des situations différentes »
  - sa traçabilité qui répond à l'exigence de la preuve, essentielle à la notion de document
- Quelles sont les conséquences pour l'objet documentaire ?

Les conséquences sont que l'objet documentaire:

1- se multiplie : le numérique génère de nouveaux objets documentaires qui sont des documents indépendants du support, des objets hybrides, des documents structurés à « la syntaxe utilisable ». Au titre de ces nouveaux objets documentaires, on pourrait d'ailleurs inclure l'internaute. En effet, il documente ses actions, laisse des traces <sup>4</sup> quand il est acteur, génère, met en ligne des contenus (photos, opinions, ou blogs) indexés par les moteurs de recherche. La donnée qu'il produit est comme un tatouage qu'il aura du mal à effacer. S'amorce là un autre débat : celui des enjeux de

---

<sup>3</sup> ISO 9000 : 2005, article 3.4.1.

<sup>4</sup> Nous faisons ici référence à la métaphore de l'antilope exprimée par Suzanne Briet – théoricienne des sciences de la documentation et une des créatrices de l'INTD- dès 1951. Capturée et enclose dans un zoo, l'antilope change de statut : elle devient un objet d'études, un objet documentaire.

l'identité numérique. D'après Olivier Ertzscheid, «les réseaux sociaux posent aujourd'hui, au sens propre, la question documentaire appliquée au facteur humain» ; [4]

2- devient accessible à différents niveaux jusqu'à la granule documentaire dont on aura du mal à retrouver la structure d'origine et à rattacher à une unité informative d'ensemble (changement d'échelle) ;

3- se complexifie : le même document (même contenu, même destinataire, même forme) change de présentation et peut être retrouvé sur plusieurs supports.

A chaque intervention de l'auteur ou du lecteur, se met en place une « chaîne de composition décomposition recomposition de l'information » [4].

Quelles conséquences l'introduction des flux électroniques a-t-elle pour la gestion d'un fonds de bibliothèque?

#### 1.1.2.1 Flux électroniques versus monographie ?

Les collections détenues par les institutions patrimoniales sont des ensembles d'objets manipulables et descriptifs. Dans les bibliothèques par exemple, c'est l'achevé d'imprimer qui clôture un état du savoir et permet sa description. Comment peut-elle passer de « l'ensemble stable et organisé d'objets finis » [1] au traitement de ces nouveaux objets de collection que sont les documents électroniques ? En fait, la bibliothèque gère depuis toujours des ressources continues. En effet, les périodiques - les "on going publications"- ont un caractère inachevé et évolutif qui les rapproche des caractéristiques des ressources électroniques.

D'après Bertrand Calenge, une autre distinction serait plus opportune, fondée sur la nature des supports (et non plus sur le support):

- le "support de type monographique" (daté et achevé) regroupant les livres mais aussi les images numériques de livres

- le "support de type périodique", en construction continue : périodiques imprimés mais aussi sites web et ressources en ligne. [1]

Cette différenciation support de type monographique / périodique nous a semblé pertinente à définir. En effet, cette réflexion sera utile lors l'application du modèle FRBR aux documents d'activité de l'IPE.

Quels sont les enjeux de cette recomposition / complexification pour les métiers de l'information documentation ?

#### 1.1.2.2 Les enjeux de la complexification du document pour les métiers de l'information documentation

Ils sont au moins doubles.

Le premier enjeu est la nécessité de mieux cerner le périmètre évolutif de l'objet documentaire pour pouvoir le décrire. Il nous faut donc d'autres termes pour mieux traduire ces frontières mouvantes.

Pour exprimer le primat du contenu sur le support, on préférera utiliser le terme d'unité documentaire.

L'unité documentaire ou entité documentaire -son intitulé dans le monde du records management-

est un « document, collection ou partie de document, quel qu'en soit le support, formant une unité indépendante d'un point de vue bibliographique.»<sup>5</sup>

L'entité documentaire s'attache à ce que l'on indexe et que l'on recherche par opposition aux unités de gestion (ce que l'on stocke, ce que l'on manipule).

Le second enjeu de l'évolution des techniques d'écriture (la dernière en date étant le numérique) est de pouvoir attacher un document à une trace, de stabiliser un contenu, d'identifier l'objet documentaire (ou plutôt son fragment, l'objet de données). En effet, les identifiants traditionnels (ISBN / ISSN/ ...) se rapportaient à une entité physique. Mais comment pourra-t-on identifier une unité d'information qui devient indépendante d'un support, dématérialisée ?

Parallèlement aux évolutions technologiques, la terminologie a dû évoluer : on passe progressivement de l'appellation de « document » à celle de « document en ligne » puis à celle de « ressource ».

## 1.2 Ressources

Sont considérées comme des ressources des « informations construites dans une logique de médiation et d'usage (réception), évolutives (susceptibles d'être mises à jour), et éventuellement adaptables (personnalisables). Leur fonction est d'être utiles et de rendre des services. » [7]

D'avantage que sur sa valeur d'usage, on peut insister sur le cadre dans lequel une ressource évolue et se transforme.

En effet, la ressource (issu de l'anglais "resource" très connoté web) est consubstantielle à son domaine d'usage, le web.

En parlant de ressources, on désigne donc implicitement les ressources électroniques -et bien sûr informationnelles !- dont on peut souligner les caractéristiques en terme de :

- volumétrie : une ressource recouvre une grande potentialité d'informations ;
- instabilité et "ubiquité": la ressource évolue en ligne et génère différentes versions intellectuelles, du draft aux versions révisées, différents formats, différentes versions linguistiques ;
- interaction : les ressources sont liées entre elles.

Une ressource est « implicitement définie comme quelque chose qui peut être nommé et identifié ». [10]

Ces ressources sont atteignables sur le web par des URL (Uniform Resource Locator).

## 1.3 Données

La donnée (data en anglais, toujours au pluriel) est le plus petit élément d'information.

Dans le vocabulaire de l'informatique, une donnée est définie comme une « représentation d'une

---

<sup>5</sup> BOULOGNE Arlette [et al.], INTD- er. Vocabulaire de la documentation.

information (élément de connaissance) sous une forme conventionnelle destinée à faciliter son traitement »<sup>6</sup>.

Donnée (s) sous-entend une notion de matière première, non traitée comme l'indique la tautologie souvent rencontrée : « données primaires ».

Une donnée est « une représentation conventionnelle, après codage, d'une information sous une forme permettant d'en faire le traitement électronique »<sup>7</sup> : il s'agit de donner à l'ordinateur des éléments faciles à traiter. Le guide de catalogage et d'indexation de l'UNESCO<sup>8</sup> définit les éléments de données comme des « éléments d'informations contenus dans une unité bibliographique (par exemple la pagination, le titre, l'auteur, etc.). »

Plus que par son codage, qui peut générer des données de formes différentes (texte, images, sons...), on peut définir la donnée comme un élément chargé de sens qui véhicule une information. Ainsi, la même information, le même terme, pourra avoir des codages informatiques différents selon le système utilisé, alors que du point de vue info documentaire il s'agit toujours de la même « donnée ». [9]

Une donnée contextuelle devient donc une information. Avec le numérique, cette information devient à la fois multimédia, composite, distribuée et portable.

## 1.4 Traiter l'information brute

La donnée est traitée afin de la rendre calculable, exploitable par les systèmes d'information. Les éléments de données structurés dans des champs sont stockés dans les bases de données : on passe de l'information primaire à l'information secondaire.

Rappelons que tout document fait l'objet de deux types de traitement : d'abord un traitement matériel puis un traitement documentaire en deux phases portant sur le contenant et le contenu du document.

1- La description bibliographique est définie dans le vocabulaire de l'ADBS comme « l'ensemble de données bibliographiques décrivant un document, servant à son identification et pouvant être structurées en suivant les prescriptions des normes nationales ou internationales. La description bibliographique est utilisée dans l'élaboration des notices bibliographiques et catalographiques. »<sup>9</sup>  
Il s'agit de décrire le document mais aussi de l'analyser.

2- L'analyse est l'opération intellectuelle visant à représenter le contenu du document sous une forme différente de sa forme originelle pour en faciliter la sélection et la consultation. Elle permet de :

---

<sup>6</sup> Journal officiel du 22/09/2000

<[http://franceterme.culture.fr/FranceTerme/recherche.html#INFO118\\_ouverture](http://franceterme.culture.fr/FranceTerme/recherche.html#INFO118_ouverture)>.

<sup>7</sup> CACALY Serge, LE COADIC Yves- François, POMART Paul- Dominique, SUTTER Eric. Dictionnaire de l'information. 3e éd. Paris, Armand Colin, 2008. 295 p. ISBN 978-2-200-35132-8

<sup>8</sup> UNESCO. Guide de catalogage et d'indexation CDS/ ISIS, p.7, consultable à l'adresse <<http://unesdoc.unesco.org/images/0009/000925/092554fo.pdf>>.

<sup>9</sup> ADBS. Vocabulaire de la documentation, consultable à l'adresse <[http://www.adbs.fr/description-bibliographique-16764.htm?RH=OUTILS\\_VOC](http://www.adbs.fr/description-bibliographique-16764.htm?RH=OUTILS_VOC)>.

- contrôler les points d'accès au document ;
- d'indexer et classer le contenu sémantique des documents.

Ces définitions sont autant d'éléments de compréhension qui nous ont aidés à poser le décor.

Il s'agit désormais de comprendre comment sont organisées ces données, sémantiquement et syntaxiquement avec les langages, les formats et de mesurer les conséquences de leur évolution.

## 2 Structurer l'information : des mutations techniques, fonctionnelles et opérationnelles

Une triple mutation est en cours:

- une mutation technique, concernant les outils, induite par les évolutions technologiques ;
- une mutation fonctionnelle concernant les usages et les besoins des usagers ;
- en conséquence, une mutation opérationnelle des règles et normes qui doivent s'adapter.

Afin de rendre compte de ces changements, nous avons choisi une approche interconnectant l'évolution des outils de structuration, des usages, des langages et des normes documentaires. [11]

### 2.1 Les supports et leurs évolutions : catalogues et bases de données

#### 2.1.1 Définitions: le catalogue et ses fonctions

Comment retrouver l'information stockée? [12]

Rappelons qu'un catalogue est une représentation sous forme de liste descriptive des documents que possède une bibliothèque, un fonds documentaire (livres, périodiques/revues, produits documentaires, mémoires, thèses, rapports...) quel qu'en soit le format mais aussi en établissant des liens vers des ressources externes.

Les fonctions du catalogue sont:

1- d'identifier, décrire, signaler une ressource

Cette fonction de signalement des données s'inscrit dans la continuité des registres anciens établis pour enregistrer des données de nature diverse (par exemple les registres des conscrits établis depuis la levée en masse de 1793).

2- de localiser le contenu final afin de vérifier la disponibilité du document. Cela permet de gérer la circulation des collections de la bibliothèque.

L'identification du document est donc au service de la localisation.

L'unité d'information dans le catalogue est déterminé par le support : on décrit le contenant puis éventuellement le contenu.

Si les fonctions du catalogue semblent pérennes, l'outil évolue.

## 2.1.2 Rappels historiques : les catalogues sur fiche

Pour bien comprendre la genèse du modèle conceptuel FRBR, il est utile de brièvement retracer l'historique des catalogues et de passer en revue ses "pères fondateurs".

Dans un catalogue ancien, c'est la fiche (ou la notice, l'index) qui fait référence au livre, au document et permet sa localisation.

Les premiers catalogues imprimés signalent les documents sur un registre. Au début du XX<sup>ème</sup> siècle, les fiches imprimées sont rangées dans des tiroirs : ce sont les catalogues sur fiches. Signalons d'ailleurs que la bibliothèque de l'IPE conserve ses anciennes métadonnées à portée de main puisqu'il suffit d'ouvrir le tiroir meuble pour effleurer les fiches soigneusement conservées.

Pour rendre possible une recherche sur plusieurs champs, la même fiche est imprimée avec comme intitulé, en « vedette », le titre du livre / de l'auteur/ le sujet. On aboutit donc à des catalogues par sujet et à des catalogues thématiques (auteur, titre), avec des listes constituant autant d'angles d'approche pour accéder au document. Ces listes données sur les données sont les ancêtres des métadonnées. Un de ces premiers catalogues à visée encyclopédique est la Bibliographie universelle des belges Otlet et La Fontaine (16 millions de fiches).

Avec ces catalogues sur fiches, la médiation d'un professionnel est indispensable pour que l'utilisateur des bibliothèques publiques accède à l'information.

Depuis le XIX<sup>ème</sup> siècle, des bibliothécaires catalogueurs ont engagé une réflexion sur les usages et les évolutions possibles de leur outil de travail.

Quelques-uns de ces "totems" sont des précurseurs du modèle FRBR :

→ Anthony Panizzi, bibliothécaire/conservateur du British Museum, (et par ailleurs correspondant de Mérimée pour le projet de remaniement de la bibliothèque impériale) a l'intuition de l'importance de l'*œuvre*<sup>10</sup>.

→ Cutter, bibliothécaire américain, définit ainsi les objectifs du catalogue<sup>11</sup> :

- permettre à un individu de trouver un ouvrage dont il connaît l'auteur, le titre ou le sujet
- de montrer les documents que la bibliothèque détient (pour un auteur donné, sur un sujet donné ou pour un genre de littérature donné)
- d'aider au choix d'un ouvrage en fonction de son édition ou de son thème (littéraire ou traitant d'un sujet particulier). »

Les tâches que les FRBR fixent aux utilisateurs sont dans la lignée des objectifs du catalogue décrits par Cutter en 1876. [13]

→ Ranganathan, bibliothécaire indien, connu pour sa classification à facettes, énonce dès 1931 Five Laws of Library Science : en résumé, le lecteur ne doit pas avoir à demander à consulter un document pour savoir s'il répond ou non à ses besoins. [12]

Cette fonction avant-gardiste du catalogue sera formalisée dans le modèle FRBR.

→ Enfin, plus récemment, Lubetzky s'est attelé à simplifier les règles de catalogage et a fait émerger la notion d'*œuvre* dans ses *Cataloging Rules and Principles*.

---

<sup>10</sup> *Catalogue of printed books in the British Museum.*

<sup>11</sup> *Rules for a Dictionary Catalog.*

Dans les années 1960, l'augmentation du volume des informations, conjugué aux progrès techniques des technologies de l'information et de la communication, entraîne une évolution des outils traditionnels.

Cette évolution n'est pas une révolution. En effet, si les outils changent, les données restent : ainsi, les fiches des catalogues papier sont intégrées dans les bases de données.

### 2.1.3 Support du catalogage, la base de données

C'est l'époque du développement de l'informatique documentaire<sup>12</sup>. Rappelons sommairement les enjeux de l'informatisation des opérations classiques de traitement documentaire :

- la gestion des acquisitions : gestion des commandes (suggestions, commande, réception, vérification, comptabilité) ;
- la gestion des périodiques : abonnement, bulletinage ;
- la mémorisation de l'information ;
- traitement de l'information ;
- l'accès au document primaire ;
- la recherche ;
- la diffusion : circulation interne, gestion des prêts, produits documentaires, statistiques... [17]

Des logiciels de traitement documentaire permettent de suivre le cycle de vie de l'information.

Une partie de ces informations est stockée dans des bases, développées pour gérer, retrouver et exploiter les données.

Que sont les bases de données ?

D'après le vocabulaire de la documentation de l'ADBS, les bases de données sont un « ensemble de données structuré, généralement en champs, organisé en vue de son utilisation par des programmes correspondant à des applications distinctes (gestion, recherche, tri, cartographie, etc.). Ce regroupement structuré de données, géré par un système de gestion de base de données (SGBD), se réalise de manière à faciliter l'évolution indépendante des données et des programmes.

Quelles sont les conséquences du transfert des données des catalogues imprimés et des catalogues sur fiches vers le support informatique ?

Le mode de description des catalogues sur fiches et de structuration des données a été conservé. On est toujours dans le triptyque :

- fichier bibliographique
- fichier d'autorité (donnant l'accès – sujet, auteur....- au document)
- données d'exemplaires.

Si la structure des données est restée la même, l'informatisation a permis de multiplier le nombre

---

<sup>12</sup> D'après la définition d'Arlette Boulogne, c'est « l'informatique appliquée au document. Essentiellement axée sur la recherche de documents dans les années 1960, puis sur la gestion et l'organisation, l'informatique documentaire a évolué en prenant en compte les fonctions de production et de diffusion des documents ».

d'index et de rendre les catalogues interrogeables, à la fois par les listes mais aussi par mots clés. [16] L'utilisateur peut désormais combiner plusieurs critères de recherche.

On est passé en quelques décennies du catalogue papier aux catalogues en ligne - l'OPAC (Online Public Access Catalog) généralisés dans les années 1980- et aux portails qui ne cessent de s'enrichir et de gagner en fonctionnalités.

L'automatisation dans la production et le traitement des données a permis de gérer ce développement toujours croissant des volumes d'information.

L'offre de bases de données s'est diversifiée.

L'évolution des technologies et la naissance de l'informatique du contenu ont permis d'intégrer de nouvelles fonctionnalités : accès centralisé à des ressources stockées dans des bases de données hétérogènes, accès direct aux ressources numériques en texte intégral, gestion du contenu.

Ainsi une bibliothèque numérique est constituée de données structurées (notices documentaires classiques) et de données plein texte peu ou pas structurées (accès au document en texte intégral). [17]

Ces données d'identification servent la plupart du temps à retrouver, classer et gérer les objets contenus dans le système. Dans le monde des documents, les données d'identification sont souvent appelées métadonnées.

Dans un contexte de croissance des volumes d'information, une normalisation de la description bibliographique au niveau international est devenue nécessaire.

## 2.2 Décrire et enregistrer : normes et formats

Une norme est définie par l'ISO comme un « document établi par consensus et approuvé par un organisme reconnu, qui fournit, pour des usages communs et répétés, des règles, des lignes directrices ou des caractéristiques, pour des activités ou leurs résultats, garantissant un niveau d'ordre optimal dans un contexte donné. » Ces caractéristiques de validation et d'obligation la différencient du standard, terme anglais prêtant à confusion par sa bivalence : norme/ standard.

### 2.2.1 La normalisation de la description bibliographique

Cette "mise au pas" a été conduite par l'organisme international représentant les bibliothèques et les professions de l'information : l'IFLA (*International Federation of Library Associations and Institutions*).

Dans le cadre du programme CBU (*Contrôle bibliographique universel*), l'IFLA a impulsé la mise en place d'un ensemble de principes de catalogage afin d'établir des normes internationales sur la présentation des descriptions bibliographiques, le contenu à leur donner et le format pour les échanger.

Le contexte est encore celui des catalogues sur fiches et des contenus textuels imprimés. Deux dates essentielles peuvent être rappelées :

- 1961 : publication des « Principes de Paris », base à la normalisation internationale du catalogage
- 1971 : publication de l'ISBD (*International Standard Bibliographic Description = Description bibliographique internationale normalisée*) qui fixe la présentation des notices et définit tous les éléments devant y figurer, leur ordre de transcription ainsi que le système de ponctuation. Ce modèle est centré sur l'objet publié et sa gestion une fois celui-ci publié (localisation, entrée,



sortie). On se doute qu'avec un tel système, les contenus informationnels des objets sont très faiblement représentés. [22]

En France, l'application des ISBD s'est fait à travers les fascicules de catalogage AFNOR (dites normes AFNOR).

Mais de nombreux pays, y compris européens (Allemagne...), ont choisi de suivre les règles décrites par les AACR (Anglo-American Cataloguing Rules/ *Règles de catalogage anglo-américaines*).

Au total, les normes françaises de catalogage sont très nombreuses et éclatées: on dénombre onze normes sur la description bibliographique, quatre normes sur la forme et la structure des vedettes, une norme sur le choix des accès à la description bibliographique et une norme sur le classement bibliographique. [14]

Adaptées au temps du catalogue sur fiches, ces normes sont en décalage avec les besoins simples de l'utilisateur : accéder simplement au contenu.

Nous avons vu que le contenu et la présentation de la notice sont normalisés. Pour enregistrer ces notices bibliographiques, un format informatique est nécessaire.

## 2.2.2 Enregistrer les notices : les formats documentaires informatiques

En 1965 est mis au point le format MARC (Machine Readable Catalog). C'est le format le plus utilisé dans les bibliothèques mais avec des spécificités, quasiment chaque bibliothèque nationale générant son MARC.

Devant cette « babélisation des formats<sup>13</sup> », un essai de convergence est initié par l'IFLA avec des formats d'échange universel : UNIMARC (UNiversal MARC) puis MARC 21. L'objectif est que chaque bibliothèque conserve son format de travail (en respectant le cadre de l'ISBD) et échange ses données avec le format UNIMARC.

UNIMARC prend en compte tous les supports : des champs spécifiques prévoient la description et l'accès aux ressources numériques.

D'autres formats existent.

Ainsi, le CCF, Common Communication Format, est utilisé par le Centre de documentation de l'IIEP. Le CCF a été défini par l'UNESCO en 1984 comme format d'échange pour les services « secondaires » gérant des données bibliographiques. Cela répondait au besoin de disposer d'un format compatible avec les formats MARC pour réutiliser les données des bibliothèques. Mais ces formats « ne convenaient pas pour la gestion interne, qui nécessitait une description de contenu plus détaillée que dans les notices MARC ». [15]

Ces formats, finement structurés mais à la structure rigide, ont pour vocation l'échange entre les données.

Mais, ils ne sont utilisables que dans l'univers des bibliothèques ce qui limite leur intérêt dans un contexte d'échange des informations entre institutions patrimoniales (archives, centres de documentation, musées).

De plus, ils ne peuvent pas être utilisés pour tous les types de ressources publiés sur le web.

---

<sup>13</sup> La formule a été empruntée à Marie-Renée Cazabon.

Enfin, se pose la question de leur exploitation : leur implémentation est souvent faite dans des langages propriétaires, ce qui les isole des évolutions technologiques.

Donc, les "anciens" formats ne sont sans doute pas optimaux pour intégrer la description de ressources électroniques, abondantes et protéiformes.

Pour pouvoir échanger avec d'autres communautés mais aussi pour permettre que l'information produite « en interne » soit utilisées dans des applications extérieures, les bibliothèques ont besoin d'encoder autrement qu'en MARC l'information bibliographique.

Dans le contexte d'une profusion de données sur le web, la nécessité de décrire, identifier, ordonner les ressources afin de les retrouver est renforcée.

## 3 Décrire et identifier pour trouver et obtenir sur le web

Les plateformes web donnent accès à un réseau d'informations numériques. Pour décrire mais aussi structurer et gérer les données véhiculant ces informations, les standards internationaux utilisent un outil compatible avec le développement du web qui remet en question le monopole du MARC: les métadonnées.

### 3.1 Décrire, identifier sur le web : les métadonnées

Après une définition des métadonnées et une délimitation de leur périmètre d'influence, nous présenterons une typologie des métadonnées et leur processus de création.

#### 3.1.1 Contexte et périmètre des métadonnées

Métadonnée (anglais Metadata element = métadonnée), métadonnées (anglais Metadata), littéralement, "données sur les données" : ce sont des données servant à décrire d'autres données (primaires) : les métadonnées sont donc une information secondaire. D'après le vocabulaire de la documentation, c'est un « ensemble structuré de données créées pour fournir des informations sur des ressources électroniques ».

Les métadonnées sont un mot récent (1969) issu du secteur de l'informatique pour désigner dans le monde de l'information documentation une réalité bien antérieure au web: par exemple, une table des matières est une donnée d'accès au contenu d'un document.

Du monde physique, ce terme est désormais utilisé pour désigner l'information descriptive à propos de ressources du Web.

Notre objectif n'est pas de rentrer dans la complexité des métadonnées mais de présenter quelques éléments basiques de compréhension pour la suite de notre propos (notamment sur les formats de métadonnées à relier aux évolutions récentes vers le web sémantique). [20] [21]

L'objectif est donc de décrire.

Mais décrire quoi ?

On a déjà pris note de la diversification et la complexification de l'objet documentaire qui ne désigne plus seulement un document (quelque soit son support) mais aussi des activités, des personnes.

Décrire pour qui ? Les métiers de l'informatique, de l'information documentation, de l'enseignement, de la géographie..., toutes les communautés d'utilisateur utilisent désormais des métadonnées.

L'environnement de production et d'utilisation de l'information change. La tendance est au décloisonnement des outils de production des données qui se détachent de leur corps de métier originel. Comme le constate Sylvie Dalbin [23], « on assiste à un effort visible de mutualisation et de rationalisation de référentiels métiers commun ».

### 3.1.2 Typologie des métadonnées

La métadonnée se situe à un double niveau, faisant référence à la fois au contenant et au contenu (description). [24] [28]

On peut les répartir en trois catégories principales:

1- Les métadonnées descriptives sont au cœur des métiers de l'information et de la documentation.

Elles représentent le contenu de la ressource et sont dans la lignée des notices des catalogues bibliographiques. Un catalogue est donc typiquement un ensemble de métadonnées descriptives.

Elles ont donc pour fonction :

- de décrire ;

Certaines métadonnées figurent dans le document (titre, auteur, support...). L'utilité est alors de permettre l'indexation de la ressource et de l'extraction (automatique dans le cas de moteurs de recherche) de l'information contenue dans mais aussi sur le document.

- de localiser : par exemple, l'url est une métadonnée de localisation d'un document électronique.

2- Les métadonnées de structure sont utiles pour l'utilisateur final.

En effet, elles décrivent la structure logique du document : les relations d'organisation, l'articulation entre un/ des fichiers constitutifs du document.

Ces métadonnées sont intéressantes car elles peuvent permettre la navigation : elles font le lien entre jeux de métadonnées et objets documentaires (est-ce une collection, un élément de...).

3- Les métadonnées administratives concernent davantage les administrateurs de bases de données et les archivistes / records managers, en particulier les métadonnées de préservation liées à la gestion du cycle de vie du document, sa création, sa mise à jour et son archivage.

On peut par exemple gérer l'historique d'un fichier numérique en récapitulant ses différentes transformations) [21] Les métadonnées de droit régissent les usages et gèrent la permission d'accès au document numérique : cela en fait un filtre entre la description et l'accès au contenu de la ressource. En cela, la métadonnée se différencie de la notice traditionnelle qui est un pont direct vers le document.

4- On pourrait ajouter à cette liste les métadonnées sur les métadonnées produites (qui les a créées, quand, etc.) : les « méta métadonnées ». [22]

Quels sont les liens entre métadonnée et la ressource décrite ?

- externes aux ressources, c'est-à-dire stockées en dehors, sous forme de catalogues ou de bases de données.

Par exemple, une notice catalographique est séparée du document qu'elle décrit ; la liaison avec le document est assurée par une métadonnée de localisation (cote, URL...).

- internes :

- soit encapsulées dans la ressource, par exemple dans les images numériques fixes (les champs EXIF des images JPEG), dans les applications bureautiques sous format Word ou pdf (propriétés). Elles sont alors consubstantielles à la ressource qu'elles identifient et gèrent à la fois.
- soit englobantes, pouvant inclure la ressource elle-même (entièrement ou partiellement). C'est-à-dire que contrairement aux données traditionnelles des catalogues, elles fournissent à l'utilisateur un service intégré : la description mais aussi l'accès à la ressource ... à condition toutefois que l'URL fonctionne.

Les métadonnées sont donc utiles à plusieurs niveaux :

- 1- en fournissant des renseignements structurés sur des données (dans la continuité des informations textuelles associées aux produits documentaires traditionnels : les tables des matières, les index, les glossaires, les modes d'emplois ou les notices bibliographiques).
- 2- en renseignant sur elles-mêmes (par exemple fournissent des données techniques sur leur utilisation).

Comment mettre en œuvre les métadonnées ?

### 3.1.3 Comment écrire la métadonnée pour décrire la ressource ?

Les métadonnées sont appliquées à des types de ressources diversifiées dans des domaines différents : les ressources documentaires ne peuvent pas être décrites de manière uniforme.

On peut distinguer les métadonnées généralistes et les métadonnées par types de ressources :

#### 3.1.3.1 Utiliser des métadonnées généralistes

Le format Dublin Core, créé en 1995, est un cadre pour l'échange des ressources numériques. Dublin Core établit un set de quinze paires "propriétés-valeur"<sup>14</sup>. Par exemple la propriété relation permet d'établir un lien vers une ressource liée et de reconstituer des relations entre les documents.

L'objectif de ce format est de définir un ensemble de métadonnées communes à diverses communautés, assez simple pour que des non-spécialistes puissent les produire et assez structurées pour faciliter la recherche et la récupération des ressources par les crawlers.

---

<sup>14</sup> Normalisées : ISO 15836-2003

### 3.1.3.2 Utiliser des métadonnées par types de ressource [19]

Les outils de catalogages et d'indexation des données via les métadonnées ne sont plus réservés au catalogueur expert.

De nouveaux producteurs d'information (informaticiens, géographes,...) utilisent les métadonnées pour créer et retrouver facilement l'information. Cette multiplication et démocratisation des outils - désormais à disposition de tous les producteurs d'information - remet d'ailleurs en cause les pratiques et le statut d'expert des professionnels des bibliothèques et de l'information documentation.

Chaque ensemble de métadonnées s'appuie sur un modèle métier sous-jacent aux institutions patrimoniales (archives, bibliothèques, musées), mais aussi enseignement, géographie... Pour décrire ces données de domaines différents, on se doute que les métadonnées ne peuvent pas être uniques.

D'autres langages ont été élaborés en dehors du monde des bibliothèques. Ils servent à décrire différents types de documents, par exemple :

- les archives : EAD (Encoding archival description) ;
- le texte : TEI (Texte encoding initiative) ;
- les thèses : TEF (Thèses électroniques françaises) ;
- ou à plus large échelle les ressources liées à l'éducation: LOM (Learning Object Metadata) utilisé dans le domaine pédagogique.

Des formats de données permettent d'écrire la métadonnée d'où une imbrication possible des schémas de métadonnées pour représenter des 'objets' différents. Par exemple, les thèses électroniques françaises utilisent à la fois : [29]

- Dublin Core comme format de métadonnées descriptives ;
- Dublin Core étendu pour les métadonnées administratives ;
- METS pour les métadonnées de droit ;
- CINES pour les métadonnées de conservation

Certains formats de description ont d'ailleurs été initiés par les acteurs commerciaux, souvent à l'avant-garde des innovations. Par exemple, le format ONIX sert à diffuser des données enrichies avec des résumés, images de couverture....

La multiplication des objets documentaires et des métadonnées très diverses a rendu nécessaire l'utilisation de modèles et schémas de métadonnées<sup>15</sup>. Ces ensembles de métadonnées sont normalisés dans une exigence d'interopérabilité. L'objectif est à la fois d'échanger mais aussi de réutiliser des ensembles de données communes.

Ainsi, les formats d'échange des données évoluent et se multiplient. Il en va de même pour les usages des utilisateurs.

---

<sup>15</sup> Un schéma de métadonnées définit l'ensemble des éléments de données, l'agencement des métadonnées entre elles et les règles pour les utiliser.

## 3.2 Trouver et obtenir sur le web : évolution des usagers et des usages

Le catalogue a dû s'adapter à différentes plateformes technologiques mais surtout se transformer en fonction des besoins exprimés par les usagers.

### 3.2.1 Le contexte de l'évolution du comportement des usagers

Depuis les années 1990, les bibliothèques et centres de documentation sont passés d'une situation de monopole de l'information documentaire à une situation de concurrence. En effet, d'autres sources d'information sont préférées par les utilisateurs. Des enquêtes récentes révèlent que le catalogue n'est plus « l'outil principal d'accès aux ressources documentaires et informationnelles ». [32]

Pourquoi ?

Quelques pistes d'explications sont proposées :

- les changements induits par la concurrence des moteurs de recherche : « accès convivial, rapide et immédiat à une variété croissante de ressources documentaires ». On pourrait ajouter la facilité : un clic dans l'onglet de recherche rapide permettant de lancer une requête.
- l'augmentation des sources d'information : c'est le dilemme exhaustivité versus pertinence. Or, l'utilisateur des bibliothèques ne réclame pas l'exhaustivité : il est abreuvé d'une surcharge de résultats qu'il n'a pas le temps de dépouiller. Il veut aller à l'essentiel.
- les nouveaux usages des utilisateurs avec le développement d'Internet : habitude de naviguer sur le web et de récolter ce que l'on y trouve (sérendipité versus esprit critique et/ou sélection du contenu).
- la concurrence des catalogues marchands en ligne qui se placent dans une autre logique. Leur objectif, utilitariste, est de vendre plus donc les recherches sur leur site ne produisent jamais de silence, générateur de frustration pour l'acheteur.
- la construction de nouveaux « schémas mentaux » du lecteur engendrés par de nouveaux modes de lecture dans la logique de l'hypertexte, de l'association de contenu (et non de la juxtaposition) donc l'utilisateur a besoin d'informations interconnectées.

De nouvelles problématiques émergent. Comment captiver l'attention de l'utilisateur client, versatile et exigeant ?

Pour cela, il faut déjà repérer quels sont les exigences et des besoins nouveaux exprimés par l'utilisateur.

### 3.2.2 De nouvelles exigences

Rappelons que l'accès au contenu d'un catalogue reflète la manière dont les données sont entrées : or, nous avons vu que le modèle des ISBD décrit peu le contenu. Il suppose chez l'utilisateur des prérequis (connaissance des éléments fondamentaux de l'objet qu'il souhaite obtenir : titre, auteur, collection...) qu'il doit constituer avant de consulter le catalogue.

Plusieurs rapports sur l'avenir du catalogue pointent les insuffisances des catalogues actuels par rapport aux besoins des utilisateurs finals.

Le rapport d'OCLC (Online Computer Library Center) sur les Catalogues en ligne, paru en 2009, permet de catégoriser les doléances des utilisateurs et des bibliothécaires. [30] L'enquête récemment menée sur les bibliothèques du Québec complète cette étude et souligne que la recherche par sujet est le premier type de recherche dans les catalogues de bibliothèque : il faut donc garder la valeur ajoutée des langages documentaires. [32]

Ces enquêtes ont été menées à l'initiative de grandes bibliothèques mais la synthèse que nous avons faite de leurs résultats est transposable à d'autres catalogues.

Quelles sont donc les améliorations préconisées par les utilisateurs et les professionnels de l'information?

### 3.2.2.1 Les exigences des utilisateurs pour l'amélioration des catalogues

Les améliorations peuvent être classées selon quatre grands axes.

1- en terme d'interrogation du catalogue :

- disposer d'une seule interface de recherche versus complexité des interfaces de recherche ;
- consulter des documents en texte intégral ;
- retrouver les fonctionnalités de l'environnement marchand habituel ;
- disposer d'un outil de recherche bibliographique permettant de découvrir d'autres ressources ce qui reviendrait à mieux intégrer les outils documentaires avec d'autres sources d'information (dépôts institutionnels...) ;
- retrouver l'univers graphique auquel l'internaute est habitué avec les moteurs de recherche : une interface intégrant des images (quatrième de couverture...).

2- en terme d'accès aux données :

On peut imaginer le désarroi du lecteur dont la requête génère des centaines de notices : son besoin n'est pas de connaître la description matérielle des documents, ni les multiples mentions de publication mais simplement d'accéder à un exemplaire.

Sans surprise sont donc plébiscités :

- l'accès au document, « la fourniture » (savoir dans quelle bibliothèque le document est disponible) ;
- la facilité d'accès (classement thématique/ alphabétique) ;

Beaucoup d'utilisateurs réclament un catalogue plus intuitif. La navigation à facettes par exemple permet d'affiner les résultats de recherche.

- la pertinence de la réponse ;
- l'immédiateté des réponses ;
- des informations interconnectées: l'idéal serait la notice faisant des relations entre des données différentes (auteurs / sujet / ...) et davantage de liens vers des contenus et textes intégraux en ligne.

3- en terme d'enrichissement du catalogue

- un enrichissement des contenus (ajout de résumés, extrait, table des matières...) ;

Les « utilisateurs finaux accordent une place prédominante aux données d'enrichissement (tables des matières, résumés, etc.) et aux liens vers des contenus en ligne (textes et supports). »

- une participation de l'internaute à l'enrichissement du catalogue ;

Une indexation des ressources par les usagers a été mise en place dans certaines bibliothèques.

Le rapport Tessier encourage d'ailleurs ces pratiques collaboratives : « Les internautes peuvent en effet contribuer à enrichir une ressource (par exemple une photo ou un ouvrage) par des liens, des

ressources associées (par exemple des thèses sur l'ouvrage ou l'auteur) ou des commentaires. » [33]

4- en terme de services : en signalant les manifestations, les expositions liées à un auteur, un thème ou la géo localisation d'autres bibliothèques.

### 3.2.2.2 Les améliorations préconisées par les bibliothécaires [30]

Sont par exemple réclamés par les professionnels des bibliothèques :

- la fusion des notices en double
- l'enrichissement des notices de catalogue avec des données de table des matières, des résumés et des extraits
- l'ajout de notices dans d'autres langues

L'objectif est d'aller vers un catalogue portail en créant des liens vers des sites web, des bases de données en ligne et des documents numériques.

Si l'on devait dresser la fiche d'identité du catalogue idéal pour les usagers et les producteurs de métadonnées, on pourrait exiger un outil combinant:

- une correspondance de la notice bibliographique avec les besoins des utilisateurs, comme le préconisent les FRBR) ;
- un contenu plus valorisé que le contenant (support) ;
- un contenu dense (des informations bibliographiques sur les auteurs mais aussi un résumé, un index) ;
- un contenu organisé permettant à l'utilisateur de mener une navigation à facettes ;
- des contenus interconnectés ;
- des services.

### 3.2.3 Les catalogues en évolution : d'un outil local à un outil global

« Support du catalogage et de la circulation des documents, le système de gestion de bibliothèque - système permettant au minimum le catalogage des documents et la gestion de leur circulation - a tout de même connu des évolutions fonctionnelles régulières » [27]

En effet, pendant longtemps, les catalogues publics de bibliothèques ont servi à réserver l'accès aux ressources détenues « physiquement » par la bibliothèque. Les ressources en ligne étaient signalées à part. L'utilisateur devait donc interroger plusieurs applications afin de repérer et obtenir le document voulu.

Les catalogues sont désormais presque tous des OPAC qui ont évolué dans leurs fonctionnalités.

Parmi les évolutions récentes, rappelons l'effort dans la présentation qui favorise les parcours, des fonctionnalités de navigation accrues (OPAC à facettes), des notices enrichies.

On est passé d'un catalogue comme un outil local (qui permettait de savoir dans quelle bibliothèque se trouve tel exemplaire) à un catalogue global.

En effet, les catalogues ont suivi et accompagné les mutations du web.



Ils se sont adaptés à l'évolution des pratiques de recherches documentaires tout en tenant compte des évolutions du web :

- du web de documents : la logique de l'hypertexte permet d'atteindre une page Web depuis un pointeur présent dans une autre page Web. L'utilisateur navigue à travers un réseau de documents, les pages Web. Les bibliothèques se sont approprié le web pour valoriser leurs collections. [38]

- du web social qui met en scène de nouveaux usages de l'internaute, acteur du web. C'est la prolifération des blogs, de l'information devenue flux, de l'appropriation (folksonomies), et des services personnalisés (customisation).

L'OPAC dit de "nouvelle génération" intègre les principes du web social / collaboratif. En effet, la bibliothèque dite 2.0 permet d'une part la contribution des usagers, avec la possibilité d'enregistrer des commentaires ou d'attribuer notes et tags aux documents référencés ou d'enrichir la notice avec des contenus variés. D'autre part il amène la bibliothèque à l'utilisateur en l'intégrant aux réseaux sociaux et à son environnement de travail.

Désormais l'OPAC assure souvent une fonction de portail. Il intègre les ressources en ligne produites par la bibliothèque mais également des métadonnées qu'elles ne détiennent pas mais vers lesquelles elles donnent un accès. [16]

### 3.2.3.1 Changer les pratiques : mutualiser, partager

Trop de catalogues étaient indépendants, gérés en autonomie pour un usage local. Le partage des données structurées fait pourtant partie des missions des bibliothèques : elles ont développé des protocoles (Z 39.50 puis OAI-PMH) pour améliorer l'interopérabilité. L'harmonisation technique a conduit à la mise en place de plateformes logicielles. L'objectif était d'aider les bibliothèques à partager rapidement plus de métadonnées et à améliorer la recherche et la fourniture de documents à l'échelle du Web.

Le mode de fonctionnement des bibliothèques a été ré évalué. Les bases de données très volumineuses, d'envergure à la fois nationale et internationale, contiennent des notices créées et consultées par des milliers de bibliothèques. Les bibliothèques s'impliquent dans des programmes de catalogage pour participer à la constitution du catalogue.

Cette mutualisation du travail de catalogage répond à des exigences :

- de rentabilité économique et de rationalisation des pratiques : besoin de réduire les coûts de catalogage en évitant que les mêmes tâches de catalogage ne s'accomplissent à plusieurs endroits simultanément ;
- de simplification du catalogage (catalogage "minimum") afin de ne pas crouler sous une masse éditoriale en augmentation constante ;
- de cohérence entre les catalogues ;
- de service aux lecteurs : une seule plateforme leur permet d'accéder à une recherche multi catalogues.

On se place désormais dans l'optique d'une diffusion "universelle" des notices bibliographiques.

Le catalogue collectif, outil d'identification et de localisation des documents, répertorie sur un support unique, les collections de plusieurs organismes.

C'est le cas par exemple du catalogue du SUDOC (Système Universitaire de Documentation). Ce catalogue collectif accessible sur Internet est réalisé par les bibliothèques et centres de documentation de l'enseignement supérieur et de la recherche. Chaque bibliothèque universitaire

catalogue ses documents dans une base de données centrale. Les notices bibliographiques sont liées à de nombreuses notices d'autorité qui sont récupérées automatiquement. Chaque exemplaire est finement localisé (bibliothèque, salle, cote), ce qui permet d'avoir un état assez précis de ce que possèdent les bibliothèques universitaires françaises.

Le catalogue comprend plus de neuf millions de notices bibliographiques.

L'accès est simplifié pour l'utilisateur : une interface unique ouvre l'accès à :

- une présentation permettant d'affiner la recherche selon le type de publications (livres, thèses, revues, ressources électroniques, documents audiovisuels, microformes, cartes, partitions, manuscrits et livres anciens...);

- un choix par langue ;

- un choix par pays de publication.

La recherche d'une référence précise peut être longue. Les doublons générés lors de la recherche d'une œuvre classique sont nombreux. Les œuvres d'un même auteur sont présentées dans un ordre peu compréhensible côté lecteur : le tri des résultats par titre aboutit à une liste classée par titre propre des documents. Dans la même liste, cohabitent des œuvres de nature et supports différents (livre, film...). Mais, surtout, les œuvres ne sont pas reliées entre elles : le catalogue est statique.

