

Diplôme de conservateur de bibliothèque

Mémoire d'étude / décembre 2016

## **Les outils de découverte en bibliothèque universitaire**

**Soledad Lida**

Sous la direction d'Anna Svenbro  
Adjointe au responsable du Service informatique - Bibliothèque  
Interuniversitaire de Santé (Paris)



## Remerciements

*Ma gratitude va tout d'abord à ma directrice, Anna Svenbro, pour son soutien, ses conseils, sa disponibilité et son suivi attentif tout au long de la préparation de ce mémoire.*

*J'adresse également mes remerciements les plus chaleureux à tous les professionnels de l'enseignement supérieur et de la recherche qui ont très aimablement pris le temps de répondre à mes questions et m'ont fourni la matière de ce travail : Laurent Aucher, Vincent Baas, Aurélie Bec, Dorothee Benhamou-Suesser, Anne-Claire Bernaudin, Sonia Bouis, Raphaële Bussemey, Emilie Cosson, Brigitte Crayssac, Yohann Davrieux, William Demet, Loïc Ducasse, Gianna Giorgi, Stéphane Gully, Perrine Helly, Johnatan Joly, Thomas Jouneau, Cyril Leroy, Laurent Lhuillier, Laurent Nabias, Pascale Nalon, Claude Niederlender, Ewa Nieszkowska, Anne-Sophie Pascal, Laurent Quinson, François Renaville, Gaël Revelin, Dominique Rouger, Laurent Schmitt, Julien Sicot, Yves Tomic, Michaël Villanova et Carole Vinot.*

*Je remercie vivement aussi Mathieu Ponthenier (d'Ebsco) et Pierre Persinette-Gautrez (d'Ex Libris – Proquest) pour le temps qu'ils m'ont gracieusement accordé et les réponses qu'ils ont apportées à mes interrogations.*

*Un grand merci également à Bruno Béguet, Sandrine Berthier, Floriane Berti, Valérie Bertrand, Delphine Coudrin, Thierry Fournier, Emilie Liard et Elisabeth Noël pour les différents éléments d'information qu'ils m'ont fournis.*

*Un grand merci aussi à la Kohateam – l'équipe de gestion du SIGB Koha du SCD de Saint-Étienne – dont les discussions et les échanges n'ont cessé de nourrir ma réflexion.*

*Ma reconnaissance va également à Loïc Ducasse dont les remarques et les commentaires ont été très éclairants sur un grand nombre de questions.*

*Ce travail doit aussi beaucoup à Dominique Rouger pour ses abondantes observations et ses analyses précieuses à bien des égards.*

*À Sandrine Mousin, Jean-Marc Parédès, Frédérique Hauville, Nicolas Beudon ainsi qu'à Dana, j'adresse toute ma gratitude pour leurs relectures attentives... et leur patience.*

*Y además, gracias a Fernando por la traducción al inglés y la fiebre del ajedrez.*

### ***Résumé :***

Développés dans les dix dernières années, les outils de découverte rassemblent au sein d'un même index des données provenant de sources diverses (catalogue local, bases de données, archives ouvertes, bases de connaissance etc.), afin de permettre d'interroger de manière unifiée un large éventail de ressources de niveau universitaire. Reçus de prime abord avec intérêt par les professionnels des bibliothèques, ces outils suscitent pourtant de nombreuses interrogations quant à leur modèle économique et à leur positionnement par rapport à d'autres systèmes et logiciels de bibliothèque. Adoptés par plusieurs bibliothèques universitaires françaises, leur mise en œuvre a pu impliquer une forte complexité, avec des résultats parfois très éloignés de l'idéal de l'outil tout-en-un. En ce qui concerne les données d'utilisation, celles-ci apparaissent pour l'heure contrastées et hétérogènes, aussi bien du côté du public que des utilisateurs professionnels. Tandis que plusieurs bibliothèques expérimentent des solutions alternatives, des efforts sont également en cours pour rendre les outils de découverte plus efficaces, en améliorant notamment la qualité des métadonnées indexées.

### ***Descripteurs :***

Catalogues de bibliothèques en ligne -- France

Logiciels documentaires

Sources d'information électroniques

Interfaces utilisateur (informatique)

Bibliothéconomie -- Innovations technologiques -- France

**Abstract :**

Having been developed during the last decade, discovery tools include a central index with data derived from different sources (local catalog, databases, open access repositories, knowledge bases etc.), in order to make possible a thorough search in only one instance across a wide variety of academic resources. Initially welcomed with great interest by library professionals, these tools eventually posed several questions regarding their economic model and their relationships with other library systems. After being adopted by several academic libraries in France, their implementation often brought about technical complications and results at times quite distant from those expected. The observed data, from user's point of view, allows so far a wide range of conclusions. While other university libraries experience some alternative solutions, there are also current efforts to increase the efficiency of discovery systems, in particular improving the quality of indexed metadata.

**Keywords :**

Online library catalogs -- France

Information retrieval--Computer programs

Electronic information resource searching.

