

# L'indexation matière en transition

---

*De la réforme de Rameau à l'automatisation*

## Introduction

*Étienne Cavalié. Bibliothèque nationale de France.*

Ceci n'est pas un manuel d'indexation à l'usage des catalogueurs. Il ne convient pas d'en écrire un en ce moment, car son contenu risquerait d'être périmé très rapidement.

Les transformations en cours sur les questions d'indexation invitent plutôt à nous intéresser justement à cette actualité, afin de donner à voir un processus en train de se réaliser, afin de dresser un panorama plus complet que les diverses présentations publiées ici ou là.

Commençons par esquisser les contours de cette actualité. Suite à la publication en 2010 d'un nouveau code de catalogage, RDA (*Resource Description and Access*) conçu pour se substituer aux AACR, la France a pris le temps de la réflexion sur son éventuelle adoption. Le choix de n'adopter en l'état ni RDA ni, quelques années plus tard, Bibframe, format de catalogage conçu en 2013 par la Bibliothèque du Congrès, s'est traduit par le projet de concevoir une *transposition* de RDA, et de refuser, selon l'expression devenue fameuse de Philippe Le Pape, l'idée d'un « grand soir catalographique »<sup>1</sup> au profit de nombreux petits matins. À partir de 2015, la France s'engage dans la Transition bibliographique<sup>2</sup>, à travers notamment ses deux agences bibliographiques, la Bibliothèque nationale de France (BnF) et l'Agence bibliographique de l'Enseignement supérieur (Abes).

Ce projet de Transition bibliographique vise à concevoir un code de catalogage (nommée ensuite RDA-FR) qui ne constitue pas de régression par rapport aux règles déjà en vigueur dans l'Hexagone, et qui permette d'implémenter le modèle FRBR, conçu par un groupe de travail de l'IFLA dans les années 1990 et enrichi ensuite des modèles FRAD et FRSAD<sup>3</sup>. L'objectif de cette introduction n'est pas de rappeler ce que contiennent ces modèles de structuration de l'information bibliographique, mais un chapitre y pourvoit et présente le modèle qui a repris, harmonisé et supplanté ces trois-là, à savoir le modèle LRM (*Library Reference Model*), publié en 2017 et diffusé depuis 2018 sous l'appellation de modèle IFLA-LRM.

La Transition bibliographique a aussi pour ambition, outre d'être une source de consignes de catalogage pour le traitement courant, de reprendre les données bibliographiques existantes

---

<sup>1</sup> Philippe Le Pape, *Il n'y aura pas de grand soir catalographique : aujourd'hui, c'est déjà demain*, Support de présentation présenté à Journée ADBU, Montpellier, 13 février 2015, [En ligne] <[http://adbu.fr/wp-content/uploads/2015/02/Ph\\_Le\\_Pape\\_ADBU\\_ABES\\_040215.pdf](http://adbu.fr/wp-content/uploads/2015/02/Ph_Le_Pape_ADBU_ABES_040215.pdf)> (Consulté le 21/02/2019).

<sup>2</sup> Le site <<https://www.transition-bibliographique.fr>> lui est consacré.

<sup>3</sup> *Cataloguer aujourd'hui : Identifier les œuvres, les expressions, les personnes selon RDA-FR*, Claire Toussaint dir., Paris, Éditions du Cercle de la Librairie, 2018 (Bibliothèques).

pour les mettre au niveau du code. Étant donné l'écart entre les données existantes et la cible (encore incomplètement définie), cela se traduit par une succession de jalons à poser progressivement.

Une des ambitions du modèle LRM est de faciliter la diffusion des métadonnées des bibliothèques dans le web, en adoptant une manière de concevoir l'information qui soit plus conforme aux pratiques des autres secteurs d'activité. Le format MARC est trop spécifique aux bibliothèques, et sa manière de structurer l'information en zones, sous-zones, indicateurs, positions fixes, doit être abandonnée si l'on souhaite pouvoir fournir des données que d'autres corps de métier puisse réutiliser sans passer par une formation préalable de plusieurs mois.