La simplicité d'utilisation, le contenu enrichi (images de couverture...), les possibilités de personnalisation des requêtes et de participation de l'utilisateur dans la logique du web collaboratif (ajout d'information : commentaires, notation...) sont quelques uns des points forts de ce catalogue nouvelle génération. Les utilisateurs sont reliés aux bibliothèques et aux librairies possédant un exemplaire d'une œuvre. Donnant accès aux multiples langues, éditions, exemplaires reliés à une œuvre, ce catalogue à rebond est dans la logique FRBR.

Les catalogues se sont donc adaptés au web : ils ont intégré certaines des exigences des usagers et de nouvelles fonctionnalités.

Cependant, malgré les évolutions engagées, les OPAC restent encore linéaires dans leur structure. Les résultats des recherches restent souvent présentés de manière désorganisée et mettent trop en avant des informations matérielles portant sur les éditeurs et les dates de publication. [31]

On peut aussi constater que l'affichage des données et la navigation dans les catalogues ne sont pas toujours adaptés aux nouveaux besoins des utilisateurs.

Il devient donc nécessaire de modifier et d'améliorer les interfaces mais aussi de structurer différemment les données sur lesquelles s'appuient la navigation et l'affichage.

En effet, les catalogues reflètent la façon dont les données sont "entrées". Or le primat est donné actuellement donné à l'exemplaire, dont on décrit les multiples *manifestations* (au sens FRBR du terme), aux dépens de l'analyse de l'œuvre.

Donc, en même temps et même prioritairement à l'accessibilité de leurs données, les institutions patrimoniales devront repenser leurs pratiques de catalogage.

La nouvelle problématique est d'inscrire les catalogues dans un environnement numérique pour atteindre d'autres utilisateurs et valoriser les données bibliographiques : devenir visible sur le web.

### 3.2.3.2 Devenir visible sur le web

OCLC est parti d'études menées sur les pratiques de recherche des lecteurs. Le constat est sans appel : le moteur de recherche est plus utilisé que les catalogues des OPAC. OCLC a d'abord conclu des accords avec les moteurs de recherche et des librairies en ligne afin que les notices de son catalogue mondial unique (« Worldcat ») soient récupérées, référencées donc trouvables. Puis il a mis à disposition sur le site Worldcat<sup>16</sup> sa base bibliographique de plusieurs millions de notices. Son objectif est de « mettre les bibliothèques à l'échelle du web ».

Mais surtout, il faut rendre visible cette information qui reste interne à la sphère des catalogues de bibliothèques. Or l'utilisateur habitué au moteur de recherche n'a pas toujours le réflexe de passer par la case catalogue. Mais tant que les pages de réponses de catalogue ne seront pas moissonnées par les moteurs de recherche, les catalogues resteront du domaine du web « caché » ou « web profond ».

Le défi est désormais d'ouvrir les silos de données pour rendre visibles les données des catalogues qui ne sont actuellement pas trouvables sur le web, donc pas indexables par les crawlers des moteurs de recherche. Comme le résume Olivier Ertzscheid, [25], « sans l'incontournable visibilité que leur confèrent aujourd'hui les moteurs de recherche, les catalogues de chaque bibliothèque si achevés et si fournis soient-ils sont condamnés à rester hors de portée de clic de ceux pour qui ils sont bâtis : leurs usagers ».

Et pour éviter cela, l'objectif est de mettre les données non seulement « sur » mais « dans » le web. Aujourd'hui, les catalogues sont sur le web.....mais leurs données ne sont pas exploitables par les moteurs de recherche : comment "libérer" les données des catalogues, les rendre utilisables DANS le Web ?

Pour cela, les standards traditionnels évoluent, par exemple vers le Web sémantique.

## 4 Mettre les données dans le web

Comment valoriser les données qui intéressent d'autres utilisateurs au-delà du pré carré des bibliothèques et centres de documentation et assurer leur visibilité dans le web ?

Nous dresserons d'abord un panorama des outils (langages, normes) permettant cette évolution.

### 4.1 L'évolution des langages

#### 4.1.1 Les langages de description des formats de documents : HTML, XML

Le langage HTML (Hypertexte Markup Language) est un langage qui permet d'afficher de l'information à l'aide de balises dont le nombre est limité. Il exprime davantage la forme que le contenu et est interprété par le navigateur.

---

<sup>16</sup> <<http://www.worldcat.org>>

Rappelons qu'un document HTML repose sur des données non structurées. Il est composé de deux types d'éléments : le contenu effectif du document (texte, images, sons,...) et les balises qui sont des codes déterminant la façon dont le contenu sera affiché et interprété par les navigateurs.

Les métadonnées, éléments descriptifs du document électronique, font partie de ce document. Mais, le contenu et la structure du document étant imbriqués, l'information codée en HTML ne peut être réutilisée.

C'est pourquoi un autre langage a été développé, le XML (eXtensible Markup Language : langage à balises étendu ou langage à balises extensible), que nous présentons car c'est un des langages sur lequel est basé le RDA Toolkit.

Ce format de données est devenu une norme du W3C (World Wide Web Consortium) en 1998. Ce langage permet d'une part de définir des formats de documents comportant des règles de grammaires strictes et de créer des documents respectant ces formats. Il permet d'autre part un balisage des données avec des balises ouvrantes et fermantes qui expriment uniquement le contenu et qui sont extensibles.

Les balises et la structure d'un document XML sont de préférence "déclarées" dans une DTD - Document Type Definition- qui permet d'ordonner les éléments et les relations entre eux (un peu comme les formats MARC/ UNIMARC).

XML permet donc de structurer intellectuellement les documents du web de façon très souple. Un autre de ses avantages est qu'il s'agit d'un outil non propriétaire, s'inscrivant dans la logique du libre préconisée par l'UNESCO.

#### 4.1.2 Ouvrir les métadonnées MARC au web avec le XML

Le format MARC utilisé dans les bibliothèques ne s'est pas momifié. Il s'est adapté aux évolutions technologiques.

Deux pistes sont possibles pour ouvrir les métadonnées MARC au web:

- une conversion du format MARC en langage XML. C'est le MARCXML utilisé par l'ABES comme format pivot pour alimenter le catalogue SUDOC avec des données en provenance d'autres applications ;
- une saisie des métadonnées en XML pour générer du MARC [38]

Des schémas de métadonnées ont été développés pour saisir les données sans passer par l'étape MARC : MODS et MADS.

MODS est un schéma XML de métadonnées descriptives de ressources électroniques qui peut s'appliquer à d'autres documents. Il est développé depuis 2002 par la Bibliothèque du Congrès. Ce modèle est plus riche que celui du Dublin Core et plus proche des modèles de structuration des informations bibliographiques utilisés par les bibliothèques, tout en étant plus facilement accessible par d'autres communautés (chercheurs, éditeurs par exemple), sa syntaxe étant plus simple que celle de MARC 21 ou que celle de l'UNIMARC. [18]

Un grand pas est ainsi franchi vers le web de données.

Mais, malgré ces adaptations techniques, on peut souligner que l'on reste en interne, enfermé dans un format propre au monde des bibliothèques et de la documentation.

Parallèlement, une modernisation des règles de catalogage a été entreprise.

## 4.2 La modernisation des règles de catalogage

Le rapport de l'IFLA sur les FRBR souligne dès 1998 le besoin « d'adapter les règles et les pratiques de catalogage aux mutations induites par l'apparition de nouvelles formes de publication électronique et par les nouvelles possibilités d'accès à l'information via un réseau. » [49]

En effet, la communauté du catalogage est arrivée à un croisement : après quarante années de création de notices dans un environnement à dominante imprimé, on passe à une nouvelle norme de contenu qui va au delà du domaine des bibliothèques, vers un monde constitué d'objets numériques et de métadonnées multiples [39].

Les règles de catalogage doivent rattraper les évolutions des catalogues et les formaliser : un changement profond est en cours dans les principes et les codes de catalogage.

Quelques jalons chronologiques permettent de retracer les étapes de la modernisation des règles de catalogage :

- 1998 : publication par l'IFLA des FRBR (Functional requirements for bibliographic records = *Spécifications fonctionnelles des notices bibliographiques*) : une nouvelle analyse des données bibliographiques à l'aide du modèle entité relation est proposée. [49]

- 2007 : publication de l'ISBD consolidé dont l'objectif principal est de « donner les directives permettant un catalogage descriptif compatible à l'échelle mondiale, afin de faciliter l'échange international de notices bibliographiques parmi la communauté internationale des bibliothèques et de la documentation (y compris les producteurs et les éditeurs) ».

Il essaie de décrire le document dans la logique et la conformité du modèle conceptuel FRBR dans la perspective d'un futur code international de catalogage. Mais le rapport souligne les difficultés de correspondance entre les termes du modèle FRBR (qui appartient « à un haut niveau d'abstraction ») et la terminologie ISBD.

- 2009 : publication des « Principes internationaux de catalogage » (PIC). [26]

Ces principes doivent guider l'élaboration des codes de catalogage pensé pour un catalogue en ligne : ils remplacent donc les « Principes de Paris ». Ils s'appuient sur le modèle FRBR.

L'utilité des PIC est clairement affichée : «développer le partage des données bibliographiques et d'autorité au niveau international » avec comme objectif principal le « confort de l'utilisateur ».

- 2009 : publication du nouveau code de catalogage RDA (Resource Description Access) visant, non seulement à remplacer les AACR (règles anglo-américaines de catalogage), mais aussi à devenir le code international de référence.

En effet, RDA, ce code de catalogage qui n'en porte pas le nom, a une visée plus universelle que le monde des bibliothèques. Il est ouvert vers les autres institutions culturelles.

Publié en juin 2010, RDA n'est disponible qu'en ligne, sous la forme d'un site web payant, RDA Toolkit, mieux adapté au contexte des catalogues.

RDA est en quelque sorte l'application des FRBR sous forme de code de catalogage des FRBR : il est articulé autour de la structure FRBR mais aussi sur les FRAD et les Principes internationaux de catalogage.

RDA introduit une nouvelle approche du catalogage : [37] [39]

- il s'agit d'une part de « penser autrement » l'information bibliographique à partir des modèles FRBR / FRAD ;

- il s'agit d'autre part de « créer autrement » l'information bibliographique.

On se dirige vers de nouvelles méthodes de catalogage par dérivation/récupération de métadonnées existantes (prise en compte de l'informatisation généralisée des catalogues et des outils de catalogage) et vers une nouvelle structure des catalogues avec une notice par entité FRBR. On peut souligner l'importance croissante et la généralisation des notices d'autorité.

Cette norme traite du contenu de l'information bibliographique. Elle s'affranchit du mode de présentation et des formats utilisés pour encoder l'information.

Actuellement les discussions sont menées sur l'adoption de RDA dans les bibliothèques françaises: adopter RDA est un moyen d'insérer les catalogues sur le Web de données et donc d'assurer leur avenir.

Ces questions sont examinées depuis le printemps 2010 dans le cadre de l'AFNOR par deux instances : un groupe technique, chargé d'évaluer son intérêt et ses modalités d'adoption, et un groupe stratégique qui avalise.

Mais RDA n'est qu'une option parmi d'autres : ainsi, l'Italie a élaboré un nouveau code de catalogage fondé sur le modèle FRBR, REICAT (Regole Italiane di Catalogazione) publié en 2009.

[42]

Les bibliothèques ont récemment adapté leurs outils normatifs aux mutations induites par le numérique. L'architecture des catalogues est amenée à évoluer afin de s'orienter vers le web de données.

L'enjeu est désormais d'être des acteurs du web sémantique, de porter leurs métadonnées dans le web.

Pour cela, il faut appréhender les outils et langages permettant techniquement de porter les données sur le web.

## 4.3 Vers le Web de données par le web sémantique

### 4.3.1 L'évolution vers des données liées

Le web de données n'est pas une idée neuve : l'expression web de données (traduction de « *Web of data* ») a été lancée dans les années 1990 par Tim Berners-Lee. [36]

Dans le web de données, on change d'échelle : ce ne sont plus les documents (pages HTML) qui sont liés -comme dans le web de documents (dit 1.0)- mais les données. C'est pourquoi on parle aussi de « linked data ».

Comment réaliser cette liaison entre données ? Grâce à un outil, le web sémantique qui est l'ensemble des technologies et de standards développés par le W3C (World Wide Web Consortium) pour construire le web de données.

Le Web sémantique part du principe que les données structurées (par exemple, les métadonnées contenues dans un catalogue de bibliothèque) sont déjà disponibles : il propose un ensemble de techniques visant à les rendre plus utilisables.

Les technologies du Web sémantique permettent de créer de telles données, d'exprimer des vocabulaires et des règles qui les décrivent, et de bâtir des systèmes capables de les manipuler dans de bonnes conditions d'interopérabilité. [18]

En quoi le web sémantique est une opportunité pour les métiers de l'information documentation?

### 4.3.2 Les enjeux du web sémantique

L'objectif est de créer un espace global de données interconnectées au delà des espaces métiers et des supports technologiques.

Les données sont décrites suivant un modèle commun et reliées par des liens, exploitables par des machines. La navigation entre objets documentaires doit se faire de façon fluide « sans nécessité de percevoir les limites des différentes bases de données ni leur format. » [35]

Pour cela, il faut permettre aux ordinateurs de comprendre les documents et d'exploiter les données structurées, par exemple les métadonnées contenues dans un catalogue de bibliothèque. En effet, ces données créées par les professionnels de l'information documentation sont à forte valeur ajoutée : elles pourront être échangées, réutilisées.

Ces données riches créées par les bibliothèques doivent être rendues visibles dans le web.

Le web de données doit en effet refléter les fonds de bibliothèques et pas seulement des données commerciales.

Pour l'utilisateur final, le web sémantique devrait apporter, en plus de l'accès aux données, les relations entre les données. C'est aussi la possibilité de réinvestir, réutiliser les données mais aussi trouver des applications.

Une des conséquences est que chaque donnée devient autonome, exploitable par les technologies du web. Un réseau global d'information qui s'étend à l'infini composé de données circulant librement est en train d'être constitué. Les initiatives sont foisonnantes et certaines seront développées dans la troisième partie de cette étude. Il est encore trop tôt pour mesurer les impacts politiques, économiques ou sociaux générés par cette libération des données.

Le web sémantique est un langage pour les machines : comme l'explique très clairement Emmanuelle Bermès, il utilise une « grammaire » (RDF) et un « vocabulaire ». [34]

Quels sont ces outils qui permettent d'identifier, décrire les données ?

### 4.3.3 Comment identifier la donnée?

Le « projet technique de standardisation des données et des relations entre les données [est] surtout basé sur un langage, RDF, dont les promoteurs œuvrent au déploiement ». [40]

RDF (Resource Description Framework) serait-elle le volapuk global entre les données ?

RDF n'est pas un format de données. C'est un cadre (framework) pour décrire des ressources, un modèle de description des données dans lequel toute ressource est identifiée par une URI.

Les URI (Uniform Resource Identifier) sont des identifiants web. Ils peuvent tout décrire y compris des entités physiques.

Le principe de RDF est celui d'une syntaxe : des déclarations (« assertions ») sont faites sur les ressources sous la forme Sujet - Verbe – Complément, comme dans une phrase simple. Le sens est

donné par RDF, qui le code dans des ensembles de triplets (sujet/prédicat/objet).

Avec RDF, dans un document, on part du principe que :

- des "choses" particulières (des gens, des pages Web...) sont des sujets. Les choses ont toujours un type (une classe) ; [34]
- le prédicat permet d'exprimer les propriétés des "choses" ou les relations des choses entre elles ;
- l'objet peut être un texte ou une URI.

Grâce à cette décomposition en un triplet, on identifie bien de quoi on parle (le sujet, la ressource), quelle propriété on lui attribue (le prédicat) et quelle valeur a cet attribut (la valeur, l'objet). Il reste à donner un nom à ces différentes choses.

Le sujet et l'objet sont chacun identifiés par un Identifiant de Ressource Universelle (URI), exactement comme dans les liens sur une page Web.

Les URL, Uniform Resource Locators, sont des URI qui identifient la ressource par le moyen d'y accéder (« l'adresse »). Les verbes (prédicats) sont également identifiés par des URIs, ce qui permet à chacun de définir un nouveau concept en définissant une URI "quelque part" sur le Web. Il est nécessaire d'utiliser une URI différente pour chaque concept spécifique.

RDF est une infrastructure qui permet l'encodage, l'échange et la réutilisation des métadonnées structurées dans un environnement en ligne. Cette structure se révèle être une façon naturelle de décrire la plupart des données traitées par les machines.

On peut aussi écrire ces triplets en utilisant la syntaxe XML.

C'est un moyen d'échange et de développement des métadonnées.

Cette sémantique (attributs, propriétés) ainsi que le fait d'éclater l'objet à décrire en trois éléments (sujet, prédicat, objet) est à rapprocher de la structure du modèle conceptuel FRBR.

RDF permet de représenter l'information sous forme de graphes des données reliées et d'aboutir à un réseau global d'informations.

L'intérêt est de représenter les données contenues dans les notices : les données sont ouvertes et exploitables selon les besoins des utilisateurs.

Des questions, latentes à cet espace de libre circulation, restent encore en suspens:

- des questions de droit : pourra-t-on "tatouer" la donnée pour savoir qui l'a créée?
- des questions de normalisation et de contrôle : qui va contrôler la validité des données au niveau national et international ? Qui va s'assurer de leur pérennité ?

Sur la base du langage RDF se sont ensuite développés des vocabulaires spécifiques destinés à des applications particulières : ces vocabulaires servent à décrire les données et les relations.

#### 4.3.4 Les vocabulaires du web

Comme nous l'avons déjà vu, les métadonnées ne sont plus l'apanage des bibliothèques mais sont générées par différents métiers. Les vocabulaires et règles ont toujours existé mais l'informatique en a multiplié la nature et le nombre.

Pour développer l'interopérabilité des données des bibliothèques sur le Web, le W3C a dressé une liste de vocabulaires /ontologies dont voici quelques exemples :

- le FOAF (Friend of a friend) décrit les personnes et les relations entre personnes (concept du web 2.0) tout en utilisant les technologies du web sémantique. Ce vocabulaire permet donc de



traiter les données personnelles et de les échanger. Les classes (sujet) et les propriétés (prédicats) sont déclarées dans des vocabulaires pour être réutilisées (par exemple OWL, RDFS).<sup>17</sup>

- le SKOS (Simple Knowledge Organisation Systems ou Système simple d'organisation des connaissances) est un vocabulaire RDF permettant de décrire les classifications et les thésaurus. Les concepts du langage d'indexation Rameau ont par exemple déjà été traduits en SKOS ainsi que les autorités sujet de la Library of Congress. De même, les archives de France développent un projet de mise à disposition d'un thésaurus dans les normes du web sémantique. D'autres exemples sont développés dans la troisième partie.

- TEF (Thèses électroniques françaises) qui regroupe les descriptions et les informations de gestion ;

- EAD (Encoded archival description) qui permet l'encodage des instruments de recherche archivistiques (cf Calames) ;

- METS (Metadata encoding and transmission standard description) pour les collections d'objets numériques.

Ces technologies permettent une plus grande souplesse pour la réutilisation des données que les outils traditionnels.

Parallèlement au « déploiement de la logique des métadonnées, on assiste à un effort visible de mutualisation et de rationalisation de référentiels métiers communs ». [23]

Comment échanger des données de moins en moins bibliographiques mais émanant d'institutions différentes donc de plus en plus hétérogènes? Comment faire communiquer deux bases de données, deux systèmes d'information différents ?

Pour être interopérables, les systèmes d'informations utilisent des outils de médiation (ontologies, modèles de référence/ référentiels) auxquels on rattache les données.

Des ontologies, qu'est-ce à dire ?

Ce terme d'essence philosophique a été adopté par les chercheurs du web et de l'intelligence artificielle. Les définitions d'ontologie ne sont pas toujours alignées, ce qui complique la compréhension et rend l'ontologie difficile à différencier d'une base de connaissances.

Nous comprenons l'ontologie comme un vocabulaire commun pour ceux qui ont besoin de partager l'information dans un domaine. Elle inclue des définitions lisibles en machine des concepts de base de ce domaine et de leurs relations. [18]

Un type d'ontologie caractéristique du Web possédera une taxonomie et un ensemble de règles d'inférence :

- la taxonomie définit des classes d'objets et les relations entre eux ;

- les règles d'inférence fonctionnent avec un mécanisme d'héritage (conséquence logique), par exemple dans le CIDOC CRM que nous détaillerons plus loin.

La modélisation du savoir sous forme de « concepts » pour faciliter le partage et apporter des solutions aux problèmes de volume, de complexité et d'hétérogénéité croissants des informations est une approche nouvelle. La modélisation est indépendante de tout format de données, de toute syntaxe.

---

<sup>17</sup> C'est d'ailleurs un faux débat d'opposer le web dit 2.0 au web dit 3.0 car les outils du web sémantique servent au web social.

Le modèle de référence FRBR vise à permettre la modélisation des métadonnées. L'enjeu est d'engager une réflexion sur toutes les utilisations possibles des métadonnées et de structurer différemment les données sur lesquelles s'appuient la navigation et l'affichage.

## 5 Un outil de médiation des données : le modèle conceptuel FRBR

L'IFLA s'est doté en 1998 d'un modèle conceptuel de représentation des données bibliographiques dénommé FRBR (Functional Requirements for Bibliographic Records). Ce modèle, qui fait appel à l'appariement « entité-relation », a été élaboré entre 1990 et 1997 à l'intention d'un public élargi (bibliothécaires, éditeurs, libraires, responsables de la gestion des bases de données ...). [49]

### 5.1 Présentation des FRBR

#### 5.1.1 Définition

D'abord comprendre l'acronyme FRBR : à quoi correspondent ces « fundamentals requirements for bibliographic records » ? La traduction parle de « spécifications fonctionnelles des notices bibliographiques », c'est-à-dire :

- d'établir un niveau minimal de fonctionnalité et de spécification<sup>18</sup> de données ;
- de rationaliser le travail de catalogage en disposant d'une « norme de niveau minimal », en établissant pour les notices un ensemble de données communes à diverses communautés, un « standard » dans la lignée du Dublin Core allégé pour les métadonnées.

Les FRBR sont donc un modèle (et pas un code ni une norme) préconisant les fonctions d'une notice bibliographique de base quel que soit le format utilisé (UNIMARC, MARC21...).

Un modèle conceptuel, qu'est-ce donc ?

Un cadre conceptuel est un "canevas" qui guide la construction d'un dispositif et des bases des systèmes d'information à associer, qui permet de les documenter [23] et d'élaborer un outil opérationnel qu'une machine pourra traiter.

Les modèles sont utilisés dans différents métiers, contextes et particulièrement dans les sciences de l'information pour représenter, expliciter un système en fonction d'un point de vue particulier: dans notre cas comprendre ce que signifie cataloguer un document. Les FRBR peuvent être considérés comme un modèle d'un modèle si on considère qu'une notice est une représentation

---

<sup>18</sup> On peut d'ailleurs se demander comment « spécifier » (terme un peu martial) sans tomber dans la norme. Le terme de « préconisations » serait peut-être plus neutre s'agissant d'une démarche volontaire.

d'un document. Pour résumer, FRBR est une représentation et une description simplifiée de l'univers bibliographique. [45]

### 5.1.2 Un cadre théorique

Il s'agit d'un cadre théorique pour :

- faire comprendre « l'essence même de ce sur quoi la notice bibliographique est censée renseigner » [49] en structurant autrement les données ;
- faire adhérer le besoin réel de l'utilisateur aux éléments de données contenus dans une notice : on reprend une des lois de Ranganathan sur le catalogue « Le lecteur ne doit pas avoir à demander à consulter un document pour savoir s'il répond ou non à ses besoins ».
- réduire les coûts de catalogage en rationalisant les opérations de saisie

Ce cadre se veut le plus large possible : il s'applique à tous les types de documents que l'on peut trouver dans un centre de ressources (textuel, cartographique, audiovisuel...), conservé sur tous supports (papier, disque optique...), utilisant tous types de techniques d'enregistrement (acoustique, numérique...).

Signalons que si la force d'un modèle théorique est de permettre la compréhension d'entités complexes, le risque est de rester confiné dans des préoccupations trop abstraites et déconnectées de la réalité bibliographique.

### 5.1.3 Au centre du modèle : l'utilisateur

Dans la lignée de Cutter, l'utilisateur est au centre du modèle.

Paradoxalement, aucune étude n'a été menée sur l'utilisateur pendant la période de gestation du modèle. L'avis d'experts en catalogage connaissant les pratiques de recherche des usagers a seulement été utilisé. [31]

Quelles fonctions doit remplir la notice bibliographique pour qu'elle soit utile à l'utilisateur ?

Les exigences croissantes de l'utilisateur (habitué aux recherches sur le web) et ses besoins sont pris en compte : pour améliorer sa satisfaction, il est essentiel de comprendre, d'analyser les « tâches » qu'il met en œuvre lors de la consultation d'un catalogue.

Quelles sont ces fonctions? Une donnée se justifie par son utilité : elle doit permettre à l'utilisateur de trouver/ identifier / choisir obtenir :

- 1- utiliser la donnée pour trouver
- 2- identifier une entité (vérifier la correspondance entre une description et un document)
- 3- sélectionner l'entité la plus adaptée aux besoins : par exemple une version linguistique
- 4- utiliser la donnée pour accéder à l'entité

On pourrait rajouter une dernière fonction : la mise en relation par la navigation. [66]

Cette démarche fonctionnelle de caractérisation de la notice, d'adéquation entre des besoins et l'obtention d'un résultat est essentielle et d'ailleurs rappelée dans les PIC de 2009 [26] pour un catalogue efficace et performant.

Ces opérations correspondent, interface lecteur, à :

- la saisie des critères de recherche : le premier réflexe pour trouver une œuvre est d'écrire le titre correspondant puis le nom de l'auteur
- après cette première recherche, la pertinence des résultats va être évaluée : identifier les plus pertinents des cinquante résultats générés : vérification de l'intitulé exact du titre et éventuellement relancer une requête avec le titre exact et le nom de l'auteur
- l'affinage des résultats
- l'obtention de l'exemplaire

Pour plus de détails, on pourra se référer à l'exemple du Portrait de Dorian Gray développé par Maxwell [57].

FRBR est un outil pour comprendre la structure d'une notice mais aussi l'utilité des données d'un catalogue.

Une particularité est l'utilisation d'un vocabulaire normé (entités, attributs,...) pour permettre l'« intercompréhension » entre communautés et systèmes d'information.

L'analyse des objets documentaires et les relations qui les unit (comme dans un système de gestion de bases de données relationnelles) est essentielle.

FRBR est donc un cadre général qui définit les éléments [minimum] d'une notice bibliographique vitaux/essentiels pour l'utilisateur et les relations entre ces éléments.

Pour réaliser cet "audit" des systèmes d'information bibliographique, un vocabulaire normé (entité, attributs...) est nécessaire. Il permet « l'intercompréhension » dans l'analyse des données. Mais avant tout, c'est une structure qui permet d'organiser les données.

## 5.2 La structure du modèle

La pratique catalographique est d'identifier une unité d'information [un document] par des données physiques: l'exemplaire nous donne des précisions sur son aspect matériel, sa publication et nous permet de remonter à son contenu intellectuel (titre...). Pour cela, il utilise la structure développée par l'ISBD qui découpe la page en éléments signifiants. On peut dire que, jusqu'à présent, « tout le catalogage a été effectué au niveau de la manifestation ». [59]

Ce catalogage permettait d'être en adéquation avec un fonds des bibliothèques mais aussi des librairies/ éditeurs qui commercialisent les exemplaires.

Les FRBR, eux, partent du CONTENU, des « objets » que décrivent les données en commençant par l'*œuvre*, nœud central, pour descendre vers l'exemplaire. D'après le texte des FRBR, « l'analyse s'attache non pas à des données particulières mais aux choses que décrivent les données » [49].

Pour représenter autrement des données pertinentes pour l'utilisateur de notice bibliographique (qu'il soit catalogueur ou lecteur), le modèle FRBR définit un cadre, des « entités » (« objets fondamentaux ») dont la pierre angulaire est l'*œuvre*.

### 5.2.1 Définition des entités et articulation de l'arbre FRBR

Quels sont ces objets fondamentaux identifiés comme pertinents? Quels liens relient ces entités ? Ces entités FRBR de base sont les éléments des métadonnées.

Elles sont réparties en trois groupes :

- des données structurelles
- des données qui sont le point essentiel des notices d'autorité :
  - \* des données de responsabilité (personne / collectivité)
  - \* des données de sujet (concept / objet / événement)

### 5.2.1.1 Les entités structurelles du groupe 1

Elles rassemblent les produits d'une activité (intellectuelle ou artistique) nommés ou décrits dans les notices bibliographiques. Elles sont structurées en niveaux hiérarchiques.

Les FRBR suivent la logique de l'utilisateur qui recherche (en général) une œuvre et non une édition particulière. La démarche est donc inverse de l'analyse bibliographique classique: on part non pas de l'exemplaire pour accéder à l'œuvre mais de l'œuvre, de laquelle part une relation en cascade vers les autres niveaux.

Avec les FRBR, la notice bibliographique traditionnelle est éclatée entre quatre niveaux de description : œuvre / expression / manifestation / item.

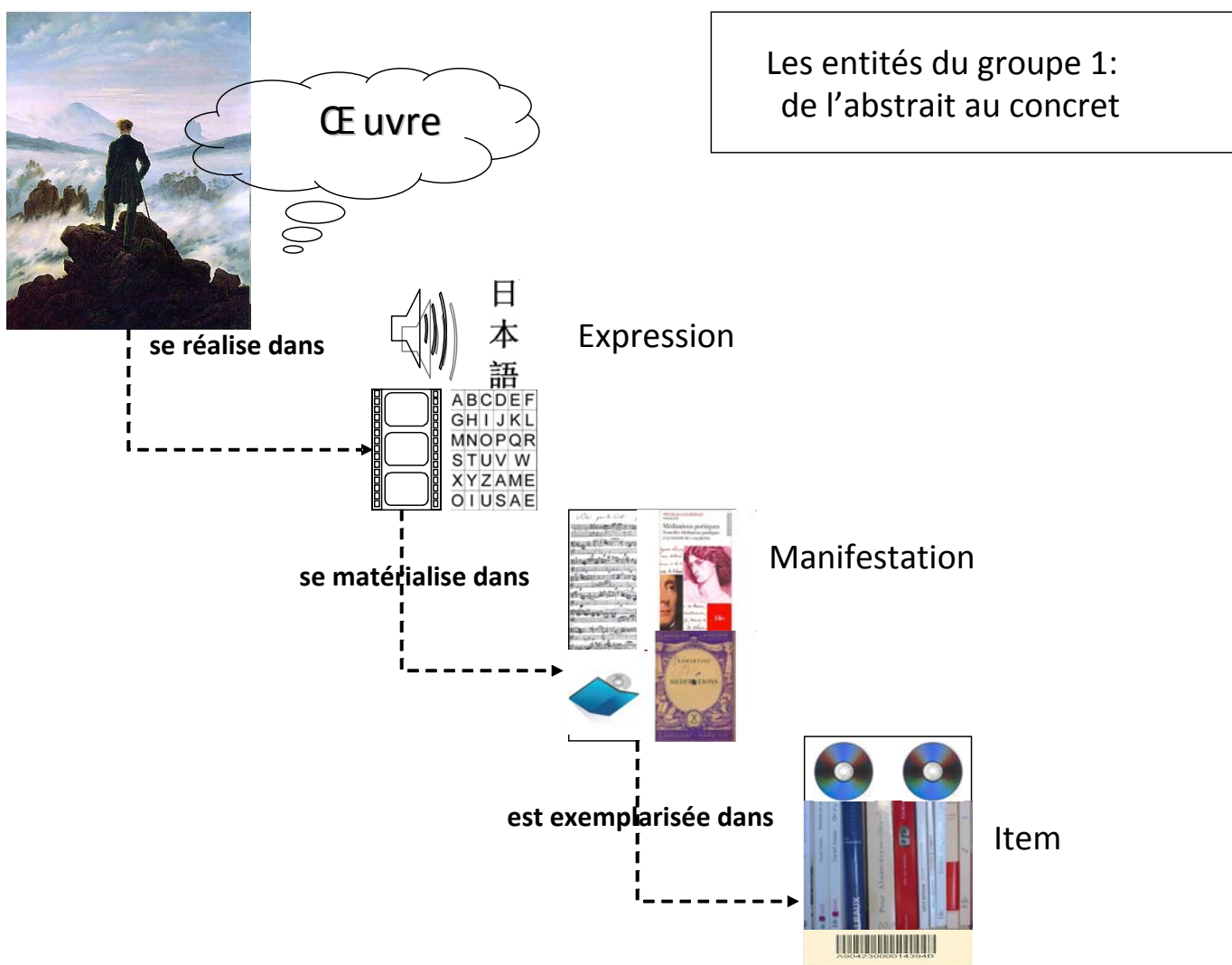


Figure 1 Les entités du groupe : de l'abstrait au concret

Qu'est-ce que l'**Œuvre** dans la terminologie FRBR ? « C'est une création intellectuelle ou artistique déterminée » [49], une abstraction qui peut être commune à plusieurs contenus. Cette idée, intuition non copyrightée d'un créateur, peut sembler difficile à saisir tellement elle est évanescente. On peut cependant l'imaginer le contexte de sa création, par exemple le cheminement intellectuel du génie romantique qui a présidé au Méditations poétiques.

Fort heureusement, ce brainstorming créatif se réalise (intellectuellement ou artistiquement) dans une **expression** : une "forme" quelle qu'elle soit (sonore, visuelle, écrite, gestuelle...) bref, une trace qui permettra de l'interpréter, de l'adapter.

Une **expression** est donc une forme épousant les différents états d'un texte : traductions, versions... Elle permet de définir l'*œuvre* et de l'exprimer.

Les entités *manifestation* et *item* sont plus conformes à la tradition de catalogage, au contenant décrit dans les notices (pavé ISBD) : elles expriment la présentation matérielle du document.

La **manifestation** est la matérialisation de l'*expression* de l'*œuvre*, c'est-à-dire sa publication, quelque soit le support : manuscrits, livres, périodiques, affiches, films, multi support... [49]

L'utilisateur, lui, attend d'obtenir une instance unique d'une *manifestation*, c'est-à-dire un *item*.

L'*item* est donc le bout de la chaîne, l'objet conservé (par la bibliothèque, le Centre de ressources).

Si on résume avec une approche bottom-up les entités du groupe 1, retenons déjà que pour avoir accès à une *œuvre*, il faut disposer de l'*item* d'une *manifestation* concrétisant une *expression* de cette *œuvre*.

#### 5.2.1.2 Illustration

Les FRBR ER ayant été conçus pour le monde des bibliothèques, nous allons développer à titre d'illustration de cette théorie un exemple qui part d'un livre.

Imaginons donc un roman qui a donné lieu à un film et a également inspiré une adaptation théâtrale. L'*œuvre* regroupe sous sa bannière les *œuvres* qu'elle a générés :

Œuvre 1 Effroyables Jardins, roman de Michel Quint

Expression 1 le texte original en français

Manifestation 1 le livre édité en 2000 aux éditions Joëlle Losfeld

Manifestation 2 le livre édité en 2004 chez Gallimard...

Expression 2 la traduction en anglais In Our Strange Gardens

Œuvre 2 Effroyables Jardins, film de Jean Becker

Expression 1 la version française

Manifestation 1 la version DVD parue en 2003

Expression 2 la version anglaise

Expression 3 la version anglaise sous titrée en français...

Œuvre 3 Effroyables Jardins, pièce de théâtre mise en scène par Marcia de Castro

Expression 1 l'interprétation de Michel Sauzet...

Signalons que le titre du roman peut lui-même être relié au poème d'Apollinaire, *Les grenadines repentantes*<sup>19</sup>.

---

<sup>19</sup> In *Calligrammes, poèmes de la paix et de la guerre (1913-1916)*.

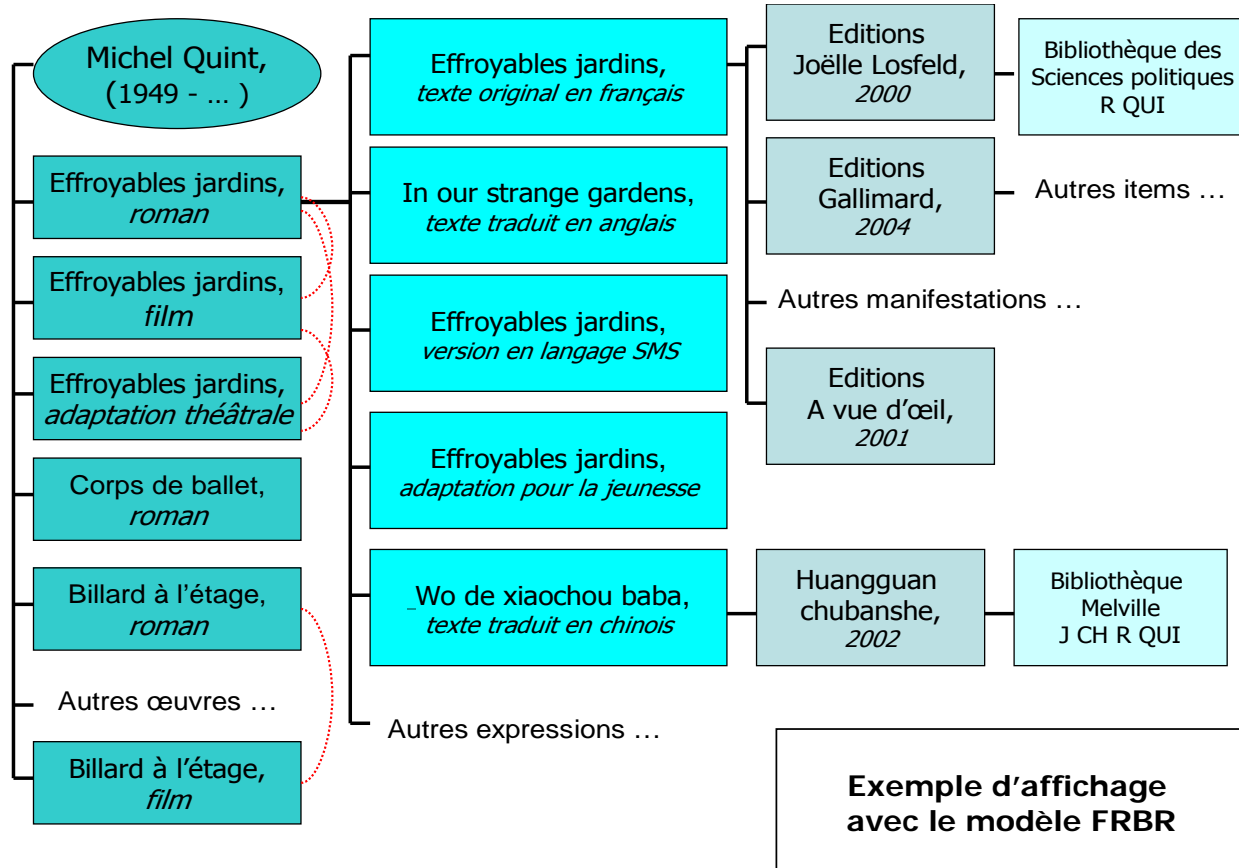


Figure 2 Exemple d'affichage avec le modèle FRBR

A partir du noyau de l'œuvre, un regroupement hiérarchique permet :

- de présenter de manière synthétique l'œuvre et les entités dérivées (c'est la "collocation", au sens de regroupement de toutes les productions d'une œuvre, d'une personne...);
- de visualiser les relations/ interrelations entre notices : dans notre exemple inabouti, cinq instances d'expressions réalisent une instance d'œuvre ;
- d'enrichir le contenu par des renseignements complémentaires sur d'autres œuvres liées, d'une nature autre que textuelle (pièce de théâtre, film, musique de film, spectacle de cirque, tableau, œuvre d'art...) dans l'optique (utopique ?) d'un catalogue sans frontière entre objets documentaires.

Grâce à la structure FRBR, on dispose d'une vue panoramique de l'œuvre et de ses produits dérivés. Il n'est plus besoin de rebondir de fiche en fiche, ni surtout de sortir du catalogue pour prendre connaissance des éléments d'information.

### 5.2.1.3 Approfondissement des entités du groupe 1

Sans vouloir se lancer dans l'herméneutique des FRBR, quelques remarques pour souligner des points jugés essentiels ou d'interprétation délicate dans ce modèle. Ces remarques concernent le premier groupe d'entités, le plus important, qui correspond à la description proprement dite des documents, le « pavé ISBD », plus les données locales.

- De l'utilité de l'entité œuvre :



Signalons tout d'abord que l'*œuvre* FRBR est le reflet d'un environnement culturel occidental qui y projette ses représentations de la "création" (par exemple, un Japonais n'a pas la même perception de l'*œuvre* ni de l'*expression*). C'est un peu gênant pour un modèle qui se veut universel et au-delà des différences culturelles.

A quoi sert l'*œuvre* si elle ne se réalise que dans l'*expression* ? Ne peut-on pas faire l'économie de ce niveau abscons pour passer directement à l'*expression*?

Il est pourtant indispensable d'identifier l'*œuvre*, même si ce niveau semble flou, car isoler l'*œuvre* permet :

- de répondre au questionnement de l'utilisateur qui n'a souvent qu'une idée vague de ce qu'il recherche
- d'analyser un contenu avant de s'occuper du contenant (édition, exemplaire...)
- de déployer la palette des *expressions* d'une *œuvre* : toutes les révisions d'un texte ou ses traductions, les versions linguistiques d'un film, les interprétations d'un concerto... et de quantifier les relations entre œuvre et expression(s) [de 1 à n].
- d'établir des liens indirects entre plusieurs *expressions* d'une même *œuvre* lorsqu'il est impossible quand il n'est pas possible de relier les *expressions* entre elles: par exemple quand on ne sait pas quel texte a servi de base dans le cas d'une traduction
- de prendre conscience de la NATURE des relations entre une *œuvre* et une autre *œuvre* mais entre une *œuvre* et une autre entité.

On peut remarquer que le choix du terme *œuvre* dans le modèle FRBR prête à confusion avec son utilisation dans le langage naturel où l'œuvre sous-entend une production.

- Des frontières entre œuvre et expression :

La frontière entre *œuvre* et *expression* est perméable et parfois difficile à apprécier.

- d'après le rapport FRBR, les traductions sont l'*expression* d'une *œuvre* mais ce n'est pas si simple. Ainsi on pourrait considérer que la version française de La Plaisanterie de Kundera, originellement mal traduite en français et en anglais et reniée par son auteur, est une traduction libre donc une autre *œuvre* et nullement une expression du texte original. Par contre les traductions issues de l'*œuvre* en français de Kundera ou revues avec son accord sont bien des *expressions*.