Library catalogs and users.

User interfaces (Computer systems)

**Droits d'auteurs**

Cette création est mise à disposition selon le Contrat : « **Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France** » disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.



# Sommaire

<b>SIGLES ET ABREVIATIONS .....</b>	<b>11</b>
<b>INTRODUCTION.....</b>	<b>13</b>
<i>La nébuleuse discovery .....</i>	<i>13</i>
<i>Une réponse au double défi du web et de la documentation numérique 15</i>	<i>15</i>
<i>Des attentes initiales très fortes... et de nombreux questionnements... 16</i>	<i>16</i>
<i>Méthodologie.....</i>	<i>19</i>
<b>I. CONTEXTE ET ENJEUX.....</b>	<b>21</b>
<b>1. A la recherche d'un outil centralisateur .....</b>	<b>21</b>
A. <i>La multiplication d'outils de signalement et de gestion de ressources numériques .....</i>	<i>21</i>
B. <i>Les moteurs de recherche fédérée .....</i>	<i>22</i>
<b>2. Les grandes catégories d'outils de découverte .....</b>	<b>23</b>
A. <i>Les interfaces de découverte .....</i>	<i>23</i>
B. <i>Les services de découverte à index général centralisé .....</i>	<i>24</i>
C. <i>Des outils mixtes : une interface et un index d'origines différentes .....</i>	<i>25</i>
<b>3. Quel modèle économique pour les services de découverte à index centralisé? .....</b>	<b>26</b>
A. <i>Un marché fortement concurrentiel, évolutif et anglo-saxon.....</i>	<i>26</i>
B. <i>Les flux des données.....</i>	<i>29</i>
C. <i>Le modèle SaaS.....</i>	<i>33</i>
D. <i>La maintenance .....</i>	<i>36</i>
<b>4. Les spécificités du paysage français.....</b>	<b>40</b>
A. <i>Outils propriétaires, outils libres, outils mutualisés .....</i>	<i>40</i>
B. <i>En amont de l'installation .....</i>	<i>41</i>
C. <i>Les besoins en termes de ressources humaines .....</i>	<i>44</i>
<b>II. LES SERVICES DE DECOUVERTE A L'EPREUVE DE LEUR IMPLEMENTATION.....</b>	<b>47</b>
<b>1. Appréhender l'index général centralisé.....</b>	<b>47</b>
A. <i>L'index général centralisé et ses inconnues.....</i>	<i>47</i>
B. <i>La granularité de l'indexation .....</i>	<i>49</i>
C. <i>Des variations selon les éditeurs, les langues et les disciplines..</i>	<i>50</i>
D. <i>Adapter l'index général centralisé à la bibliothèque .....</i>	<i>52</i>
<b>2. L'intégration avec les systèmes locaux .....</b>	<b>53</b>
A. <i>L'intégration avec le SIGB .....</i>	<i>53</i>
B. <i>L'intégration avec la(les) base(s) de connaissance.....</i>	<i>59</i>

C.	<i>L'intégration avec les systèmes d'authentification</i>	63
D.	<i>L'articulation avec d'autres bases</i>	65
<b>3.</b>	<b>Une recherche unifiée ?</b>	<b>66</b>
A.	<i>Une recherche unifiée, des métadonnées distinctes</i>	66
B.	<i>Les facettes</i>	71
C.	<i>L'algorithme de pertinence</i>	77
D.	<i>Des interfaces multiples</i>	81
<b>III.</b>	<b>LES OUTILS DE DECOUVERTE ET LEUR UTILISATION</b>	<b>85</b>
<b>1.</b>	<b>Du côté du public</b>	<b>85</b>
A.	<i>Quel public cible ?</i>	85
B.	<i>Outils de découverte et information literacy</i>	86
C.	<i>Une évaluation de l'utilisation encore balbutiante</i>	94
<b>2.</b>	<b>Du côté des utilisateurs professionnels</b>	<b>103</b>
A.	<i>Un bilan mitigé</i>	103
B.	<i>Accompagner l'évolution des pratiques professionnelles</i>	105
<b>IV.</b>	<b>PERSPECTIVES</b>	<b>107</b>
<b>1.</b>	<b>Des alternatives aux interfaces commerciales : les outils libres</b>	<b>107</b>
A.	<i>La mise en place</i>	107
B.	<i>Les moyens mobilisés</i>	111
<b>2.</b>	<b>Des alternatives aux index centralisés</b>	<b>112</b>
A.	<i>Paramétrer un résolveur de liens avec Google Scholar</i>	112
B.	<i>Intégrer des API d'autres bases à un catalogue public</i>	113
C.	<i>Exploiter les données de Crossref ?</i>	115
D.	<i>Des corpus et des réseaux : le dispositif CERCLES</i>	115
<b>3.</b>	<b>Des index de découverte à la curation des données</b>	<b>117</b>
A.	<i>Des bonnes pratiques à l'Open Discovery Initiative</i>	117
B.	<i>Quid d'un index de découverte en open access ?</i>	125
C.	<i>Des données de qualité pour les outils de découverte</i>	126
D.	<i>La curation de données au niveau d'une bibliothèque</i>	133
	<b>CONCLUSION</b>	<b>135</b>
	<b>SOURCES</b>	<b>137</b>
<b>1.</b>	<b>Bibliothèques</b>	<b>137</b>
<b>2.</b>	<b>Experts dans le domaine des métadonnées</b>	<b>139</b>
<b>3.</b>	<b>Fournisseurs</b>	<b>139</b>
	<b>BIBLIOGRAPHIE</b>	<b>141</b>
<b>1.</b>	<b>Informatique documentaire : contexte et évolutions</b>	<b>141</b>
<b>2.</b>	<b>Normalisation, curation et échanges de métadonnées</b>	<b>142</b>