Or le vocabulaire d'indexation matière le plus répandu en France, le langage Rameau, maintenu par le Centre national Rameau localisé à la BnF, connaît des problèmes comparables. Certes, depuis plusieurs années, Rameau est diffusé « dans le web de données » (via le site [data.bnf.fr](http://data.bnf.fr)) suivant l'ontologie SKOS suffisamment répandue. Mais sa structuration, le principe de vedettes pré-construites, les conditions d'utilisation (certains concepts ne peuvent s'appliquer qu'aux guerres, ou aux maladies), les règles de construction des chaînes d'indexation – bref, le manuel d'utilisation de Rameau de plusieurs centaines de pages, est un obstacle à une plus large adoption. Pourtant sa vocation à proposer un vocabulaire encyclopédique contrôlé pourrait satisfaire bon nombre de besoins, dans des domaines parfois très éloignés des bibliothèques ; et les technologies du web de données faciliteraient une utilisation complémentaire en ayant recours à des vocabulaires plus spécialisés.

Mais la syntaxe Rameau est trop dépendante de strates historiques, ou de règles d'usages par thématiques. Elle n'est pas assez prédictive. Une ambition de diffusion plus large de ce vocabulaire doit donc passer préalablement par une réforme pour harmoniser les règles qui l'organisent, et en réduire considérablement le nombre. C'est ainsi que s'engage en 2016 une réflexion sur la réforme de Rameau, dont les premiers résultats commencent à être sensibles depuis 2018, et plus encore en 2019.

La motivation première de cet ouvrage, ce fut donc la réforme du vocabulaire Rameau, très largement utilisé dans les bibliothèques françaises pour l'indexation signalétique. Un des objectifs, essentiel, de ce livre est donc de donner une meilleure lisibilité à cette réforme en cours de réalisation. Mais cette lisibilité se conçoit dans un contexte plus général, car les initiateurs de cette réforme, qui en ont défini la trajectoire, réfléchissent dans un environnement conceptuel et technologique en mutation. Et c'est ainsi que, de fil en aiguille, a été définie la structuration de cet ouvrage, des conditions d'indexation aux conditions d'utilisation, de manipulation, de cette indexation.

La première partie présente le contexte en mutation de l'indexation matière aujourd'hui. La réflexion menée par Bruno Bachimont sur les typologies d'accès aux éléments de sens dans un document, et ses relations avec l'indexation, pourraient servir à profit de grille de lecture à l'ensemble des chapitres suivants. Françoise Leresche présente ensuite la manière dont le nouveau modèle « IFLA LRM » (*Library Reference Model*), qui se place dans la prolongation du modèle FRBR (*Functional Requirements for Bibliographic Records*), traite de la question de l'indexation, en la plaçant dans la cohérence d'ensemble de ce modèle de description de ressources bibliographiques.

Deux chapitres présentent la réforme Rameau, sa méthodologie de déploiement, par Florence Ménard, et ses objectifs, par Ewa Nieszkowska-Serlan. La Classification décimale Dewey est évoquée (Jean Maury), car si elle n'a pas une actualité aussi brûlante pour les bibliothèques françaises, son propre cheminement et ses potentialités, ainsi que sa forte utilisation dans les bibliothèques françaises, méritent que l'on s'y attarde.

Le présent ouvrage cherche à déterminer à quoi (et à quelles conditions) peut servir l'indexation matière. Les deux parties suivantes explorent les enjeux de l'indexation matière, autour de deux axes. Dans le déploiement de services et d'interfaces qui exploitent cette information, il ne s'agit pas tant de décrire en quelques termes bien sentis la quintessence absolue d'un document (fantasme absolu du bibliothécaire : résumer l'ouvrage dans son indexation), mais de positionner, d'inscrire un document au sein d'autres documents, pour faire exister la collection ; ou de l'inscrire dans un contexte de services. Ces deux aspects, collections et services, ne sont évidemment pas antinomiques, mais nous permettent de structurer les chapitres suivants selon ces deux axes.