- un résumé d'une œuvre n'est pas l'expression d'une œuvre mais devient une nouvelle œuvre. Ce point pourrait être nuancé en fonction de l'auteur du résumé (résumé d'éditeur ou résumé d'auteur).

- un arrangement est a priori l'expression d'une *œuvre* : cependant, si l'arrangement est poussé à un « degré significatif », cela devient une autre *œuvre*. Le catalogueur aurait donc une liberté d'interprétation en contradiction avec le cadrage voulu par le modèle.

- *Remarques complémentaires* :

- Les exemples d'*œuvre* choisis dans le manuel de l'IFLA sont indiscutablement des œuvres phares, des chefs d'œuvre universels (Roméo et Juliette, Jules et Jim, L'art de la fugue...). Il aurait été intéressant de choisir également des œuvres moins connues et plus sujettes à interprétation : par exemple, l'assemblage de deux œuvres est-elle toujours une œuvre ? Le modèle FRBR orienté objet en affinant la notion d'œuvre nous semble résoudre cette absence de précision.

- On pourrait se demander quelle serait l'*expression* de l'enseignement de Pythagore : son œuvre étant réservée à des initiés et non transmissible par écrit, peut-on parler de réalisation intellectuelle dans ce cas-là ? On passerait directement d'une *œuvre* à une / des *manifestation* (s), matérialisation dans un support.

- Comment considérer les versions en Braille de textes : comme la *manifestation* d'une *expression* d'une *œuvre* ou comme une nouvelle *expression* ?

D'autres sujets de discussions (concernant les ressources continuatives, les structures plurales...) peuvent être trouvés sur la liste FRBR/ IFLA [50] qui héberge les échanges entre spécialistes du modèle, et dans les écrits du spécialiste des FRBR, Patrick Leboeuf. [51] [52] [53]

#### 5.2.1.4 Les entités de responsabilité et de sujet

Elles permettent d'obtenir des renseignements complémentaires sur les entités du groupe 1 et de répondre aux questions sur l'auteur et le sujet.

- Qui a fait l'œuvre / expression / manifestation / item ?

Le groupe 2 réunit les personnes, les collectivités (et les noms de famille) responsables d'une entité du groupe 1 : responsabilité du contenu intellectuel ou artistique, de la production matérielle et de la distribution, ou de la conservation de ces produits (relations d'appartenance, de production...). On peut remarquer que le groupe 2 peut être rapproché de FOAF.

- De quoi parlent l'œuvre / expression / manifestation / item / personne / collectivité ?

Le groupe 3 fédère les entités constituant le sujet d'une œuvre:

*Concept* : une notion abstraite ou une idée (romantisme,...) ;

*Objet* (réalité tangible : Trocadéro, Statue de la Liberté...);

*Événement* (fait accompli ou survenu : Seconde Guerre Mondiale, Grande Peste...);

*Lieu* (coordonnées topographiques : le désert d'Atacama, Sapporo...).

## 5.2.2 Les relations entre entités

Pourquoi est-il important d'établir des relations entre entités?

L'attitude de recherche classique de l'utilisateur est de formuler une requête à l'aide de champs. Cela permet d'obtenir des renseignements quasi exhaustifs sur un document particulier. Cette pratique laisse l'utilisateur confiné dans le cadre de la description " plate " de l'œuvre.

A contrario, à l'aide des relations, l'utilisateur peut « naviguer » autrement dans l'univers représenté dans une bibliographie, un catalogue ou une base de données. Expliciter les relations entre entités permet d'interconnecter les informations et donc, en fournissant des informations supplémentaires, d'enrichir un contenu, de " dynamiser " le catalogue. « Les relations doivent aider à faire le lien entre les entités et permettre à l'utilisateur de trouver l'information qu'il cherche. Dans un catalogue, l'utilisateur doit pouvoir passer d'une notice à des notices connexes en se servant des relations ». [62]

Les relations fondamentales sont établies:

- entre entités bibliographiques (réalisée / matérialisée / représentée par)

- entre entités bibliographiques et entités de responsabilité

Elles indiquent la logique liant les entités bibliographiques et permettent de les co-localiser.

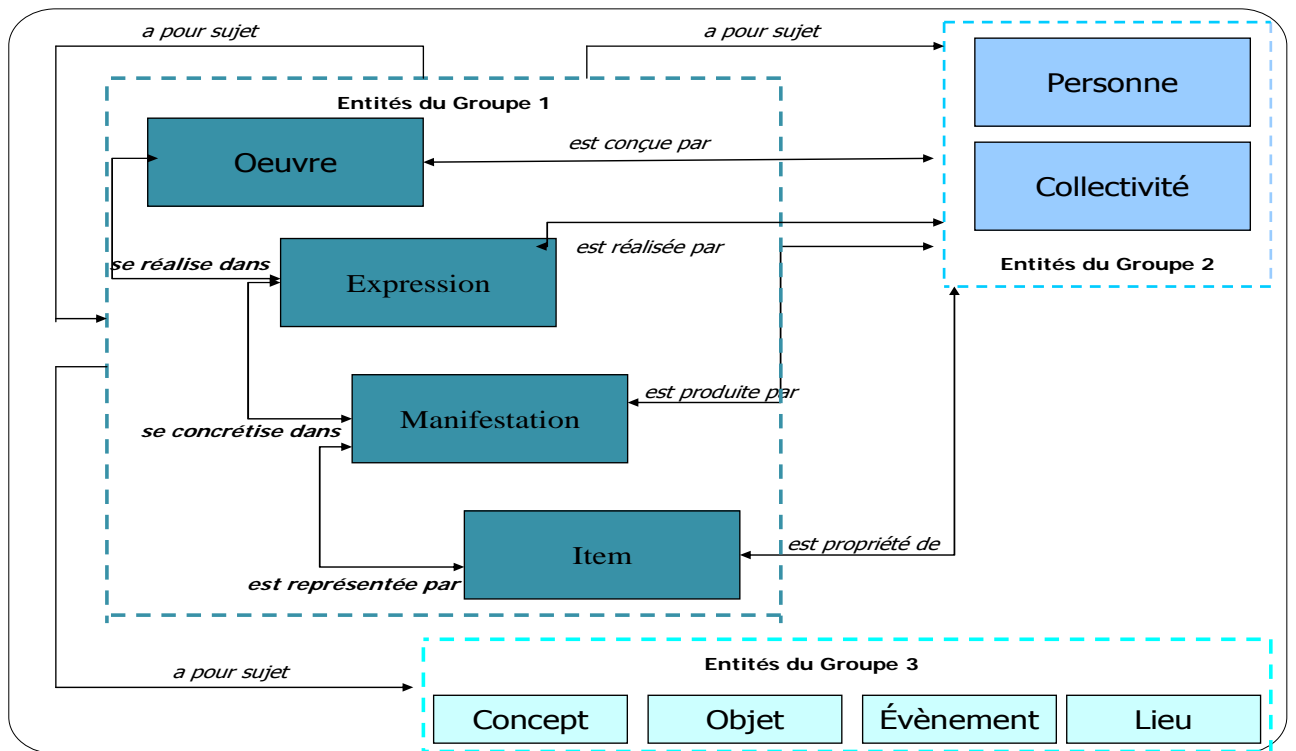


Figure 3 Entités FRBR et principales relations verticales

Les relations de sujet sont étudiées dans FRSAR et ne figurent pas ici.

Différentes familles de relation existent: de suite, de relation, supplément, de tout à partie sont des exemples de relations « horizontales » d'œuvre à œuvre.

L'objectif n'est pas ici de faire une typologie approfondie des relations mais, par souci de clarification, d'en différencier quelques grandes familles:

- les relations de dérivation (versions, éditions, traductions, révisions) ;
- les relations du tout pour la partie (sélection d'anthologies, collections, séries, chapitres...) ;
- les relations d'accompagnement (supplément, complément, index...) ;
- les relations séquentielles (parties d'une collection, d'une monographie en plusieurs volumes) ;
- les relations entre ressources ayant des caractéristiques communes (auteur, éditeur, titre, sujet...).

Dans le cas d'un volume d'une trilogie, l'œuvre représente alors un élément d'une œuvre de plus vaste dimension » (relation de tout à partie).

Mais le " jeu" de poupées russes n'est pas fini : on peut déployer d'autres relations en combinant les entités du groupe 1 avec les autorités et les accès sujet.

L'objectif est d'obtenir un niveau minimal de données utiles pour l'utilisateur : le niveau de complétude de la notice peut ensuite varier.

En poussant les potentialités du modèle entité relation à l'extrême (c'est-à-dire en reliant l'œuvre à la hiérarchie d'œuvres et aux neuf autres entités), on pourrait obtenir un réseau de relations réticulaires quasi inextricable.

## 5.2.3 Les attributs

A chaque entité sont associés des éléments d'information dénommés « attributs » dans le vocabulaire spécifique de la modélisation. Ces caractéristiques permettent d'identifier l'entité mais aussi de distinguer les exemplarisations d'une même entité.

Parmi les dizaines d'attributs proposés et non exhaustifs, une sélection des caractéristiques des entités du groupe 1 les plus pertinentes pour un fonds documentaire est ici proposée:

Tableau 1: Sélection de caractéristiques identifiant les entités structurales

Œuvre	Expression	Manifestation **	Item
Titre*	Titre*	Titre*	Identifiant****
Forme	Forme	Mention de responsabilité	Provenance
Date	Date	Mention d'édition	Etat matériel
Autre caractéristique distinctive	Langue***	Mention de collection	Annotations/inscriptions
Complétude visée	Résumé du contenu	Type de support	Restrictions de communication
Public visé	Contexte de l'expression	Configuration requise (ressource électronique)	...
...	Aptitude de l'expression à recevoir un développement	Mode d'accès (ressource électronique accessible à distance)	
	Aptitude de l'expression à recevoir une révision ...	Adresse d'accès (ressource électronique accessible à distance)*** ...	
		Identifiant (ISSN, ISBN...)	

\* On peut remarquer que l'attribut titre (nom de contenu) se retrouve à trois niveaux d'entités (œuvre / expression / manifestation). Le titre d'expression n'est pas explicité par le modèle.

\*\* D'autres attributs de manifestation prennent en compte les spécificités des œuvres autres que bibliographiques : musicales / enregistrements sonores (attributs de manifestation : largeur de sillon, mode de gravure...), images projetées (polarité, génération, format de présentation), cartographiques...

\*\*\* Il semble étonnant que l'attribut de langue figure au niveau de l'expression et non de l'œuvre. En effet, la langue de la création est inhérente à l'œuvre (on pense, on chante en une langue) et non à l'expression.

\*\*\*\* L'adresse d'accès à la ressource électronique (URL) se trouve au niveau de la manifestation et non de l'item [49].

Est-ce là une façon de régler le traitement des publications électroniques natives, un peu transparent dans le rapport de l'IFLA ? La solution proposée est de mettre au niveau manifestation les nombreuses variations de formats (PDF/ HTML) matérialisant la même expression.

#### 5.2.4 Les identifiants des entités

Se pose la question essentielle des identifiants pour les entités qui constituent l'ossature du modèle FRBR, en particulier pour les entités abstraites - *l'œuvre* mais aussi *l'expression* - et les entités du groupe des responsabilités.

D'après Pisanski, un « système d'identifiants distincts pour les entités devrait aider les utilisateurs à trouver leur chemin, et cela indépendamment de la façon d'interpréter l'univers bibliographique ».

[60]

Or, le développement des technologies de l'information et de la communication crée les conditions d'un partage efficace et rentable du travail, même à une échelle mondiale.

Des identifiants reconnus internationalement sont nécessaires pour :

- afficher l'information de façon bien identifiable ;
- identifier de façon unique les instances.

Ils sont la « clé d'une intégration réussie des données bibliographiques dans le web sémantique » dans la perspective d'une utilisation par les bibliothèques du web de données. Le projet VIAF (Virtual International Authority File), lancé en 2003, avait justement pour objectif d'identifier les agents en leur attribuant une URI.

Mais, comme le souligne Pisanski lors du congrès de l'IFLA en 2010, « la plupart des outils d'identification des entités FRBR à une échelle mondiale ne sont pas encore en place ». [60]

## 5.3 L'utilité du modèle FRBR<sub>ER</sub>

On pourrait trouver paradoxal qu'après quelques décennies de catalogage dans un monde numérique un modèle se soucie enfin de la structuration et de l'utilité d'une notice bibliographique. Mais les FRBR offrent bien plus.

Ce modèle permet d'une part de formaliser des concepts afin d'autre part d'améliorer le produit final, le catalogue.

### 5.3.1 Le modèle entité relation, un support de réflexion théorique

Comme le souligne Françoise Leresche, les FRBR sont « au centre des travaux sur les catalogues ». [56]

En quoi ce cadre de référence pour comparer des données qui « peuvent ne pas être structurées de la même façon » peut-il susciter une révolution culturelle dans le monde du catalogage ? [54]

Les pistes de réflexion se présentent à plusieurs niveaux :

- une réflexion sur l'essence même du catalogue de bibliothèque, ses fonctionnalités, ses objectifs mais aussi ses limites / frontières [39] dans un monde numérique où désormais les métiers, les informations s'imbriquent ;

- une explicitation de la structure du catalogue ;

- une réflexion sur le contenu de l'information bibliographique ;

Que mettre dans la notice ?

Comment restructurer les données ?

Le but n'est plus d'être exhaustif dans la description, mais d'être pragmatique : il faut reconsidérer le niveau minimal de description, celui qui est satisfaisant pour l'utilisateur, et s'en satisfaire.

Comment approfondir l'*œuvre*, jusqu'alors sous valorisée dans les catalogues aux dépens des mentions de publication et des exemplaires ?

Comment éclater les données traditionnellement enfermées dans une notice monolithe pour les répartir dans des cases interconnectées ?

- une réflexion sur les règles de catalogage et la possibilité de créer des formats qui rendront compte des entités, des attributs et des relations entre entités.

Plus globalement ce modèle conduit à une mise à plat de l'information bibliographique et de la valeur scientifique de l'information produite.

On met désormais au second plan les types de support (imprimés ou ressources en ligne). Les relations entre les objets d'information et les métadonnées des déclarations entre entités deviennent fondamentales ainsi que les identifiants.

D'autres questions connexes au catalogage mais plus globales peuvent être soulevées. Comment identifier le producteur d'une donnée ? Quels rôles doivent jouer les agences bibliographiques nationales dans la production des métadonnées ? Quel avenir pour la profession ? Va-t-on vers un rapprochement du métier vers celui de gestionnaire de base de données ?

### 5.3.2 Le modèle entité relation, une utilité pratique ?

Le but est d'améliorer la notice bibliographique (le produit), mais aussi les opérations de catalogage et le produit final (le catalogue). [45]

Quels sont les conséquences prévisibles de ce modèle orienté utilisateur ?

#### 5.3.2.1 Les conséquences pour l'utilisateur final

L'approche « orientée utilisateur » préconisée par le modèle est nouvelle ; les FRBR formalisent après coup les pratiques de l'utilisateur.

La nouvelle organisation de l'information bibliographique selon le modèle FRBR vise notamment à améliorer la présentation des catalogues au niveau de la recherche et de l'affichage. [65] [68] [69]

##### ➤ AU NIVEAU DE LA RECHERCHE :

- une amélioration de la recherche dans les catalogues en ligne : la focalisation sur l'usager est l'occasion de concevoir des catalogues en ligne qui correspondent mieux à la manière de chercher des utilisateurs.

Collocaliser les données permet de multiplier les points d'accès pour les utilisateurs.

- une navigation plus facile donc une recherche plus efficace grâce aux liens entre métadonnées ; Cette structuration optimise l'efficacité des tâches utilisateur listées plus haut : chercher pour trouver et obtenir.

- une recherche facilitée : l'intuitivité du modèle a été démontrée.

##### ➤ AU NIVEAU DE L'AFFICHAGE DES RESULTATS

Les résultats sont plus lisibles et "intelligents", au sens étymologique du terme de lier ensemble pour créer du sens car :

- regroupés au niveau de l'œuvre :

Toutes les ressources qui se rattachent à une même œuvre (traductions, éditions critiques, enregistrements sonores, etc.) sont centralisées sur un écran unique. Le nombre de réponses à une requête est donc réduit et la redondance est limitée.

En cela, les FRBR sont une réponse à la « publication multiple et à la description des versions dans le domaine mouvant et multiple qu'est le Web »<sup>20</sup>.

- mieux structurés :

Les réponses sont organisées dans un ordre logique.

La structure de graphe de relations entre données rappelle le réseau du world wide web et la navigation par hyperliens. On passe du stade du document à celui de l'information dont la granule la plus fine est la donnée.

Le catalogue n'est donc plus dans une logique linéaire héritée des catalogues imprimés et des catalogues sur fiches : il deviendrait un ensemble d'informations interconnectées.

Quelles sont les conséquences prévisibles pour le gestionnaire des bases de données de ces changements dans la structuration de l'information bibliographique ?

---

<sup>20</sup> Emmanuelle Bermès. Figoblog <<http://www.figoblog.org/document594.php>>

### 5.3.2.2 Les conséquences pour le gestionnaire du système d'information

Les conséquences sont prévisibles à plusieurs niveaux, d'abord dans la production de l'information bibliographique et, corollairement, dans l'évolution du métier. [47]

- Dans la production de l'information bibliographique :
  - une meilleure compréhension entre catalogueursCela est rendu possible par l'harmonisation des termes qui sont normés (même appellation) et ont la même signification pour tous : la marge d'interprétation lors du catalogage se veut réduite. Par exemple, on ne retrouve plus les approximatifs « basé sur » / « traduit par ».
- une meilleure organisation des métadonnées donc une limitation des doublons ;
- un catalogage plus facile à terme si on évolue vers un partage ou une copie des métadonnées ;
- une prise en compte de nouveaux supports ressources électroniques ;
- une ouverture prévisible vers les applications du web sémantique et donc une interopérabilité avec les données produites par d'autres communautés.

Un des apports essentiels de ce modèle conceptuel est de permettre l'échange avec d'autres communautés et de ne pas rester cloisonné au cercle des bibliothèques/ documentation. En effet, d'autres pratiques, d'autres réflexions venues du monde de l'édition (éditeurs, diffuseurs, libraires), de l'informatique ou des sciences peuvent enrichir le débat.

Les données générées pourront être récupérées par d'autres producteurs, par des utilisateurs. Ce changement « d'aire géographique » offre à terme la possibilité de relier des données qui ne sont pas dans l'univers bibliographique mais contenues dans d'autres réservoirs (données statistiques, géographiques, juridiques...), d'intégrer non seulement de la documentation mais aussi de l'information.

- une remise en cause prévisible des pratiques, d'un savoir faire métier.
- Les FRBR ne sont pas une révolution car on ne crée pas de nouvelles données mais on les dispose autrement.
- Le catalogueur « atomise » la notice traditionnelle et analyse les liens et la nature des liens entre une œuvre et les autres entités. Il « tisse » un réseau de relations inédites. On pourrait presque rapprocher son attitude du peintre cubiste qui déconstruit l'objet traditionnel (la notice) pour le représenter sous d'autres facettes afin de mieux le révéler au public.
- Cela entraîne un changement d'échelle du catalogage vers l'infiniment ( ? ) petit. Quelle granularité de l'information veut-on obtenir? Jusqu'où arrêter le dépouillement ?

L'application des FRBR nécessitera un effort d'adaptation pour le catalogueur car repenser les éléments fondamentaux d'une notice, c'est aller contre la tradition catalographique (admise, enseignée) et remettre en cause des pratiques quasi normées.

C'est principe des vases communicants : on veut faciliter l'accès aux données pour l'utilisateur mais a contrario la tâche peut être compliquée pour le catalogueur. Une bonne culture générale et une bonne connaissance du fonds à indexer semblent des prérequis pour relier les entités. Une connaissance technique des formats du web dans un environnement numérique en mutation semble nécessaire, ainsi que la collaboration avec d'autres communautés, d'autres métiers.

Avec la multiplication des métadonnées, des outils faciles à utiliser par les non professionnels, peut grandir dans la profession un sentiment de dépossession d'un savoir faire métier. Patrick Leboeuf



posait dès 2004 la question fatale : dans « la perspective du catalogage automatique des ressources électroniques et plus généralement de tous les types de documents via leur version numérique de production, existe-t-il encore un *jour d'après* pour les catalogueurs ». [52]

Suffira-t-il « un jour » d'appuyer sur un bouton pour extraire d'une notice des données automatiquement FRBRisées ? Cette heure n'est pas encore venue. Il nous semble au contraire que la réflexion engagée grâce aux FRBR est une occasion pour le catalogueur : il pourra se positionner non plus en expert des normes mais montrer sa valeur ajoutée en terme d'organisation des informations, d'interconnexion des données et, à un niveau plus général, dans la construction d'ontologie (s).

À plus grande échelle, des conséquences sont prévisibles pour l'ensemble des institutions patrimoniales. Par exemple, « bien que les FRBR soient d'une application limitée pour les archives », les changements dans les normes de bibliothèques (c'est-à-dire l'application par les RDA des FRBR) ont un impact sur les pratiques de description archivistiques. [41]

Le modèle conceptuel FRBR entité relation reste cependant incomplet : d'une part il ne traite pas le contenu des données d'autorité et d'autre part il reste trop replié sur le monde des bibliothèques. C'est pourquoi un enrichissement du modèle et des adaptations ont été menés.

## 5.4 L'élargissement : les modèles liés

Sur le socle de FRBR entité relation, une réflexion commune a été menée entre acteurs culturels à la fois sur les données d'autorité et sur l'élargissement du modèle FRBR à des données issues d'autres communautés.

### 5.4.1 FRAD et FRSAD

#### 5.4.1.1 Les données d'autorité

Les données d'autorité sont les points d'accès contrôlés et toute autre information dont se servent les institutions pour co-localiser les œuvres d'une personne, d'une famille ou d'une collectivité particulières ou les différentes éditions d'un titre. Les points d'accès contrôlés comprennent les formes autorisées et les variantes de forme du nom rassemblées par les catalogueurs pour identifier une entité.

Dans la continuité des FRBR, le groupe de travail FRANAR (*Functional Requirements and Numbering of Authority Records*) a défini les données d'autorité dans le modèle FRAD (2009). L'objectif de FRAD (*Functional requirements for authority data*) est d'identifier et faire une analyse détaillée des entités représentées par les points d'accès dans les notices bibliographiques.

Ses missions sont triples :

- définir les données des entités bibliographiques Personne, Collectivité, Famille (et non pas les points d'accès déjà représentés dans FRBR) ;
- « étudier la faisabilité d'un Numéro international normalisé pour les données d'autorité (ISADN) » et les modalités de sa mise en œuvre ;
- « servir de lien officiel avec l'IFLA et travailler sur les fichiers d'autorité avec d'autres groupes intéressés par ce sujet ». [43]

Ces entités Personne / Collectivité sont connues par leurs noms dont les attributs sont par exemple :

- pour les personnes : date, titre, domaine d'activité, affiliation, adresse langue, biographie ...
- pour les collectivités : lieu, date, langue, adresse, ...
- et pour les relations entre personnes, familles, collectivités et œuvres

Ils constituent la base pour les points d'accès contrôlés.

#### 5.4.1.2 Les autorités sujet

Pour rendre compte de la classification, le groupe FRSAR en entamé en 2007 ses travaux pour les notices d'autorité matière. Le rapport final a été publié en 2010.

FRSAD (*Functional Requirements for Subject Authority Data*) modélise le contenu des entités du groupe 3 avec :

- *thema* : c'est la "super classe" de toutes les entités qui peuvent être le sujet d'une œuvre ;
- *nomen* : tout signe ou séquence de signes (caractères alphanumérique, symboles, sons, etc.) par lequel un *thema* est connu.

La question est désormais de savoir si tous ces modèles liés sont compatibles entre eux. Par exemple, on peut noter que pour l'entité *œuvre*, les attributs sujet, lieu d'origine de l'*œuvre* et histoire ont été ajoutés à ceux déjà retenus dans les FRBR.

Ces initiatives ne vont-elles finalement pas nuire à la volonté d'harmonisation et de cohérence de l'univers bibliographique ?

Les FRBR des bibliothèques ont été rapprochés d'autres modèles, du modèle CRM développé dans les musées et du modèle orienté objet.

#### 5.4.2 Vers un formalisme orienté objet

Dès 2003, un groupe d'harmonisation des FRBR / CIDOC CRM a été constitué pour rapprocher les deux modèles. L'objectif est la mise au point d'un outil conceptuel permettant d'intégrer information bibliographique et muséographique.

L'objectif est de traduire le modèle FRBR en méthodologie orientée objet (nommé FRBR<sub>oo</sub>) pour :

- l'enrichir des concepts du CRM ;
- faciliter l'interopérabilité sémantique des structures informationnelles des bibliothèques et des musées.

Le résultat est la parution en 2008 de la version complète des FRBR selon un formalisme orienté objet : FRBR<sub>oo</sub>, modèle qui « ne se substitue pas aux FRBR mais permet de se rapprocher du web sémantique. » [38]

Ces modèles seront présentés plus en détail dans la deuxième partie concernant l'expérimentation des FRBR.

## 5.5 Du modèle théorique aux applications

Les FRBR restent un cadre de réflexion en particulier sur la façon de déverrouiller les métadonnées des catalogues de bibliothèque, de sortir du « modèle métier » [23].

Ce modèle « ne pousse pas l'analyse jusqu'au degré qu'il faudrait atteindre dans le contexte d'un modèle de données intégralement développé », [49] donc il laisse toute liberté pour encoder les données. Mais, comme le fait remarquer Patrick Lebœuf, les FRBR n'étant pas un modèle de données, comment pourraient-ils donc être « implémentés » ? [54]

Les éléments de réflexion présentés dans cette section ne prennent pas en compte la nouvelle structure de catalogue préconisée par le code de catalogage RDA, encore en discussion<sup>21</sup>.

### 5.5.1 Mettre en œuvre les FRBR dans les catalogues?

La nécessité de mettre le modèle FRBR au cœur des règles de catalogage est un véritable point de convergence entre acteurs de la profession.

Mais comment mettre en place le modèle FRBR? Comment répartir les données structurées dans les niveaux qui peuvent intéresser les utilisateurs pour les mettre sur le web de données?

Mettre en œuvre les FRBR qu'est-ce au juste ?

Cela reviendrait à « éclater » le bloc notice entre différents niveaux d'information :

- l'*œuvre* à peu près assimilable à une notice d'autorité
- la *manifestation* à la notice bibliographique
- l'*item* à la notice d'exemplaire

On a insisté sur l'intérêt des FRBR qui permettent un regroupement par *œuvre* dans les catalogues. Mais cela engendre aussi beaucoup de questions pratiques : combien de notices obtenir pour une *œuvre* ? Ne va-t-on pas vers une multiplication des métadonnées ? On passerait d'une notice par document à une notice par entité. Cela implique un changement de démarche, de ne plus partir de la description de la *manifestation* basée sur l'*item*.

Si tout le monde s'accorde sur la nécessité de passer à « une utilisation plus intelligente des millions de notices bibliographiques, d'autorité et de données locales qui existent en MARC21, de manière à améliorer la conception des systèmes et de FRBRiser l'affichage des OPAC et les index », la mise en pratique est délicate. [58]

Comment récupérer les millions de métadonnées existant en format MARC ? Quelle méthode inventer pour reconstituer des données existantes dans de nouveaux formats ?

On ne peut pas reprendre au cas par cas toutes les données des catalogues : ce serait un travail coûteux en opposition avec les objectifs de réduction du coût du catalogage du modèle développé par l'IFLA.

Peut-on envisager la transformation automatique de la structure actuelle de nos données en une autre structure sans trop de pertes d'information ?

---

<sup>21</sup> Pour une présentation détaillée de RDA, on pourra se référer par exemple aux travaux de Françoise Leresche cités dans la bibliographie ainsi qu'au site du groupe de travail technique chargé d'évaluer l'intérêt et les modalités de l'adoption de RDA dans les bibliothèques françaises. *rda-en-france.enssib.fr*

On entend beaucoup parler de FRBRisation des catalogues, de catalogue ou système FRBRisé. Les fournisseurs de logiciels de bibliothèques commencent à se positionner sur le modèle FRBR. Un livre blanc au titre évocateur (les [FBBR sans attendre](#)) vient même d'être publié sur la question en septembre 2011. [44]

Mais qu'entend-on par FRBRisation ? C'est le « processus d'extraction des entités FRBR des données bibliographiques existantes » [70]. On parle aussi de FRBRisation pour nommer les interfaces des catalogues de bibliothèques regroupant les listes de résultats par ordre logique d'entités du groupe 1.

Le néologisme FRBRisation traduit-il vraiment l'application du modèle FRBR conduisant à une transformation des catalogues ?

Il nous semble qu'existe actuellement une confusion entre la FRBRisation des interfaces et la FRBRisation des DONNEES.

Ce que l'on considère actuellement comme des systèmes FRBRisés et des prototypes n'en sont pas toujours : ils restent trop dépendant des pratiques anciennes et de données uniquement bibliographiques pour l'identification des œuvres. En effet, il serait bénéfique d'intégrer les données d'autorité pour la représentation et l'identification des œuvres. [68]

Peut-on appliquer, « implémenter » les FRBR ?

À l'origine, le modèle a été conçu pour être purement conceptuel. Néanmoins, il a inspiré des réalisations pratiques, notamment pour améliorer l'affichage des résultats dans des bases très volumineuses comme FRBR Display Tool de la Bibliothèque du Congrès.

Une étude est actuellement menée à partir de deux catalogues basés sur les FRBR (Fiction Finder d'OCLC et le catalogue de la Bibliothèque d'Australie) pour évaluer leur adéquation avec les tâches de l'utilisateur définies par le modèle [69]. Les résultats de cette étude permettront d'identifier les options d'interface les plus adaptées pour une future implémentation des FRBR.

Quels choix et quelles modalités d'intégration ? [67]

D'après les auteurs du Livre blanc sur les FRBR [44], deux options sont envisageables :  
- une transformation des « SIGB actuels pour y ajouter la notion d'œuvre. Puis répercuter cette modification dans les interfaces de consultation des catalogues : on pourrait parler de «FRBRisation» des notices bibliographiques. L'œuvre devient un simple enrichissement de la manifestation.

- une réécriture de nouvelles applications « en adoptant une modélisation FRBR de la base de données et en mettant en œuvre les recommandations du RDA pour la codification des données. L'œuvre devient le cœur du processus descriptif. »

On peut au final se demander si les éditeurs de SIGB ne confondent pas FRBRisation des interfaces, parfois appelée « recherche à facette fondée sur le modèle FRBR », avec FRBRisation des données.<sup>22</sup>

A ce sujet, on pourrait d'ailleurs rappeler l'avertissement lancé dès 2003 par Patrick LeBœuf : la terminologie abstraite des FRBR « ne devrait pas être purement et simplement incorporée telle quelle aux ISBD et aux règles de catalogage » [51] Pour ne pas galvauder le sens de *l'œuvre*, il aurait fallu mettre en place une passerelle entre la terminologie FRBR et celle des règles de description/ catalogage.

---

<sup>22</sup> Couperin <<http://www.couperin.org/fr/presentation/nos-cooperations-nationales-et-internationales/icolc/item/550-congres-icolc-printemps-2008-a-san-francisco>>

Afin de présenter l'intérêt des FRBR de manière plus concrète que par l'étude du texte de l'IFLA, il nous semble essentiel de l'illustrer avec des catalogues inspirés du modèle FRBR.

Nous trouvons plus approprié de parler de mise en œuvre du modèle FRBR et de "catalogue inspiré du modèle FRBR" dans les expériences présentées ci dessous.

## 5.5.2 Des exemples de catalogues inspirés du modèle FRBR

Pour mieux comprendre l'intérêt des FRBR pour un catalogue, notre démarche a été de rencontrer sur le terrain les pionniers français de la FRBRisation d'abord à la Cité de la Musique puis à la Gaîté lyrique.

### 5.5.2.1 Le catalogue bibliographique de la Cité de la Musique

La Cité de la Musique a mis en ligne un catalogue bibliographique inspiré des FRBR. Au centre du catalogue, est placée l'œuvre musicale. Elle est accessible à partir du titre français le plus utilisé et provient d'une reprise de l'existant : l'autorité titre (TUM = Titre Uniforme Musical). Ces fichiers d'autorité ont permis de relier chaque entité à la notion centrale d'œuvre.

L'utilisateur fait une requête simple, par exemple sur un requiem.

Il aboutit à une liste de résultats triés par pertinence dans laquelle il sélectionne l'œuvre de son choix.

- Trier par
- 1) ~~Berlioz, Hector / Grande messe des morts , H 75~~
  - 2) Brahms, Johannes / Un requiem allemand , Op. 45
  - 3) Schütz, Heinrich / Musikalische Exequien , Op. 7, SWV 279-281
  - 4) Fauré, Gabriel / Requiem , Op. 48
  - 5) Verdi, Giuseppe / Requiem
  - 6) Mozart, Wolfgang Amadeus / Requiem , Ré mineur, KV 626
  - 7) Charpentier, Marc-Antoine / Messe des morts à quatre voix , Ré mineur, H 7
  - 8) Liszt, Franz / Transcription pour piano du "Confutatis" et du "Lacrymosa" du Requiem de Mozart , S 550
  - 9) Charpentier, Marc-Antoine / Messe pour les trépassés , H 2
  - 10) Weill, Kurt / Das Berliner Requiem
  - 11) Zeisl, Eric / Requiem ebraico
  - 12) Britten, Benjamin / War requiem , Op. 66

Oeuvres

Manifestation

### Notice de l'oeuvre sélectionnée

#### Titre

Grande messe des morts / Hector Berlioz  
fin mars-juin 1837 / H 75

#### Auteur(s)

Berlioz, Hector (1803-1869), compositeur

#### Sur l'oeuvre

Versions révisées en 1852 et 1867.-

#### Sur la création de l'oeuvre

Créée à Paris, Saint-Louis des Invalides, le 5 décembre 1837.  
Première édition : Paris, M. Schlesinger, 1838

#### Sources

BN

#### Document(s) numérique(s)

##### CD Audio(s)

- Requiem, Op.5 "Messe des morts" / Hector Berlioz ; Peter Schreier, ténor ; Choeur et orchestre symphonique de la radiodiffusion bavaroise ; direction : Charles Munch
- Requiem : Hostias / Hector Berlioz ; Orchestre de l'Opéra de Paris ; Choeurs de la Radio Française ; dir. Hermann Scherchen, 3 min 59 sec

##### Partition(s)

- Grande Messe des morts, "Requiem" / Hector Berlioz ; Edited by Jürgen Kindermann, 1 conducteur (147 p.)

#### Document(s)

##### Article(s)

- Les grands requiem et leur message / par José Bruyr

Après sélection d'une oeuvre (dans notre exemple la Grande Messe des Morts de Berlioz), il accède à ses *manifestations* (nommées ici "document", en langage non FRBR) qui peuvent être matérialisés sur différents supports (partition, article, CD audio, vidéo...). Une *oeuvre* peut aussi être reliée à des interprétations d'oeuvres enregistrées lors de concerts. A partir du noyau titre uniforme d'oeuvre, l'*oeuvre* est contextualisée par les informations sur sa création, ses versions.

A partir de cette page, l'utilisateur peut :

- explorer les autres oeuvres associées au requiem ;
- rebondir sur le compositeur pour recueillir des informations complémentaires sur cette Personne et pas seulement en tant qu'auteur mais aussi de sujet d'une oeuvre (on accède par exemple à une biographie de Berlioz).
- décider d'en savoir plus sur chaque matérialisation, par exemple sur l'article de la revue Musica documentant les grands requiem sur les détails de la partition Requiem : les interprètes, l'éditeur, la composition des chœurs...
- et accéder aux exemplaires :

**Titre**

Grande Messe des morts, "Requiem" / Hector Berlioz ; Edited by Jürgen Kindermann

**Auteur(s) / Interprète(s)**

Berlioz, Hector (1803-1869)

Kindermann, Jürgen

Macdonald, Hugh

**Editeur**

Kassel : Bärenreiter, cop. 1978

**Description matérielle**

1 conducteur (147 p.)

**Notes**

Préface en anglais, français et allemand

**Informations sur l'oeuvre**

Grande messe des morts

ténor1

**Choeurs**

choeur\_mixte1 - STB. 80 sopranos, 60 ténors et 70 basses (!),  
choeurs à 3 voix divisées.

**Type d'ensemble**

orchestre\_symphonique1. très grand orchestre symphonique, avec  
quatre "petits orchestres" de cuivres

0102556

Items

**Exemplaires****Localisation**

Médiathèque

**Cote**

500.E BER:PP:39749

La Médiathèque n'a pas gardé l'entité expression car :

- l'implémentation est difficile à appliquer (grande complexité pour l'utilisateur et mise en place qui aurait généré un niveau d'arborescence "exponentiel") ;
- la nature et le volume du fonds sont peu adaptés.

Cependant l'expression s'exprime à l'affichage dans les interprétations, arrangements, les partitions.

Concernant les documents numériques, l'implémentation a été possible grâce aux enregistrements faits à la Cité de la Musique et à la Salle Pleyel pour la musique classique et pour le jazz et grâce aux sources (ouvrages, connaisseurs et programmes).

Au total, l'utilisateur (en particulier l'auditeur des concerts de la Cité de la Musique et de la salle Pleyel) a un accès plus simple à des œuvres documentées, bien structurées et reliées à d'autres supports. De plus, des passerelles sont tissées entre l'œuvre musicale et l'objet muséographique conservé dans les collections du Musée de la Musique.

La visibilité du catalogue de la Cité de la Musique sort renforcée de la FRBRisation.

En back office, il a fallu procéder à une reprise d'un existant : récupérer les [anciennes] notices au format MARC et compter sur la valeur ajoutée des catalogueurs - leur connaissance du fonds - pour retrouver tous les documents concernant une œuvre musicale et faire les liens.



### 5.5.2.2 D'autres expériences de catalogues inspirés du modèle FRBR

#### 5.5.2.2.1 En France : le catalogue du Centre de ressources de la Gaîté lyrique

Ouvert depuis mars 2011, le Centre de ressources de la Gaîté lyrique a créé ex-nihilo un catalogue inspiré de la réflexion FRBR. Ce catalogue sera mis en ligne à l'automne 2011. Le savoir est organisé dans de nouveaux domaines décloisonnant les disciplines (corps et nouvelles technologies, littérature numérique, réseaux sociaux et nouvelles formes du web, street art...). L'important est que le public qui n'est pas forcément un adepte des bibliothèques accède aux objets documentaires.

Ce catalogue, qui va s'enrichir progressivement de ressources issues des événements, est déconcertant au premier abord par les thèmes abordés, la simplicité de l'interrogation (pas de pré requis) et l'interface de recherche très visuelle.

On va à l'essentiel sans se préoccuper d'orthodoxie bibliothéconomique. Sont privilégiés le confort et l'accès de l'utilisateur aux ressources, tout simplement.

À part dans la démarche orientée utilisateur, où sont les FRBR ?

L'*œuvre* est l'événement auquel sont associées des ressources en ligne captées d'événements. Les entités FRBR retenues sont l'*œuvre* qui peut avoir plusieurs *expressions-manifestations*. Une manifestation peut se matérialiser en plusieurs documents distincts de nature différente (*item*).

Comment s'organise la chaîne d'indexation ? Pour chaque nouvelle *expression / manifestation* est créée une notice en fonction de la nature de la ressource (par exemple, une notice « musique enregistrée » pour un concert à la Gaîté Lyrique), puis le champ titre uniforme est renseigné : y figurent l'œuvre du créateur / création des autorités = auteur(s), et l'événement associé. Enfin, sont créées les liens sémantiques, préalables à l'étape de l'exemplarisation.

Sont ajoutées des ressources associées (liens internet / liens boutique et agenda). Les notices sont connectées par des relations logiques (est une représentation de/ est représenté par / fait partie de / contient / cite / est cité / fait référence à / est référencé dans). Des liens convergent vers l'œuvre (par exemple, un document imprimé, un article fait référence à, cite, fait partie de, commente ... une œuvre).

Ainsi, en interrogeant le catalogue, on peut obtenir une notice avec les rubriques titre /auteur /date /événement /mot-clé / résumé/ durée / langues / exemplaire /note générale mais aussi accéder à d'autres ressources (Ouvrages / Catalogues / Vidéos / Articles, textes / Photographies /liens web).

Ce n'est pas une application à la lettre des FRBR mais ce catalogue réconcilie le public avec les fonctions essentielles du catalogue : chercher pour trouver.

En France, la « FRBRisation » des catalogues institutionnels (BnF, Sudoc) est en cours d'étude.

#### 5.3.2.2.2 OCLC et WORLDCAT

FictionFinder est un prototype développé dans le catalogue OCLC Worldcat. L'utilisateur a la possibilité de faire une recherche par auteur, titre, sujet, résumé, ISBN ... mais également chercher ou naviguer par genre ou personnages. Une recherche débouche non pas sur une liste de notices bibliographiques, mais sur une liste d'œuvres. Une indexation très fine permet

d'accéder aux notices bibliographiques d'œuvres de fiction, d'e book, de matériel audio. OCLC se sert d'un algorithme pour convertir les données existantes et FRBRiser l'affichage. Les résultats sont organisés par œuvre/ expression : langues /manifestation. Par exemple, en lançant une requête sur « Les trois mousquetaires », on obtient l'affichage de toutes les éditions du roman sous tous les formats disponibles et avec les *items*.

#### 5.5.2.3 Tableau de synthèse des expérimentations FRBR

Il s'agit d'un document de travail élaboré pour mieux comprendre le fonctionnement concret d'un catalogue FRBRisé.

PAYS	INSTITUTION	SOLUTION LOGICIELLE	REMARQUES	Liens vers les documents explicatifs
FRANCE	Cité de la Musique <a href="http://mediatheque.cite-musique.fr/masc/?url=/clientbooklineCIMU/toolkit/p_requests/default-concert.htm">http://mediatheque.cite-musique.fr/masc/?url=/clientbooklineCIMU/toolkit/p_requests/default-concert.htm</a>	Logiciel Aleph	Pas de niveau expression	<a href="http://www.opsys.fr/POD/index.aspx?IdPage=214">http://www.opsys.fr/POD/index.aspx?IdPage=214</a>
FRANCE	Gaîté lyrique catalogue en ligne exemple <a href="http://folio.cecilepondard.com/lagaite/piste08bis/ress_bio.html">http://folio.cecilepondard.com/lagaite/piste08bis/ress_bio.html</a>			<a href="http://www.fulbi.fr/sites/default/files/disciullo.pdf">http://www.fulbi.fr/sites/default/files/disciullo.pdf</a>
FRANCE	Base de données Electre  NB : Pour enrichir les données du catalogue Sudoc, l'ABES étudie actuellement leur mise en conformité avec le modèle FRBR.			