3.	<b>Outils de découverte : études générales .....</b>	<b>145</b>
4.	<b>Utilisation, <i>Information literacy</i> et études d'utilisabilité.....</b>	<b>149</b>
5.	<b>Sitographie des éditeurs d'outils.....</b>	<b>153</b>
6.	<b>Sitographie des outils de découverte en BU .....</b>	<b>153</b>
7.	<b>Etudes de cas sur des outils spécifiques .....</b>	<b>156</b>
	<b>ANNEXES.....</b>	<b>161</b>
	<b>TABLE DES MATIERES.....</b>	<b>171</b>



## *Sigles et abréviations*

- A&I : Abstract et Indexation  
ABES : Agence Bibliographique de l'Enseignement Supérieur  
ACEF : Association des clients d'Ex Libris France  
AFNOR : Association Française de Normalisation  
API : Application Programming Interface  
AUSIDEF : Association des Utilisateurs SirsiDynix en France  
BnF : Bibliothèque Nationale de France  
BU : Bibliothèque universitaire  
CAS : Central Authentication Service  
CCTP : Cahiers de Clauses Techniques Particulières  
CMS : Content Management System  
CNIL : Commission Nationale Informatique et Libertés  
COMUE : Communauté d'Universités et d'Établissements  
COUNTER : Counting Online Usage of NeTworked Electronic Resources  
Couperin : Consortium unifié des établissements universitaires et de recherche pour l'accès aux publications numériques  
CSS : Cascading Style Sheets  
DC : Dublin Core  
DOAJ : Directory of Open Access Journals  
DOI : Digital Object Identifier  
EAD : Encoded Archival Description  
ENT : Environnement Numérique de Travail  
ERMS : Electronic Resource Management System  
FRBR : Functional Requirements for Bibliographic Records  
FTP : File Transfer Protocol  
GED : Gestion Electronique de Documents  
http : HyperText Transfer Protocol  
ICOLC : International Coalition of Library Consortia  
ISTEX : Initiative d'excellence de l'Information Scientifique et Technique  
LSH : Lettres et Sciences Humaines  
MARC : MACHine Readable Catalog  
NFAIS : National Federation of Abstracting and Information Services  
NISO : National Information Standards Organization  
OAI-PMH : Open Archive Initiative Protocol for Metadata Harvesting  
OCLC : Online Computer Library Center  
OPAC : Open Public Access Catalog

ORCID : Open Researcher and Contributor ID  
PMID : PubMed Identifier  
PRES : Pôle de Recherche et d'Enseignement Supérieur  
RAMEAU Répertoire d'Autorité-Matière Encyclopédique et Alphabétique Unifié  
RDA : Resource Description and Access  
RDF : Resource Description Framework  
RFID : Radio Frequency IDentification  
SaaS : Software as a Service  
SCD : Service Commun de Documentation  
SGBD : Système de Gestion de Base de Données  
SGBm : Système de Gestion de Bibliothèques mutualisé  
SHS : Sciences humaines et sociales  
SICD : Service interétablissement de coopération documentaire  
SIGB : Système Intégré de Gestion de Bibliothèque  
SRU : Search/Retrieve via URL  
SRW : Search/Retrieve Web service  
SSO : Single Sign-On  
STM : Sciences, Techniques, Médical  
SUDOC : Système Universitaire de DOCumentation  
URI : Uniform Resource Identifier  
URL : Uniform Resource Locator  
VIAF : Virtual International Authority File  
XML : eXtensible Markup Language

# INTRODUCTION

---

## La nébuleuse *discovery*

La notion de *découverte* - traduction, plus ou moins heureuse, de l'anglais *discovery* - désigne dans le domaine des sciences de l'information la capacité à trouver une information pertinente sur un sujet donné.