Sur l'indexation comme moyen d'inscrire un document au sein d'une collection, François Pichenot confronte les principes et la réalité, en posant une question essentielle : que contiennent réellement les données d'indexation d'un catalogue de bibliothèque ? Quelle proportion de la collection est indexée ? Avec quel(s) vocabulaire(s) ? Il déploie et documente ainsi à l'usage du lecteur une méthodologie d'analyse de ces données, qui posent les questions de base nécessaires pour réfléchir ensuite à une offre de services.

Les deux chapitres suivants viennent rendre compte de réalisations ou d'expérimentations à la Bibliothèque nationale française : sur les métadonnées du catalogue d'abord ; sur les collections numérisées ensuite, par Jean-Philippe Moreux. Dans les deux cas, on en appelle aux possibilités de l'automatisation pour essayer de réduire la part invisible des ressources, celles qui ne sont pas indexées et qui se comptent par milliers, dizaines de milliers, centaines de milliers.

Le chapitre de David Guillemin sur l'offre de formation en France vient faire le point sur la manière dont l'indexation est enseignée au sein des formations universitaires diplômantes. Car les pratiques professionnelles, la manière dont on peut concevoir une offre de services, dépend en bonne part de la formation reçue, et de ce qu'on a entendu et appris des objectifs d'une « bonne » indexation.

La dernière partie enfin entrecroise plusieurs enjeux sur l'indexation vis-à-vis de ses utilisateurs finaux : la question de l'indexation collaborative, par Hélène Cavalié ; une proposition d'interface s'efforçant de rendre manipulable la richesse sémantique du référentiel Rameau, dans le catalogue de la bibliothèque municipale de Fresnes, par Pierre Bournerie ; une réflexion conduite par Bernard Strainchamps sur indexation et médiation, inspirée notamment par son expérience acquise dans la mise en place de la plateforme bibliosurf.com ; une interface destinée aux gestionnaires de collection (qui sont eux aussi des utilisateurs légitimes de leur propre catalogue) adossée à la classification Dewey pour faciliter la mise en œuvre de la politique documentaire, par Géraldine Geoffroy ; la contribution de Raphaëlle Lapôtre enfin, qui pose un certain nombre de questions sur la visualisation des données d'indexation, et propose des pistes de réflexions pour dépasser certaines impasses.

Les chapitres relèvent de plusieurs types, allant de la contribution intellectuelle au retour d'expérience, en passant par le panorama de l'existant. À travers cette diversité des interventions, l'ensemble de l'ouvrage s'efforce de répondre à un certain nombre de questions, ou au moins de proposer des éléments de réflexion à ses lecteurs.

En amont d'une question qui serait « à quoi sert l'indexation matière ? », une autre est posée : une bibliothèque peut-elle encore avoir la maîtrise de son indexation matière ? À l'heure où tout est mis en œuvre pour diminuer la charge de travail représentée par le catalogage local, tant au niveau des agences bibliographiques (qui diffusent leurs métadonnées), des fournisseurs de notices (qui vendent les leurs) ou des établissements eux-mêmes (qui veulent consacrer une plus grande part de leurs forces vives à la médiation, notamment), est-il encore possible d'avoir une politique autour de l'indexation matière ? Si une bibliothèque peut, pour chaque document, récupérer des notices quelque part, comment peut-elle maîtriser le contenu de ces notices ? Et si elle ne le peut, comment pourra-t-elle construire une offre de services adossée à un contenu qu'elle ne maîtrise pas ?

Il y a plusieurs réponses à apporter. Naturellement, une bibliothèque peut choisir son puits de données. Certaines bibliothèques préfèrent se passer des notices de la Bibliographie nationale française parce qu'elles n'ont pas l'usage d'une indexation aussi complète, et que des notices plus « simples » leur conviennent tout aussi bien, surtout si elles peuvent les obtenir dans des délais plus courts. L'analyse détaillée du catalogue de Roubaix permet de s'arrêter aussi sur deux éléments importants. D'abord, il faut connaître non seulement ses collections, mais aussi le contenu de son catalogue pour construire des modalités pertinentes de recherche et de navigation. Il ressort en effet qu'il y a plusieurs types de mots-clés dans un catalogue donné : ceux qui viennent combler un manque (pour que dans une bibliothèque donnée, une recherche sujet sur « Astéroïdes » ou sur « Guerre de Sept Ans » fournisse toujours au moins un livre) ; ceux qui facilitent les rebonds dans le catalogue ; ceux qui permettent de constituer des corpus. Un deuxième élément essentiel qui ressort de ce chapitre, c'est qu'il faut avoir des compétences en interne pour réaliser cette analyse. Certes, les tâches de catalogage en tant qu'activité de description documentaire nécessitant une connaissance des règles de catalogage, du format de catalogage, des règles d'indexation (et de la prodigieuse syntaxe Rameau), décroissent dans les bibliothèques. Mais cela ne signifie pas qu'on doive se passer de compétences sur les métadonnées : compétences de traitement et d'analyse, pour en extraire des informations, et pour concevoir des services.