*Tableau 2 Synthèse sur l'implémentation des modèles conceptuels FRBR entité relation et orienté objet*

ITALIE	<p>REICAT (nouveau code de catalogage)</p> <p>OPAC SBN (National Library Service)  <a href="http://www.sbn.it/opacsbn/opac/ccu/base.jsp">http://www.sbn.it/opacsbn/opac/ccu/base.jsp</a></p>	Logiciel Aleph	Création d'un titre uniforme UNIMARC pour chaque œuvre Reicat appelle « publications » les manifestations	<p><a href="http://www.ifla.org/files/hq/papers/ifla75/135-trombone-fr.pdf">http://www.ifla.org/files/hq/papers/ifla75/135-trombone-fr.pdf</a></p> <p><a href="http://www.docstoc.com/docs/16876773/Cataloguing-with-XML-RDF-and-FRBR">http://www.docstoc.com/docs/16876773/Cataloguing-with-XML-RDF-and-FRBR</a></p>
ITALIE	<p>Prototype : IFPA ISIS  IFPA 2</p> <p><a href="http://pclub3.ts.infn.it/frbr/FRBR.htm">http://pclub3.ts.infn.it/frbr/FRBR.htm</a>  <a href="http://pclub3.ts.infn.it/frbr/wwwisis/FRB2.01/FORM.HTM">http://pclub3.ts.infn.it/frbr/wwwisis/FRB2.01/FORM.HTM</a> (démonstration)</p>	CDS/ ISIS	Basé sur une structure MARC Travaux de Sturman <i>Implementing the FRBR Conceptual Approach in the ISIS Software Environment: IFPA</i> (ISIS FRBR Prototype Application) cf <i>Cataloging &amp; Classification Quarterly</i> , 1544-4554, Volume 39, Issue 3 & 4, 2005, Pages 253 – 270.	<p><a href="http://pclub3.ts.infn.it/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/FinalDownload/DownloadId-5C17B402AF78E85A4A68795214E5BD07/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/frbr/IFPA-BRAZIL-051104a.pdf">http://pclub3.ts.infn.it/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/FinalDownload/DownloadId-5C17B402AF78E85A4A68795214E5BD07/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/frbr/IFPA-BRAZIL-051104a.pdf</a></p>
SUEDE	<p>Libris (National library of Sweden Swedish Union Catalogue) comprend environ 175 bibliothèques qui utilisent Integrated Library System (ILS) pour le catalogage = cela permet une méta recherche dans différentes bases de données et catalogues</p> <p>Base BURK</p>	MARC 21 transformation avec RDF server wrapper → les données sont chargées dans un « triple store » pour permettre la recherche (avec SPARQL)	Every record in the LIBRIS database gets assigned one or more FRBR-keys, these keys are the normalized concatenations of an author and the original title.	<a href="http://www.kb.se/dokument/Libris/artiklar/Project%20report-final.pdf">http://www.kb.se/dokument/Libris/artiklar/Project%20report-final.pdf</a>
FINLANDE	CULTURE SAMPO Portail qui donne accès à des	Modèle CIDOC CRM Schémas de métadonnées +		

	contenus culturels hétérogènes (objets culturels + ressources de références)	ontologies de domaine (ONPI)		
NORVEGE	Pode project	Notices MARC (NORMARC) converties en notices FRBR	Utilisation des travaux de Trond Aalberg § FRBRiser avec Koha (en norvégien) <a href="http://www.bibpode.no/?q=node/11">http://www.bibpode.no/?q=node/11</a>	<a href="http://www.bibpode.no/blogg/?p=1388">http://www.bibpode.no/blogg/?p=1388</a> <a href="http://www.bibpode.no/blogg/?p=1775">http://www.bibpode.no/blogg/?p=1775</a>
NORVEGE	BYBSIS <a href="http://www.bibsys.no/english/pages/index.php">http://www.bibsys.no/english/pages/index.php</a>	Bybsysmarc		
SLOVENIE	Slovenian National Bibliography (projet TNU commun avec la Norvège et la Suède)	Notices MARC		
ARGENTINE	Le prototype Litero a été développé par l'INMABB, Instituto de Matemática de Bahía Blanca (Conicet / Universidad Nacional del Sur). <a href="http://litero.inmabb-conicet.gob.ar/demo/browse/?mfid=46&amp;source=prev&amp;index=access_points_all">http://litero.inmabb-conicet.gob.ar/demo/browse/?mfid=46&amp;source=prev&amp;index=access_points_all</a> <a href="http://code.google.com/p/litero/">http://code.google.com/p/litero/</a> OPAC MARC <a href="http://catalis.uns.edu.ar/">http://catalis.uns.edu.ar/</a>	La base de données est <b>Maleta, qui succèderait à CDS/ISIS.</b> MARC 21 <a href="http://maleta.org/Doc/OverView">http://maleta.org/Doc/OverView</a> <a href="http://maleta.org/Doc/CDS">http://maleta.org/Doc/CDS</a>  Logiciel open source, <b>bases CDS ISIS, moteur WXIS</b>		<i>A s'inscrire à la liste de discussion <a href="mailto:litero@googlegroups.com">litero@googlegroups.com</a></i>  <a href="http://www.bn.gov.ar/descargas/catalogadores/ponencias/251109_06a.pdf">http://www.bn.gov.ar/descargas/catalogadores/ponencias/251109_06a.pdf</a>
ARGENTINE	CAYCIT (Centro Argentino de Información Científica y Tecnológica) CONICET <a href="http://www.caicyt.gov.ar/ccpp">http://www.caicyt.gov.ar/ccpp</a>	<b>ISIS</b> <a href="http://www.amanza.com.ar/norcurri/porque.htm">http://www.amanza.com.ar/norcurri/porque.htm</a>		<a href="http://eprints.rclis.org/bitstream/10760/11007/1/Manzanos-El%20impacto%20de%20FRBR.pdf">http://eprints.rclis.org/bitstream/10760/11007/1/Manzanos-El impacto de FRBR.pdf</a> <a href="http://www.ibiblio.org/rliit/docs/encuentr">http://www.ibiblio.org/rliit/docs/encuentr</a>

				<a href="#">os/2008/ponencias/FRBR.pdf</a> <a href="http://eprints.rclis.org/archive/00012534/02/Manzanos-FRBR_Impact.pdf">http://eprints.rclis.org/archive/00012534/02/Manzanos-FRBR_Impact.pdf</a>
ARGENTINE	Phoenix public library <a href="http://www.phoenixpubliclibrary.org/default.jsp">http://www.phoenixpubliclibrary.org/default.jsp</a>	ENDECA		<a href="http://www.nctpg.org/programs/past_programs/2007/Ross_Pub_Lib_FRBRization_NCTPG_407.ppt">http://www.nctpg.org/programs/past_programs/2007/Ross_Pub_Lib_FRBRization_NCTPG_407.ppt</a>
BRESIL	LexML Portail du Sénat fédéral (information juridique et législative) <a href="http://projeto.lexml.gov.br/documentacao/resumo-em-frances">http://projeto.lexml.gov.br/documentacao/resumo-em-frances</a>	Logiciel libre XTF (eXtensible Text Framework - avec Lucene)	FRBR <sub>oo</sub> s'inspire des REITA (Italie)	<a href="http://blog.law.cornell.edu/voxpath/2010/10/15/lexml-brazil-project/">http://blog.law.cornell.edu/voxpath/2010/10/15/lexml-brazil-project/</a>
ETATS-UNIS	Library of Congress FRBR Display Tool		FRBR Display Tool allows libraries to display their resources by clustering bibliographic records according to the FRBR model	<a href="http://www.loc.gov/marc/marc-functional-analysis/tool.html">http://www.loc.gov/marc/marc-functional-analysis/tool.html</a>
	VTLS Library Catalog Virtua	Système de gestion bibliothéconomique Virtua	Mapping de notices en MARC 21 → FRBR	<a href="http://frbr.bnportugal.pt/documentos/The_vision_of_software_vendor.ppt">http://frbr.bnportugal.pt/documentos/The_vision_of_software_vendor.ppt</a>
	Indiana University : catalogue musical Scherzo (TEST) <a href="http://webapp1.dlib.indiana.edu/scherzo/">http://webapp1.dlib.indiana.edu/scherzo/</a>		Transformation de notices MARC → FRBR métadonnées converties en XML →	<a href="http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/admin/news.shtml">http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/admin/news.shtml</a> <a href="http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/metadata/catalogingTool/">http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/metadata/catalogingTool/</a>
ORGANISATIONS INTERNATIONALES	FAO <a href="http://www.slideshare.net/faoaim">http://www.slideshare.net/faoaim</a>	CDS ISIS + EIMS-CDR (développé par la FAO pour leur workflow de		<a href="http://aims.fao.org/community/pages/functional-requirements-bibliographic-">http://aims.fao.org/community/pages/functional-requirements-bibliographic-</a>

ES	<a href="#">s/frbrfao-model-implementation-and-benefits</a>	publication) avec Oracle 9 "customization of data input and output interfaces" à l'Institut for Computer and Information Engineering de Pologne + FAO		<a href="#">records-frbr</a> Contacts : Claudia.Nicolai@fao.org; Imma.Subirats@fao.org; Stephen.Katz@fao.org
	AKOMANTOSO / UNDESA <a href="#">United Nations Department of Economic and Social Affairs (UNDESA)</a>	Inspiration FRBR + application Xml		<a href="http://schemadocs.akomantoso.org/schema-release/november-2009-and-march-2009/?searchterm=FRBR">http://schemadocs.akomantoso.org/schema-release/november-2009-and-march-2009/?searchterm=FRBR</a> <a href="http://schemadocs.akomantoso.org/schema-release/november-2009-and-march-2009/?searchterm=FRBR">http://schemadocs.akomantoso.org/schema-release/november-2009-and-march-2009/?searchterm=FRBR</a>
	VIRTUA ILS <a href="http://www.vtls.com/products/virtua">http://www.vtls.com/products/virtua</a>	Développé par l'IFLA avec la Bibliothèque du Congrès et la bibliothèque nationale du Canada  Exemple de développement ? <a href="http://ils-gateway.pnm.my/">http://ils-gateway.pnm.my/</a>  Libdb...		<a href="http://www.vtls.com/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/FinalDownload/DownloadId-9EC4D0C18F6506963D9E0A50ABBD7360/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/media/en-US/brochures/vtls_virtua.pdf">http://www.vtls.com/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/FinalDownload/DownloadId-9EC4D0C18F6506963D9E0A50ABBD7360/75B81CB3-28EF-4B4B-9F83-6643D7CE0E53/media/en-US/brochures/vtls_virtua.pdf</a>
AUSTRALIE	Austlit : australian literature ressource <a href="http://www.austlit.edu.au/">http://www.austlit.edu.au/</a>  Australian music center <a href="http://www.australianmusiccentre.com.au/">http://www.australianmusiccentre.com.au/</a>	Le logiciel AustLit est une implémentation maison basée sur le modèle FRBR et fortement influencé par les travaux d'ABC Harmony et d'INDECS.  Le logiciel a été développé en interne. Il s'agit d'une implémentation du modèle FRBR.	Il est centré sur la notion d'œuvre ; affichage de toutes les expressions et de toutes les manifestations sur une page web.  Notices en MARC XML. Insertion d'un niveau subwork	<a href="http://www.austlit.edu.au/about/technicalPlatform">http://www.austlit.edu.au/about/technicalPlatform</a>  <a href="http://alia.org.au/publishing/alj/54.1/fulltext/ayres.html">http://alia.org.au/publishing/alj/54.1/fulltext/ayres.html</a> <a href="http://www.docstoc.com/docs/30459356/Case-Studies-in-Implementing-FRBR">http://www.docstoc.com/docs/30459356/Case-Studies-in-Implementing-FRBR</a> <a href="http://www.australianmusiccentre.com.au/about/websitedevelopment#f">http://www.australianmusiccentre.com.au/about/websitedevelopment#f</a>

			correspondant à individual work » (cf FRBR <sub>oo</sub> ) Création d'une entité « honours », « parties » and « articles »	
TAIWAN	National Palace Museum in Taipei <a href="http://www.npm.gov.tw/en/home.htm">http://www.npm.gov.tw/en/home.htm</a> ...		2002: peu d'intérêt car avant FRBROO et pas de référence au CIDOC CRM « FRBR-based metadata format”	An Application Practice of the IFLA FRBR Model: A Metadata Case Study for the National Palace Museum in Taipei <a href="http://www.citi.sinica.edu.tw/papers/sophy/185-F.pdf">http://www.citi.sinica.edu.tw/papers/sophy/185-F.pdf</a>



En disposant ce panorama global des systèmes se disant FRBRisé, il nous a été possible :

- de distinguer les modèles totalement expérimentaux (prototypes) des modèles plus avancés sur la voie FRBR ;
- de comprendre que FRBRisation des interfaces ne signifie pas FRBRisation des données ;
- de dégager les expériences menées avec le logiciel libre ISIS utilisé par l'UNESCO (résultats surlignés) ;
- mais aussi, grâce à un focus par aire géographique, de mettre en valeur des cultures de l'information différentes : pays pionniers versus pays conservateurs.

Après avoir présenté la palette d'outils théoriques (formats, langages, modèles conceptuels...) permettant de porter les données sur le web, nous allons aborder la réalité, le terrain de notre expérience : un Centre de documentation positionné dans un contexte international dont le fonctionnement est bon mais qui dispose d'un logiciel devenu obsolète et d'objets documentaires complexes.

## Deuxième partie

Application des FRBR au cadre d'une  
organisation internationale :

l'expérience menée sur les bases  
internes de l'IIPE UNESCO

Dans cette partie, nous mesurerons l'applicabilité du modèle conceptuel FRBR à la réalité du terrain : les bases internes de l'IIPE.

Afin de mieux comprendre notre terrain d'expérience, nous présenterons d'abord les spécificités de l'Institut international de planification de l'éducation et des documents gérés par le Centre, l'architecture des bases de données, les outils et pratiques de catalogage. Un audit sommaire des besoins exprimés nous permettra ensuite de dégager les objectifs de l'expérience.

## 1 Présentation de l'IIPE

### 1.1 L'Institut international de planification de l'éducation

#### 1.1.1 Une organisation internationale, l'IIPE

L'IIPE (Institut international de planification de l'éducation) est un centre de formation et de recherche spécialisé dans la planification et la gestion de l'éducation. Il a été créé par l'UNESCO en 1963 dans le contexte de la décolonisation.

Les programmes de l'IIPE répondent aux objectifs de développement décidés au niveau international et, notamment, aux Objectifs du millénaire pour le développement (OMD). L'Éducation pour tous (EPT) est au centre des actions de l'Institut, guidant ses activités et son appui aux pays.

L'IIPE est formé de deux établissements : le siège à Paris et, depuis 1997, un bureau régional pour l'Amérique latine à Buenos Aires. À Paris, l'équipe de 80 personnes est internationale et interdisciplinaire (économistes, sociologues, statisticiens, démographes, experts en technologie éducative, analystes des politiques, planificateurs..).

Cette structure est dotée de statuts et d'un fonctionnement. C'est le conseil d'administration – Governing Board- qui fixe la politique globale de l'IIPE et décide de la nature de ses activités dans le cadre de la politique générale de l'Unesco. Il arrête le programme de travail et le budget de l'Institut et s'assure de la mise en œuvre des activités prévues<sup>23</sup>.

#### 1.1.2 Les missions de l'IIPE

La mission générale est la promotion de l'enseignement et de la recherche en matière de planification de l'éducation dans le cadre du développement économique et social. Pour cela, l'IIPE assure:

- la formation des planificateurs et des gestionnaires aux techniques d'analyse et de planification

---

<sup>23</sup> Pour plus de renseignements, on se reportera au site de l'IIPE :

<<http://www.iiep.unesco.org/en/abouthiep/governing-board.html>>

La formation est assurée par diverses "formules" : cours annuels, cours intensifs de courte durée, enseignement à distance.

- la recherche : Les chercheurs et les experts sont amenés à effectuer des missions qui donnent lieu à des rapports détaillés.

- la coopération internationale : d'une part par l'assistance technique (expertise avec des programmes sur mesure pour aider les gouvernements confrontés à une crise), d'autre part par la promotion du travail en réseau et l'organisation de forums sur les politiques éducatives. Dans le cadre des Objectifs du Millénaire pour le développement<sup>24</sup> et des objectifs de l'Education pour tous, le 8<sup>ème</sup> Plan à moyen terme (2008-2013) insiste sur la réduction de la pauvreté et particulièrement sur le continent africain.

### 1.1.3 Ses spécificités

Son financement : Les dotations de l'UNESCO, qui représentaient près de la moitié du financement il y a 15 ans, n'en représentent désormais plus qu'un sixième<sup>25</sup>.

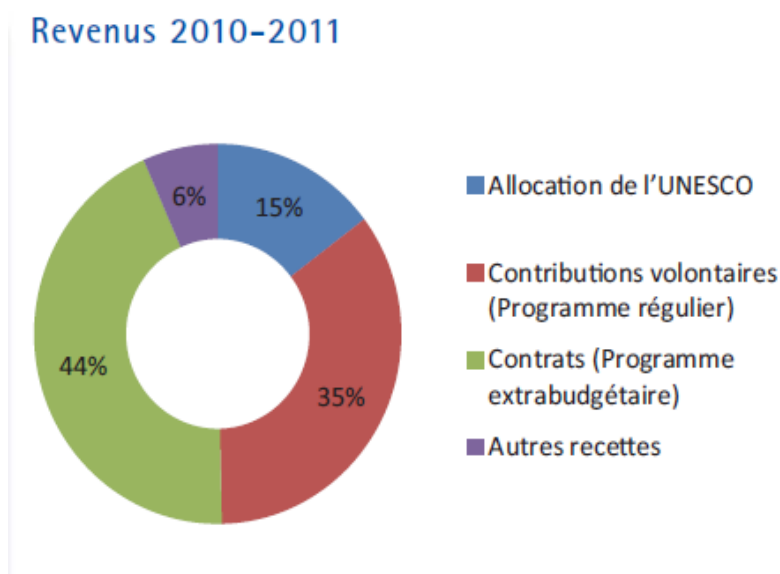


Figure 4 : Les ressources de l'IIEP en 2010-2011

A contrario, les contributions volontaires des Etats membres augmentent : on compte en particulier un petit nombre de donateurs individuels très généreux<sup>26</sup>.

La contrepartie est la nécessité de rendre compte auprès de ces financeurs de l'évolution des activités de l'Institut.

<sup>24</sup> Les Objectifs du Millénaire pour le développement :  
<<http://www.unesco.org/new/fr/education/themes/leading-the-international-agenda/education-for-all/education-and-the-mdgs/>>

<sup>25</sup> L'IIEP en action 2010-2011 :  
<[http://www.iiep.unesco.org/fileadmin/user\\_upload/About\\_IIEP/pdf/Activities\\_overview\\_2011\\_FR\\_web.pdf](http://www.iiep.unesco.org/fileadmin/user_upload/About_IIEP/pdf/Activities_overview_2011_FR_web.pdf).

<sup>26</sup> Evaluation of the UNESCO International Institute for Educational Planning, p.105 et 171.  
<<http://unesdoc.unesco.org/images/0014/001449/144954E.pdf> >

Une triple temporalité :

- un rôle prospectif : fixer les tendances de demain en matière de planification  
En effet, les chercheurs repèrent les nouveaux défis liés aux mutations économiques, sociales, politiques et leur impact sur la planification de l'éducation ;

- un rôle de producteur grâce à la diffusion, par l'unité des publications, des travaux des chercheurs de l'Institut ;

En effet, l'atelier de l'IIPE a pour fonction l'impression de la production des ressources émanant des services, à la fois documents internes et publications. Seule l'impression des couvertures des publications a été externalisée. Les chercheurs et enseignants de l'IIPE envoient leur production à l'atelier, unité chargée de la reprographie. Mais les versions papier sont de plus en plus remplacées par des versions électroniques.

- un rôle patrimonial / historique de la bibliothèque qui possède un fonds unique sur les politiques éducatives.

Pivot de l'institut, le centre de ressources reflète ce rôle d'expertise de l'IIPE dans la planification de l'éducation.

## 1.2 Le Centre de documentation de l'IIPE

Rappelons qu'un système d'information est un ensemble d'éléments en interaction et formant un tout organisé et cohérent, mis en œuvre pour permettre l'accès à l'information. [Boulogne] Ce contenu global (les informations, les fonctions et les services) est déterminé et mis en œuvre grâce à des moyens humains et techniques.

### 1.2.1 Le personnel du Centre de documentation

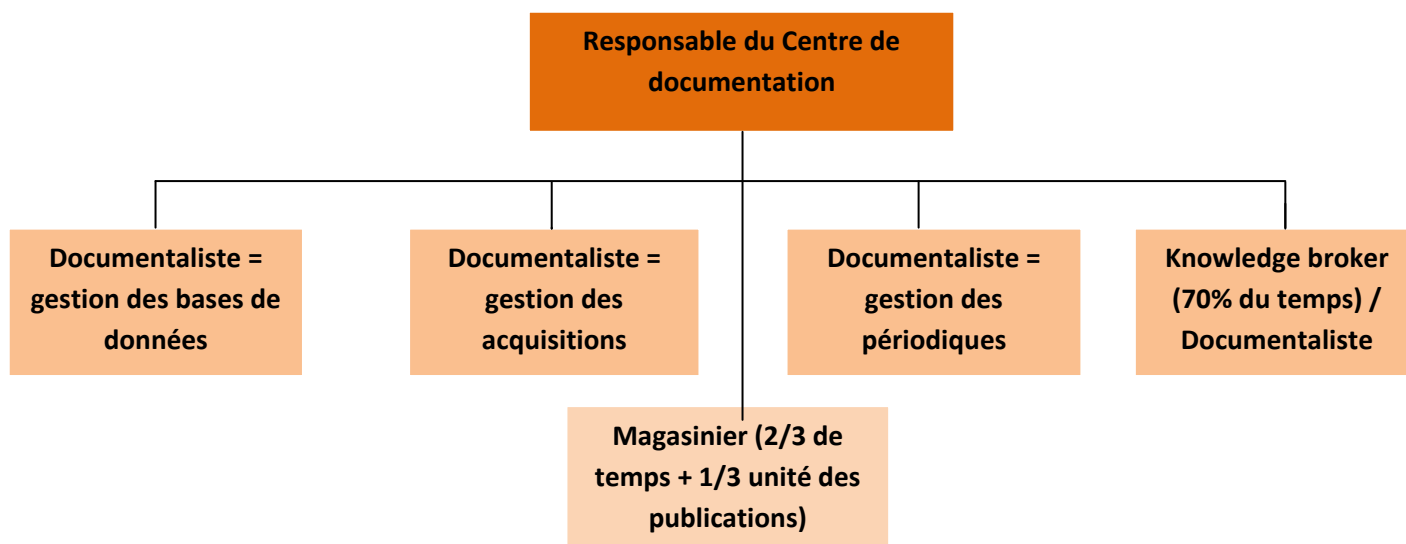


Figure 5 : L'organisation fonctionnelle du Centre de documentation de l'IIPE

Les documentalistes gèrent l'ensemble de la chaîne documentaire et sont également spécialisées dans un domaine de compétence : bases de données, acquisitions, périodiques. Elles sont très expérimentées et, bénéficiant régulièrement de formations sur les nouveaux outils, elles sont à la pointe dans leur domaine.

Le cœur de métier de la documentaliste chargée de la plateforme HIV and AIDS Education Clearinghouse (prix du meilleur travail en équipe de l'UNESCO en 2009) est atypique : elle se définit comme un « passeur », encourageant les spécialistes à échanger sur leurs pratiques, leurs connaissances, comme un « médiateur » entre le chercheur et l'utilisateur final, bref un « knowledge broker ».

La bibliothèque de l'IPE est la bibliothèque de référence dans le domaine des politiques d'éducation.

C'est un lieu mixte : à la fois espace de médiation de l'information mais aussi un espace social.

### 1.2.2 Le profil des usagers

Les utilisateurs du Centre de documentation sont :

- des experts de la planification et de la gestion de l'éducation (formateurs, chercheurs) ;
- des stagiaires en formation à l'IPE ou à l'UNESCO ;

Ils peuvent suivre différents modules (formation annuelle ; formation intensive ; formation à distance). Les stagiaires à l'année utilisent la documentation pour rédiger leur mémoire de fin d'études. Ils contribuent aussi à enrichir le fonds de la bibliothèque IPE en lui apportant la documentation sur le système éducatif de leur pays d'origine.

La bibliothèque est également un lieu de rencontre et d'échanges pour ces étudiants issus de cultures différentes. On perçoit à côté des besoins d'information classiques « l'exigence de sociabilité concrète des individus qui veulent côtoyer leurs semblables ». [3]

- des étudiants de 3<sup>ème</sup> cycle en sciences de l'éducation ou en développement.

Les lecteurs sont bilingues voire trilingues : ils peuvent consulter indifféremment une version linguistique d'un document.

## 1.3 Les fonctions documentaires

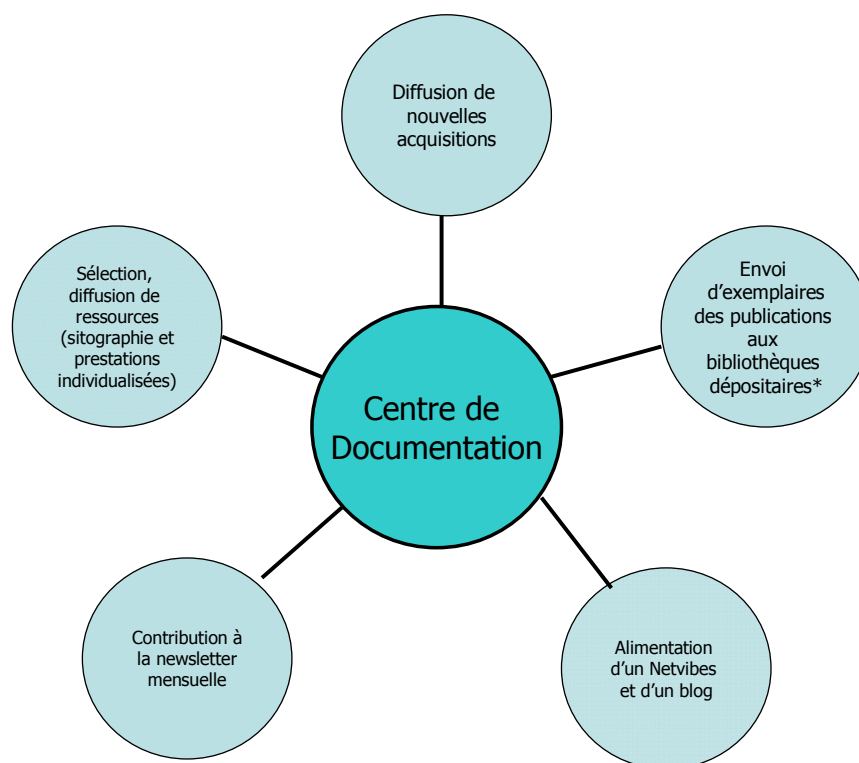
Comment fonctionne le système d'information -ensemble structuré des informations et des fonctions mise en œuvre par les activités- de l'IIPE?

Le service de documentation est d'abord un système d'information technique qui assure le traitement matériel du document protéiforme<sup>27</sup> et de données incluant les opérations de « collecte, le signalement, l'analyse, le stockage, la recherche et la diffusion de ceux-ci pour l'information des usagers »<sup>28</sup>.

Quelles sont les spécificités du Centre ?

On peut souligner :

- la diversité des services offerts aux usagers ;
- les services sont en constante évolution : en plus des traditionnels services d'information, le Centre de documentation assure sa visibilité en diffusant des prestations individualisées et en utilisant les outils 2.0 (Netvibes, blog, partage de signets...).



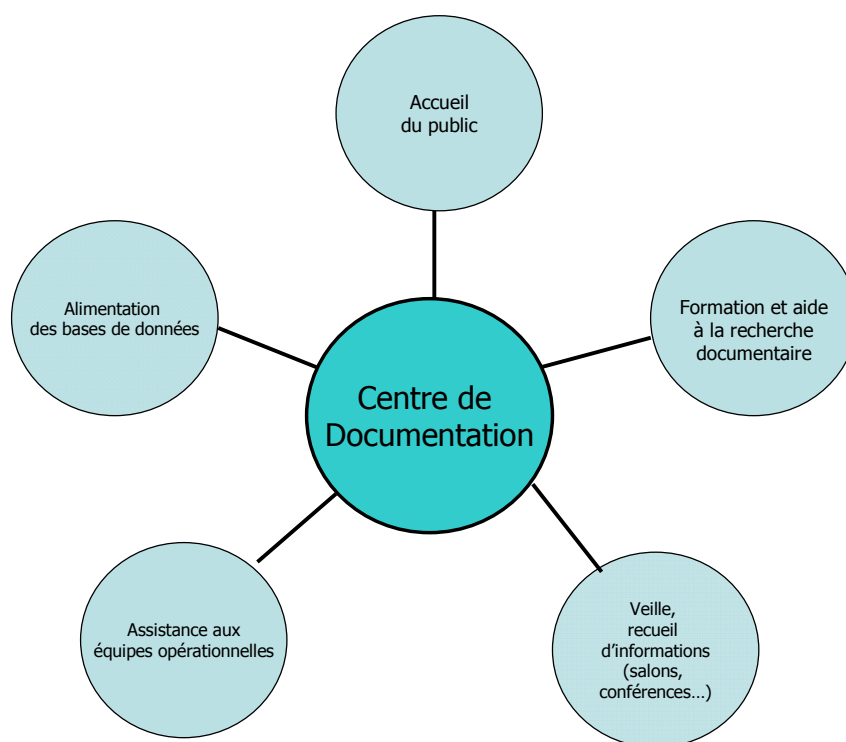
Les fonctions de push du Centre de Documentation de l'IIPE

<sup>27</sup> Tel que défini dans la partie didactique de l'étude.

<sup>28</sup> AFNOR.Vocabulaire de la documentation

\* réseau international des bibliothèques depositaires : 150 bibliothèques situées dans les ministères de l'éducation, des universités, des institutions partenaires et dans les bureaux de l'UNESCO reçoivent gratuitement des publications de l'IIEP.

- la place essentielle accordée à l'accueil et au service à l'utilisateur: les documentalistes assurent toutes une fonction de médiation et accompagnent les usagers -en particulier les stagiaires en planification et gestion de l'éducation- dans l'acquisition des savoirs et savoir faire.



Les fonctions exercées en présentiel

Quels sont les objets documentaires détenus par l'IIEP ?



## 1.4 Les documents du Centre de documentation

Le centre de ressources regroupe les documents concernant les systèmes éducatifs des différents pays, en particulier des pays en développement et toute la « littérature » disponible sur la planification de l'éducation et son contexte économique et social. La bibliothèque de l'IIPE a le même profil que la bibliothèque du Centre international d'études pédagogiques (CIEP) de Sèvres, plus axée sur la France, l'Europe et la francophonie.

Le plan de classement organise le fonds autour de grands thèmes généraux (sciences de l'éducation, contexte et planification de l'éducation, niveaux et types d'enseignement, coopération internationale...) et spécialisés (systèmes régionaux et nationaux d'éducation).

La bibliothèque comprend près de 33000 ouvrages, documents et rapports et notamment :

- une sélection de publications de l'UNESCO et des grands organismes internationaux ;
- des documents officiels et publics (plans et politiques, annuaires statistiques d'éducation...);
- des documents produits par les chercheurs (et intégrés à la base bibliographique EPIDOC) ;
- des documents à diffusion restreinte (rapports d'évaluation de programmes et de projets) ;
- 400 titres de périodiques dans les domaines de la planification de l'éducation, la politique de l'éducation, de la gestion de l'éducation et du développement.

Le fonds est multilingue : en anglais pour les deux tiers des documents, en espagnol, en français, etc.

La spécificité du fonds est l'importance de la littérature grise<sup>29</sup> : rapports, actes de congrès, séminaires...mais aussi des fragments de rapports ou de documents sans aucune mise en forme éditoriale (absence de titre, de pagination, de table des matières). [74]

Quelques caractéristiques des documents papiers conservés par l'IIPE :

➤ Sur le fond :

On peut souligner la rareté voire l'unicité de certains documents qui ont été détruits ou perdus dans leur pays d'origine.

Ces documents peuvent être :

- des originaux ;
- des doubles (exemplaire conservé par l'expéditeur) ;
- des copies, reproductions, totale ou partielle (extraits).

➤ Sur la forme : ces documents sont souvent bruts, sans mise en forme : c'est un contenu sans contenant élaboré.

Certains documents de cette base sont fragiles et doivent être préservés. C'est pourquoi une importante politique de numérisation (en particulier des plans et politiques de l'éducation) a été menée par le service.

Les versions électroniques des documents peuvent être téléchargées sur le site au format PDF. Le but de cet accès est offrir un accès universel aux travaux et réflexions des chercheurs.

---

<sup>29</sup> D'après la définition de l'AFNOR : « tout document dactylographié ou imprimé produit à l'intention d'un public restreint, diffusé hors des circuits commerciaux de l'édition et en marge des dispositifs de contrôles bibliographiques ».

La tendance est à l'augmentation du nombre de documents en liaison avec la multiplicité croissante des acteurs du monde du développement.

L'accès aux documents se fait via les bases de données.

## 1.5 Les bases de données de l'IPE

Les bases de l'IPE sont alimentées avec le logiciel de gestion documentaire CDS ISIS. C'est le logiciel de l'UNESCO pour la gestion documentaire.

Cet open source multi langue a été développé pour le système Windows sous le nom de Winisis ; son développement s'est arrêté en 2005 avec la version 1.5.

Les bases sont publiées sur le site web de l'IPE (front office) via le module d'interface web Genesis.

Les bases Planipolis et Clearinghouse sont développées en PHP/ My SQL. Le logiciel de gestion documentaire CDS ISIS gère plusieurs bases de données.

### 1.5.1 Les bases internes

On peut distinguer la base des activités et la base des documents d'activité.

#### 1.5.1.1 La base de signalement des activités du Governing Board (GB)

Qu'est-ce qu'une activité IPE ? C'est un ensemble d'actions planifiées, coordonnées, réalisées, et évaluées, par une ou plusieurs personnes appartenant à une ou plusieurs équipes, selon les différents axes ou modes opératoires définis par le plan à moyen terme. Les modes opératoires sont : la recherche, la formation, l'assistance technique, la coopération, les publications.

La base GB enregistre les activités du plan à moyen terme.

Elle comporte deux types de fiches :

- D'une part des fiches « Projet » (environ 50 fiches), au contenu de nature générale, contrôlé par le Centre de documentation, Les projets de l'IPE sont définis et planifiés par le Plan à moyen terme<sup>30</sup>, et comportent les grands objectifs stratégiques par domaines d'action : par exemple « P8-002-Ministries, Support Ministries of Education in the preparation, costing, implementation and monitoring of sector plans, linked to national poverty reduction strategies » est le 2ème projet du 8ème Plan à moyen terme.

Le contenu de ces fiches est stable pour la durée du Plan à moyen terme.

---

<sup>30</sup> Plan à moyen terme de l'IPE 2008-2013 : <<http://unesdoc.unesco.org/images/0019/001905/190586f.pdf>>

- D'autre part, dans le cadre de ces projets sont menées des activités qui donnent lieu à des fiches. Par exemple, dans le cadre du P8-002 a été menée l'activité « North Sudan Education Sector Reform: Capacity Development in Education Sector Planning ».

Ces fiches « Activité » (environ 1500) sont des records<sup>31</sup>, saisis presque exclusivement en anglais par les membres des équipes IPE dans des champs pré déterminés.

L'accès au formulaire est contrôlé.

Les champs de la fiche sont de deux sortes :

- les champs fermés (sous forme de case à cocher ou de liste d'autorité) : pays, partenaires et bailleurs de fonds par exemple

- les champs libres avec du contenu rédactionnel et informationnel : titre de l'activité, responsable, descriptif de l'activité, résultats escomptés et productions prévues, calendrier et budget prévisionnel, indicateurs de suivi et d'évaluation.

Les champs indexation sujet et thèmes du plan de classement sont remplis par le Centre de documentation.

La fiche activité est un document qui engage la responsabilité de son producteur et qui est soumis à une procédure de validation : la mention de validation est un des champs obligatoires. Cette base GB sert au suivi et à l'évaluation des activités de l'IPE : c'est un outil indispensable pour la rédaction du rapport d'activités annuel présenté au conseil d'administration. Une version publique de la base GB a été proposée jusqu'en 2007 (dernière année du 7<sup>ème</sup> PMT).

On ne peut que constater l'hétérogénéité des contributeurs et des pratiques de saisie et d'indexation dans la base d'activité : ainsi une activité peut être renseignée de façon elliptique ou très détaillée. Les documentalistes exercent a posteriori un contrôle qualité sur les données saisies, les transformant ainsi en informations.

Les activités sont codées. Ce code est intégré dans les notices INTEPI générées par l'activité. En effet, les activités IPE donnent lieu à des productions intellectuelles qui sont intégrées dans la base interne nommé INTEPI.

#### 1.5.1.2 La base des documents internes : INTEPI

La provenance des documents de la base interne varie :

➤ quant à leur nature et leur condition de diffusion :

- les documents administratifs : les rapports de mission (BTOR)

- la littérature grise : travaux de recherche, contributions à des séminaires extérieurs, contributions à des publications extérieures / pédagogie interne : matériaux de cours, supports de cours, mémoires des stagiaires, cours intensifs, cours à distance / autres : ateliers, débats...

➤ quant à leur support : des documents papier archivés au Centre de documentation/ des "documents numériques natifs" stockés sur le serveur.

Pour être saisi dans la base, l'auteur ou le responsable d'une ressource (cas du directeur de mémoire responsable des mémoires des stagiaires) doit être un personnel de l'IPE.

---

<sup>31</sup> Définition de *Records* donnée par le Livre blanc Records management : « documents créés, reçus et préservés à titre de preuve et d'information par une personne physique ou morale dans l'exercice de ses obligations légales ou la conduite de son activité ».

Les champs de la base interne sont volontairement alignés sur ceux de la base d'activité. En effet, c'est l'aboutissement d'une réflexion menée par l'équipe du centre. L'objectif à plus long terme est de mettre en commun les ressources internes en y incluant notamment une base interne de gestion des contacts (noms de personnes et d'organismes liés à l'IIPE) comportant des données pouvant être reliées à la base INTEPI.

On retrouve donc :

MFN 9753		Format: ISIDOR	
Document Code:	BTOR/2011/03/PM, P8-002-Ministries BTOR/2011/03/ES IIEP/TASP/PM/11.097 IIEP/S.309		
Team Code:	TASP		
Activity Code:	TASP-000211		
Title:	BTOR - Sudan - 2011/03 - Maalouf, Philippe; Suso, Emmanuelle		
Date:	2011/03		
Research:	P8-002-Ministries, Support Ministries of Education in the preparation, costing, implementation monitoring of sector plans, linked to national poverty reduction strategies		
Countries:	Sudan		
Authors:	Maalouf, Philippe Suso, Emmanuelle		
Institutional Authors:	UNICEF Sudan. Federal Ministry of General Education, FMOGE		
Meetings:	IIEP/S.309: Workshop on Policy for the Education sector in Sudan. Khartoum, 8-16 March.		
File location:	<a href="http://btor.iiep-unesco.org/btor/BTOR009753.pdf">http://btor.iiep-unesco.org/btor/BTOR009753.pdf</a>		
Language Code:	eng		
Document Content:	bto		
Type of record:	d		

Le BTOR contient le code Activité, et le code du Séminaire pour lequel il a été rédigé.

Cette base est accessible sur Internet par mot de passe.

## 1.5.2 Les bases en libre accès : la bibliothèque virtuelle

Les [bases du Centre de documentation](#) sont accessibles en ligne.

- EPIDOC est la base des documents bibliographiques.

Elle comprend environ 30 000 notices rédigées en trois langues (anglais, espagnol, français). L'interface de consultation du catalogue permet à l'internaute de sélectionner des documents, de faire des rebonds sur les descripteurs (navigation en cascade), d'accéder au texte intégral d'un document...

Une collection de 400 titres de périodiques (base PERISCOPE) est accessible par le catalogue en ligne EPIDOC.

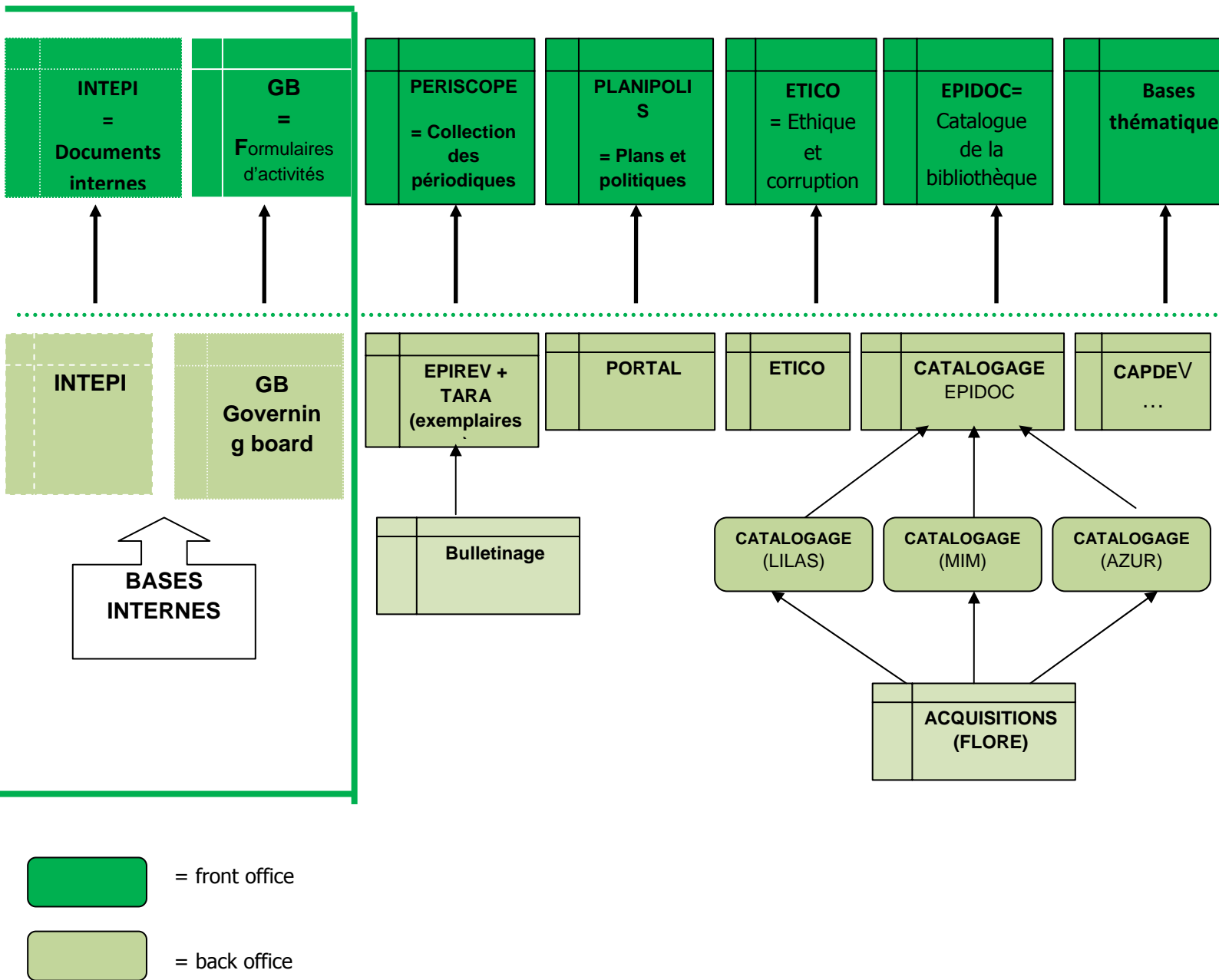
- PLANIPOLIS est la base des plans et des politiques d'éducation.