Cette notion a été mise en avant pour qualifier des outils, dont le nom anglais *discovery tools* a été traduit littéralement par *outils de découverte*. Depuis que l'expression est apparue, au tournant de l'année 2010, il est devenu courant de parler d'outils de découverte dans le monde des bibliothèques. Pour autant, l'objet désigné par ces termes est rarement délimité avec précision, laissant à chacun le loisir de lui attribuer des contours différents. La définition qu'en donne Tosca Consultants dans son enquête de 2016 sur les logiciels métiers destinés aux bibliothèques reflète bien le périmètre très large des réalités recouvertes par cette expression :

« service de recherche en ligne conçu pour une utilisation par le public, proposant une recherche à la Google, la navigation à facettes, des possibilités de rebond sur des termes de recherche apparaissant au sein d'un nuage de mots et des fonctions Web 2.0 ; cet outil sollicite simultanément le catalogue de la bibliothèque, les bases de métadonnées relatives aux documents numériques produits par la bibliothèque ou sa tutelle et les bases de métadonnées relatives aux ressources numériques acquises par la bibliothèque<sup>1</sup>. »

Or les outils de découverte ne forment pas un ensemble homogène. Il ne s'agit pas d'une catégorie d'outils mais de plusieurs catégories d'outils, rassemblées sous une même étiquette, ce qui peut être source de confusion, y compris pour les professionnels eux-mêmes qui les mettent en œuvre. Au sein de cet ensemble on peut distinguer<sup>2</sup> :

- Les interfaces de découverte (en anglais *discovery interfaces*) : ce sont des interfaces qui permettent d'interroger de manière conjointe des ressources de différentes provenances (systèmes locaux, ressources en ligne) mais qui ne contiennent pas elles-mêmes nativement des métadonnées.
- Les services de découverte à index général centralisé (en anglais *index-based discovery services*) : ils sont constitués d'une interface de découverte et d'un index général centralisé rassemblant les métadonnées provenant de sources diverses et le cas échéant de systèmes locaux.
- Les outils mixtes : ils combinent une interface d'un fournisseur et un index général centralisé d'un autre fournisseur, mis en œuvre conjointement mais sans être nativement liés.

---

<sup>1</sup> TOSCA CONSULTANTS. Les logiciels métier destinés aux bibliothèques. Dans : *Tosca consultants* [en ligne]. 11 mars 2016. [Consulté le 10 juillet 2016]. Disponible à l'adresse : <http://toscaconsultants.fr/les-logiciels-metier-destines-aux-bibliotheques>.

<sup>2</sup> Pour ce qui concerne le développement de ces définitions voir : BREEDING, Marshall et National Information Standards Organization. *The future of library resource discovery: a white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee*. Baltimore : NISO, 2015. [Consulté le 21 février 2016]. Disponible à l'adresse : [http://www.niso.org/apps/group\\_public/download.php/14487/future\\_library\\_resource\\_discovery.pdf](http://www.niso.org/apps/group_public/download.php/14487/future_library_resource_discovery.pdf).

En anglais, on trouve également l'expression *discovery layer* qu'on pourrait traduire par *brique de découverte* et qui désigne, par la métaphore de la strate superposée, le fait que ces interfaces ou ces services viennent s'ajouter à des systèmes d'information existants (et parfois complexes) au sein des bibliothèques.

D'un point de vue chronologique, les premiers outils à se développer ont été les interfaces de découverte. Apparues dans les années 2006-2007, elles ont eu vocation à remplacer l'OPAC de SIGB vieillissants, c'est pourquoi elles ont pu être appelées parfois « OPAC de nouvelle génération ». Qu'elles aient été le produit de fournisseurs de logiciels ou d'initiatives *open source*, elles se caractérisent par :

- une interface avec une barre de recherche unique qui permet de chercher par défaut dans un grand nombre de ressources disponibles ;
- un moteur d'indexation qui indexe les données provenant de différentes sources locales (SIGB, entrepôt OAI, fichiers chargés manuellement etc.) ;
- l'interopérabilité<sup>3</sup> possible avec le SIGB pour permettre l'affichage des informations des exemplaires et du dossier lecteur ;
- l'interopérabilité possible avec un résolveur de liens pour permettre l'accès au texte intégral à partir de références trouvées dans la liste de résultats ;
- des résultats présentés par ordre de pertinence ;
- une navigation par facettes ;
- des fonctionnalités sociales web 2.0.

C'est un peu plus tard, au tournant des années 2010, que les index généraux centralisés se sont constitués. Il s'agit là de produits fournis uniquement par des sociétés commerciales. Ce sont des index de masse qui couvrent un grand ensemble de contenus universitaires et de recherche, et qui peuvent notamment être alimentés par :

- les métadonnées et/ou le texte intégral d'éditeurs commerciaux ;
- les métadonnées de bases de références bibliographiques (abstract et indexation) ;
- les métadonnées et/ou le texte intégral d'entrepôts en *open access*.

La constitution de ces index a permis l'apparition des deux autres catégories d'outils - les services de découverte à index général centralisé et des outils mixtes. Ces outils proposent des interfaces de découverte, telles que décrites précédemment, mais avec, en plus, un index général centralisé intégré qui peut être interrogé par le moteur d'indexation au même titre que les ressources locales.