Une autre manière de répondre à la question de l'impossible politique d'indexation porte sur une meilleure maîtrise et compréhension des outils disponibles pour enrichir les métadonnées déjà existantes. Les chapitres de Jean-Philippe Moreux, de Bernard Strainchamps, ainsi que celui d'Hélène Cavalié et ma propre contribution, en évoquent plusieurs. La plupart relèvent en partie d'expérimentations, ou de projets qui ne sont en l'état pas réutilisables clé en main par tout établissement. Cela ne signifie pas que ce soit impossible, et ces chapitres permettent de rendre compte de ce qui se prépare, pour inviter à s'y préparer.

Il convient de terminer sur une précision importante : quel est le sens du terme « indexation » dans ces pages ?

Ce livre n'a pas vocation à construire une théorie nouvelle de l'indexation, il se veut pour l'essentiel plus pragmatique. Rappelons aussi, comme cela est dit plus haut, qu'il a été suscité notamment par l'actualité autour de la réforme Rameau et de l'implémentation du modèle LRM dans le cadre de la Transition bibliographique en France. On parlera donc essentiellement de l'indexation matière comme pratique professionnelle classique consistant à ajouter à la description matérielle (transcription) d'un document un certain nombre de termes visant à décrire son ou ses sujets, généralement en s'adossant à des référentiels comme Rameau et Dewey.

Néanmoins les différentes contributions peuvent évoquer avec des indexations autres que « matière », qui viennent apporter des précisions sur une ressource susceptibles tout autant de servir de point d'accès à cette ressource (ce qui est le rôle d'un index). L'indexation en texte intégral peut également être convoquée.

La porosité de ces différentes acceptions de « l'indexation » se révèle en effet fortement au fil de ces pages, du point de vue des usages comme de celui des outils. Initialement, l'ajout de mots-clés dans les notices (décrivant des livres imprimés, par exemple) venait compenser le fait qu'il n'était pas possible de mettre le texte intégral dans le catalogue. Or il apparaît manifeste que l'accès du texte intégral (ou au moins le résumé, la quatrième de couverture) au sein de l'outil numérique vient aider l'indexation, que celle-ci soit réalisée humainement ou informatiquement.

En se concentrant sur l'indexation matière, on aurait couru le risque de conduire une réflexion trop étroite, en ignorant les interfaces (où les facettes Sujet, Domaine ou Thème côtoient les dates et les types de documents) et les données (où les référentiels pour alimenter les notices coexistent, indépendamment de leur nature).

L'indexation matière demeure donc bien le cœur de l'ouvrage, mais d'autres modalités d'indexation sont mobilisées quand cela a paru pertinent pour alimenter la réflexion générale.

Le présent ouvrage est destiné tout autant aux gestionnaires de collections, catalogueurs, administrateurs de logiciels de bibliothèques, qu'aux directeurs d'établissement, pour permettre à tous de savoir quels sont les projets à l'œuvre aujourd'hui, quelles sont les évolutions prévues, et quelles questions cela peut poser sur les compétences, sur l'organisation du travail, ou encore sur les objectifs d'un établissement. Les responsables de formations aux métiers des bibliothèques, et les formateurs eux-mêmes, sont évidemment les bienvenus dans ces pages également.

L'espoir est naturellement aussi d'alimenter la culture professionnelle de quiconque travaille dans le monde des bibliothèques, de la documentation, des sciences de l'information.