L'augmentation de la numérisation des plans et politiques a permis à la base Planipolis d'étendre son public et d'atteindre 1000 000 visiteurs par mois en 2011 (contre 15 000 en 2008).

- Des bases thématiques issues des programmes de recherche de l'Institut (par exemple ETICO : éthique et corruption dans l'éducation).

- HIV and AIDS Education Clearinghouse est le service d'échange d'informations sur le SIDA et l'éducation. Il offre différents services : accès à la bibliothèque virtuelle (plus de 7 000 références dans quatre langues), lettre d'information, alertes, forum, liens...  
Les bases de données en ligne sont mises à jour tous les mois.  
Les différentes bases de données ne sont pas reliées entre elles.

Figure 6 : Schéma simplifié de l'organisation des bases de données de l'IPE



## 1.6 Outils et pratiques de catalogage

Les outils et les pratiques mis en place s'appliquent aux différentes bases.

Rappelons que les normes sont les AACR et que la structuration des notices bibliographiques dans WINISIS a été réalisée en utilisant les principes du CCF (Format commun de communication).

Un même langage documentaire est utilisé pour le fonds de l'IPE.

Les documents sont indexés à l'aide du thésaurus multilingue de l'UNESCO. La valeur ajoutée des langages documentaires en termes de contrôle terminologique n'est bien sûr plus à démontrer ; on peut cependant souligner la révision régulière de ce thésaurus en fonction des demandes des documentalistes afin de correspondre au mieux à l'évolution des sujets traités. Par exemple, les nouvelles problématiques sur le genre nécessitent de nouveaux descripteurs (harcèlement sexuel / homosexualité /...) qui seront prochainement intégrés.

Il n'existe pas actuellement à l'IPE de notices d'autorité décrivant les collectivités et les personnes. Une base d'autorités pour les collectivités auteur a existé jusqu'en 2000, elle comportait environ 2000 notices. Mais elle n'était pas exploitée de façon rentable pour justifier le maintien de sa mise à jour (pas de liens le combinant avec les notices bibliographiques).

Les documents sont organisés avec le plan de classement du Centre de documentation.

Les pratiques de catalogage sont le reflet des caractéristiques des collections à traiter (littérature grise, fonds multilingue).

Elles sont marquées par le pragmatisme et le souci de l'utilisateur.

L'objectif est de renseigner au mieux le chercheur d'informations en lui fournissant des documents pertinents et non redondants, par exemple en évitant qu'il retombe sur le même document dans plusieurs versions linguistiques.

La notion de service est donc essentielle.

Ce catalogue orienté usager doit répondre à des défis :

- comment cataloguer la littérature grise dans un monde numérique?
- comment cataloguer mieux les documents en différentes langues ?

Le manuel de catalogage de l'UNESCO prévoit le traitement des langues de la façon suivante : un code de langue identifie sous forme d'une étiquette numérique la langue de l'œuvre traitée. Les indications d'autres versions linguistiques sont entrées dans un autre champ<sup>32</sup>.

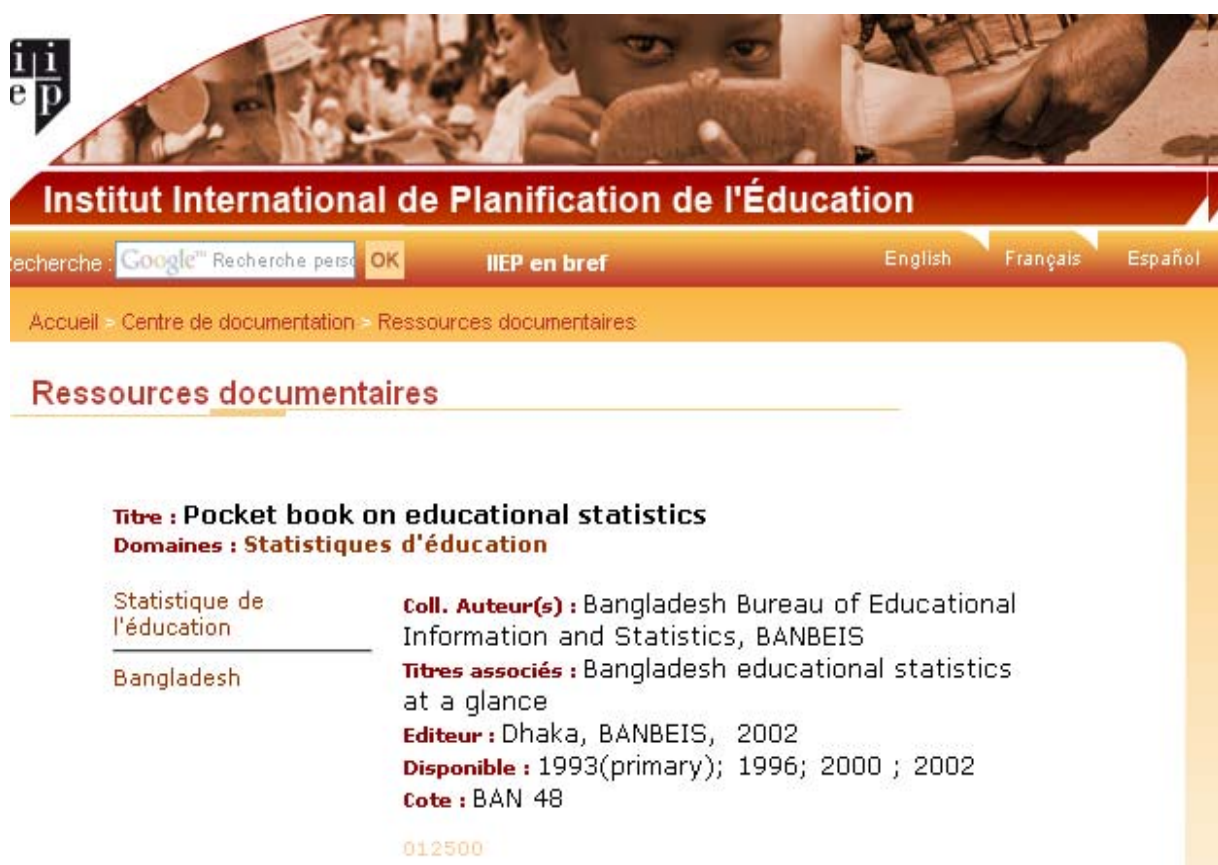
Mais, pour éviter le bruit documentaire (la présence répétée du même document dans plusieurs versions linguistiques dans la liste des résultats affichés suite à une interrogation de la base) le Centre de documentation a fait le choix de saisir dans une seule notice les différentes langues : l'anglais d'abord en tant que titre principal, suivi du français puis de l'espagnol, du portugais ensuite et finalement des autres langues en caractères latins dans un champ répétitif « titre en autre langue ». Ainsi, à l'affichage, n'apparaissent que les champs se rapportant à une même langue.

---

<sup>32</sup> Guide de catalogage et d'indexation CDS/ISIS UNESCO, p.10.

<<http://unesdoc.unesco.org/images/0009/000925/092554fo.pdf>>

Le même souci de simplification a été appliqué au niveau des manifestations : les différentes éditions d'un même document sont regroupées dans une seule notice, ainsi que les différentes années dans le cas de séries, malgré les possibles variations de titre. Par exemple, les annuaires statistiques d'éducation : le titre ancien de la collection continue d'apparaître dans la notice mais dans un champ spécifique ('Titre associé'), ainsi que les années disponibles rassemblées dans un même champ.



The screenshot shows the IIEP website interface. At the top, there is a banner with the IIEP logo and a photograph of children. Below the banner, the text 'Institut International de Planification de l'Éducation' is displayed. A search bar contains 'Google Recherche perso OK'. Navigation links include 'IIEP en bref', 'English', 'Français', and 'Español'. A breadcrumb trail reads 'Accueil > Centre de documentation > Ressources documentaires'. The main heading is 'Ressources documentaires'. The record details are as follows:

<b>Titre :</b> Pocket book on educational statistics	
<b>Domaines :</b> Statistiques d'éducation	
Statistique de l'éducation	<b>Coll. Auteur(s) :</b> Bangladesh Bureau of Educational Information and Statistics, BANBEIS
Bangladesh	<b>Titres associés :</b> Bangladesh educational statistics at a glance
	<b>Editeur :</b> Dhaka, BANBEIS, 2002
	<b>Disponible :</b> 1993(primary); 1996; 2000 ; 2002
	<b>Cote :</b> BAN 48
	012500

Notons que malgré les complications survenues lors de l'automatisation du prêt -il a bien fallu rattacher à chaque exemplaire un code-barres, le choix de départ a quand même été maintenu.

## 2 Identification des besoins

### 2.1 Les constats

Rappelons que les objets documentaires de l'IIEP émanent de producteurs divers et sont enregistrés sur différentes bases :

- fiches d'activité de la base GB ;



- documents administratifs et littérature grise de la base INTEPI ;
- publications de la base EPIDOC.

L'architecture actuelle de gestion des bases de données est complexe.

Chaque base de données est accessible à travers un formulaire de recherche. Autrement dit, pour effectuer une recherche, l'utilisateur ne peut interroger plusieurs index simultanément. Il doit utiliser un formulaire de recherche par base : des opérations multiples sont donc nécessaires pour recueillir et assembler des données disséminées.

Cela engendre une perte de temps et une impression de complexité. Des améliorations récentes ont déjà été mises en œuvre : simplification des interfaces de recherche (boîte de recherche à la Google), rubriques d'aides, rebond entre notices. Il est nécessaire d'améliorer la navigation (pas de liens entre les bases) ainsi que l'affichage des données.

Pour cela, il faut structurer autrement les données mêmes sur lesquelles s'appuient la navigation et l'affichage et simplifier le schéma informatique actuel de gestion des bases de données.

Des réflexions ont déjà été engagées par le Centre de documentation :

- sur le changement de logiciel documentaire (comment remplacer CDS ISIS ?) et sur une autre organisation des bases ;
- sur la fusion à terme des bases dans une base commune INTEPI/ GB ;
- plus en amont, sur la nécessité de réorganiser l'étape de collecte de l'information.

## 2.2 Les objectifs

L'utilisateur doit pouvoir accéder de manière groupée à des informations interconnectées jusqu'alors éparpillées dans différentes bases de données.

Deux objectifs majeurs ont conduit le Centre de documentation à mener cette réflexion :  
- un enjeu documentaire:

Comment valoriser le fonds documentaire ? Comment donner de la valeur ajoutée aux métadonnées de qualité produites par le Centre en les exploitant de façon optimale ?

En résumé, il s'agit de mieux exploiter les informations gérées par les différentes bases en recoupant les données disponibles, mettre en relation des données hétérogènes, faciliter l'accès à l'information par une meilleure identification des ressources (accès centralisé ou démultiplié ?). De plus, par la mise en commun des ressources des différentes bases (internes et externe), une exploitation optimale des documents est attendue, passant par la multiplication des points d'accès, et l'harmonisation des données de nature hétérogène provenant de ces différentes bases.

- un enjeu institutionnel : à travers les documents produits par l'IIPE, il s'agit de retracer les activités de l'IIPE, leur historique et leur évolution (par pays, par thème, par auteur, par type de document...), de permettre leur suivi et évaluation.

Il faut aussi rendre l'information accessible à tout moment au public cible (direction, administration, chercheurs, personnel enseignant et du Centre de documentation) l'information. Les garanties d'authenticité, de fiabilité, d'intégrité et de traçabilité doivent être assurées pour les documents issus de la base GB.

La gestion des flux de documents produits en interne pourra ainsi être facilement partagée entre les différents utilisateurs.

Il s'agit de :

- factoriser les informations communes à plusieurs ressources tout en évitant les doublons : cela implique de cataloguer finement pour ne pas enregistrer le même document (s'il existe déjà) et de recouper les données (notamment linguistiques) ;
- réduire en les organisant le nombre de réponses à certaines requêtes pour un résultat de recherche plus pertinent ;
- préparer et ouvrir les données catalographiques aux applications du Web sémantique.

Comment donner à l'utilisateur la possibilité d'interroger des bases de données corrélées ? Faute d'enquête menée auprès de cobaye, nous avons essayé de nous mettre à la place de l'utilisateur qui ne se préoccupe pas de connaître la construction, le fonctionnement mais qui veut des résultats pratiques à utiliser.

Partir d'une interface d'affichage réunissant les informations utiles pour l'utilisateur en les classant a permis de concrétiser l'objectif à atteindre : regrouper sur une même plateforme les ressources issues de différentes bases mais se rattachant à une même activité.

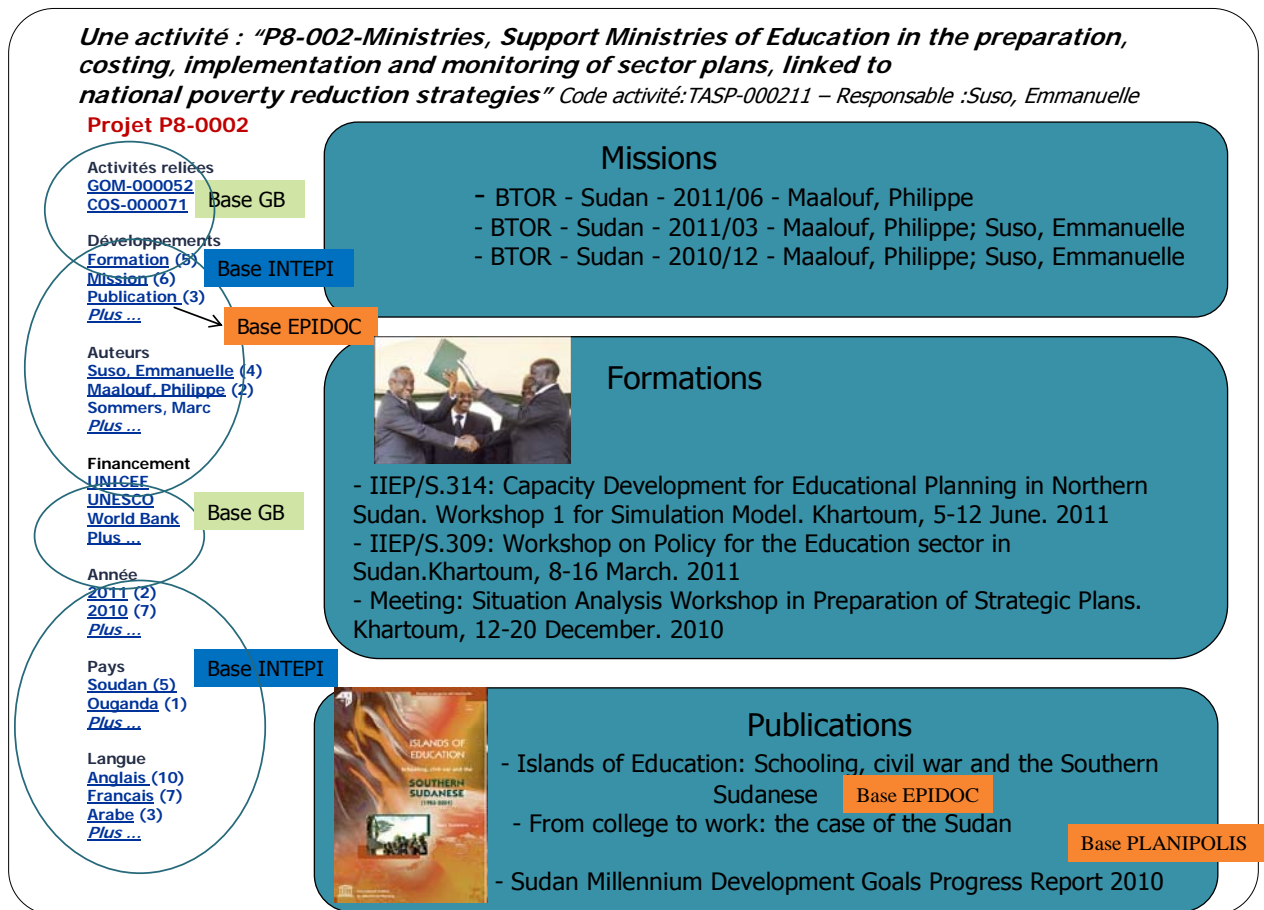


Figure 7 : Représentation de l'interface de résultat "idéale" regroupant les ressources des trois bases.

On part d'une activité du Plan à Moyen Terme (dans cet exemple P8-002) à laquelle on rattache d'autres activités, des rapports de missions menées dans le cadre de cette activité et les produits (formations et publications) générés par cette activité.

La présentation à facettes permet d'organiser les rubriques (par champs activité/ auteurs/ financeurs/ pays concernés...) puis d'affiner la recherche en sélectionnant un intitulé précis (principe de la navigation en cascade). En effet, on s'appuie sur des données bien indexées, ce qui permet d'extraire et présenter ces champs à l'utilisateur pour qu'il affine ses résultats.

Corrélativement s'est engagée une réflexion sur un moyen homogène de traiter les éditions électroniques: en effet, malgré le dépôt légal des publications électroniques, on assiste à une massification de ces ressources versatiles (versions multiples...). Mieux vaut donc gérer l'anarchie tôt afin de prévenir le chaos.

Afin d'améliorer l'accès aux contenus, il faut réorganiser le contenu : les données. C'est le but de l'expérimentation des modèles FRBR.

## 3 Expérimentation des modèles entité relation et orientés objet : le parcours

L'expérience a été menée avec le modèle conceptuel ER, de conception récente et développée par la communauté internationale des bibliothèques à destination des bibliothèques puis avec un modèle orienté objet.

### 3.1 Scénario 1 : expérimentation du modèle FRBR<sub>ER</sub>

Le modèle FRBR<sub>ER</sub> est d'abord expérimenté au système de production, de gestion et d'utilisation des données du fonds IPE.

Se pose d'emblée la détermination du point d'accès pour regrouper les informations provenant des différentes bases de données. En effet, l'objectif est de modéliser les trois bases (base des activités, base des documents internes, base bibliographique) en mettant au centre les activités de la base Governing Board.

Mais quelle unité documentaire choisir ?

Cette étape du choix de l'œuvre est fondamentale car « pour toute tentative de FRBRisation, le point de départ est l'identification des notices bibliographiques qui représentent une Œuvre et ensuite l'identification, à l'intérieur de ce groupe de notices, des Expressions et Manifestations potentielles. L'identification des Œuvres est l'étape la plus cruciale, car elle implique l'ensemble de la base de données et délimite toutes les étapes suivantes ». [58]

L'hypothèse de départ a été de considérer les projets du Plan à moyen terme et les activités comme des *œuvres*, en s'appuyant sur la définition des entités d'ensemble et des entités composantes, et de leurs relations entre tout et partie (une activité IPE représentant alors une partie du tout représenté par un projet).

Le Plan à moyen terme est alors considéré comme une « super œuvre » à partir duquel s'articulent plusieurs niveaux hiérarchiques d'œuvres, allant des projets, sous lesquels sont rattachées les activités des équipes IPE, générant des œuvres constituées par les séminaires, les cours, les rapports de mission, les publications. <sup>33</sup>

Ce point d'entrée par les activités aurait permis de tracer le travail des équipes dans différents pays et d'y rattacher les documents de la base interne ainsi que les productions bibliographiques et atteint l'objectif de relier les différentes bases.

Une autre hypothèse forgée dans le même temps a été d'aborder le Plan à moyen terme comme une collection définie dans un sens large, d'après la définition donnée par Bertrand Calenge d'« ensembles d'objets manipulables et descriptibles ») en considérant :

- chaque projet comme un volume d'une collection scindée en

---

<sup>33</sup> « C'est-à-dire qu'en toute logique l'entité *œuvre*, par exemple, peut constituer une combinaison d'œuvres individuelles réunies par un éditeur scientifique ou un compilateur sous la forme d'une anthologie, ou bien un ensemble de monographies réunies par un éditeur commercial pour constituer une collection, ou bien une réunion de documents privés articulés en un seul fonds par un dépôt d'archives. » [49]

- activités (rattachées à un projet comme une sous collection) générant des
- productions

Pour tester l'utilité finale de cette hypothèse, on a compté pour chaque niveau hiérarchique le nombre de "notices" rattachées. A l'usage, ce niveau hiérarchique s'est avéré trop lourd à gérer avec trop de fiches activités rattachées à des projets.

Le choix de l'*œuvre* s'est finalement porté sur l'activité. En effet, c'est l'échelon le plus fin et le « noyau dur » de la base GB, qui permet de montrer toutes les facettes de l'activité. L'activité concrétise le projet : le projet dans sa finitude (planifié sur la durée du PMT de cinq ans) pourrait alors être envisagé comme une collection fermée ou une monographie en plusieurs volumes.

Pour prendre en compte la diversité des « documents », l'*œuvre* a été subdivisée en différents niveaux. Finalement, l'arborescence obtenue s'est révélée très développée.

Par exemple, l'activité Équité et accès à l'enseignement postsecondaire engendre des développements à plusieurs niveaux :

- des œuvres générales : séminaires, cours intensif, mission, œuvres pour publications, œuvres pour diffusion interne ;
- ces œuvres générales pouvant se ramifier en un niveau plus fin

Cette hypothèse aborde donc les projets comme une collection, avec pour composantes les activités. Cet échelon plus fin concrétise le projet - noyau dur de la base GB - et permet de montrer toutes les facettes de l'activité.

Mais considérer une activité comme une *œuvre* FRBR serait déformer la définition de l'*œuvre* au sens FRBR (« entité abstraite »).

Le postulat de départ était donc faux, la définition de l'*œuvre*, comme entité, ne pouvant recouvrir celle d'activité, somme d'actions concrètes.

Toutes ces expérimentations ont conforté l'idée que le modèle entité relation est inapproprié pour les activités. Il a donc fallu changer de piste : le choix a été d'abandonner la modélisation de l'activité IIPÉ mais de poursuivre la modélisation du fonds des documents des bases interne et bibliographique avec les FRBR<sub>ER</sub>.

Première étape : reprendre les notices et identifier les entités.

Les « objets » décrits par les données catalographiques ont été repris et structurés autrement : La première étape est l'identification de l'*œuvre*, nœud central, qui se réalise dans l'*expression*, se concrétise dans la *manifestation* et est représentée dans l'*exemplaire*.

L'œuvre est ici un projet du Plan à Moyen terme ayant généré une activité (le policy forum) qui a à son tour produit une contribution à un séminaire.

**OE1.1 Projet P8-009-Dialogue, Policy dialogue**

**OE1.1.1 Activité COF-000099 – Policy Forum "Challenges of Financing Basic Education: Revisiting Solutions Involving the Private Sector"**

**oe1.1.1.1.1 S.299/contribution de M. Itoua "Contribution of the FAPE (Fédération africaine des associations de parents d'élèves et d'étudiants) to the Education Policy Forum".**

**e1** S.299/Itoua Le texte original en anglais

**m1** Le document S.299/Itoua en version électronique

**d1** Le document S.299/Itoua en version électronique stocké sur le serveur IIEP ([http://btor.iiep-unesco.org/doc/SEM299/SEM299\\_Itoua\\_eng.pdf](http://btor.iiep-unesco.org/doc/SEM299/SEM299_Itoua_eng.pdf))

**d2** Le document S.299/Itoua en version électronique stocké sur le serveur de l'équipe

**m2** Le document S.299/Itoua en version papier

**d1** L'exemplaire stocké au Centre de documentation

**d2** L'exemplaire stocké dans les archives des secrétaires

**e2** S.299/Itoua La traduction en français : "**Contribution de la FAPE ((Fédération africaine des associations de parents d'élèves et d'étudiants) au Forum des politiques éducatives"**

**m1** Le document S.299/Itoua en version électronique

**d1** Le document S.299/Itoua en version électronique stocké sur le serveur IIEP ([http://btor.iiep-unesco.org/doc/SEM299/SEM299\\_Itoua\\_fre.pdf](http://btor.iiep-unesco.org/doc/SEM299/SEM299_Itoua_fre.pdf))

**d2** Le document S.299/Itoua en version électronique stocké sur le serveur de l'équipe

**m2** Le document S.299/Itoua en version papier

**d1** L'exemplaire stocké au Centre de documentation

**d2** L'exemplaire stocké dans les archives des secrétaires

**oe1.1.1.1.2 S.299/Contribution de M. Patrinos**

**e1** S.299/Patrinos Le texte original en anglais

**e2** S.299/Patrinos La traduction en français

**OE1.2 P8-009-Advocacy, Advocacy activities with development aid partners**

Quelle est l'utilité de cette reconstruction?

Dans cet exemple, et à ce stade d'une notice FRBRisée a minima, l'accès unique par l'*œuvre* fournit à l'utilisateur des données présentées un ordre logique :

- une déclinaison d'*expressions* (les différentes versions linguistiques d'un document mais aussi les versions abrégées, révisées ...);
  - des documents périphériques de nature et formats différents qui ont été collectés (photos prises au cours d'un séminaire, vidéos, présentations powerpoint...) et reliés aux entités;
  - des précisions sur le lieu de stockage des exemplaires que ce soit en version papier (stockage sur les rayonnages, dans les archives, dans un autre service) ou en version électronique (stockage sur un serveur interne);
- et donc enrichit l'accès à un contenu en le présentant sous plusieurs formes, langues, support, contributeurs...

« L'arbre FRBR » -c'est-à-dire le déploiement des niveaux- peut encore se ramifier mais il faut savoir jusqu'à quel point détailler les entités et leurs relations. Cela dépend des besoins des utilisateurs du centre.

Cette première étape de détermination des entités du groupe 1 n'est qu'un préalable : seront ensuite établies les relations entre entités et inter entités qui permettront de connecter les données (cf infra § Les relations).

Puis ont été définies les entités FRBR des autres groupes.

➤ Les entités de responsabilité :

L'entité *Collectivité* est fondamentale dans notre étude. Elle correspond à un organisme, un groupe d'individus ou à des organismes agissant en tant que :

- partenaire intellectuel d'une activité (Collaborateur / Co-auteur)
- partenaire financier (bailleur de fonds)
- organisateur d'un développement généré par une activité

L'entité *Personne* peut être:

le responsable d'une équipe, d'une activité, d'un développement généré par une activité. l'auteur (ou le co-auteur) d'une production issue d'un développement généré par une activité (document de recherche, rapport de mission, contribution à un séminaire/ une publication extérieures, publication IIPÉ).

Pour les personnes, membres du personnel IIPÉ, et collectivités, il n'existe pas de notices d'autorité pour l'instant, mais simplement une liste.

➤ Les entités de sujet :

Les *concepts*, « notion abstraite ou idée », peuvent recouper les thèmes du plan de classement et les entrées du thesaurus de l'UNESCO utilisées pour l'indexation matière. Les *objets*, « réalité matérielle », ne figurent pas dans la base interne. Les *événements* sont définis dans les FRBR de façon très approximative (« fait accompli ou déjà survenu »). Ils n'ont pas été retenus comme sujets d'une œuvre.

Enfin, les *lieux* comme sujets d'une œuvre peuvent permettre de se référer au cadre géographique d'une activité ou constituer le sujet d'une activité.

### ➤ Les attributs

La deuxième étape est de constituer au niveau de chaque entité une liste des caractéristiques de l'objet documentaire décrit (« attributs » d'après la terminologie FRBR) puis de vérifier leur pertinence grâce au mapping entre les champs des différentes bases de données.

Ont été listés :

- d'une part les attributs inhérents à chaque entité (titre, mention de responsabilité, date, langue, pagination, la restriction de communication de certains documents confidentiels...);
- d'autre part les attributs forgés : les identifiants normalisés (par exemple ISBN et ISSN pour la manifestation) ou créés par le Centre de documentation (la cote, le code du document pour l'item).

### ➤ Les relations

La partie la plus ardue du processus a été de déterminer les relations qui qualifient la nature des liens existant entre les entités et retenu les plus pertinentes. Nous ne présentons ici que quelques unes des relations entre entités du groupe 1 qui soulignent la spécificité du fonds interne de l'IIFE.

Le panel des types de relation d'*œuvre* à *œuvre* du modèle FRBR est très étendu: relation de suite, complément, résumé, adaptation, transformation, imitation...

Le modèle fait une différence entre l'*œuvre autonome* (indépendante : un sommaire, une traduction libre..) et l'*œuvre référentielle*, consubstantielle à une autre œuvre<sup>34</sup> ]; par exemple, les index, le guide du professeur. Le catalogueur doit bien connaître son fonds pour apprécier le degré d'autonomie d'une œuvre et la distinction se fait parfois au cas par cas : par exemple, savoir si les matériaux de cours proposés dans une formation ne peuvent être utilisés qu'avec le recours à l'œuvre liée.

Mais nous avons surtout utilisé des relations hiérarchiques (de tout à partie) car elles nous ont permis de superposer les niveaux d'œuvre. Par exemple, les différents matériaux d'un cours (supports, exercices) entretiennent le plus souvent une relation de partie dépendante d'une *œuvre* en plusieurs parties.

Les contributions à des séminaires représentent des *œuvres* autonomes (indépendantes) avec une relation de partie indépendante d'une *œuvre* en plusieurs parties.

Pour les *expressions* de la même *œuvre* ont été identifiées des relations de traduction et de révision. Par exemple, si l'auteur d'un document (exposé, cours) présenté dans un séminaire retravaille son texte et le publie ensuite comme article dans un périodique ou une monographie, cela devient une autre *expression* avec une relation de révision.

Pour les *manifestations*, nous avons jugé utile la relation de présentation alternative entre *manifestations* de la même *expression*, dans le cas d'un article dont la publication serait prévue simultanément dans plusieurs pays et/ou dans plusieurs revues, ou de la même contribution présentée à des séminaires différents.

La relation de reproduction a été utilisée entre une *manifestation* et un *document* pour les cas fréquents de fichiers numérisés à partir d'une édition papier. Ce fichier numérisé peut d'ailleurs être découpé (relation de tout à partie) en unités intégrées dans d'autres documents et de les tracer.

Puis ont été établies les relations de « *responsabilité* » qui existent entre les entités du deuxième

---

<sup>34</sup> « Si étroitement associée à l'autre œuvre en relation qu'elle a peu de valeur hors du contexte de cette autre œuvre » (p.61 du rapport IFLA).



groupe et les entités du premier groupe (de qui émane l'*œuvre*, par qui est réalisée une *expression*, produite une *manifestation*, détenu un *document*).

Les relations de « *sujet* » (groupe 3) avec les entités des groupes 1 et 2 ont été abordées grâce aux récentes extensions et développements du modèle FRBR, le FRAD (*Functional Requirements for Authority Data*) et le FRSAD (*Functional Requirements for Subject Authority Data*). Ils donnent des pistes pour intégrer le thésaurus multilingue de l'UNESCO et le plan de classement de l'IIEP.

La modélisation en FRBR<sub>ER</sub> a permis d'engager un travail de réflexion sur les documents du système d'information : recherche de documents similaires pour établir des relations (par exemple : regroupement des activités à la thématique proche, reliés à des travaux des équipes, vérification des *expressions* (traduction, révision) des actes de colloque, de séminaire), vérification des droits d'utilisation et des conditions d'accès des documents numérisés. Ce travail de correspondance a ainsi conduit à dresser une typologie plus fine des documents de la base interne. Il a aussi permis de mettre en évidence les doublons: un document, signalé dans la base interne, par exemple un article de périodique, pouvant aussi être accessible dans la base bibliographique.

La modélisation en ER est donc satisfaisante pour les documents de la base interne et le catalogue bibliographique. Elle permet notamment de mieux gérer les œuvres en version multiple (fonds multilingue) et garantirait donc à l'utilisateur un accès plus pertinent à l'information. En effet, les collections constituées d'œuvres "s'exprimant" de manière diverse (publications dans différentes éditions par différents éditeurs et différents supports) retirent le plus de bénéfice de l'expérience FRBR. [68]

Cependant le modèle FRBR entité relation, adapté aux fonds des bibliothèques, n'est pas conçu pour modéliser le processus des activités, cœur du système d'information de l'IIEP.

Nous avons alors envisagé une approche structurant la mémoire informationnelle de l'IIEP et prenant appui sur les activités et leurs acteurs. Pour intégrer la dimension temporelle des activités IIEP, nous avons opté pour une modélisation objet.

## 3.2 Scénario 2 : une démarche objet

La démarche objet permet de sortir du modèle bibliographique entité relation et s'ouvrir à d'autres communautés.

### 3.2.1 Etape 1 : Modélisation des activités en CIDOC-CRM

Qu'est-ce que le CIDOC-CRM ? Ce modèle de référence (CRM) a été conçu par le Comité international pour la documentation (CIDOC) pour le patrimoine culturel. Le modèle, dont les travaux ont débuté en 1995, est devenu en 2006 la [norme ISO 21127](#).

Ce modèle sémantique est une ontologie de l'information relative au patrimoine culturel, c'est-à-dire une formalisation des relations qui unissent les concepts fondamentaux de ce type d'information. La présentation en est basée sur l'approche « orientée objet ». [71] [73]

L'approche par événement, centrale dans cette ontologie, nous a poussés à l'envisager pour modéliser les activités IPE, en suivant les conseils d'Anila Angjeli et Patrick LeBoeuf.

Comment procéder ?

Il nous a fallu nous familiariser avec la terminologie de CIDOC CRM: comprendre ce qu'est une classe, une propriété, un domaine, une cible.

La syntaxe s'articule ainsi : la classe source / domaine (assimilable au sujet) → la propriété (assimilable au verbe) → la classe cible (assimilable à l'objet).

Une classe peut être déclarée comme sous classe d'une autre classe : chaque instance de cette sous classe hérite des propriétés de la classe. La palette des classes et des propriétés développées par le CRM est conçue pour des objets muséographiques dans le but de les contextualiser autant que de les décrire.

Une instance exemplarise une classe : par exemple, la Contribution n°1 au séminaire sur les politiques d'éducation est une instance de la classe F14 Œuvre individuelle. Une classe déclare une propriété: une activité peut avoir lieu (P7), consister en (P9), motiver / être motivée par (P17).

Ces propriétés du CIDOC correspondent aux relations des FRBR. Il faut respecter le vocabulaire mais également les conventions de notation : l'identifiant (une lettre E pour entité et P pour propriété et un chiffre) précède la classe.

Il faut souligner le rôle central :

- d'une part du site web du CIDOC CRM dont les vidéos expliquent remarquablement comment construire les schémas de modélisation et modéliser l'information dynamique (par exemple les séminaires) [73].

Sont particulièrement centraux pour notre modélisation les tutoriels concernant l'activité, l'intervalle de temps et le lieu.

- d'autre part, des expériences déjà menées concernant l'information du secteur public qui nous ont confortées dans la faisabilité de cette modélisation. [72]

Comment aborder la modélisation de l'activité IPE avec le CIDOC CRM ? La structure du CIDOC CRM permet d'abord de rendre compte de la place centrale du plan à moyen terme, autour duquel gravitent les projets et leurs activités IPE. En effet, l'entité

temporelle *E7 Activity* -qui a une existence bornée (début/fin) - permet de hiérarchiser les différents niveaux d'œuvre, depuis le plan à moyen terme, en passant par les projets, puis les activités IIPE.

Le CIDOC permet ensuite d'introduire la notion d'événement de l'activité IIPE. Grâce au CRM, on peut modéliser les processus : workflow de production d'une activité (et sa validation), de production d'une publication, de stockage d'une ressource électronique et de documentation des processus.

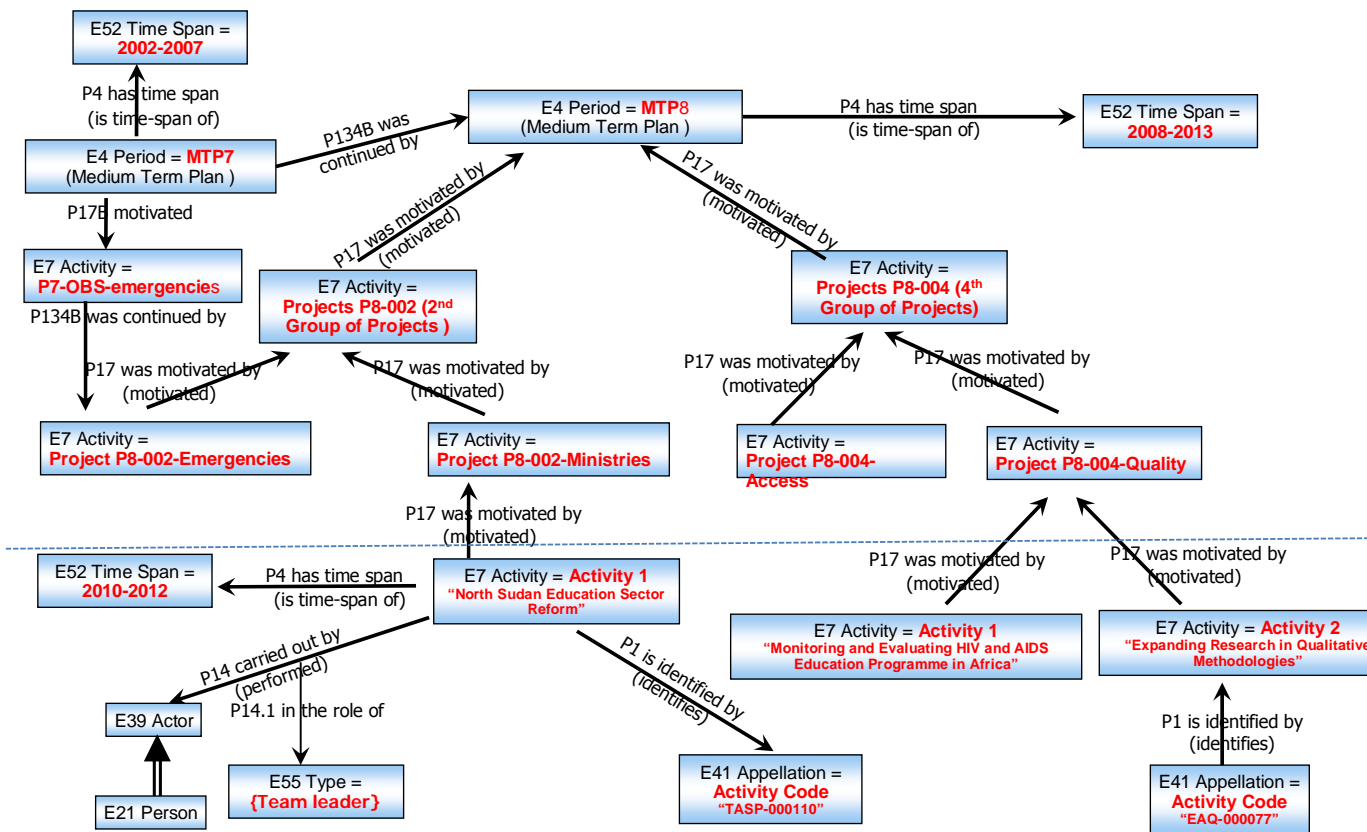


Figure 8 : Modélisation avec le CIDOC CRM

Une activité IPE est facilement contextualisable à l'aide des entités temporelles et des entités « persistantes »<sup>35</sup>, indépendantes du temps (les « choses, les personnes). Par exemple, lorsqu'un séminaire ou un colloque ont lieu, on peut préciser :

- l'intervalle de temps pendant lequel il s'est tenu (classe *E52 Localisation dans le temps*<sup>36</sup>) et la localisation dans l'espace (*E53 Place*) qu'identifie *E48 Place Name* mais aussi
- les agents impliqués dans son déroulement : *E39 Actor*, *E21 Person*, *E74 Group*,
- et même préciser leur fonction par la propriété *P14. I. in the role of* associée à *E55 Type*.

Le CIDOC CRM permet donc de représenter avec quelques propriétés l'activité dans ses aspects spatio-temporels.

Il permet aussi de retracer finement des changements concernant un groupe et ses acteurs. En utilisant les classes *E85 Joining* / *E86 Leaving* et de leurs propriétés, on indique un changement d'instance dans la classe *E21 Person* au sein d'une groupe (*E74 Groupe*), cela au cours d'une période donnée (*E52 Intervalle de temps*) : changement de membre dans l'équipe, changements de responsable d'une activité, cas fréquents dans notre institution. (Figure 3)

<sup>35</sup> Notons que les classes des entités temporelles ne peuvent être reliées à celles des entités persistantes : elles sont « disjointes ».

<sup>36</sup> Traduction de la classe *E52, Timespan*, proposée par M. LeBœuf.

Il est également possible de rendre compte d'un changement du lieu où se déroule une activité.

Schématiser les informations et travailler sur leur structuration nous a aussi permis de mettre à plat la signification de chaque activité. Nous avons procédé à un travail quasi sémantique pour qualifier ce qui est évident ou implicite, pour chaque classe et ses relations. La présentation en graphe du CIDOC CRM offre une vision globale des classes et de leurs relations.

A ce stade, nous avons souhaité mesurer les conséquences de la coexistence de deux modèles FRBRER et CIDOC CRM.

Après avis auprès d'experts de la Bibliothèque nationale de France, ont été pointés les risques d'incompatibilité de ces deux modèles, et d'incohérence dans le traitement des données pour une future implémentation.

Nous avons donc finalement opté pour la reformulation orientée objet du modèle FRBR. En effet, à la croisée des univers bibliographiques (modèle FRBR) et muséographiques (CIDOC CRM) a été créé en 2008 un hybride : le FRBR<sub>OO</sub> « qui s'appuie sur la démarche du CRM, met l'accent sur les processus ». [77]

L'objectif est de construire avec un modèle unique la passerelle permettant de relier les activités aux productions et documents qu'elle génère.

### 3.2.2 Etape 2 : expérimentation avec FRBR<sub>OO</sub>

La démarche a d'abord été de prendre connaissance des réflexions du groupe de travail international sur l'harmonisation des modèles FRBR et CIDOC CRM pour approfondir la modélisation en FRBR<sub>OO</sub>.