---

<sup>3</sup> Le dictionnaire *Larousse* définit l'interopérabilité comme : « Capacité de matériels, de logiciels ou de protocoles différents à fonctionner ensemble et à partager des informations ». LAROUSSE, Éditions. *Définitions : interopérabilité - Dictionnaire de français Larousse* [en ligne]. [s. d.]. [Consulté le 26 novembre 2016]. Disponible à l'adresse : <http://www.larousse.fr/dictionnaires/francais/interop%C3%A9rabilit%C3%A9/43787>.

## Une réponse au double défi du web et de la documentation numérique

L'apparition des outils de découverte a répondu à plusieurs problématiques, parmi lesquelles on peut d'abord citer les difficultés de signalement de la documentation numérique. Celle-ci n'a pas cessé de prendre une place de plus en plus importante dans les collections des bibliothèques universitaires depuis le début des années 2000. Or les SIGB classiques n'étaient pas conçus pour répondre aux problèmes posés par les ressources numériques : gérer des modes d'accès complexes, des mises à jour continues des données, celles-ci étant désormais susceptibles d'être modifiées à tout moment sans information préalable, et une granularité d'information beaucoup plus fine<sup>4</sup>. Cela a pu se traduire par un signalement très dispersé de la documentation numérique à travers de nombreux outils : catalogue de la bibliothèque, site web, listes alphabétiques de titres, de plateformes d'éditeurs ou de regroupements thématiques etc. Le résultat était parfois peu lisible pour l'utilisateur final qui était en outre obligé de refaire à chaque fois sa recherche *da capo* dans chaque ressource. Les moteurs de recherche fédérée ont tenté d'apporter une réponse à ce problème en permettant d'interroger plusieurs ressources en même temps, mais ils comportaient de nombreuses imperfections<sup>5</sup>. Il y avait donc une demande en matière d'outils qui permettraient une recherche unifiée et fluide dans les ressources numériques.

Ce besoin d'une recherche unifiée correspond également au mode de fonctionnement des moteurs de recherche du web, notamment Google et Google Scholar qui ont très rapidement conquis le monde académique, y compris ses utilisateurs les plus avancés (enseignants, chercheurs, doctorants)<sup>6</sup>. Pour les bibliothèques, l'enjeu était de mettre à disposition de l'utilisateur des outils qui correspondent à ses pratiques courantes, au lieu de demander à celui-ci de s'adapter aux contraintes des outils de bibliothèque. En effet, les OPAC traditionnels pouvaient poser de réels problèmes d'utilisabilité, exigeant des utilisateurs d'être capables de manipuler des interfaces plus ou moins complexes, avec des index, des syntaxes ou des vocabulaires contrôlés spécifiques qui dans la plupart des cas leur étaient complètement étrangers. L'objectif était donc de proposer une interface documentaire dont les utilisateurs pourraient se servir comme de n'importe quel autre moteur de recherche ou site web, à savoir une recherche initiale tous mots qui s'affinerait par la suite, si nécessaire, à travers des facettes. L'expression anglaise *web-scale discovery services* (outils de découverte à l'échelle du web) qui est utilisée parfois pour qualifier les services de découverte à index général centralisé rend bien compte de cette dimension et de l'ambition de rivaliser avec les grands acteurs du web.

Cette généralisation de l'usage des moteurs de recherche s'est accompagnée d'une attente de l'utilisateur vis-à-vis de l'accès au contenu en texte intégral :

<sup>4</sup> BERTHIER, Sandrine. *Le SIGB: pilier ou élément désormais mineur de l'informatique documentaire ?* Villeurbanne, Rhône, France : [s. n.], 2012. [Consulté le 23 février 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/60267-le-sigb-pilier-ou-element-desormais-mineur-de-l-informatique-documentaire.pdf> ; COUDRIN, Delphine. Les dispositifs de signalement en université. Dans : *Intégrer des ressources numériques dans les collections*. Villeurbanne, France : Presses de l'enssib, 2014, p. 112-125. YONGMING WANG et DAWES, Trevor A. The Next Generation Integrated Library System: A Promise Fulfilled. *Information Technology & Libraries*. Septembre 2012, Vol. 31, n° 3, p. 76-84.

<sup>5</sup> Voir *infra* chapitre I, 1, B « Les moteurs de recherche fédérée ».

<sup>6</sup> GALLEZOT, Gabriel. *Enquête sur les Pratiques Informationnelles des Chercheurs : Quelques Résultats* [en ligne]. [s. d.]. [Consulté le 23 mars 2016]. Disponible à l'adresse : <http://urfistinfo.hypotheses.org/1901>.