« La tâche principale du Groupe d'harmonisation était d'exprimer le modèle FRBR avec les concepts, les outils, les mécanismes et les conventions de notation fournis par le CIDOC CRM et de cette façon réaliser une extension au CIDOC CRM qui soit pleinement interopérable. » [77]

Quelles sont les caractéristiques de cet outil conceptuel permettant d'intégrer information bibliographique et information muséographique ?

Dans la continuité du CIDOC CRM, les notions de temps et d'évènements, centrales pour les activités -modélisables ainsi dans leur temporalité- sont au cœur des FRBR<sub>OO</sub>.

A partir du noyau de départ de la modélisation en CIDOC CRM, un « mapping » vers le modèle FRBR<sub>OO</sub> a été possible grâce au tableau de correspondance/ d'équivalence des entités, attributs et relations FRBR<sub>ER</sub> / FRBR<sub>OO</sub>.

En effet, aux attributs des entités et aux relations du FRBR<sub>ER</sub> correspondent les « propriétés » qui expriment les relations entre les classes du modèle FRBR<sub>OO</sub>.

Quelques exemples d'alignement effectués entre les deux modèles :

- l'attribut FRBR<sub>ER</sub> de manifestation « adresse d'accès » (code alphanumérique, par exemple l'adresse URL permettant l'accès à une ressource électronique) appartient à la section section 4.4.38 des FRBR<sub>ER</sub>.<sup>37</sup>

<sup>37</sup> Il est relié à la classe F3 Manifestation Product par la propriété P3 a pour note {P3.1 has type E55 Type = "Access address"} du E62 Chaîne de caractère.

- l'attribut d'expression « Résumé de contenu » appartient à la section 4.3.9 des FRBRER. Le tableau de correspondance relie cette section à F24 Expression par la propriété P148 a pour composant F2 Expression {P2 has type E55 Type = "Summary"} : la relation se fait au niveau de l'expression comme dans les FRBR<sub>ER</sub>.

Le plus des FRBR<sub>OO</sub> est une qualification précise de la nature du résumé : synopsis, notes, table des matières, ... (**Figure 10**).

- l'attribut « identifiant » se situe dans le modèle <sub>ER</sub> à la fois au niveau de la manifestation (attribut inhérent à une publication identifiée par son ISBN / ISSN) et au niveau du document (attribut forgé par le Centre de documentation : cote, code-barres).

Dans les FRBR<sub>OO</sub>, on a la possibilité d'attribuer au niveau de la Manifestation singleton un identifiant forgé par le Centre de documentation pour ces productions à caractère unique. (**Figure 10**)

La richesse de la taxonomie FRBR<sub>OO</sub> a été considérée comme un atout : les classes d'objets et les relations entre eux sont nombreuses.

En effet, par rapport aux FRBR<sub>ER</sub>, la notion d'*œuvre* est clarifiée et éclatée en de multiples sous-classes qui permettent de la préciser et de l'affiner l'*œuvre*.

L'*Œuvre* (F1) ne prend son sens qu'une fois identifiée à ses sous-classes. Déroulons la panoplie des *œuvres* déclarée par FRBR<sub>OO</sub> et utilisées pour modéliser les activités IIPÉ: « trois sous-classes de l'*œuvre* correspondent à chacune des trois acceptions possibles de la notion d'*œuvre* dans FRBR<sub>ER</sub> ». [76]

- **F14 L'Œuvre individuelle.** Elle comprend les concepts présents dans une seule Expression. Nous considérons par exemple un rapport de mission de l'IIPÉ comme une instance de F14 Œuvre individuelle : en effet, ce document administratif n'est ni adaptable ni traduisible, et qu'il se réalise dans une expression réputée complète (F22 Expression formant un tout)<sup>38</sup>. Mais plus généralement, F14 Œuvre individuelle correspond à un état des connaissances de l'*œuvre* à un instant donné.

- **F15 L'Œuvre complexe.** Elle est complexe dans sa construction : elle rassemble des concepts présents dans plusieurs expressions, par exemple des traductions, des révisions, des adaptations. Grâce à ce regroupement de versions, elle permet la collocation.

- **F16 L'Œuvre réceptacle.** Cette classe abstraite réunit plusieurs Expressions dans une nouvelle Expression qui les englobe.

Cette classe très générique est une coquille vide qui abrite :

- F17 L'Œuvre agrégative. Elle ressemble des concepts qui sont par essence la sélection et/ou l'arrangement d'Expressions issues d'une ou plusieurs œuvres : les mélanges, les anthologies, les actes de séminaires par exemple.

- F19 L'Œuvre éditoriale (« somme de l'œuvre autoriale et de la totalité du paratexte éditorial » [76] par exemple, les publications de l'IIPÉ sont une instance de cette classe. De plus, les FRBR<sub>OO</sub> rendant explicite la contribution intellectuelle de l'éditeur, les publications des actes d'un séminaire de l'IIPÉ peuvent être modélisées. (**Figures 8 et 10**)

---

<sup>38</sup> Signalons que les deux libellés identiques « is realised » pour les propriétés R3 (domaine F1 Work) et R9 (domaine F14 Individual Work) s'appliquant à la même cible F22 Self Contained Expression pourrait porter à confusion

Comment établir le lien entre les classes d'œuvre (cf. supra) du FRBR<sub>00</sub> et l'activité E7 du modèle CIDOC ? C'est par la classe F27 *Conception d'œuvre* et les propriétés P17 *Motivé par* et R16 *Initié par* que ce pont va être établi. (Figure 8)

Au cœur du modèle, se trouve le mécanisme d'héritage : toute classe est déclarée comme sous-classe d'une autre classe, dont elle hérite les propriétés. Cela permet de ne pas répéter les propriétés et offre la possibilité de faire des raccourcis. Par exemple, F22 *Expression* formant un tout est super classe de F24 *Publication Expression*, on peut relier une instance de F24 à F22 par la propriété de R14 *Incorporate* : en effet le domaine de R14 est la classe F22 et sa cible est F24. Grâce à ce mécanisme, nous n'avons pas eu besoin de répéter la classe F24 dans notre exemple pour indiquer que le texte global du séminaire (dans sa mise en forme définitive) fixé par l'éditeur intègre un ou plusieurs textes qui lui ont été fournis par un ou des auteurs. (Figure 8)

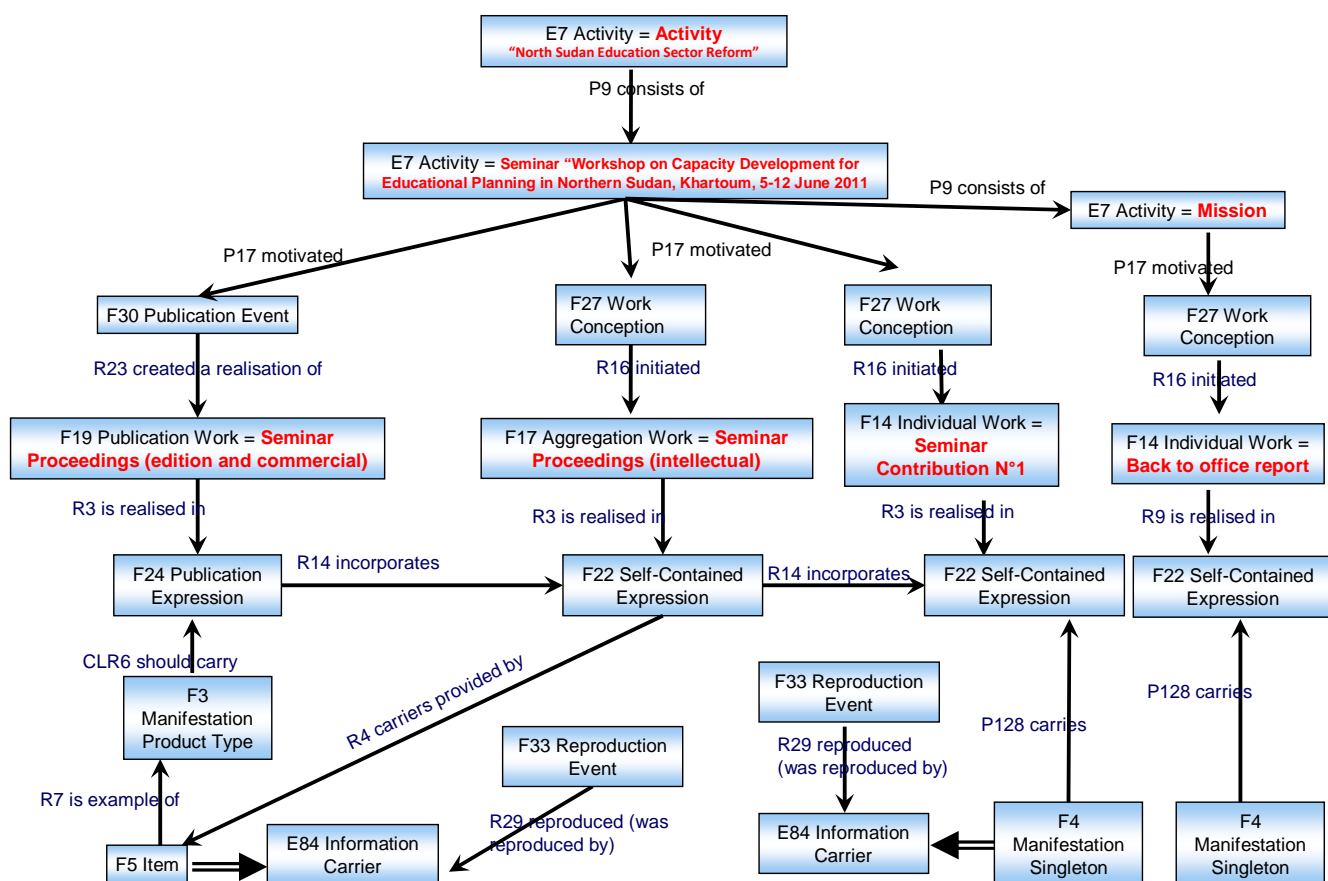


Figure 9 : De l'activité CRM aux œuvres FRBR<sub>00</sub>

Les FRBR<sub>00</sub> montrent la possibilité de modéliser les traductions et les révisions par la propriété de dérivation (R2 est dérivé de) qui associe une instance de F1 Œuvre à une autre instance de F1 Œuvre. Puis c'est au niveau de l'expression que la langue de la traduction est précisée. (Figure 9)

Au niveau de l'expression, la classe F2 est définie comme « tout ensemble de signes qui sert à

véhiculer les concepts d'une œuvre ». Ses sous classes permettent de modéliser le tout et la partie.

La sous-classe F22 Expression formant un tout est complétée par la sous-classe F23 Expression Fragment (associée à la propriété R15 à un fragment). Elle pourrait être utilisée pour une modélisation rétrospective du fonds interne de l'IIEP. En effet, le fonds IIEP comporte des versions partielles de texte (sous forme manuscrite ou dactylographiée principalement) : paragraphes ou extraits de documents d'un rapport en interne, extraits d'articles de périodique ou de monographie à paraître par exemple ; ou bien des documents dont il ne reste qu'un fragment pour cause de dégradation matérielle. Actuellement ces mentions apparaissent en note dans les notices (si cela est jugé utile). Retracer l'histoire de ces textes pourrait présenter un intérêt, par exemple en les rattachant aux œuvres auxquelles elles font référence quand cela est possible.

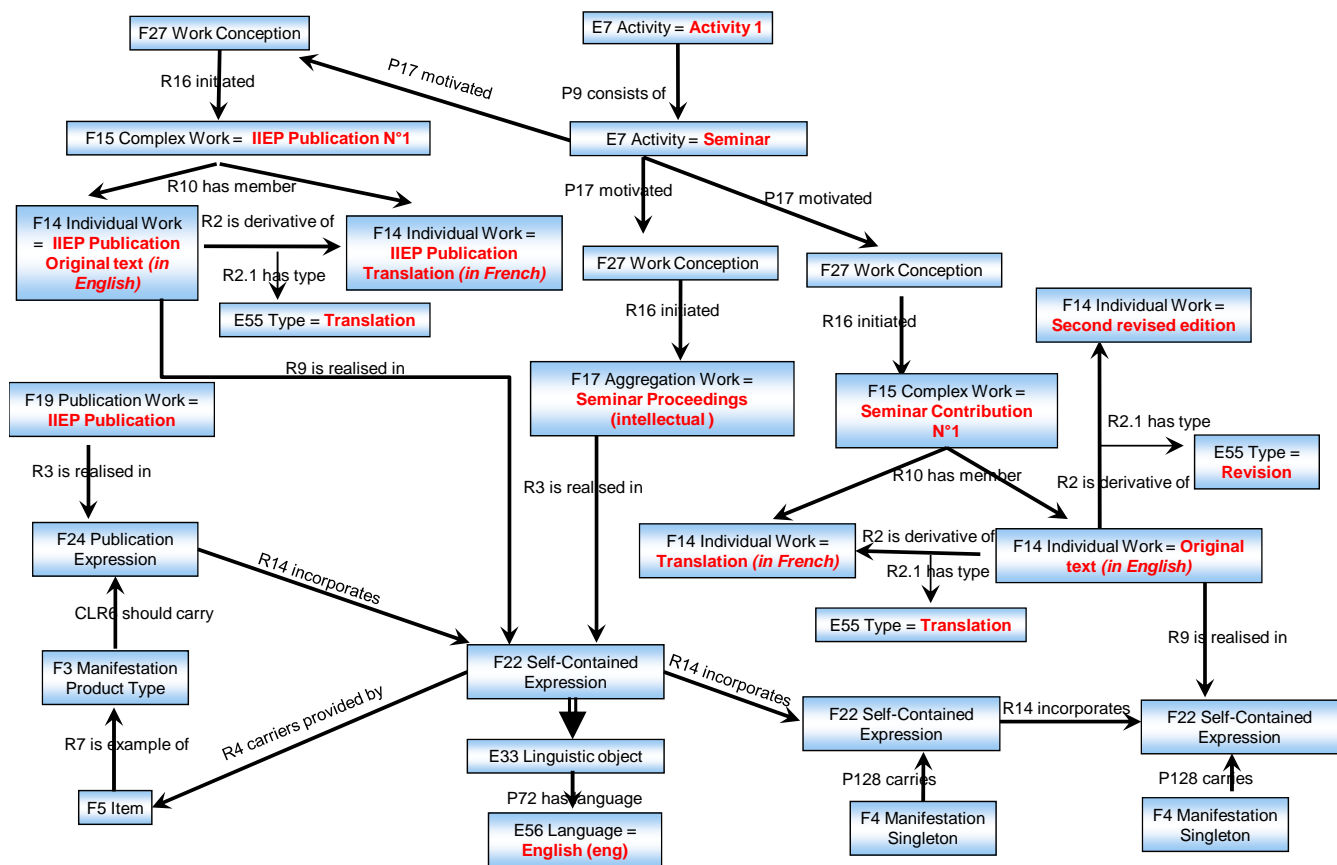


Figure 10 : Œuvre individuelle, agrégative et traductions

Autre apport des FRBR<sub>00</sub>, la manifestation est redéfinie et dédoublée par les classes :

- Manifestation Type de Produit ayant des caractéristiques communes partagées par au moins deux documents débouchant sur l'Item, document produit à au moins deux exemplaires.
- Manifestation Singleton permettant de distinguer des productions uniques (par exemple un brouillon de cours, un rapport manuscrit) des productions éditoriales. (Figure 10)

Il est également possible de modéliser la politique de numérisation du fonds de l'IIEP grâce à la classe F33 Reproduction provided Event, qui s'applique aussi bien à des documents uniques (F4



Manifestation Singleton) qu'à un exemplaire d'une production éditoriale (F5 Item), les deux étant des instances de E84 Information Carrier. (Figure 8)

Après avoir identifié les classes dont nous avons besoin pour représenter notre fonds, il nous a fallu modéliser tous les éléments susceptibles de leur être attribués.

- En pratique, toute donnée peut être traitée avec deux classes « outils » servant à :
- catégoriser (E55 Type) ; cette classe E55 Type permet aussi d'introduire des points d'accès et des vocabulaires contrôlés ;
  - identifier et coder (E41 Appellation et sous classes / E42 Identifier).

Dans le schéma développé par l'IIFE et présenté ici (Figure 11), les données d'autorité (personnes, collectivités, et sujets) ont été modélisées par l'attribution d'un identifiant, processus qui présente un grand intérêt et présenté en annexe des FRBR<sub>00</sub>. Nous n'avons pas jugé utile ici de modéliser le processus lui-même d'attribution d'un identifiant mais cela est tout à fait possible.

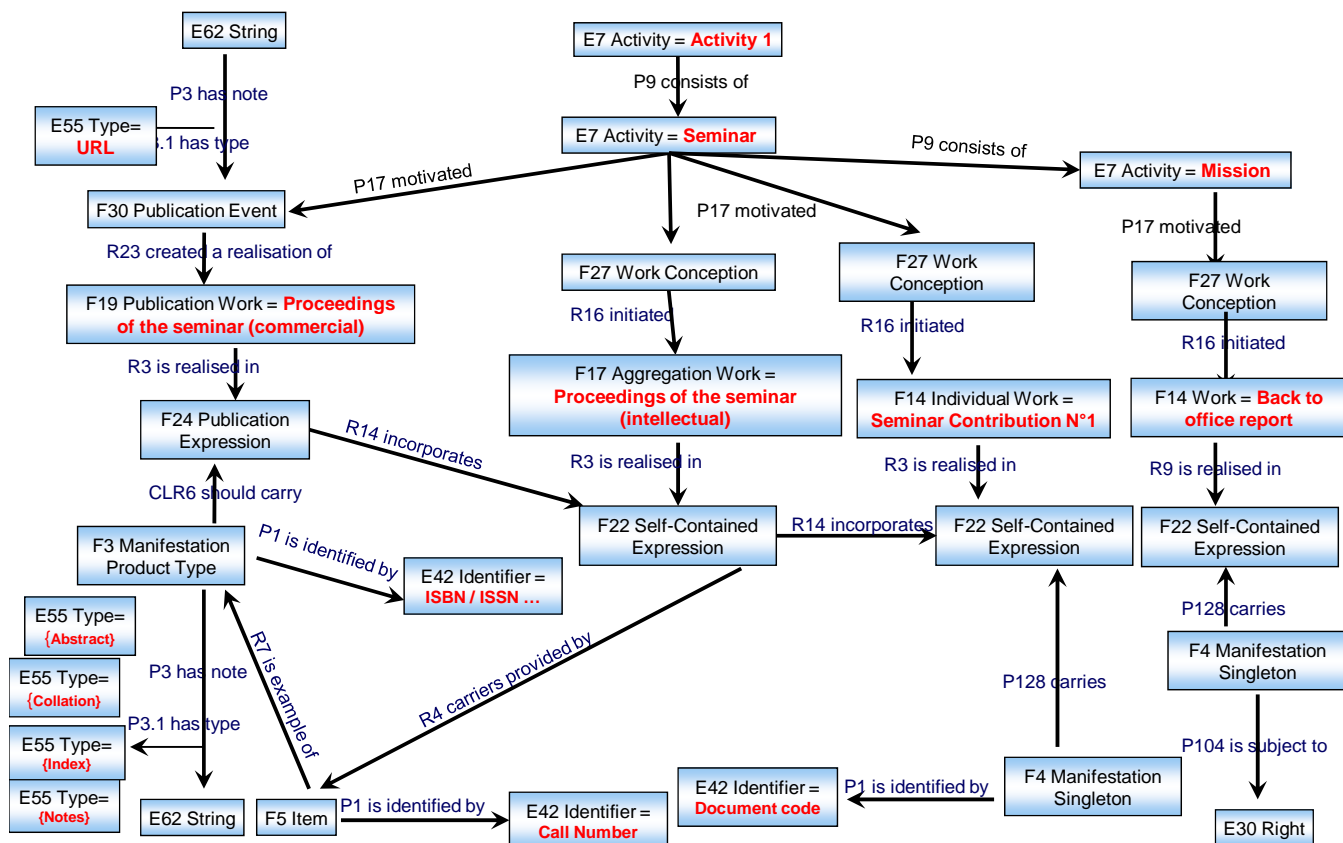


Figure 11 : Publication et identifiants

Pour modéliser la partie descriptive des documents (au niveau manifestation), la classe E62 Chaîne de caractères peut aussi être utilisée. Elle permet de « stocker » des données brutes ou des informations descriptives : par exemple des notes, des résumés, des paginations, ou des URL. On peut la lier à E55 Type par le biais de la propriété P2 a pour type (sous propriété de P3 a pour note).

L'intérêt de continuer à modéliser le fonds IPE (interne et bibliographique) avec les FRBR<sub>OO</sub> est la prise en compte tous les stades d'une activité et d'un document, donc de rendre compte du workflow de production et de validation.

Grâce à la modélisation, la passerelle peut enfin être établie entre l'univers statique des bibliothèques et l'univers du records management et de tous les environnements pour lesquels le « cycle de vie » est un événement primordial. [23]

L'intérêt est aussi de produire des données structurées et contrôlées, ces données pouvant être traduites en langage informatique (avec les triplets RDF) et mises sur le Web de données.

Dans ce domaine, de nombreuses évolutions voient le jour, avec la création de référentiels de valeurs, de jeux de métadonnées, qui offrent désormais des URI pérennes qu'il devient possible d'intégrer à des catalogues bibliographiques. Cette possibilité d'implémentation est très motivante.

Cependant, ce modèle orienté objet nécessite un grand investissement en temps. La période de déchiffrement est très longue. Il faut s'habituer à la complexité du vocabulaire utilisé pour qualifier les œuvres. Les exemples fournis dans le manuel FRBROO pour expliquer les sous classes d'une œuvre ne sont pas toujours éclairants : ainsi, il ne nous a pas été utile de savoir que « The concept for the layout created by printer Guido Morris for the text of Michael Hamburger's English translation of 12 poems by Georg Trakl for publication in 1952 [an F19 Publication Work] » était un Container Work.

En plus d'une simplification des exemples, une traduction en français du modèle <sub>OO</sub> serait appréciée et indispensable pour populariser le modèle.

Prendre conscience de l'importance fondamentale du tableau de traduction des FRBR<sub>ER</sub> en <sub>OO</sub> (p.73 et al.) de la version 1.01 des FRBR<sub>OO</sub> aurait été un gain de temps.

De même, il faut savoir que plusieurs chemins sont possibles pour modéliser : il faut en choisir un et s'y tenir.

Il a fallu tout expliciter, ne rien oublier dans les éléments décrits (ce qui gonfle le volume de chaque donnée).

Mais après avoir atteint une bonne maîtrise de son formalisme, FRBR<sub>OO</sub> permet de choisir "à la carte" ce que l'on souhaite expliciter tant le modèle présente de possibilités (modélisation de l'élaboration d'une œuvre, de la contribution intellectuelle de l'éditeur, d'un spectacle...).

Au terme de cette expérience, nous avons pu nous rendre compte que l'on ne pouvait modéliser l'ensemble d'un fonds documentaire avec un modèle conçu pour le monde des bibliothèques. En effet, on ne peut faire rentrer dans le moule FRBR<sub>ER</sub> des objets documentaires comme les activités, dont la caractéristique essentielle est la temporalité. L'orientation vers un formalisme orienté objet a alors permis de partir sur de nouvelles pistes : le modèle FRBR<sub>OO</sub> permet de penser à une mise des données sur le web sémantique mais, dans sa grande richesse, il s'est révélé très complexe et chronophage.

Dans ces conditions, nous étudierons les différentes orientations à prendre par le Centre de documentation de l'IPE ainsi que les axes de vigilance à privilégier.

## Troisième partie

### Les préconisations dans le contexte de l'IPE

# 1 Les FRBR : un moyen de structurer l'information

## 1.1 Retour sur expérience

La construction d'une ontologie (au sens d'une représentation la plus objective possible du système d'information) a demandé comme prérequis de bien définir les types de « choses », donc de faire un audit du vocabulaire pour lever les ambiguïtés, de constater les écarts de représentation puis de construire des passerelles. L'interopérabilité n'est en effet possible qu'une fois établi que les concepts implicites sont bien les mêmes.

Bâtir un modèle conceptuel a aidé à comprendre, expliquer et analyser : modélisation avec un référentiel sémantique qui permet d'établir un OUTIL, un langage commun entre des spécialistes d'un domaine et ceux qui seront chargés de le mettre en œuvre (informaticiens)

Formaliser nous a permis de découvrir des blocages, des incohérences et quelques redondances entre les documents internes [10].

Les documents d'information présentant l'architecture et le fonctionnement des différentes bases faisaient défaut. Ce projet a été l'occasion pour l'équipe du Centre de mettre en place ces documents explicatifs.

Cette expérience au long cours a mis en avant la nécessité d'expertise du catalogueur : il faut avoir une connaissance approfondie des bases pour être capable de choisir avec discernement les données à relier.

Il faut aussi avoir une connaissance globale de la "culture" de l'institution pour savoir quels sont ses besoins immédiats mais aussi à plus long terme et pour cela instaurer un dialogue entre les différents services.

Il est nécessaire de bien connaître l'ampleur du projet à entreprendre et pas sous estimer le temps nécessaire à une modélisation des documents (qui est vraiment travail à temps plein). Il faut savoir trouver le bon niveau de description :

- ne pas être trop précis au risque de se perdre dans la masse de données et de perdre du temps. Mais plus la structuration des données est précise, plus le taux de satisfaction de l'utilisateur dans la découverte des ressources sera élevé ;
- ne pas être trop général au risque d'être inexact.

Nous avons parfois versé dans l'hyper technicité pour décrypter un modèle orienté objet très complexe aux dépens de l'analyse des besoins.

Il faut insister sur un point de méthode utile pour ceux qui voudraient se lancer dans une entreprise de modélisation. Il est nécessaire de prendre d'abord la mesure des besoins réels des utilisateurs: une enquête permettant de cerner les attentes de chacun et une étude sur la façon de chercher l'information doit être lancée au préalable.

## 1.2 L'avenir des FRBR à l'IPE

Cette étape de modélisation de la réalité du fonds documentaire a été essentielle. Quel avenir désormais pour une application des FRBR?

Le modèle FRBR est-il adapté pour les besoins de l'Institut ? Est-ce une réponse possible permettant de constituer un socle commun pour l'interrogation des bases ? Peut-on l'adopter tel quel ou bien doit-on attendre ?

### 1.2.1 Expérimenter le modèle FRBR<sub>oo</sub> à l'IIFE

C'est la piste intégrale : développer le modèle <sub>oo</sub> pour les bases internes et le catalogue bibliographique. Ce choix est possible à condition de ne pas coller au modèle très complexe des FRBR orienté objet et de n'en retenir que les éléments essentiels pour l'IIFE.

Quel calendrier ? L'intégration en FRBR<sub>oo</sub> peut être planifiée pour le prochain plan à moyen terme (2013-2018).

Ce délai permettrait :

- de réfléchir à une solution logicielle en remplacement de CDS ISIS et corrélativement à l'avenir des bases internes (aller vers une fusion à terme des bases GB et INTEPI) ;
- d'engager une réflexion de fond sur l'organisation du système de connaissances à l'IIFE.

Pourquoi le modèle FRBR<sub>oo</sub> ?

Il permet un traitement homogène entre les bases donc un alignement entre les données de l'IIFE mais permet d'envisager aussi des relations avec des données issues d'autres institutions. Il est directement axé vers le web de données : l'information bibliographique est exploitable par les applications du web sémantique.

Débuterait alors une deuxième phase, la phase d'application, d'ingénierie informatique. Or, nous avons déjà vu le risque de confusion entre FRBRisation des données et des interfaces. En effet, comme le fait remarquer Sylvie Dalbin, « le passage entre un modèle conceptuel orienté métier et un modèle conceptuel orienté données informatiques reste le point d'achoppement dans cette démarche qui va d'une réalité à une application ». [23]

On ne pourra faire l'économie d'une expertise pour le déploiement du modèle orienté objet : il conviendra de s'appuyer sur la rédaction d'un cahier des charges technique. Cela permettra notamment de vérifier la faisabilité du projet et mesurer le rapport coût bénéfice, l'investissement à la fois financier mais aussi en temps nécessaire pour le mener à bien.

En effet, bien que conscients qu'un résultat durable s'appuie souvent sur un effort de longue haleine, on peut rester perplexe pour la suite du projet au vu de l'investissement en temps déjà déployé pour cette expérience.

## 1.2.2 Cataloguer le fonds avec le code de catalogage RDA

Cette voie moyenne concerne a priori la base bibliographique mais pourrait être envisagée pour les autres documents IIPE.

En effet, le nouveau code de catalogage RDA articulé autour des FRBR entité relation présente des avantages :

- il remplace les AACR et se situe donc dans la lignée du catalogage à l'IIPE ;
- il offre une totale liberté pour présenter et encoder les données : c'est une norme de contenu qui ne préconise donc aucun mode de présentation "matériel" (ISBD ou autre) ni d'encodage (peu importe le format) ;
- surtout, il va bien plus loin que le FRBR : RDA ouvre vers une « évolution de la structure des catalogues qui se rapprocheraient des bases de données orientées-objet » ;
- il est conçu pour l'environnement numérique et le web : l'objectif est de permettre aux bibliothèques d'aller sur le web de données ;
- il a l'ambition de décloisonner les institutions patrimoniales et d'être utilisé par d'autres communautés (archives et musées notamment) : il « inscrit le catalogage dans une circulation de métadonnées impliquant plusieurs acteurs : créateurs des ressources, éditeurs, etc. et permet la récupération et le partage des métadonnées entre différentes communautés d'utilisateurs. »

Mais un catalogage avec RDA soulève bien des interrogations en particulier la question du rétrospectif : comment procéder alors à une conversion du fonds ? Quelle compatibilité entre les données héritées, récupérées et les données créées façon FRBR ?

Des tests ont été menés aux Etats-Unis. Au vu des résultats, les trois bibliothèques nationales américaines ont pris la décision d'adopter RDA à partir du début 2013.

En France, des négociations sont en cours sur l'adoption de RDA : actuellement une étude sur l'adéquation de RDA aux besoins est menée dans le cadre de l'AFNOR. Le comité stratégique a estimé que les bibliothèques françaises ne pouvaient pas appliquer RDA en l'état.

L'adoption de RDA implique en effet une forte évolution des systèmes de gestion des bibliothèques.

Différents scénarios d'implémentation sont à l'étude pour gérer la transition :

- plus au moins respectueux de la structure FRBR ;
- proches du modèle de données actuel (ou même en régression par rapport à la majorité des catalogues actuels car ne prévoyant pas de liens entre notices bibliographiques et notices d'autorité.

Il faudra donc continuer de suivre les travaux de la commission de FRBRisation et attendre les décisions finales du groupe stratégique pour savoir si la France va d'abandonner l'ISBD, adopter RDA et selon quel échéancier. [79] [87]

Entre alignement sur les pays européens et voie française, "attentisme" bien coordonné et adoption de RDA en l'état, le débat est ouvert tout en gardant à l'esprit la priorité : la continuité de service public dans les bibliothèques.

Pour l'IIPE, une piste sera déjà d'étudier la faisabilité d'une conversion des notices en procédant à un mapping du format CCF au FRBR.

### 1.2.3 Vers un autre système de gestion intégrant le modèle FRBR ?

Même si l'IIPE n'est a priori pas concernée par les SIGB, une piste est de surveiller l'évolution des logiciels de bibliothèques. D'après l'enquête 2011 réalisée par Tosca Consultants, les nouveaux produits sont rares mais les entreprises s'apprêtent à développer de nouveaux systèmes de gestion: mode hébergé, support des FRBR et probablement de RDA, banalisation de la gestion des ressources numériques... [88]

Un renouvellement de l'offre au cours des trois ou quatre années à venir est prévu avec :

- une orientation vers des solutions en mode hébergé ;
- l'impulsion donnée par le code de catalogage RDA.

On pourra donc se renseigner :

- sur les fonctionnalités de logiciels comme Aleph, qui a permis de mettre en œuvre le modèle FRBR<sub>ER</sub> à la Cité de la Musique ou e Paprika, qui prétend anticiper l'évolution du catalogage en intégrant le modèle FRBR pour « améliorer la qualité des catalogues en établissant des liens entre les entités catalographiques et donc la qualité de l'interface publique<sup>39</sup> » ;
- sur les fonctionnalités de certains modules « Linked open data » mis en place par des CMS, par exemple Drupal<sup>40</sup> ;
- rester en veille sur les fonctionnalités de PMB qui a servi à intégrer un modèle entité relation à la Gaîté lyrique.

Mais avant tout, surveiller l'évolution des logiciels de traitement documentaire pour étudier comment remplacer ISIS.

Il ne faudrait pas que l'expérience de modélisation occulte les autres moyens de structurer l'information. Prioritairement au lancement d'une étude sur la partie informatique, il est essentiel d'engager une réflexion sur la gestion des connaissances.

---

<sup>39</sup> <<http://www.decalog.net/produit/e-paprika>>

<sup>40</sup> <<http://web-semantique.developpez.com/tutoriels/lpc/modules-drupal/>>



## 2 Mettre en place d'autres moyens de structurer l'information : le knowledge management

La gestion des connaissances est un « mode de capitalisation, organisation, utilisation et partage systématique des savoirs et des savoir-faire des employés, clients et partenaires d'une entreprise dans le but de l'aider à atteindre ses objectifs en terme de productivité et de qualité. »  
« Elle recouvre toutes les opérations d'identification, de collecte, de formalisation, de mémorisation et de diffusion des informations utilisées et produites au sein de l'organisme ».

### 2.1 La méthode

Dans une première phase, nous avons identifié les différents documents produits en interne et leur processus de création, validation.

## 2.1.1 Typologie des documents IPE reçus et enregistrés

*Tableau 3 Typologie des documents de l'IPE*

Code des documents	Nom des documents	Modalités de dépôt (papier ou numérique/ systématique ou aléatoire)	Parution (planifiée/ irrégulière)	Procédure d'information du Centre de documentation	Producteurs	Métadonnées/ règles de nommage	Catalogage/ Indexation	Observations (degré de priorité)
COP	Contributions à des publications extérieures	Dépôt obligatoire -publication électronique : OK -publication papier : à améliorer	Irrégulière	Par les chercheurs Par le centre (veille interne)	Chercheurs		Base bibliographique et base interne Indexé par les documentalistes	A améliorer +
Publications IPE	Fundamentals of educational planning (collection de l'IPE)	Régi par l'unité des publications qui envoie version papier + PDF	Irrégulière	Notification par mail	Chercheurs	Puis circuit UNESDOC qui charge le PDF à partir des notices IPE	Base bibliographique Indexé par les documentalistes	OK
COS	Contributions à des séminaires extérieurs	Facultatif Pas de dépôt systématique / perte d'info support de présentation (format.ppt)	Irrégulière	Par les contributeurs/ par le centre de doc (veille interne) Déduction d'un COS par les TOR/ BTOR	Chercheurs/ enseignants		Base interne Indexé par les documentalistes	A améliorer ++
ITC	Intensive training course	Dépôt obligatoire -publication électronique : OK -publication papier : à améliorer	Irrégulière	Réception par centre de doc d'une fiche synthétique envoyée par	Chercheurs/ enseignants	Fiche attribuant un numéro d'ordre à l'évènement	Base interne Indexé par les documentalistes	A améliorer +++

				l'unité des publications Déduction d'un ITC par les TOR				
SEM	Séminaire	Dépôt obligatoire -publication électronique : OK -publication papier : à améliorer	Irrégulière	Réception par Centre de documentation d'une fiche synthétique envoyée par l'unité des publications Déduction d'un SEM par les TOR	Chercheurs/enseignants	Fiche attribuée un n° d'ordre à l'évènement	Base interne Indexé par les documentalistes	A améliorer +++
BTOR	Rapports de mission	Dépôt obligatoire : documents validés et visés par le directeur de l'IPE	Dépend des fréquences des missions	Envoi par email par le secrétariat de la direction (IPE/ nom du programme/ initiales du missionné/ année/ N° chrono)	Missionnés	1- Code chronologique attribué par les secrétaires 2- Nommage attribué par le centre de doc	Catalogué par les documentalistes mais pas indexation matière	OK
Docs IPE sur le site web	Matériels de formation /EQA							
Autres documents : prg docs confidentiels	Documents de programme						Base INTEPI	

## 2.1.2 Analyse sommaire du fonctionnement du circuit de l'information

Dans une deuxième phase, nous avons procédé à une analyse très sommaire du fonctionnement du circuit de l'information.

Le principal changement intervenu récemment concerne la dématérialisation progressive de la chaîne physique de production des documents.

A l'ère de l'imprimé roi, l'objet d'information traditionnel était conçu selon une chaîne de production et de diffusion spécifique. En effet, la production des chercheurs et enseignants était systématiquement envoyée à l'atelier de publication de l'IPE pour impression de leurs cours et travaux. Chaque document était ainsi reproduit et une copie automatiquement adressée au Centre de documentation. Puis le circuit de production s'est progressivement modifié : les versions électroniques ont peu à peu remplacé les versions papier.

Cela a entraîné une perte progressive de la chaîne de production des documents internes, depuis leur sortie de l'atelier de reprographie jusqu'au Centre de documentation.

La conséquence pour le fonds documentaire est que beaucoup de documents ne sont plus récupérés automatiquement.

Cette analyse a permis d'identifier un dysfonctionnement dans le processus documentaire : l'étape de la collecte des documents (fonction entrée) est à améliorer.

Comment faire pour que ces documents numériques produits par les équipes IPE soient identifiés, récupérés et indexés par le Centre de documentation ?

Des initiatives ont déjà été prises par le Centre pour collecter les documents, par exemple en préparant des dossiers sur le serveur commun afin que chaque équipe y verse ses contributions.

Il faudra donc réfléchir à mieux gérer les connaissances tout en conciliant des approches métiers différentes.

La troisième phase consiste à proposer des solutions pour capitaliser les savoirs, pour préserver, partager et valoriser l'information produite en interne.

## 2.2 Des solutions pour capitaliser les savoirs

Pour favoriser la gestion des connaissances, peuvent être mises en place des solutions intégrant des outils techniques, des ressources humaines et s'appuyant sur des normes.

Plusieurs axes sont possibles pour renforcer les outils et pratiques collaboratives. Mettre en place des dispositifs d'information en gardant à l'esprit la nécessité d'articuler plusieurs points de vue : organisationnel (stratégique), technique, orienté utilisateur mais aussi producteur de l'information.

1- La plateforme de gestion et de collaboration des documents Sharepoint, actuellement sous utilisée par l'Institut, peut être un moyen d'améliorer la collecte des documents.

Le Sharepoint permet de visualiser les étapes du cycle de vie du document, en particulier ses modifications, et de travailler à plusieurs sur un même document. Ces fonctions de workflow et de travail collaboratif pourraient rendre service aux collaborateurs du Centre de documentation qui sont la plupart des producteurs d'information.

Pour cela, il faudrait -sous l'égide du Centre de documentation- organiser des séances de formation en partenariat avec le service informatique : cela permettrait aux utilisateurs de s'approprier l'outil et de prendre l'habitude de verser les documents créés dans la bibliothèque de documents. Un accompagnement au changement et une communication sur les bénéfices engendrés par ces nouvelles pratiques doivent être envisagés.

2- Prospecter "gratuitement" dans les outils du marché proposant une nouvelle organisation des connaissances et tester leurs réelles fonctionnalités.

Développé par la société Cogniva, ISIS (Integrated Semantic Information Services) est une infrastructure informationnelle basée sur un modèle à facettes servant à organiser les documents : facette de contenu (description du contenu du document : lieu, temps, thème, personne, ...), facette de contexte (description du lien entre le contenu d'un document et les fonctions et activités de gestion et d'exploitation), facette d'organisation du travail.

Cela permet de disposer d'un outil commun à la fois pour les besoins d'uniformité de la gestion documentaire et souple pour organiser différents types d'information et de connaissance. L'intérêt est que l'outil ISIS automatise certaines tâches de gestion de l'information, dont le processus de création de métadonnées.

3- S'appuyer sur les ressources humaines en interne en renforçant la collaboration avec l'unité des publications grâce, par exemple, à la mise en place de métadonnées communes des documents produits.

Mais un défi doit être relevé. Il faudrait que les producteurs de contenus inscrivent leurs métadonnées en respectant les standards disponibles.

Pour cela, l'IPE peut prospecter pour :

- trouver des fonctionnalités d'annotation de documents numériques afin de favoriser le partage de connaissances
- trouver quel schéma de métadonnées est le plus adapté pour les besoins des équipes (LOM, Dublin Core...)
- s'intéresser au consortium Open Archives : les données visées sont notamment toute la littérature grise scientifique, les thèses, mais aussi les pré-prints, ces articles qui seront publiés des mois après dans les grandes revues scientifiques mais dont l'échange est nécessaire immédiatement à la poursuite des travaux et au développement de la connaissance.

4- Faire expliciter les connaissances tacites, immatérielles et les savoir faire internes au Centre de documentation, notamment en termes de gestion des documents, de nommage des dossiers et de processus de saisie. Cela passe par exemple par la rédaction de notes de procédures, de guides de bonnes pratiques. Ces outils de travail pourraient être centralisés dans un dossier commun et servirait de référence pour les autres équipes de l'Institut pour harmoniser les pratiques.

5- Comment améliorer les flux internes de production et de circulation des documents ? Pour aider à formaliser et à mettre en œuvre la politique de gestion des documents d'activité, on pourra s'appuyer sur la récente série des normes ISO 30300 concernant le management des documents d'activité. [80]

Il s'agit de mettre en place et faire respecter une démarche de records management<sup>41</sup>, de gestion des documents conformément aux règles et procédures établies, et en particulier élaborer un plan de classement des activités afin de représenter l'IPE par ses activités. Ce document de référence serait un support préalable à la mise en place d'un référentiel de classement de d'archivage des documents.

Cela permettrait :

- d'indiquer les catégories documentaires inhérentes à chaque activité ;
- de fournir au personnel un cadre pour la création de documents répondant aux procédures de l'IPE ;
- d'assurer que les processus de travail produisent des documents constituant bien le support des activités, et sont bien porteurs de ce que l'on attend d'eux.

Ces fonctions de gestion et d'administration des données relèvent du Centre de documentation : cela revient à renforcer et élargir la fonction contrôle qualité du Centre de documentation.

## 3 Attendre ... tout en se préparant au web de données

Il convient d'assurer un traitement homogène des données du fonds bibliographique, des documents internes et des activités.

Pour cela, il faut veiller en priorité à :

- revoir la construction du titre de l'activité : réfléchir sur les éléments indispensables du titre (à relier avec l'objectif, le nom du lieu...);
- mener un travail sur les autorités et créer des notices d'autorités pour les noms de personne et les organismes.

En effet, rappelons que les notices d'autorité établies pour les titres correspondent majoritairement à l'entité *Œuvre* du modèle FRBR.