celui-ci doit se faire de la manière la plus directe possible. À l'heure où Google Scholar peut permettre de trouver un article en texte intégral en ligne en quelques clics, demander à l'utilisateur de faire une recherche dans le catalogue, de vérifier l'état de collection de la revue papier, puis, si le numéro souhaité ne s'y trouve pas, de refaire sa recherche successivement dans une base de revues ou dans une base de données, pour au final demander l'article par le prêt entre bibliothèques, devient difficilement acceptable. De même, le mouvement en faveur de l'*open access* a pu lui aussi renforcer l'attente vis-à-vis du texte intégral, en élargissant la masse de documentation numérique de niveau académique accessible sur le web, par exemple à travers des archives ouvertes ou des répertoires de publications en accès ouvert comme DOAJ. Désormais, tout obstacle à l'accès au texte intégral peut être une source de frustration pour l'utilisateur final et toute référence présentée doit s'accompagner, à chaque fois que cela est possible, d'un accès au contenu par le chemin le plus court. Pour reprendre les mots de Romain Wenz, « Google propose d'inverser complètement la logique du catalogue. Il s'agit de chercher les informations dans le document lui-même, puis d'avoir affaire aux données de description pour valider la ressource. »<sup>7</sup> En d'autres termes, la recherche de références bibliographiques seules semble révolue, ce que le slogan anglo-saxon a désigné par les termes de « *discovery and delivery* » (découverte et accès). L'OCLC avait pris acte de cette exigence en 2009 : « pour beaucoup d'utilisateurs finaux, si elle ne donne pas accès à quelque chose dont il/elle a besoin, la découverte seule est une perte de temps »<sup>8</sup>. C'est à cette nouvelle articulation entre recherche bibliographique et accès au contenu que les outils de découverte ont voulu répondre.

Conjointement, d'un point de vue professionnel, ces outils sont apparus au moment où la structure des données des catalogues traditionnels rencontrait de grands besoins de mutation. Mis au défi du web, les catalogues ont entamé une longue mue pour quitter leurs silos et adopter une syntaxe de données indexable par les moteurs de recherche et interrogeable dans le web sémantique. La transition vers le modèle FRBR et le code de catalogage RDA, n'est certes pas le sujet de ce travail, mais on ne peut pour autant ignorer que ce souhait général d'évolution était aussi présent en arrière-plan lorsque les outils de découverte se sont développés.

## **Des attentes initiales très fortes... et de nombreux questionnements**

Au regard de ces éléments, on comprend que le développement des outils de découverte ait suscité des attentes très fortes chez les professionnels<sup>9</sup>, qui y ont vu parfois la solution prodigieuse qui allait (enfin) permettre de :

- fournir un outil de recherche qui réponde aux pratiques courantes des usagers ;

---

<sup>7</sup> WENZ, Romain. *L'avenir des catalogues*. [s.l.] : ENSSIB, 2009. [Consulté le 30 avril 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/21205-l-avenir-des-catalogues.pdf>, p.37.

<sup>8</sup> « For many end users, without the delivery of something he or she wants or needs, discovery alone is a waste of time ». OCLC. *Online catalogs: what users and librarians want: an OCLC report*. Dublin, Ohio : OCLC, 2009. Disponible à l'adresse : <http://www.oclc.org/content/dam/oclc/reports/onlinecatalogs/fullreport.pdf>, p. 20.

<sup>9</sup> Voir notamment YONGMING WANG et DAWES, Trevor A. The Next Generation Integrated Library System: A Promise Fulfilled. *Information Technology & Libraries*. Septembre 2012, Vol. 31, n° 3, p. 76-84.



- proposer un point d'accès unique aux collections de la bibliothèque, tous supports, toutes ressources, tous niveaux de granularité confondus ;
- améliorer la visibilité des ressources numériques, parfois sous-utilisées, et accroître à terme leur consultation ;
- simplifier et rendre fluide l'accès à la documentation numérique ;
- renouveler l'image de la bibliothèque, comparable à n'importe quel site web.

Répandus d'abord dans le monde anglo-saxon, les outils de découverte gagnent progressivement les bibliothèques françaises. Ils sont plus d'une trentaine en fonctionnement aujourd'hui dans les bibliothèques d'enseignement supérieur sans compter les installations en cours<sup>10</sup>. Les premières installations datent de la fin des années 2000, mais on dénombre aussi de nombreuses installations récentes (datant de 2015 ou postérieures) et plusieurs projets en cours. Les interventions consacrées à ces outils dans des congrès professionnels, la journée d'étude organisée par le consortium Couperin en 2012, l'enquête menée par celui-ci en 2013, ainsi que l'étude par l'ABES consacrée à un éventuel outil de découverte national<sup>11</sup>, témoignent du bon accueil et de l'intérêt que les professionnels de l'enseignement supérieur français ont réservé à ces outils.

Néanmoins, le fait que certaines bibliothèques ayant expérimenté un outil de découverte aient fait le choix de ne pas les reconduire, peut soulever des interrogations quant à la plus-value de ces outils. Ainsi, l'Université de Bretagne Occidentale a préféré, après une enquête auprès du public, ne pas prolonger l'utilisation du produit qu'elle a expérimenté en 2014-2015<sup>12</sup>. D'autres bibliothèques universitaires, comme celles de l'Université Rennes 2 ou de l'Université de Savoie Mont Blanc, ont choisi de mettre un terme à l'utilisation d'un outil pour le remplacer par un autre<sup>13</sup>. D'autres encore, comme la bibliothèque de l'université d'Utrecht, ont pu faire, après analyse du paysage économique et des ressources offertes par la bibliothèque, le diagnostic qu'un outil de découverte n'était plus nécessaire pour son établissement<sup>14</sup>. Si l'on ajoute à cela que des voix se sont élevées dans le monde des bibliothèques pour aborder le sujet

<sup>10</sup> *Outil de découverte — Bibliopedia* [en ligne]. [s. d.]. [Consulté le 19 janvier 2016]. Disponible à l'adresse : [http://www.bibliopedia.fr/wiki/Outil\\_de\\_d%C3%A9couverte](http://www.bibliopedia.fr/wiki/Outil_de_d%C3%A9couverte).

<sup>11</sup> ABES. *Outils de découverte : signaler grâce à BACON* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 20 juin 2016]. Disponible à l'adresse : [http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES\\_2016\\_session\\_parallele\\_1\\_06\\_BACON\\_OUTILS\\_DE\\_DECOUVERTE\\_VD](http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES_2016_session_parallele_1_06_BACON_OUTILS_DE_DECOUVERTE_VD) ; COUPERIN. *Discovery-tools - Journée d'étude Discovery tools libres du 24 janvier 2012* [en ligne]. [s. d.]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.couperin.org/groupe-de-travail-et-projets-deap/acces-aux-ressources-cat/discovery/item/954-journee-detude-discovery-tools-libres-du-24-janvier-2012> ; COUPERIN. *Discovery-tools - Analyse de l'enquête 2013* [en ligne]. [S. l.] : Couperin, 2013. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.couperin.org/groupe-de-travail-et-projets-deap/acces-aux-ressources-cat/discovery/item/1083-analyse-de-l-enquete-discovery-tools-2013>. ABES. Vers un outil de découverte national. Dans : *Fil ABES* [en ligne]. 23 novembre 2012. [Consulté le 6 avril 2016]. Disponible à l'adresse : <http://fil.abes.fr/2012/11/23/vers-un-outil-de-decouverte-national/>.

<sup>12</sup> HELLY, Perrine. Catalogue+ : arrêt du service. Dans : *Ubodoc* [en ligne]. 9 juillet 2015. [Consulté le 25 mars 2016]. Disponible à l'adresse : <http://ubodoc.univ-brest.fr/catalogue-arret-du-service/>.

<sup>13</sup> *Outil de découverte — Bibliopedia. Op. cit.*

<sup>14</sup> KORTEKAAS, Simone. Thinking the unthinkable: a library without a catalogue - Reconsidering the future of discovery tools for Utrecht University library. Dans : *LIBER* [en ligne]. 4 septembre 2012. [Consulté le 20 mars 2016]. Disponible à l'adresse : <http://libereurope.eu/blog/2012/09/04/thinking-the-unthinkable-a-library-without-a-catalogue-reconsidering-the-future-of-discovery-tools-for-utrecht-university-library/>. KORTEKAAS, Simone et KRAMER, Bianca. Thinking the unthinkable – doing away with the library catalogue: Based on a paper presented by Simone Kortekaas at the 37th UKSG Annual Conference, Harrogate, April 2014. *Insights: the UKSG journal* [en ligne]. Novembre 2014, Vol. 27, n° 3, p. 244-248.

avec une nuance plus sceptique<sup>15</sup> ou encore alerter sur un possible effet de mode<sup>16</sup>, un tel revirement depuis les attentes initiales a de quoi susciter des interrogations.

C'est pourquoi il a semblé pertinent de se pencher sur les problèmes concrets qui se posent aux bibliothèques universitaires, lorsqu'elles mettent en place ces outils. Ces problèmes peuvent être regroupés en quatre grands ensembles :

Le premier ensemble de problèmes concerne d'abord le positionnement de ces outils dans le paysage des logiciels de bibliothèques et en particulier les modèles économiques sur lesquels ils reposent. Touchant à la fois à la documentation numérique et à l'informatique documentaire, les outils de découverte se trouvent à l'intersection des éditeurs de logiciels, des agrégateurs de métadonnées, des éditeurs fournisseurs de données et des bibliothèques qui achètent les services des uns et des autres. Plus particulièrement, les services de découverte à index général centralisé sont produits par des sociétés qui sont également par ailleurs éditrices de logiciels et agrégatrices de données. Quels sont les tenants et les aboutissants de cette concentration verticale ? Quelle marge de manœuvre peut-il rester aux bibliothèques dans un marché si concentré où des acteurs assument plusieurs rôles simultanément ? Dans quelle mesure le développement d'interfaces de découverte *open source* ainsi que des initiatives en faveur de la mise à disposition des données ouvertes sous licence libre, peuvent-ils faire évoluer la situation ?