### 3.1 Mettre en place des identifiants

Les initiatives en cours sur les identifiants sont autant d'outils qui peuvent être étudiés puis utilisés:

#### 3.1.1 Les initiatives à l'échelle internationale

- L'initiative VIAF et ISNI

Concernant les fichiers d'autorité personne et collectivité, nous avons brièvement abordé le projet VIAF (Virtual International Authority File). Son objectif est de faciliter le partage des données d'autorité, de

---

<sup>41</sup> Le records management a pour objet « l'ensemble des documents « à archiver », c'est-à-dire les documents que l'organisme aura décidé de préserver à titre de preuve ou en raison de leur valeur informationnelle : il s'agit des documents essentiels ou utiles dans la conduite de ses activités ».

réduire les coûts de catalogage, de simplifier le contrôle d'autorité (création et maintenance) dans un environnement international.

VIAF apparie les notices d'autorité de plusieurs grandes bibliothèques produites avec des normes, des langues et des écritures différentes, et les combine en un service d'autorité unique. L'objectif à long terme du projet VIAF est de combiner les formes autorisées de noms en provenance d'un grand nombre de bibliothèques nationales ainsi que d'autres sources majeures d'information, pour les fondre en un service partagé d'autorités au niveau mondial.<sup>42</sup>

VIAF fournit des données d'autorités dans des langues, écritures et formes appropriées aux besoins des utilisateurs divers. Les notices VIAF sont dotées d'une URI et rendues en RDF pour faciliter leur réutilisation. L'idée est de reprendre les URI des agents identifiés dans leur base de données. A titre d'exemple, l'identifiant VIAF pour l'IPE est VIAF ID:154839678.

S'appuyant sur le dispositif VIAF, le Code international normalisé des noms (International Standard Name Identifier, ISNI – ISO 27729), en cours d'élaboration, vise à identifier les acteurs impliqués dans la chaîne de création, production, gestion, édition et distribution des contenus, et à harmoniser leurs appellations. La conception de l'identifiant ISNI reste dans une logique très proche de la gestion des fichiers d'autorité, et les données des bibliothèques seront d'ailleurs au centre de cette gestion des identifiants.

- Dans l'optique d'une collaboration renforcée avec l'unité des publications, il est possible de mieux rendre visible les publications des chercheurs grâce à l'outil JRC Names et au projet ORCID.

- « JRC-Names » est un outil d'identification multilingue des noms de personnes et d'organismes récemment élaboré par le « Joint Research Centre » de la Commission européenne. Cet outil est librement téléchargeable ce qui s'inscrit dans la philosophie préconisée par l'UNESCO.

On peut ainsi disposer de notices d'autorité sujet enrichies, avec le nom des chercheurs, leur profil, leurs publications.

- le projet d'identification unique des chercheurs : cette initiative est menée par [ORCID](#) (Open Researcher & Contributor ID) à destination des chercheurs. L'objectif, soutenu par plus de 140 institutions (sociétés savantes, éditeurs, organismes de recherche, universités), est de créer un réseau mondial des chercheurs.

Cela permettrait d'intégrer au projet du Centre de documentation les enseignants chercheurs en leur offrant la possibilité d'augmenter leur réputation professionnelle.

- On pourra aussi s'appuyer sur MADS pour créer des données d'autorité en XML sans passer par l'étape MARC. Ce schéma de métadonnées MADS (Metadata Authority Description Schema) est opérationnel pour les agents (personnes, collectivités) mais aussi les entités nommées (sujet, noms géographiques...).

---

<sup>42</sup> <[http://archive.ifla.org/IV/ifla72/papers/123-Bennett\\_trans-fr.pdf](http://archive.ifla.org/IV/ifla72/papers/123-Bennett_trans-fr.pdf)>

### 3.1.2 A l'échelle de la France

On peut se servir des autorités de la base SUDOC (Système universitaire de documentation). L'ABES (Agence bibliographique de l'enseignement supérieur) propose depuis 2010 en RDF/XML plus de deux millions d'entités décrites par ses notices d'autorité pour « promouvoir l'utilisation des autorités Sudoc au-delà du Sudoc ». L'IPE aurait tout intérêt à intégrer ces données et métadonnées produites par l'application IdRef qui permet d'interroger et de consulter les autorités du SUDOC et de récupérer :

- soit la notice d'autorité entière ;
- soit la forme normalisée du nom (personne, collectivité, géographique, famille...) avec son identifiant (pérenne bien sûr).

Chaque entité accessible par IdRef est associée à un identifiant, constitué du numéro de la notice. Par exemple, l'IPE est associée au lien permanent 026439034.

L'intérêt de ces projets est qu'ils libèrent les fichiers d'autorité des catalogues : les fichiers d'autorité deviennent alors des répertoires ouverts, échangeables, exploitables.

## 3.2 Suivre les initiatives open data des institutions culturelles

Cette piste n'est certes pas prioritaire pour l'IPE mais les évolutions des initiatives open data auront à terme des retombées sur le positionnement de l'Institut.

A l'échelle européenne, a été lancé fin 2008, le projet [Libris](#) : le catalogue collectif des bibliothèques suédoises a publié en 2008 l'intégralité de ses données bibliographiques dans le Web de données.

L'internaute a accès sur la même plateforme à des métadonnées descriptives (notices « traditionnelles ») mais aussi à des données enrichies sur les auteurs (en lien avec Dbpedia) et à une interface FRBRisée (organisation des résultats par éditions, langues...).

En France, dans le domaine des bibliothèques, les premières expérimentations concrètes portant les données dans le web ont été menées à l'initiative de l'ABES<sup>43</sup>.

A l'initiative de la Bibliothèque nationale de France, un projet d'envergure vient d'être lancé. Il peut être replacé dans le contexte d'ouverture des données publiques françaises.

---

<sup>43</sup> Inclusion de données en RDFa dans le catalogue [Calames](#)<sup>43</sup>, le projet STAR, le portail des thèses.



### 3.2.1 Le projet data.bnf : diffuser les données de la BnF sur le Web

Le projet data.bnf articule des contenus - ceux de la BnF mais aussi extérieurs à la BnF -, des services et des liens autour d'une « unité documentaire ».

L'objectif est de fédérer des données et les relier pour accéder à des connaissances. Le corpus initial porte sur des œuvres et des auteurs majeurs pour lesquels il existe de nombreuses notices dans les différentes bases de la BnF (autorités déjà créées). Le projet s'appuie sur les notices d'autorité du catalogue général qui permettent d'identifier les données et de les relier de manière précise et stable. « C'est donc le mécanisme utilisé pour la production des catalogues qui permet de produire ces pages Web, en particulier les liens aux autorités»<sup>44</sup>.

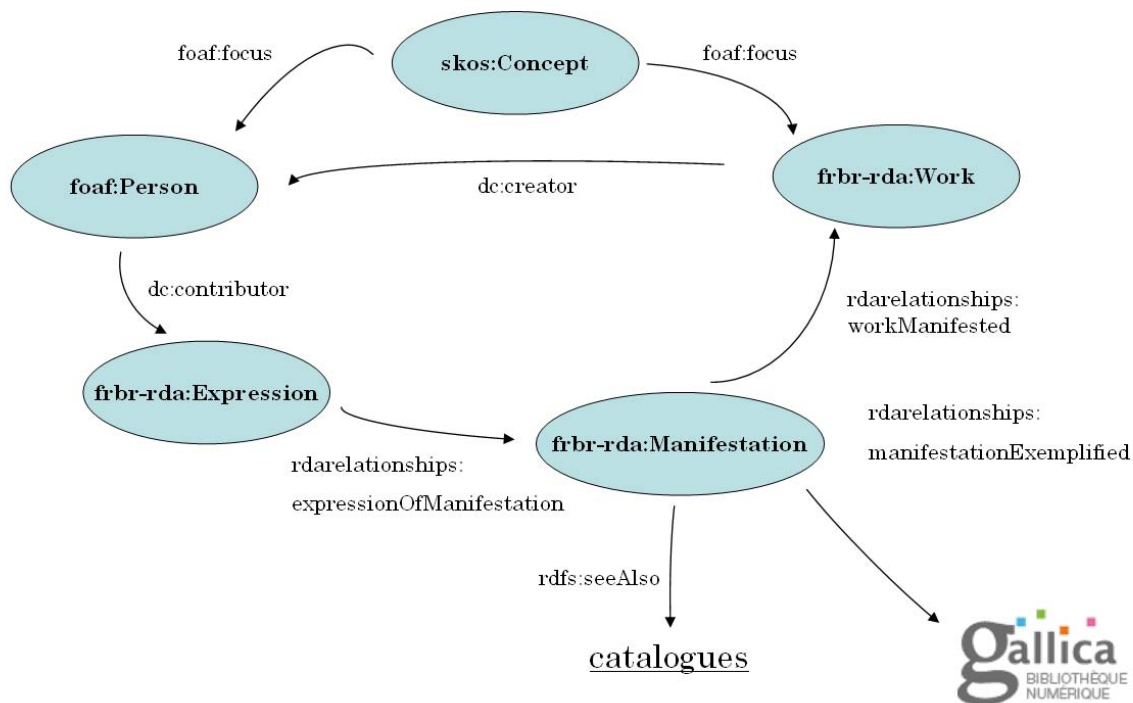
Le principe est l'extraction de données (les index d'ouvrages de référence de la littérature classique française : Lagarde et Michard) pour les transformer en pages web. Aux fiches de référence sur les auteurs et les œuvres sont reliées des données de réservoirs différents.

Comment a-t-il été possible de relier des données de réservoirs différents (catalogue général, archives et manuscrits, Gallica...)?

Les données décrivant des entités (œuvres, auteurs) ont été reliées suivant le modèle FRBR. Les fiches de références sur les auteurs et les œuvres ont été créées automatiquement en regroupant des données en différents formats normalisés (MARC pour les données bibliographiques et d'autorité, EAD pour les inventaires d'archives et les manuscrits, Dublin Core pour les documents numériques).

---

<sup>44</sup> <<http://multimedia.bnf.fr/lettres/produits/produits25.htm>>



[Source : BnF<sup>45</sup>]

Le produit a été implémenté avec le logiciel libre Logilab<sup>46</sup> : et développé avec la plateforme d'application du web sémantique Cubicweb.

Quels sont les résultats ?

L'utilisateur:

- dispose d'une interface unique pour d'accès aux ressources de la BnF (ressources issues de différents catalogues mais aussi des liens et des services) : il peut naviguer entre les œuvres d'un auteur et ses "activités", enrichir son parcours de lecture avec des documents numérisés
- peut réutiliser des données brutes exposées en RDF (dans le cadre technique du Web sémantique).

L'institution :

- gagne une meilleure visibilité de ses données sur le Web (amélioration du référencement) ;
- un effet de longue traîne : sont désormais indexées par Google les références data.bnf sur des auteurs et œuvres très spécialisées (par exemple le très obscur Adson de Montier-en-Der) ;
- un accès aux œuvres numérisées qui sortent ainsi de l'ombre.

Hors du domaine de la culture et des bibliothèques, des initiatives vers le web de données sont menées.

<sup>45</sup> <<http://data.bnf.fr/semanticweb>>

<sup>46</sup> <<http://www.logilab.org/>>

### 3.2.2 Décloisonner les institutions du savoir

Comme dans le domaine de l'expérimentation des FRBR, les pays anglo-saxons et scandinaves sont pionniers dans les initiatives « open data ». Leur démarche est à la fois associative et gratuite.

En France, l'ouverture des données publiques est récente, officialisée par le décret du 26 mai 2011<sup>47</sup>. L'objectif est la mise à disposition des citoyens des données publiques par l'intermédiaire d'un portail unique qui sera mis en service à l'automne 2011. Une licence ouverte a déjà été mise en place afin de faciliter et encourager la réutilisation des données publiques.

Les initiatives opendata sont déjà nombreuses (par exemple l'ouverture des données publiques de la Ville de Paris ou en région la Banque numérique du Savoir d'Aquitaine -BNSA) et se multiplient.

L'objectif du projet du W3C « Linking Open Data » est de placer sur le Web des données structurées en RDF. On peut reprendre la typologie proposée par le blog Les Petites Cases du Savoir pour disposer une vision globale des données actuellement « libérées » :

- des données géographiques et statistiques, par exemple avec le système d'information géographique GéoNames. L'espace RDF de l'INSEE présente les données géographiques que l'INSEE met à disposition aux formats standards du web sémantique<sup>48</sup>.

On peut observer que les fichiers Régions et département sont les instances (les exemplarisations) des classes définies dans l'ontologie.

- des données généralistes issues d'usuels (dictionnaires, encyclopédies en ligne) : l'initiative la plus aboutie étant Dbpedia issue de Wikipédia ;

- des données scientifiques, etc.

Datalift, décrit comme un « catalyseur pour le web de données » est un projet de recherche expérimentale financé par l'agence nationale de la recherche. L'objectif est de développer une plateforme pour publier et interconnecter des jeux de données structurées issues de différents formats (XML, CSV, ...) sur le web de données.

Datalift publie d'une part des jeux de données provenant de partenaires universitaires (INRIA, LIRMM), institutionnels (IGN, INSEE) et de MONDECA qui a déjà développé le projet Vega de l'INRA<sup>49</sup>. Datalift propose d'autre part un ensemble d'outils facilitant le processus de publication de jeux de données.

---

<sup>47</sup> Décret n° 2011-577 relatif à la réutilisation des informations publiques détenues par l'Etat et ses établissements publics administratifs ([mission Etalab](#)).

<sup>48</sup> <<http://rdf.insee.fr/>>

<sup>49</sup> Plateforme s'appuyant sur les technologies du web sémantique : des triplets RDF, Dublin Core pour les méta-données génériques, Skos pour la formalisation des thésaurus, Foaf pour décrire les relations entre individus qui établit un lien entre les données.

### 3.3 D'autres pistes tous azimut

Peuvent être envisagée d'une part une collaboration entre grandes institutions internationales, d'autre part une intégration à de nouveaux modèles de données.

#### 3.3.1 Collaborer avec la FAO

En effet, les profils de deux institutions sont similaires : des institutions internationales, des centres de publication et un intérêt commun ... pour les FRBR.

Une collaboration est envisageable à deux niveaux :

- sur leur projet d'archives ouvertes
- sur la constitution de réservoirs communs de données. A partir de Voc Bench<sup>50</sup>, à l'origine créé pour gérer le thésaurus multilingue de la FAO, d'autres vocabulaires sont intégrés. Cet outil aide les gestionnaires de l'information à construire et à structurer des modèles de connaissance. Il est particulièrement adapté à la gestion des vocabulaires multilingues qui nécessitent un environnement collaboratif réparti.

Une piste supplémentaire envisageable serait de constituer un réseau inter UNESCO, ou à une échelle plus fine inter IIFE, par exemple pour les bibliothèques dépositaires.

---

<sup>50</sup> Voc bench <<http://aims.fao.org/fr/tools/vocbench-2>>

### 3.3.2 Envisager des solutions alternatives pour mettre les données sur le web

1- Pour développer l'interopérabilité entre les métadonnées créées par différentes communautés (édition/production ; bibliothèques ; éducation), le Vocabulary Mapping Framework (VMF) permet de faire correspondre des métadonnées de différentes origines.

Ce vocabulaire couvre les entités des groupes 1 et 2 du modèle FRBR et leurs relations. Il passe par l'élaboration d'une ontologie en RDF / OWL pour permettre l'échange des informations établies selon ces normes. [56]

La finalité du projet VMF<sup>51</sup> est d'établir des correspondances entre différents vocabulaires de schémas de métadonnées, et de développer un outil libre permettant l'interopérabilité entre ces différentes communautés productrices de ces schémas.

2- S'inspirer de la modélisation des thèses françaises (TEF) selon le modèle FRBR. La recommandation TEF propose un jeu de métadonnées pour les thèses électroniques soutenues en France. « Son objectif est d'organiser de manière cohérente des métadonnées de thèse riches et normalisées, pour faciliter leur échange et leur diffusion, au niveau national ou international. »

Ces métadonnées sont hétérogènes avec des métadonnées descriptives (bibliographiques) et des métadonnées de gestion (administration, droits, conservation).

Le format utilisé est XML qui s'appuie sur le schéma METS qui permet d'articuler les différents types de métadonnées.

Les entités sont modélisées : chaque entité a ses métadonnées

- œuvre : métadonnées descriptive / administratives / droits
- versions : métadonnées descriptive / droits
- édition : métadonnées descriptive
- fichier : métadonnées de conservation

Puis elles sont structurées en XML ce qui offre des possibilités d'échange avec OAI-PMH. [29] Le protocole OAI-PMH est dans la lignée de l'open access.

Son objectif est de faire converger les données de la recherche (publications entreposées dans des archives ouvertes) et d'autres données issues notamment du domaine culturel et des bibliothèques.

Comment garantir l'interopérabilité entre ces différentes sources de données ? Par le Dublin Core simple : « Les données ainsi formatées sont moissonnées, c'est-à-dire récupérées dans les bases réparties, pour être versées dans une base commune qui servira de support à la création de nouveaux services. » [35]

3- Approfondir le modèle de données EDM (European Data Model)<sup>52</sup> actuellement développé dans le projet de bibliothèque numérique européenne Europeana (European Digital Library), il devrait être mis en œuvre en novembre 2011. Il présente une interopérabilité avec les schémas de données du CIDOM CRM.

Le "concept" de notice bibliographique est remplacé par l'EDM, un modèle de données appliqué au Thought Lab (laboratoire d'idées) d'Europeana. Cette plate forme expérimentale sémantique réunit les données de trois musées européens. C'est un modèle à graphe sémantique dont la base est une « interconnexion » entre des agrégations des ressources web (arêtes entre les notions étiquetées avec

<sup>51</sup> Vocabulary Mapping Framework <<http://www.doi.org/VMF/index.html>>

<sup>52</sup> European Data Model <<http://version1.europeana.eu/Web/lod/home>>

la relation sémantique qui permet de décrire les notions : plus le graphe est précis plus on peut faire des requêtes fines). [85]

Cette plate forme expérimentale sémantique repose sur une ontologie :

- un vocabulaire structuré créé par des professionnels de l'information;
- des relations traduites automatiquement entre des objets par des ontologies libres sur le web. Les classes et relations sont modelées au standard RDFS d'où une vision d'ensemble qui diffère de celle des bibliothèques traditionnelles.

Un réseau sémantique hiérarchique / un schéma conceptuel avec des relations intégrées  
On passe du catalogue (notices/ métadonnées) à un réseau conceptuel. Les classes (event, place, time span, thing..) et les propriétés (has occured, has type...) sont issues du CRM.

4- Intégrer certaines données produites par l'IPE (publications, travaux de recherche, base de données) sur le portail Isidore pour accroître encore la visibilité de l'Institut. ISIDORE est une plateforme de recherche permettant l'accès aux données numériques des sciences humaines et sociales (SHS). Elle s'appuie sur les principes du web de données et donne accès à des données en accès libre. Une cinquantaine d'institutions ont déjà rejoint le site notamment l'Institut de recherche pour le développement (IRD) dont les problématiques sont assez proches de celles de l'IPE. [82]

### 3.3.3 Mettre les données sur le web : la piste des éditeurs de logiciels

D'autres métiers, étrangers à la sphère culturelle, sont pionniers dans la course au web sémantique.

Ainsi l'e-commerce a adopté très tôt les technologies du web sémantique. Cette même sphère avait testé peu après sa parution le modèle FRBR et l'avait adapté, certes de manière iconoclaste et avec des arrières pensées très pragmatiques.

L'e-commerce recherche en effet le même Graal que les bibliothèques et centres de documentation: un catalogue permettant d'améliorer la visibilité des ses objets / produits auprès des partenaires en les exposant au plus près des utilisateurs : dans les pages web. Même si leur problématique tourne autour de l'augmentation du référencement et donc du chiffre d'affaire, nous jugeons utile de mettre une veille sur leurs travaux.

A l'initiative des principaux moteurs de recherche – et concurrents- (Google, Bing et Yahoo unis pour l'occasion), une nouvelle syntaxe (qui n'est pas RDF mais Microdata) et un nouveau vocabulaire (Schema.org) ont été développés pour améliorer l'indexation des sites. Schema.org propose des « informations sémantiques » à intégrer dans le code HTML des pages pour gérer les œuvres les plus courantes dans un catalogue commercial : des livres, des films, des enregistrements musicaux mais aussi, comme dans le modèle CIDOC CRM, des événements, des lieux, des personnes typés à l'aide de sous catégories.

Prenons par exemple, le schéma événement : un événement se produit à un moment donné dans un certain endroit. Il est subdivisé en « types » spécifiques : événement éducatif, événement social rappelant la structure des rubriques du catalogue de la Gaîté lyrique.

Il pourrait être utile de contacter l'éditeur de logiciel Antidot qui travaille sur les enjeux de la valorisation des bases de données relationnelles, et en particulier de leur transformation en données sémantiques (graphes RDF).

# Conclusion



L'IPE est à ce jour le seul Centre de documentation à avoir mené une expérimentation des modèles FRBR. Au terme de l'expérience FRBR, l'objectif - trouver une manière plus efficace pour que l'utilisateur accède à l'information contenue dans différentes bases de données - a été atteint.

Les scénarios proposés intègrent le modèle FRBR entité relation, ou plus encore, orienté objet : malgré sa complexité, il présente une approche plus globale des données des bases et sa mise en œuvre semble réalisable. Il s'agira ensuite de faire le choix d'une solution informatique en remplacement de l'actuel logiciel et de décider du déploiement (conserver des bases de données corrélées, choisir une plateforme unique d'interrogation...).

Au final il apparaît que l'expérience menée avec les modèles conceptuels a surtout permis de révéler des failles dans le système d'information. Si l'expérience FRBR était menée à son terme, cela permettrait de formaliser après coup l'intuition première en filigrane dans les pratiques actuelles du Centre de documentation de l'IPE : que l'utilisateur accède de manière groupée à des informations interconnectées.

Grâce à ces modèles conceptuels et aux pistes d'implémentation qui restent encore à explorer, des perspectives documentaires nouvelles s'ouvrent.

Elles nécessitent de positionner l'institution IPE dans un environnement global, constitué en réseau, et de créer des synergies avec d'autres institutions internationales, par exemple la FAO. C'est d'ailleurs une des préconisations du CIDOC CRM : « *les institutions doivent commencer à employer de nouvelles stratégies pour la création de métadonnées et de flux de travail partagés, et pour le partage des métadonnées et des outils de documentation entre les institutions* ». [81]

Quelque soit le scénario retenu et l'outil choisi, l'IPE devrait continuer à se préparer au web de données et à l'ouverture de ses données structurées et de qualité.

Mais il nous semble que préalablement à tout travail de modélisation du système d'information, il conviendra d'engager une réflexion plus globale sur les circuits de circulation de l'information (workflow) et plus généralement de gestion des connaissances afin de réorganiser la politique documentaire (pour quels besoins, quels usages ?). Cette analyse des besoins sous-entend l'implication des différents services de l'Institut (Centre de documentation, unité des publications, service informatique)

Le défi subséquent est de convaincre les décideurs des bénéfices de cette réflexion initiée sur la gestion des connaissances et la structuration de l'information et de faire comprendre que la documentation, loin de se résumer à une activité technique, constitue un des vecteurs d'amélioration de l'organisation globale de l'institution.

# Bibliographie

Les références bibliographiques sont présentées selon les normes de référence :

- Z44-005. décembre 1987. Documentation. Références bibliographiques : contenu, forme et structure.

- NF ISO 690-2 Février 1998 Information et documentation. Références bibliographiques. Documents électroniques, documents complets et parties de documents.

Les notices sont classées de façon thématique, puis à l'intérieur de chaque thème, s'il y a mention d'auteur(s), par ordre alphabétique du nom du premier auteur cité (auteurs personnes physiques ou personnes morales). Pour les références sans mention d'auteur, les notices sont classées en tête par ordre alphabétique du titre.

L'ensemble des notices est numéroté en continu. Ces numéros entre crochets correspondent à ceux figurant dans le corps du texte de ce mémoire.

Le "plan de classement" de la bibliographie est le suivant :

1 L'information numérique et ses avatars

2 Structurer l'information

3 Décrire et identifier pour trouver et obtenir sur le web

4 Mettre les données dans le web

5 Un outil de médiation des données : le modèle conceptuel FRBR

6 Application des FRBR au cadre d'une organisation internationale

7 Les préconisations dans le contexte de l'IIPE

8 Pour aller plus loin : pistes d'approfondissement et sites recommandés pour une veille sur les modèles FRBR et les problématiques du web sémantique

## 1 L'information numérique et ses avatars

[1] CALENGE Bertrand. **Bibliothèques et politiques documentaires à l'heure d'Internet**. Paris, Editions du Cercle de la Librairie, 2008. 264 p. Collection Bibliothèques. ISBN 978-2-7654-0962-5 *Mesure et décrit le « grand chambardement » qu'entraîne le web sur les ressources des bibliothèques et leur accès (collections et flux électroniques) mais aussi sur les pratiques culturelles et la conduite des politiques documentaires.*

[2] CHABIN, Marie-Anne. **Blog Impressions, expressions**. [Consulté le 15 octobre 2011].

<<http://www.marieannechabin.fr/archiver-et-apres/2-archiver-ou-conserver>>

*L'article est extrait du blog Impressions, expressions qui offre des pistes de réflexion sur fond de « diplomatie numérique et d'accumulation d'observations des pratiques en matière de gestion de l'information et d'archivage managérial ».*

*Archiver et/ ou conserver les documents et /ou les données : examen du dispositif législatif en la matière, de la perception de la notion d'archivage par la société, de l'offre d'archivage actuellement proposée par les prestataires et de ce que devrait être cette offre au vu de l'évolution des supports d'information et de l'importance des métadonnées pour bien tracer l'information à archiver.*

[3] CHABIN Marie-Anne. **Document trace et document source. La technologie numérique change-t-elle la notion de document ?** Information – Interaction - Intelligence, Volume 4, n°1, pp. 141-157. Issn 1630-649X

[4] ERTZSCHEID Olivier. **L'homme est un document comme les autres : du World Wide Web au World Life Web** [en ligne]. [Consulté le 25 septembre 2011].

Version préprint parue sur <http://archivesic.ccsd.cnrs.fr/docs/00/37/82/34/PDF/articledefOE.pdf>

*L'auteur, maître de conférences en sciences de l'information et de la communication, engage ici une réflexion sur l'évolution de la société qui place l'homme, un objet documentaire parmi d'autres, au centre du cycle documentaire.*

Olivier Ertzscheid est également l'auteur du blog <http://www.affordance.info>.

[5] GALLEZOT Gabriel. **Un cycle panoptique pour l'analyse des documents** [en ligne]. [Consulté le 22 septembre 2011].

<[http://hal.inria.fr/docs/00/06/24/33/PDF/sic\\_00001099.pdf](http://hal.inria.fr/docs/00/06/24/33/PDF/sic_00001099.pdf)>

*Maître de conférence (Université Nice Sophia Antipolis) et chercheur sur le thème « accès et recherche d'information », notamment en ligne. Cet axe de travail recouvre aussi bien des travaux sur l'usage des outils de recherche, les pratiques informationnelles ; des réflexions sur la définition, la nature, les formes du document numérique et des analyses sur la publication des connaissances.*

[6] GUYOT Brigitte. **Introduction à l'ingénierie documentaire**. Tome 1. 2009, 71 p.

[7] LAINÉ-CRUZEL Sylvie. **Documents, ressources, données : les avatars de l'information numérique**. Information – Interaction - Intelligence, volume 4, n°1, 2004, p. 105-119. Issn 1630-649X [Consulté le 15 septembre 2011].

<[http://archivesic.ccsd.cnrs.fr/docs/00/06/24/00/PDF/sic\\_00001018.pdf](http://archivesic.ccsd.cnrs.fr/docs/00/06/24/00/PDF/sic_00001018.pdf)>

[Résumé d'auteur] *En s'intéressant à la notion de document électronique, on différencie deux évolutions qui sont utiles à des titres distincts : d'une part les documents, qui ont valeur de référence,*

*dont l'administrateur veillera à préserver l'intégrité et dont le contexte de production est fortement significatif, et d'autre part les ressources, éléments évolutifs par nature, que l'utilisateur doit pouvoir s'approprier et qui doivent donc être mises à disposition de manière à ce que la production s'efface au profit d'une logique d'exploitation et de service.*

[8] PEDAUQUE, Roger T. **Le document comme contenant, contenu et médium. Les reformulations du numérique** [en ligne], Archivesic CNRS, version 3 du 08/07/2003. [Consulté le 10 septembre 2011]

<[http://archivesic.ccsd.cnrs.fr/docs/00/06/21/99/PDF/sic\\_00000511.pdf](http://archivesic.ccsd.cnrs.fr/docs/00/06/21/99/PDF/sic_00000511.pdf)>

*Réflexion du groupe de travail collectif Pédauque : il examine d'abord ce qu'est le document traditionnel, ses fonctions, et en quoi il constitue un objet de recherche fécond ; puis, dans un deuxième temps, analyse des conséquences de sa mutation vers le numérique, et la transformation subséquente des pratiques, et des questionnements des chercheurs.*

[9] PUIMATTO Gérard. **Les métadonnées : pourquoi, et pour quoi faire ?** [en ligne] [Consulté le 18 octobre 2011].

<[http://www.cndp.fr/savoirscdi/fileadmin/fichiers\\_auteurs/PDF\\_manuels/Les\\_metadonnees\\_Puimatto.pdf](http://www.cndp.fr/savoirscdi/fileadmin/fichiers_auteurs/PDF_manuels/Les_metadonnees_Puimatto.pdf)>

*Responsable d'un CDDP, Gérard Puimatto travaille sur des projets de plates-formes de contenus éditoriaux en ligne pour l'enseignement scolaire. Dans cet article, il définit et fait une présentation des métadonnées à l'intention de la communauté scolaire et des centres de documentation : nature, typologie, évolution des métadonnées sur le web et tour d'horizon des principaux formats.*

[10] VATANT Bernard. **Des métadonnées à la description des ressources : les langages du web sémantique.** In CALDERAN Lisette, HIDOINE, Bernard, MILLET Jacques, INRIA – Institut National de Recherche en Informatique et en Automatique. Métadonnées : mutations et perspectives. Séminaire INRIA 29 septembre - 3 octobre 2008 - Dijon. Paris, ADBS éditions, 2008. Chapitre 5, p. 163-194. Sciences et techniques de l'information. ISBN 978-2-84365-104

*Bernard Vatant, expert dans le domaine des langages de représentation et des technologies du web sémantique, présente la dérivation progressive de la notion de document à celle de ressource ; corrélativement, le format standard de métadonnées RDF évolue et s'étend : présentation des langages du web sémantique bâtis sur RDF ; enfin listage des tendances du web social sémantique dans lequel les ressources deviennent aussi des « objets d'échange social », remettant en question la médiation traditionnelle du documentaliste.*

## 2 Structurer l'information

[11] ANGJELI Anila. **La structuration des données: bibliothèques de composants et modèles. Données d'autorité : structuration, modélisation, interconnexion.** In AFNOR- CG 46. Journée d'étude AFNOR : référentiels, données d'autorité, thésaurus, ontologies... [en ligne]. 28 mars 2008, Paris [consulté le 30 octobre 2011] <<http://www.bnf.fr/documents/angjeli.pdf>> *Présentation très synthétique de l'utilité des données d'autorité pour interconnecter l'information et « piloter » l'accès aux ressources dans un contexte en évolution.*

[12] BERMES Emmanuelle. **Les catalogues sur le web.** In MEDIAL. Journée d'études : les catalogues nouvelle génération [en ligne]. 8 juin 2009, Nancy [consulté le 29 octobre 2011] <<http://www.slideshare.net/Figoblog/les-catalogues-sur-le-web-6112279>>  
*Tour d'horizon en images du catalogue et de ses évolutions jusqu'au web de données avec ses perspectives d'interopérabilité.*

[13] DENTON William. **FRBR and the History of Cataloging.** p.35- 57. In Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools. Westport, CT: Libraries Unlimited, 2007.  
*Historique des précurseurs (Panizzi, Cutter, Ranganathan, Lubetzky) qui ont ouvert la voie vers le modèle FRBR concernant*  
- *l'objectif des catalogues,*  
- *l'importance des besoins des utilisateurs,*  
- *l'idée d'œuvre,*  
- *la nécessaire standardisation et internationalisation.*  
*L'auteur, bibliothécaire web ('web librarian') à l'université de Toronto, est également le fondateur et animateur du blog consacré aux FRBR ([www.frbr.org](http://www.frbr.org)).*

[14] LERESCHE Françoise. **Les règles françaises de catalogage à la croisée des chemins: quelles évolutions pour quels enjeux ?** Bid textos universitaris de biblioteconomia i documentació [en ligne], 2009, numéro 22 [consulté le 28 octobre 2011]. ISSN 1575-5886 <<http://www.ub.edu/bid/22/leresche2.htm>>  
*Responsable du Pôle Modélisation fonctionnelle (Département de l'Information bibliographique et numérique de la BnF), l'auteur pilote également le groupe technique sur les FRBR et RDA.*  
*L'objectif de l'article est de faire un audit de la situation actuelle de la norme française de catalogage, afin d'envisager l'avenir des normes française de catalogage : le maintien de la référence directe aux ISBD (et par conséquent aux recommandations de l'IFLA) ou l'adoption de RDA.*

[15] MANNEHEUT Anne. **Les formats historiques de catalogage des documents : quarante ans de bons et loyaux services.** AFNOR, septembre 2009, 20 p.  
*Retrace l'évolution des formats historiques de production de données informatisées pour les catalogues de bibliothèques, leur adaptation constante aux évolutions de leur environnement et insiste sur le « gisement d'information » à récupérer des notices en MARC .*

[16] NIESZKOWSKA-SERLAN Ewa. **Catalogue, catalogues, métacatalogues.** AFNOR  
*Présentation de l'évolution des outils de signalement et des techniques permettant de rendre plus visibles les données sur le web.*

[17] RAIS Nadia. **Note technique : introduction à l'informatique documentaire : principes et applications, techniques et gestion d'une base de données.** Novembre 2009, 64 p.  
*Définition des notions de base (système d'information, SGBD, formats de fichiers), présentation des éléments constitutifs des logiciels de traitement documentaire et méthodes pour la mise en place d'une base de données documentaire.*

Le fascicule peut être consulté sur demande au secrétariat de l'INTD/CNAM.

### 3 Décrire et identifier pour trouver et obtenir sur le web

[18] BIBLIOTHEQUE NATIONALE DE FRANCE. **Catalogage et indexation** [en ligne], 25 juillet 2011. [Consulté le 22 octobre 2011]

<[http://www.bnf.fr/fr/professionnels/catalogage\\_indexation.html](http://www.bnf.fr/fr/professionnels/catalogage_indexation.html)>

*La partie professionnelle du site de la Bibliothèque nationale de France met à disposition des documents normatifs et des référentiels en particulier sur les processus de normalisation en documentation, sur les politiques de catalogage et d'indexation appliquées par la BnF dans ses catalogues, sur les différents formats utilisés par la BnF pour décrire ses collections.*

[19] **Métadonnée** [en ligne], modification le 6 avril 2009. [Consulté le 30 octobre 2011]

<<http://wiki.univ-paris5.fr/wiki/M%C3%A9tadonn%C3%A9e>>

*Cette ressource est intéressante car c'est une initiative collaborative de mise à disposition du savoir. Ce Wiki a été initié par le département TICE de l'Université Paris Descartes.*

[20] BERMES Emmanuelle. **Les métadonnées à la Bibliothèque nationale de France : une vue d'ensemble** [en ligne], 2005. [Consulté le 15 octobre 2011]

<[http://cdsweb.u-strasbg.fr/MDA/mda/ontologies/BNF\\_CDS20050513.ppt](http://cdsweb.u-strasbg.fr/MDA/mda/ontologies/BNF_CDS20050513.ppt)>

*Définition des métadonnées et panorama très complet des différentes catégories de métadonnées, des formats et protocoles d'échange.*

[21] CHABIN Marie-Anne. **Face aux métadonnées : carte d'orientation dans la jungle des métadonnées...**[en ligne], Groupe PIN, 21 juin 2011. [Consulté le 28 septembre 2011]

<<http://pin.association-aristote.fr/doku.php/public/metadonnees>>

*Travaux du groupe de pérennisation de l'information numérique (PIN) mis en place par l'association Aristote impliquée dans les développements et les nouveaux usages des technologies de l'information.*

[22] DALBIN Sylvie. **Métadonnées et normalisation.** In CALDERAN Lissette, HIDOINE, Bernard, MILLET Jacques, INRIA – Institut National de Recherche en Informatique et en Automatique. Métadonnées : mutations et perspectives. Séminaire INRIA 29 septembre - 3 octobre 2008 - Dijon. Paris, ADBS éditions, 2008. Chapitre 4, p. 113-162. Sciences et techniques de l'information. ISBN 978-2-84365-104

[23] DALBIN Sylvie. **Représentation et accès à l'information: transformation à l'œuvre.** In CALDERAN Lissette, HIDOINE, Bernard, MILLET Jacques, INRIA – Institut National de Recherche en Informatique et en Automatique. Métadonnées : mutations et perspectives. Séminaire INRIA 29 septembre - 3 octobre 2008 - Dijon. Paris, ADBS éditions, 2008. Chapitre 1, p. 9-58. Sciences et techniques de l'information. ISBN 978-2-84365-104

Une version préprint est également disponible à l'adresse <[http://www.atd-doc.com/xmedia/publications/INRIA-2008\\_DALBIN\\_Chap01.pdf](http://www.atd-doc.com/xmedia/publications/INRIA-2008_DALBIN_Chap01.pdf)>

*Présentation de la notion de métadonnées et du processus de création (outils, langage), focus sur les modèles conceptuels qui structurent les données ; pistes de réflexion pour faire évoluer la vision des professionnels de l'information documentation sur les métadonnées (approche par points de vue portés sur l'objet informationnel, approche par métiers pour changer notre représentation des ressources).*

[24] DESCHATELETS Jacinthe. **Dossier sur les métadonnées**. Mis à jour le 12 mai 2009 [Consulté le 24 septembre 2011]

<[http://bibliodoc.francophonie.org/article.php3?id\\_article=172](http://bibliodoc.francophonie.org/article.php3?id_article=172)>

*Le site Bibliodoc, élaboré avec l'appui de l'Organisation Internationale de la Francophonie (OIF), est un outil d'information et de communication pour les bibliothèques et centres de documentation de la Francophonie. Ce dossier très complet, augmenté d'une abondante bibliographie, présente les principaux formats de métadonnées ainsi que des pistes de réflexion (les métadonnées menacent-elles la survie des catalogueurs,...).*

[25] ERTZSCHEID Olivier. **Les « nouveaux » catalogues ou le catalogue en (r)évolutions**. Arabesques, n° 48, octobre - novembre - décembre 2007 [Consulté le 25 septembre 2011]

<<http://www.abes.fr/Arabesques/Arabesques-n-48>>

*Arabesques, revue de l'Agence bibliographique de l'enseignement supérieur, s'intéresse à la question de la documentation dans l'enseignement supérieur et la recherche.*

[26] INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **Principes internationaux de catalogage** [en ligne]. 2009, 15 p [Consulté le 18 septembre 2011]

<[http://www.ifla.org/files/cataloguing/icp/icp\\_2009-fr.pdf](http://www.ifla.org/files/cataloguing/icp/icp_2009-fr.pdf)>

*Le texte qui remplace les principes de Paris et qui repose sur le modèle conceptuel de l'IFLA, les FRBR.*

[27] MAISONNEUVE, Marc. **Actualité des catalogues et des systèmes de gestion de bibliothèque**. Bulletin des Bibliothèques de France BBF, 2011, p.12-13 [Consulté le 29 août 2011]

<<http://www.enssib.fr/bibliotheque-numerique/document-49521>>

Le Bulletin des Bibliothèques de France, publié par l'ENSSIB, est une revue consacrée aux bibliothèques et aux services de documentation.

*Présentation par le consultant Marc Maisonneuve de la nouvelle donne des systèmes de gestion de bibliothèque dans un environnement en mutation et des adaptations des OPAC dans ce contexte.*

[28] MOREL-PAIR Catherine. **Dublin Core : théorie et applications**. In MEDIADIX. Formation « Métadonnées », 16-17 septembre 2009, Saint-Cloud [en ligne]. [Consulté le 25 septembre 2011]

<[http://mediadix.u-paris10.fr/stockage\\_doc/metadonneesbibnum/metasDCMorelPair.pdf](http://mediadix.u-paris10.fr/stockage_doc/metadonneesbibnum/metasDCMorelPair.pdf)>

*Cette présentation, sous licence Creative Commons, présente les éléments Dublin Core et leur implémentation, les outils et les applications, en particulier dans l'optique du web sémantique.*

[29] NICOLAS Yann. **TEF, métadonnées pour les thèses numériques françaises** [en ligne]. [Consulté le 18 octobre 2011]

<<http://www.slideshare.net/ent12701/sp6-yann-nicolas>>

*Présentation du contexte, des fonctions et de l'outil TEF mais surtout de la démarche engagée jusqu'à l'implémentation finale.*

[30] ONLINE COMPUTER LIBRARY CENTER, OCLC. **Catalogues en ligne, les souhaits des utilisateurs et des bibliothécaires : rapport d'OCLC** [en ligne]. Dublin Ohio, OCLC, 2009, 24 p. ISBN: 978-1-55653-418-8 [Consulté le 30 septembre 2011]

*Le rapport d'OCLC sur les Catalogues en ligne, paru en 2009, permet de catégoriser les doléances des utilisateurs et des bibliothécaires.*



<[http://www.oclc.org/reports/onlinecatalogs/213724cfb\\_%20Online\\_Catalogs\\_Synopsis.pdf](http://www.oclc.org/reports/onlinecatalogs/213724cfb_%20Online_Catalogs_Synopsis.pdf)>

[31] PISANSKI Jan, ZUMER Maja. **Mental models of the bibliographic universe. Part 1: mental models of descriptions ; Part 2: comparison task and conclusions.** Journal of Documentation, 2010, Vol. 66, No. 5, 2010, pp.668-680.

[32] PLOURDE, Marcel. **L'avenir du catalogue** : sommaire du rapport présenté au Sous-comité des bibliothèques de la CREPUQ par le Groupe de travail sur le traitement de la documentation. Documentation et bibliothèques, juillet-septembre 2009, p.140-143.

*Etat de la réflexion sur l'avenir du catalogue piloté par le sous-comité de la Conférence des recteurs et des principaux des Universités du Québec (CRERUQ) ; résumé des travaux du groupe de travail canadien.*

[33] TESSIER Marc. **Rapport sur la numérisation du patrimoine écrit** [en ligne]. Janvier 2010, 64 p. <[http://www.ddm.gouv.fr/IMG/pdf/Rapport\\_sur\\_la\\_numerisation\\_du\\_patrimoine\\_ecrit.pdf](http://www.ddm.gouv.fr/IMG/pdf/Rapport_sur_la_numerisation_du_patrimoine_ecrit.pdf)>

*Etat des lieux sur la bibliothèque numérique ; éclairage intéressant car rappel du contexte de l'évolution des bibliothèques (évolution des usages et présentation des initiatives de mutualisation des actions des bibliothèques).*

#### 4 Mettre les données dans le web

[34] BERMES Emmanuelle. **Publier les référentiels sur le Web sémantique : tutoriel** [en ligne]. In AFNOR CG46 - BNF. Journée : référentiels et données d'autorité à l'heure du Web sémantique. mai 2011, Paris. [Consulté le 27 octobre 2011]<[http://www.bnf.fr/documents/afnor2011\\_bermes.pdf](http://www.bnf.fr/documents/afnor2011_bermes.pdf)>  
*Introduction au web sémantique (définitions, notions de base), focus sur SKOS, un référentiel pour les web de données et plus globalement réflexion sur les enjeux et nécessité de l'interopérabilité dans le web de données.*

[35] BERMES Emmanuelle. **Convergence et interopérabilité : l'apport du Web de données.** In IFLA<sup>53</sup>. 77<sup>e</sup> Meeting, 13-18 August 2011, Porto-Rico [en ligne]. [Consulté le 25 octobre 2011] <<http://conference.ifla.org/past/ifla77/149-bermes-fr.pdf>>

*Article essentiel car Emmanuelle Bermès place son intervention dans l'objectif à atteindre d'une ouverture et d'une convergence entre institutions culturelles. Présentation des moyens techniques permettant d'assurer cette interopérabilité, rôle des vocabulaires dans la mise en œuvre de l'interopérabilité basée sur les liens, exemples d'applications reliant les données (dont le Centre Pompidou Virtuel) et évocation des barrières freinant la constitution d'un espace d'information global.*

[36] BERNERS-LEE Tim, HENDLER James, LASSILE Ora, **Le web sémantique.** LACOMBE Elisabeth, LINK-PEZET Jo (trad.). La Lettre de l'URFIST de Toulouse [en ligne]. N°28 [consulté le 05 octobre 2011] p. 14-23 <<http://www.urfist.cict.fr/archive/lettres/lettre28/lettre28.pdf>>

---

<sup>53</sup> Note complémentaire : on peut signaler l'importance des comptes rendus des congrès mondiaux et assemblées générales de l'IFLA.

[37] LERESCHE Françoise. **Vers de nouveaux codes de catalogage : RDA.** In Centre de formation aux carrières des bibliothèques, CFCB. Journée d'étude FRBR, RDA, FRAD, RDF, web sémantique : les catalogues de bibliothèque changent d'ère [en ligne]. 10 décembre 2010, Toulouse. [consulté le 05 octobre 2011]

<<http://cfcb.univ-tlse2.fr/accueil-cfcb/archives/rendez-vous-professionnel-29570.kjsp>>

[38] MAUGER-PEREZ Isabelle. **Les formats à l'ère du Web: de la multiplicité des métadonnées à la convergence vers le sens ?** AFNOR, septembre 2009, 29 p.

*Présentation de la nécessité de la modélisation et des principaux schémas de métadonnées. Malgré l'émiettement apparent des vocabulaires et l'hétérogénéité des données, possibilité de convergence si les professionnels de l'information documentation acceptent de « libérer leurs métadonnées ».*

[39] MIKSA Shawne D. **Resource Description and Access (RDA) and New Research Potentials.** Bulletin of the American Society for Information Science and Technology, June/July 2009, Volume 35, Number 5, p.47-51. ISSN: 1550-8366

*Court article très documenté présentant les caractéristiques essentielles de RDA dans le contexte des nouvelles règles de catalogage [avec tableau comparatif entre AACR 2 et RDA], ainsi qu'un inventaire des nouvelles potentialités de recherche.*

[40] ZAKLAD Manuel. **Web socio sémantique et recherche ouverte d'information : le SI entre participation et contrôle.** In MONNIN Alexandre, BROUDOUX, Evelyne. Séminaire Folksonomie et tagging [en ligne]. 26 mars 2010, CNAM, Paris. [consulté le 25 octobre 2011]

<[http://dicen.cnam.fr/servlet/com.univ.collaboratif.utils.LectureFichiergw?ID\\_FICHIER=1295877018776](http://dicen.cnam.fr/servlet/com.univ.collaboratif.utils.LectureFichiergw?ID_FICHIER=1295877018776)>

## 5 Un outil de médiation des données : le modèle conceptuel FRBR

[41] Journée d'études « En avant avec les autorités », Washington, 9 août 2010. Bulletin sur les ressources archivistiques numériques, n°41, octobre 2010, p.11 [consulté le 20 octobre 2011]

<<http://www.archivesdefrance.culture.gouv.fr/static/4217>>

*L'article sur « Le contrôle bibliographique des documents d'archives souligne l'impact des normes de bibliothèques sur les archives » et la portée des évolutions récentes FRBR (et de leur application dans RDA) dans la description bibliographique des documents d'archive.*

[42] AFNOR GE 6. **Groupe technique sur l'adoption de RDA en France.** [consulté le 18 octobre 2011]

<<http://rda-en-france.enssib.fr/>>

*Ce groupe de travail technique est chargé d'évaluer l'intérêt et les modalités de l'adoption de RDA dans les bibliothèques françaises. Même si une partie du site est réservée aux membres du groupe de travail, les ressources proposées en accès libre sont conséquentes (lien vers les tests entrepris par la Library of Congress sur l'adoption de RDA ; vers le site francophone de RDA...).*

[43] ANGJELI Anila. **FRBR et les modèles liés** [en ligne]. In FULBI : Journée de travail sur RDA, 16 juin 2011, Ivry. [consulté le 10 novembre 2011]

<<http://www.fulbi.fr/sites/default/files/angjeli.pdf>>

*Anila Angjeli (Bibliothèque nationale de France) est présidente du groupe d'expert AFNOR « Données d'autorité ». Dans ce document essentiel, elle explique très clairement le modèle FRBR en le resituant dans le contexte du catalogage. En 20 diapositives est abordé l'essentiel des enjeux et de la structure du modèle. Dans la suite de la présentation sont développés les autorités (domaine des FRAD) et les sujets (domaine des FRSAD).*

[44] CAZENAVE Jean, DUCHARME Christian. **Livre Blanc : le FRBR sans attendre**. Decalog, septembre 2011. [consulté le 5 novembre 2011]

< <http://www.decalog.net/files/LivreBlancFRBR.pdf>>

[45] CARLYLE Allyson. **Understanding FRBR as a Conceptual Model: FRBR and the Bibliographic Universe**. In: Library Resources & Technical Services. 2006, Vol. 50, No. 4, pp. 264-273. ISSN: 0024-2527

*Vulgarisation du modèle conceptuel en insistant sur les entités fondamentales du groupe 1, en particulier les niveaux les plus abstraits que sont l'oeuvre et l'expression. S'appuie sur des exemples extérieurs au catalogue (L'amour comme oeuvre) pour faire comprendre les relations et rendre la modélisation FRBR moins abstraite. Remplace le modèle FRBR dans la continuité d'autres modèles de catalogage. Positionne les FRBR face au défi de l'implémentation et souligne la nécessité d'établir des frontières nettes entre entités oeuvre et expression/ expression et manifestation.*

[46] Centre de formation aux carrières des bibliothèques Midi-Pyrénées et Languedoc-Roussillon, CFCB. **Journée d'étude « FRBR, RDA, FRAD, RDF, web sémantique : les catalogues de bibliothèques changent d'ère»**. Toulouse, 10 décembre 2010 [en ligne]. Toulouse: Centre de formation aux carrières des bibliothèques Midi-Pyrénées et Languedoc-Roussillon. [consulté le 05 octobre 2011]

<[http://cfcb.univ-tlse2.fr/28391004/0/fiche\\_pagelibre/&RH=ACCUEIL\\_CFCB?RF=Archives-cfcb](http://cfcb.univ-tlse2.fr/28391004/0/fiche_pagelibre/&RH=ACCUEIL_CFCB?RF=Archives-cfcb)>

*Avec notamment les contributions de spécialistes des FRBR (Françoise Leresche et Philippe Le Pape) et du web sémantique (Emmanuelle Bermès).*

[47] DICKEY Timothy. **FRBRization of a Library Catalog: Better Collocation of Records, Leading to Enhanced Search, Retrieval, and Display**. Information Technology & Libraries; Mar2008, Vol. 27 Issue 1, p23-32. ISSN: 0730-9295

*Insiste sur les avantages des FRBR pour les OPAC, en particulier des bibliothèques musicales ; rappel de la structure des niveaux d'entités. Illustré par des exemples de catalogues frbrisés.*

[48] Fédération des Utilisateurs de Logiciels pour Bibliothèques, Documentation & Information, FULBI. **Journée de travail sur RDA**. Ivry, 16 juin 2011 [en ligne]. FULBI, 16/06/2011. *Journée d'information sur RDA organisée par la FULBI (regroupant les associations d'utilisateurs de logiciels dans le monde des bibliothèques et centres de documentation spécialisé). En plus de la présentation d'Anila Anjeli (FRBR et modèles liés), les présentations sur RDA de Françoise Leresche (BnF) et Philippe Lepape (ABES) peuvent être téléchargées sur le site de la FULBI.*

[49] IFLA Study Group on the Functional Requirements for Bibliographic Records. **Functional Requirements for Bibliographic Records : final report** [en ligne]. 2009 <[http://www.ifla.org/files/cataloguing/frbr/frbr\\_2008.pdf](http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf) > [consulté le 10 novembre 2011]

Traduction française de la version originale (1998) disponible : Spécifications fonctionnelles des notices bibliographiques : rapport final <<http://www.bnf.fr/pages/infopro/normes/pdf/FRBR.pdf>>  
*Fondamental : Le texte des FRBR*

[50] IFLA. **Liste électronique de discussion du groupe de travail sur les FRBR** (sur inscription) [consulté le 05 juillet 2011] <<http://infoserv.inist.fr/wwsympa.fcgi/info/frbr>>  
*Permet de suivre les discussions des spécialistes sur des points encore litigieux (frontière œuvre / expression) ; à réserver, en référence à des controverses parfois byzantines, aux FRBRidoules.*

[51] LEBCEUF Patrick. **FRBR : un p'tit coin d'paradigme...** [en ligne]. avril 2003, 22 p. [consulté le 07 novembre 2011] <[http://www.d-nb.de/standardisierung/pdf/papers\\_leboeuf\\_f.pdf](http://www.d-nb.de/standardisierung/pdf/papers_leboeuf_f.pdf)>  
*Présentation très claire et pleine d'humour par le spécialiste en la matière (en 2003, Président du Groupe de Travail de l'IFLA sur les FRBR): ce que sont et ne sont pas les FRBR, historique et réalisation, les « sujets principaux »... A destination de ceux qui souhaitent approfondir leur connaissance du modèle.*

[52] LEBCEUF Patrick. **Le jour d'après : où serez-vous ?...** [en ligne] In MEDIADIX. Journée d'étude: La fin du catalogage ?! Saint-Cloud, 21 Octobre 2004 <<http://mediadix.u-paris10.fr/archivesje/leboeufweb.pdf>>  
*Article anticipant dès 2004 les problématiques d'évolution du catalogue entraînées par le modèle FRBR: les questions essentielles sur les finalités d'un catalogue et sur notre représentation de cet outil sont amorcées.*

[53] LEBCEUF Patrick. **FRBR: Hype, or Cure-All ?** Introduction. In LE BCEUF Patrick. *Functional Requirements for Bibliographic Records (FRBR): Hype, or Cure-All?* [printed text]. Binghamton, NY: the Haworth Press, 2005. ISBN 0-7890-2799-2  
[Publié également dans la revue *Cataloging & Classification Quarterly*, 2005, Vol. 39, No. 3-4. ISSN 0163-9374]  
*Cet article introductif est essentiel car il pose le décor sur les FRBR, conceptualisation de la réalité. Il explique que les imperfections logiques dénoncées dans l'analyse du modèle ne sont que le reflet des pratiques de catalogage. Le livre entier est une somme de contributions sur les problématiques théoriques (en particulier le concept d'œuvre) et son « application » engagée sur différents documents, domaines et environnements (OCLC, Isis...).*

[54] LEBCEUF Patrick. **Le modèle FRBR : tout un nouveau monde.** IFLA. 5e Réunion d'experts sur un Code international de catalogage (IME ICC 5), 14-15 août 2007, Pretoria, Afrique du Sud.  
*Essentiel : présentation limpide du modèle FRBR<sub>ER</sub> et de ses évolutions ; illustré par des exemples « d'implémentation ».*

[55] LE PAPE Philippe. **RDA : le nouveau paradigme du catalogage (1)** [en ligne]. In ADBS. FRBR, c'est quoi déjà ? 5 à 7 de l'ADBS, 29 mars 2011. [consulté le 23 octobre 2011] <<http://www.adbs.fr/rda-le-nouveau-paradigme-du-catalogage-1-frbr-c-est-quoi-deja--102461.htm?RH=1266334869518>>  
*En moins de 30 minutes, Philippe Le Pape, chargé de la mission Normalisation à l'ABES (Agence bibliographique de l'enseignement supérieur), présente très simplement l'essentiel des FRBR qu'il situe dans la continuité d'évolutions déjà engagées à l'époque des catalogues imprimés : ainsi, le répertoire*

*Diapason (disques) s'articulait déjà autour des éléments d'identification de l'œuvre avec des 'entités' et des relations.*

[56] LERESCHE Françoise. **UNIMARC dans le contexte normatif international** [en ligne]. In ENSSIB. 3e journée UNIMARC. 1er avril 2010, Villeurbanne. [consulté le 05 octobre 2011] <<http://cfu.enssib.fr/documents/publications/124838.pdf>>

[57] MAXWELL Robert M. **FRBR: a guide for the perplexed** [printed material]. Chicago, Ill.: ALA, 2007. ISBN 978-0-8389-0950-8. [consulté le 30 octobre 2011] <<http://www.scribd.com/doc/29545162/FRBR-a-guide-for-the-perplexed>>

[58] PEONAKIS Manolis, SFAKAKIS Michalis, KAPIDAKIS Sarantos. **FRBRisation : utiliser les zones de lien UNIMARC pour identifier les Œuvres** [en ligne]. In IFLA. 77th Meeting, 13-18 August 2011, Porto-Rico. [consulté le 05 novembre 2011]

< <http://conference.ifla.org/past/ifla77/187-peonakis-fr.pdf> >

*[Résumé d'auteur] Le principal objectif de cette étude est d'associer les procédures de FRBRisation à partir de MARC21 avec la sémantique du format UNIMARC et de mettre en évidence certaines différences entre les deux formats dans le contexte de la FRBRisation. Cette étude est centrée sur l'examen de la possibilité d'utiliser les zones de lien UNIMARC pour identifier l'entité Oeuvre du modèle FRBR.*

[59] PISANSKI Jan, ZUMER Maja, TROND Aalberg. **Frbérisation : vers un nouvel avenir brillant pour les bibliographies nationales** [en ligne]. In IFLA. 75th General Conference and Assembly, 23-27 août 2009, Milan, Italie. [consulté le 27 octobre 2011]

<<http://www.ifla.org/files/hq/papers/ifla75/77-pisanski-fr.pdf>>

*[Résumé d'auteur] Cette communication présente les FRBR ainsi que les possibilités et les limites des normes actuelles et des pratiques, dans la manière d'exprimer le modèle FRBR. Nous proclamons que ce sont les bibliographies nationales qui devraient profiter le plus des FRBR, puisqu'elles contiennent une grande quantité de données bibliographiques de haute qualité pour beaucoup d'œuvres en versions multiples. La communication est fondée en partie sur les résultats de la frbérification de la Bibliographie nationale slovène et de deux autres bases de données bibliographiques.*

[60] PISANSKI Jan, ZUMER Maja, TROND Aalberg. **Identifiants: franchir les barrières linguistiques.** [en ligne] In IFLA. 76th General Conference and Assembly 10-15 August 2010, Gothenburg, Sweden. [consulté le 27 octobre 2011]

<<http://www.ifla.org/files/hq/papers/ifla76/93-pisanski-fr.pdf>>

*[Résumé d'auteur] L'identification des entités bibliographiques est un élément important permettant de franchir les barrières culturelles et linguistiques dans le monde des bibliothèques car elle autorise l'utilisation et la réutilisation des données bibliographiques dans de multiples applications. L'article liste les identifiants existants susceptibles d'être utilisés pour l'identification des entités formelles et étudie la possibilité de leur utilisation à grande échelle au sein des bibliothèques et au-delà.*

[61] RIVA Pat, BELLEMARE Nathalie. **Les FRBR pour les nuls.** Argus, automne 2009, vol.38, n°2. Issn 0315-9930

*Présentation très synthétique et en français du modèle entité relation : à destination des non initiés.*

[63] RIVA Pat. **Les modèles conceptuels FRBR et FRAD** [en ligne]. 12 juin 2008, 77 p. [consulté le 05 octobre 2011]

<<http://www.rda-jsc.org/docs/riva-20080612-fr.pdf>>

*Document de présentation très complet qui donne une vue synthétique du modèle FRBR (entité/relations/ et des modèles liés); par la présidente du « FRBR Review Group» de l'IFLA.*

[64] RIVA Pat. **Les dessous de la norme RDA : les modèles conceptuels FRBR et FRAD dans l'univers bibliographique** [en ligne].12 novembre 2009, 42 p. [consulté le 05 octobre 2011]

<[http://www.asted.org/uploadedcontent/medias/content\\_1095\\_1376.pdf](http://www.asted.org/uploadedcontent/medias/content_1095_1376.pdf)>

[65] TAYLOR Arlene G. **An Introduction to Functional Requirements for Bibliographic Records (FRBR)** p.1-19. In *Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools*. Westport, CT: Libraries Unlimited, 2007.

*Article fondamental: l'auteur, professeur à l'Ecole des Sciences de l'information de l'Université de Pittsburgh, détaille les éléments basiques du modèle théorique (détails de sa structure : entités, attributs, relations, de sa finalité), avant d'en présenter les applications pratiques (un modèle pour l'organisation des ressources du web).*

[66] TILLET Barbara. **What is FRBR? A conceptual model for the bibliographic universe.** (2004). [consulté le 15 septembre 2011]

<<http://alia.org.au/publishing/alj/54.1/full.text/tillett.html>>

*En huit pages, cette synthèse limpide présente l'essentiel du modèle FRBR. Article rédigé par la responsable de la politique de catalogage à la Library of Congress, une spécialiste des FRBR.*

[67] YEE Martha. **Can bibliographic data be put directly onto the semantic web?** *Information Technology & Libraries*, juin 2009, p.57-80. Issn 0730-9295

*Après une brève introduction au web sémantique, la discussion s'engage autour des problèmes de granularité et de structure des données, du manque de normes pour un affichage efficace et une indexation des données bibliographiques sur le web. L'article rapporte ensuite une expérience : la construction d'un modèle en RDF, qui s'avère trop peu sophistiqué pour 'traduire' des données issues du modèle FRBR.*

[68] ZHANG Yin, SALABA Athena. **FRBR: from a Conceptual Model to Application and System Development.** *Bulletin of the American Society for Information Science and Technology*, August/September 2007, volume 33, issue 6, p.17-23. ISSN: 1550-8366

*L'article passe en revue des collections ayant reçu une application FRBR et fait le point sur l'état actuel [en 2007] d'avancement et les outils utilisés pour développer un système de bases de données fondé sur les FRBR.*

[69] ZHANG Yin, SALABA Athena. **User Interface for FRBR user tasks in online catalogs.** *Proceedings of the American Society for Information Science and Technology*, 2009, Volume 46, Issue 1, pages 1–4. ISSN: 1550-8390

*Une étude est actuellement menée à partir de deux catalogues basés sur les FRBR (Fiction Finder d'OCLC et le catalogue de la Bibliothèque d'Australie) pour évaluer leur adéquation avec les tâches de l'utilisateur définies par le modèle. Les résultats de cette étude permettront d'identifier les options les*

*plus adaptées pour une implémentation FRBR. Ce court article présente le contexte de l'étude, le choix et les méthodes retenus et quelques bribes de résultats.*

[70] ŽUMER, M., PISANSKI Jan, VILAR Polona, HAREJ Viktor, MERCUN Tanja, ŠVAB Katarina. **Abattre les barrières entre des pratiques anciennes et de nouvelles demandes : le prix de l'hésitation.** In IFLA 77ème meeting, Porto-Rico. 13-18 août 2011. [consulté le 26 octobre 2011] < <http://conference.ifla.org/past/ifla77/80-zumer-fr.pdf> >

*Menée par Maja Žumer, l'école slovène des Etudes sur l'information et les bibliothèques est très active dans le domaine des FRBR. Cette intervention au Congrès de l'IFLA présente une évaluation du travail déjà accompli en matière de catalogage et soulève des questions clés dans l'objectif du catalogage : développement du modèle, FRBRisation, nécessité de systèmes d'identification, interfaces d'affichage. En filigrane, les pratiques conservatrices sont fustigées et des expérimentations dans les catalogues par la communauté des bibliothèques recommandées de toute urgence.*

## 6 Application des FRBR au cadre d'une organisation internationale

[71] BIBLIOTHEQUE NATIONALE DE FRANCE (BNF). **Modèle CIDOC CRM. Conceptual Reference Model / Modèle conceptuel de référence.** 6 avril 2011. [consulté le 3 novembre 2011] < [http://www.bnf.fr/fr/professionnels/modelisation\\_ontologies/a.modele\\_cidoc\\_crm.html](http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.modele_cidoc_crm.html) >

*Description sommaire du contexte d'élaboration du modèle.*

[72] BOUNTOURI Lina, PAPTAEODOROU Christos, GERGATSOULIS Manolis. **Modelling the Public Sector Information through CIDOC Conceptual Reference Model.** On the Move to Meaningful Internet Systems: OTM 2010 Workshops. [consulté le 5 novembre 2011] < [www.cidoc-crm.org/docs/64280404.pdf](http://www.cidoc-crm.org/docs/64280404.pdf) >

*La modélisation de l'information hétérogène du secteur public: les problématiques liées à l'interopérabilité, la nécessité d'une ontologie pour représenter des enregistrements de nature différente (production, management) et les raisons du choix du modèle CIDOC CRM. Avec une figure de correspondance entre entités CRM et activités du PSI.*

[73] CIDOC. Conceptual reference model. 21 mars 2011. [consulté le 10 novembre 2011] < <http://www.cidoc-crm.org/tutorials.html> >

à compléter avec la version vidéo < [http://www.cidoc-crm.org/cidoc\\_tutorial/index.html](http://www.cidoc-crm.org/cidoc_tutorial/index.html) >

[74] DEBACKERE Marie-Claire. **Problèmes rencontrés pour obtenir la littérature grise.** 60th IFLA General Conference. August 21-27, 1994. 28 décembre 2010. [consulté le 13 octobre 2011] < <http://archive.ifla.org/IV/ifla60/60-debm.htm> >

*Malgré son ancienneté, cet article nous a permis de comprendre les spécificités du fonds de l'IPIE (principalement en termes de spécificités de publication, de collecte d'information et de volume) puis de prendre connaissance des initiatives développées pour en faciliter l'accès grâce aux nouvelles technologies (bases de données, réseaux).*

[75] International Working Group on FRBR and CIDOC CRM Harmonisation. **FRBR object-oriented definition and mapping to FRBRER (version 1.0.1).** 2010.

<[www.cidoc-crm.org/docs/frbr\\_oo/frbr\\_docs/FRBRoo\\_V1.0.1.pdf](http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V1.0.1.pdf)>

*Le texte du modèle orienté objet*

[76] LEBŒUF Patrick. **Modélisation conceptuelle de l'information bibliographique et muséologique: CIDOC CRM et FRBRoo**. Documentation et bibliothèques, 2009, vol. 55, n° 4, p.209-214.

*Rappel du contexte (FRBR version d'origine) puis présentation du modèle conceptuel établi pour les musées, du travail mis en œuvre pour aboutir à une ontologie commune, de formalisme orienté objet compatible avec le langage RDF, dans la perspective du développement du Web sémantique et interopérable avec d'autres communautés.*

[77] RIVA Pat, DOERR Martin, ZUMER Maja. **FRBRoo: offrir une vision commune de l'information produite par des institutions de mémoire**. In IFLA. 156<sup>e</sup> cataloguing meeting, 74<sup>th</sup> IFLA general conference and council 10-14 août 2008, Québec, Canada. [consulté le 15 octobre 2011]

<[http://archive.ifla.org/IV/ifla74/papers/156-Riva\\_Doerr\\_Zumer-trans-fr.pdf](http://archive.ifla.org/IV/ifla74/papers/156-Riva_Doerr_Zumer-trans-fr.pdf)>

*Présentation du contexte puis du modèle FRBRoo, des objectifs et des avantages de ce formalisme. Cet article, disponible en version française, est essentiel car il fait le pont entre le modèle entité relation et orienté objet.*

## 7 Les préconisations dans le contexte de l'IIFE

[78] ASSOCIATION DES ARCHIVISTES FRANÇAIS (AAF), ASSOCIATION DES PROFESSIONNELS DE L'INFORMATION ET DE LA DOCUMENTATION (ADBS). **Analyse de la norme ISO 15489 au regard des pratiques archivistiques françaises** par le Groupe métiers AAF-ADBS « Records management» [consulté le 3 novembre 2011]

<<http://www.adbs.fr/groupe-de-travail-aaf-adbs-records-management-29392.htm?RH=1202925005805>>

[79] Compte-rendu du Comité stratégique de RDA. 12 mai 2011, 6p. In RDA en France. [consulté le 27 octobre 2011]

<http://rda-en-france.enssib.fr/sites/rda-en-france.enssib.fr/files/CR%20RDA%20Comit%C3%A9%20strat%C3%A9gique%2025%20mars%202011.pdf>

[80] Livre blanc. **Introduction à la série de normes ISO 30300**, Système de gestion des documents d'activité Intégration du records management et perspectives d'évolution de l'ISO 15489 CN11 – Octobre 2011, 42 p.

[Résumé d'auteur] *Ce Livre blanc a pour objectif d'expliquer les perspectives d'évolution des normes sur le records management dans les prochaines années. La politique et les moyens mis en œuvre par chaque organisation doivent donc s'y adapter et, inversement, l'amélioration et la normalisation de la gestion des documents d'activité concourront à produire la valeur et la performance dans un cadre de développement durable pour chaque organisation.*



[81] BARCA Murtha. **Réflexions sur la documentation du patrimoine culturel et sa diffusion au 21<sup>ème</sup> siècle : utilisation de données et du Web pour mettre en lien usagers et collections.** CIDOC NEWSLETTER, n°1, 2010, p.3-7. [consulté le 7 novembre 2011]

<[http://cidoc.mediahost.org/newsletter\\_01\\_2010.pdf](http://cidoc.mediahost.org/newsletter_01_2010.pdf)>

*Réflexion sur l'accès des collections des catalogues sur le web, la mise au point d'outils multilingues. L'article ouvre surtout un débat sur le positionnement des institutions culturelles face aux moteurs de recherche commerciaux.*

[82] DARDENNE Nadine. **Isidore.** TGE Adonis 28 novembre 2010, modifié le 4 juillet 2011. [consulté le 7 octobre 2011]

<<http://www.tge-adonis.fr/service/isidore>>

*Présentation de la plateforme de recherche Isidore dont l'objectif est de fédérer l'accès aux ressources numériques issues de la recherche en SHS.*

[83] Food and Agriculture Organization of the United Nations (FAO). **Implementing the FAO Open Archive based on Fedora Commons and FRBR**

<<http://aims.fao.org/fr/news/implementing-fao-open-archive-based-fedora-commons-and-frbr>>

[84] GENOUD Jean-Claude, CONTESSE Eloi, CROFTS Nicholas. **Les bases de données de nouvelle génération : outils de communication, d'échange et lieu privilégié de la transversalité culturelle et patrimoniale.** (p.7-13) In Association des musées suisses, musée historique de Lausanne. Systèmes d'information et synergies entre musées, archives, bibliothèques, universités, radios et télévisions : les bases de données et les médias numériques au service des patrimoines historique, culturel, naturel et scientifique. Actes de colloque, Lausanne, 22-23 mars 2007. [consulté le 17 octobre 2011]

<[http://www.lausanne-](http://www.lausanne-tourisme.ch/Tools/GetLinkedDoc.asp?File=11634.pdf&Title=Actes+du+colloque+de+mars+2007)

[tourisme.ch/Tools/GetLinkedDoc.asp?File=11634.pdf&Title=Actes+du+colloque+de+mars+2007](http://www.lausanne-tourisme.ch/Tools/GetLinkedDoc.asp?File=11634.pdf&Title=Actes+du+colloque+de+mars+2007)>

*L'objectif du colloque est de réunir les partenaires de la culture et du patrimoine - conservateurs, archivistes, bibliothécaires - et des techniciens des systèmes informatiques pour échanger les connaissances et créer des systèmes ouverts afin de valoriser des collections.*

[85] GRADMANN, Stefan. **Visualisation, interactivité et technologies sémantiques : les nouveaux enjeux de la communication scientifique.** Conférence i-expo 2010 (10 juin 2010)

*Stefan Gradmann est spécialiste de la conceptualisation sémantique, professeur en bibliothéconomie et sciences de l'information à l'Université Humboldt de Berlin. Impliqué dans le projet Europeana, il explique à travers l'exemple de la plate forme expérimentale Thoughtlab les enjeux du web sémantique.*

[86] Journée d'étude. « Référentiels et données d'autorité à l'heure du web sémantique » : Paris, 27 mai 2011 [on line]. Paris : AFNOR/BnF. [consulté le 13 octobre 2011]

<[http://www.bnf.fr/fr/professionnels/autres\\_journees\\_professionnelles/a.referentiel\\_donnees\\_autorites\\_110527.html](http://www.bnf.fr/fr/professionnels/autres_journees_professionnelles/a.referentiel_donnees_autorites_110527.html)>

*Présentation de la normalisation, des réalisations et des usages se rapportant aux systèmes d'organisation de connaissances et à l'ouverture des données.*

[87] LERESCHE Françoise, BOURDON Françoise. **RDA in Europe**: report of the work in progress in France proposal for an EURIG technical meeting in Paris. BNF / AFNOR. Août 2011. [consulté le 27 octobre 2011]

<[http://www.sainte.org.uk/eurig/docs/BnF-ADM-2011-066286-01\\_%28p2%29.pdf](http://www.sainte.org.uk/eurig/docs/BnF-ADM-2011-066286-01_%28p2%29.pdf)>

*Présentation au groupe européen d'intérêt autour de RDA (EURIG) de la position de la France relativement à une adoption du RDA: les pistes de travail établies par le comité technique et les cinq options possibles ; les orientations retenues par le groupe stratégique.*

[88] MAISONNEUVE, Marc. **Bibliothèques : bientôt une nouvelle offre**. Livres Hebdo n° 859. 1er avril 2011. [consulté le 18 octobre 2011]

<<http://www.toscaconsultants.fr/articles/enquetetoscaconsultantslivreshebdo2011.pdf>>

*Résultats de l'enquête annuelle sur l'évolution de l'offre des logiciels conduite auprès des fournisseurs de logiciels métier pour les bibliothèques. Pour l'enquête sur l'évolution du marché, se reporter au numéro d'avril 2011 d'Archimag.*

8- Pour aller plus loin : pistes d'approfondissement et sites recommandés pour une veille sur les modèles FRBR et les problématiques du web sémantique

Scoop it. **FRBR, RDA, etc.** [en ligne]

<<http://www.scoop.it/t/frbr-rda-etc>>

*A partir d'une interface unique, ce site donne accès aux principales ressources sur les FRBR (vidéos, présentations powerpoint, blog, compte-rendu de l'IFLA...).*

BIBLIOTHÈQUES ET ARCHIVES CANADA. **FRBR et RDA** : progrès dans la description des ressources de divers formats [en ligne]. Ottawa. Avril 2010.

<<http://www.collectionscanada.gc.ca/isbe/005002-2203-f.html>>

*L'article retrace l'origine et les impacts du modèle et donne un aperçu de modèle entité relation avec des exemples. Le site collections Canada est une mine d'informations de qualité, actualisées et en français sur le catalogage.*

ERTZSCHEID Olivier. **Affordance**. <<http://affordance.typepad.com/>>

LEPAPE Philippe. **RDA@ABES**: tout ce que nous pouvons vous dire sur RDA (Ressources : description & accès).

<<http://rda.abes.fr/>>

*Mis en ligne début octobre 2011 et animé par Philippe Lepape de l'ABES, ce blog au ton humoristique permet de s'informer agréablement sur le code de catalogage RDA et la modélisation des données bibliographiques.*

MERCIER Silvère. **Bibliobsession**.

<<http://www.bibliobsession.net>>

*Le blog d'un «bibliothécaire bibliobsédé des bibliothèques».*

OGAWA Yôko. Le Musée du silence. Arles : Actes Sud, 2002. 315 p. ISBN: 978-2742754915.  
*Une mission et des objets documentaires très particuliers qu'un jeune muséographe devra collecter, identifier et mettre en scène.*

POUPEAU Gautier. **Les petites cases du savoir.**

<<http://www.lespetitescases.net/>>

*Blog plutôt axé sur les technologies du web sémantique.*

SALAUN Jean-Michel. Bloc-notes de Jean-Michel Salaün: «repérage de données sur l'économie des documents dans un environnement numérique».

<<http://blogues.ebsi.umontreal.ca/jms/>>

*"Blogue" du directeur de l'EBSI (École de bibliothéconomie et des sciences de l'information) de Montréal.*

VALDERRAMA Asunción, BOULOGNE Arlette. **Métiers et compétences.** Documentaliste-Sciences de l'Information 2009/3 (Vol. 46).

WENZ Romain. **L'avenir des catalogues** : formats, données, outils, usages. Mémoire d'étude/ janvier 2009. ENSSIB. [consulté le 5 septembre 2011]

< [www.enssib.fr/bibliotheque-numerique/document-21205](http://www.enssib.fr/bibliotheque-numerique/document-21205) >

*Mémoire d'études du promoteur du projet data.bnf*

Pour approfondir les connaissances sur les FRBR, une bibliographie très complète a été mise en ligne en 2006 sur le site de l'IFLA à l'adresse [http://archive.ifla.org/VII/s13/wgfrbr/FRBR\\_bibliography.pdf](http://archive.ifla.org/VII/s13/wgfrbr/FRBR_bibliography.pdf). Mais elle n'est pas à jour, en particulier pour les références de la formulation objet des FRBR.

# Annexes

## Annexe 1 FRBR QUIZ

Pour tester vos connaissances FRBR, le FRBR quiz.

### A- Les FRBR sont :

- 1- Une représentation simplifiée de l'univers bibliographique;
- 2- Un ensemble de règles de catalogage ;
- 3- Un outil pour comprendre la structure d'une notice.

### B- L'utilité des FRBR pour le "créateur" de métadonnées est :

- 1- Une simplification du travail de catalogage ;
- 2- Une réflexion sur des pratiques de saisie des données ;
- 3- Une complexification du catalogage car il faut éclater les données entre différentes notices.

### C- L'utilité des FRBR pour l'utilisateur est :

- 1- Une meilleure colocation des ressources ;
- 2- De permettre aux usagers de retrouver facilement l'information ;
- 3- Impossible à évaluer car il n'y a pas de catalogue FRBRisé.

### D- Implémenter le modèle FRBR revient à :

- 1- Utiliser les formats MARC ;
- 2- Utiliser le code de catalogage RDA ;
- 3- Aucun choix n'est recommandé.

### E- FRBR analyse la notice catalographique comme un ensemble d'informations portant sur quatre aspects distincts du document décrit. Dans l'exemple ci-dessous, laquelle des affirmations est fautive ?

- 1- L'idée du mémoire sur les FRBR trouve sa réalisation dans une *expression*, le texte original écrit en français qui se concrétise dans une *manifestation*, sa version numérisée publiée en ligne en 2011. Le document est représenté par l'*item* conservé au Centre de ressources documentaires de l'INTD CNAM ;
- 2- Ce mémoire est, au sens FRBR, une *œuvre* dont la version papier remise en novembre 2011 est une *expression*. La version MP3 serait une autre *expression* de cette même *œuvre*. La publication de ce mémoire est une *manifestation* ainsi que sa traduction en japonais ;
- 3- Vous êtes en train de lire un *item* d'une version numérisée qui est une *manifestation* de la 3<sup>ème</sup> version (version corrigée), c'est-à-dire l'*expression* d'une œuvre.

**F- Le nouveau code de catalogage RDA publié en 2010 est :**

- 1- Un code qui s'appuie sur l'ISBD consolidé ;
- 2- Disponible sous la forme d'un site, signe de la volonté d'inscrire le catalogue dans le monde numérique ;
- 3- Une application des FRBR toujours en discussion en France.

**G- Le modèle FRBR orienté objet est :**

- 1- Un modèle issu des FRBR, développé pour rendre interopérables les systèmes d'information des institutions culturelles (bibliothèques mais aussi musées, archives...) ;
- 2- Une révision du modèle FRBR entité relation qui s'appuie sur la démarche du CRM pour intégrer les notions de temps et d'évènement ;
- 3- Un modèle de remplacement récemment mis au point (2008) pour se substituer au modèle entité relation.

## FRBR QUIZZ ... les réponses

A- Les FRBR sont ... → Réponses 1 et 3

B- L'utilité des FRBR pour le créateur de métadonnées ... → Réponses 2 et 3

C- L'utilité des FRBR pour l'utilisateur ... → Réponses 2 et 3

D- Implémenter le modèle FRBR ... → Réponses 2 et 3 : aucun choix n'est recommandé mais une des options possibles est le catalogage avec RDA Toolkit.

E- Un exemple de notice ... → La publication relève bien de l'entité *manifestation* mais pas la traduction qui est une *expression* d'une *œuvre* : la réponse 2 est donc fautive.

F- Le nouveau code de catalogage RDA ... → Réponses 2 et 3

G- Le modèle FRBR orienté objet ... → Réponses 1 et 2

« Il faut envisager FRBRoo comme une interprétation des FRBR ; ce n'est donc ni une nouvelle version ni un modèle de remplacement. La communauté des bibliothèques se référera encore aux FRBR(ER) qu'elle continuera à utiliser pour expliquer l'univers bibliographique. D'un autre côté, FRBRoo facilite l'implémentation au moyen des outils orientés objet, particulièrement pour l'intégration de l'information hétérogène en provenance du secteur du patrimoine culturel. Il permet aussi l'interopérabilité avec d'autres ontologies fondamentales. » Source [77]

## ANNEXE 2: Hypothèse de travail : essai de FRBRisation des bases internes à l'aide du modèle conceptuel FRBR<sub>ER</sub>

L'objectif est de montrer l'utilité du modèle relationnel dans un exemple d'œuvre complexe liant les trois bases. L'*œuvre* est ici le projet du Plan à Moyen Terme ***P8-004-Access*** générant des activités (dans notre exemple l'équipe TASP-000201). Ces activités produisent à leur tour des œuvres (séminaires, documents de travail...). Seules les entités du groupe 1 ont été mises en œuvre dans cette hypothèse.

Cet essai a permis de mettre en évidence l'incapacité de faire entrer des documents complexes dans un modèle prévu pour un cadre bibliographique. C'est pourquoi la suite de l'expérience a mis en œuvre les modèles conceptuels orienté objet.

Oe1 = Projet = ***P8-004-Access: Strategies to improve access and retention of poor and vulnerable children in primary and secondary education, especially in fragile states***

Oe1.1 = Activité = TASP-000201

Oe 1.1.1 = IIEP/S.289 International Seminar on Opportunities for Change within Education Systems in Conflict and Post-Conflict Situations

→ expression 1 le texte original en anglais

→ expression 2 le texte traduit en français

→ expression 3 la version augmentée issue du séminaire

Oe 1.1.1.1- International Seminar on Opportunities for Change within Education Systems in Conflict and Post-Conflict Situations. Paris, 28-30 January 2008 /la contribution de Talbot, Christopher

Oe 1.1.1.2 S.289 / Negotiating system change: beyond opportunities and constraints: la contribution de Buckland, Peter

Oe 1.1.1.3 S.289 / Political violence against teachers in Colombia: opportunities for change in the midst of conflict? La contribution de Novelli, Mario

Oe 1.1.3 Publication issue du Séminaire 289

Expression : le texte de la publication

Manifestation 1 Nicolai, Susan / Opportunities for education sector change in conflict and post-conflict situations IIEP/ UNESCO, 2009, 262 pages, ISBN: 978-92-803-1328-4,



- Document sur les rayons
- Document en consultation
- Manifestation2: ressource électronique
- Document [http://www.iiep.unesco.org/fileadmin/user\\_upload/Info\\_Services\\_Publications/pdf/2009/Nicolai\\_Opportunities\\_for\\_change.pdf](http://www.iiep.unesco.org/fileadmin/user_upload/Info_Services_Publications/pdf/2009/Nicolai_Opportunities_for_change.pdf)

Oe 1.1.1.2.3 S.289 / Bijlsma, Sanne/ Teaching history in political environments: looking for unity in Rwandan classrooms

- Manifestation: version papier, chapitre 16, 218-229.

Oe 1.1.1.2.4 S.289 / Kirk, Jackie / Community-based education in Afghanistan: increasing access, supporting the system and improving quality?

- Manifestation: version papier Moving from innovation to policy: IRC's work with community-based education in Afghanistan, chapitre 7, p.103-114

Oe 1.1.1.2.5 S.289 / Boersema, Jacob R. / Resisting change: former Afrikaner schools and educational reform in post-apartheid South Africa

- manifestation: version papier chapitre 12, p.163-176.

Oe 1.1.1.2.6 S.289 / Heath, Liz / Power structures, politics and change in Kosovo's higher education system

Oe 1.1.1.2.7 S.289 / May, Anika / Lopes Cardozo, Mieke /Teaching for peace: overcoming division? the potential contribution of peace education in reconciliation processes in the cases of Sri Lanka and Uganda

Oe 1.1.1.2.8 S.289 / Pupovci, Dukagjin/ Teacher training in Kosovo

- Manifestation 1 : ressource électronique, p
- Manifestation 2 : version papier, chapitre 13, p.177-188.

À relier à ses travaux antérieurs (sur la base INTEPI)

- Expression1: Education in Kosova 2000/2001
- Manifestation : Prishtina, KEC, 2001
- Document : IIEP KOS 48
- Expression 2 : Arsimi në Kosovë 2000/2001...

Oe1 = Projet = P8-004-Access peut être reliée au P8-007 OBS-emergencies Research Partnership on Education in Conflict, Emergencies and Reconstruction.

### ANNEXE 3: L'approche cubiste



Si la notice FRBRisée était un chat ... l'éclairage d'Antonín Procházka.