Le deuxième ensemble de problèmes concerne les métadonnées et leur indexation. Les outils de découverte indexent des métadonnées provenant des catalogues de bibliothèque, des ressources numériques payantes, de ressources en *open access* et de systèmes locaux, ce qui implique nécessairement une grande hétérogénéité des données, beaucoup plus grande que celle qui est gérée dans un catalogue classique. Plus particulièrement, les services de découverte à index généraux centralisés sont alimentés par des données provenant d'éditeurs et d'agrégateurs multiples dont la qualité, la granularité et la complétude des informations peut être très variable : comment gèrent-ils une telle diversité ? À cette hétérogénéité, s'ajoutent des interrogations sur la couverture, la profondeur et les mises à jour des données : quelle visibilité les bibliothèques ont-elles sur la façon dont les métadonnées des éditeurs et des agrégateurs sont intégrées dans les index généraux centralisés ? Dans une perspective plus large, quel rôle peuvent jouer les agences de normalisation et les instances bibliographiques nationales (ABES, BnF) dans l'amélioration de la qualité des métadonnées indexées ?

Un troisième ensemble de problèmes concerne l'articulation de ces outils avec les systèmes locaux existants au sein de la bibliothèque. Comment les outils de découverte s'intègrent-ils aux environnements techniques locaux ? Comment peuvent-ils s'articuler avec d'autres outils quand ils ne sont pas conçus nativement pour travailler ensemble ? Le plus important de ces problèmes touche sans doute à l'articulation entre l'outil de découverte et le SIGB qui continue à gérer les tâches liées au circuit du document physique. Jusqu'où peut-on intégrer les données et les services du SIGB dans l'outil de découverte ? Comment s'effectue cette

<sup>15</sup> BREITBACH, William. Web-Scale Discovery: A Library of Babel? Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 637-645. KEYSER, Piet de. Discovery Services, the Next Tool Libraries Must Have, or... Should Discard? Dans : *E-Discovery Tools and Applications in Modern Libraries*. Hershey, PA : Information Science Reference, 2016, p. 56-65.

<sup>16</sup> TAY, Aaron. *Musings about librarianship: Four possible Web Scale Discovery Future Scenarios* [en ligne]. 7 décembre 2014. [Consulté le 9 mars 2016]. Disponible à l'adresse : <http://musingsaboutlibrarianship.blogspot.com/2014/12/four-possible-web-scale-discovery.html>.

intégration ? Quelles sont les principales difficultés que rencontrent les bibliothèques qui ont mis cela en place ? De la même manière, une bonne articulation avec le résolveur de liens et les différents systèmes d'authentification locaux est indispensable pour que l'outil remplisse convenablement sa mission de « *discovery and delivery* ». Comment font les bibliothèques pour que tout cela fonctionne ensemble et de manière fluide ?

Le quatrième ensemble de problèmes concerne l'appropriation des outils par l'utilisateur final. Ces outils ont été conçus de manière délibérément « orientée utilisateurs » dans le but de s'adapter à leurs pratiques et à leurs évolutions. Or qu'en est-il de leur prise en main concrète, de leur utilisabilité effective ? On peut s'interroger sur la manière dont ces outils, conçus dans le contexte universitaire anglo-saxon, s'acclimatent à l'environnement universitaire français. Mais de manière plus générale, on peut aussi s'interroger sur la démarche même proposée par ces outils : quels publics servent-ils ? Parmi les différents types de publics universitaires, de l'étudiant primo-entrant à l'enseignant-chercheur, lesquels sont prioritairement visés par les outils de découverte ? Quels effets les outils de découverte ont-ils sur l'*information literacy* ? Dans quelle mesure ces outils qui se présentent comme immédiatement intuitifs dispensent-ils de formation à la recherche documentaire ? Par ailleurs, de quels outils de mesure et de quelles données les bibliothèques disposent-elles pour évaluer l'utilisation qui est faite des outils de découverte et leur effet sur la consultation de la documentation numérique ?

## Méthodologie

Par rapport à ces différents ensembles de problèmes, ce travail ne vise pas à réaliser une étude comparative d'outils – pour laquelle il faudrait des moyens qui excèdent ceux dont nous avons disposé dans le cadre de cette recherche – et il n'entend en aucun cas prescrire telle solution plutôt que telle autre – ce qui relèverait d'un registre décisionnel ou normatif qui ne peut être celui d'un mémoire d'étude. Il s'agit plutôt de comprendre ce qu'a impliqué la mise en œuvre de ces outils pour des bibliothèques universitaires, en recueillant leurs avis et leurs expériences, afin de servir éventuellement à d'autres établissements envisageant de mener ce type de projet.

L'objectif étant de pouvoir connaître le point de vue des bibliothèques sur leurs outils, il a fallu d'abord constater que, s'il y a une très abondante bibliographie disponible en langue anglaise faite par et pour des bibliothécaires sur le sujet, la bibliographie en français est extrêmement restreinte. Plus précisément, nous n'avons trouvé à ce jour aucune monographie sur la question disponible en français, alors que plusieurs sont disponibles et continuent de paraître en anglais<sup>17</sup>.

---

<sup>17</sup> SPENCER, John S. et MILSON-MARTULA, Christopher (dir.). *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014 ; BREEDING, Marshall et National Information Standards Organization. *Op. cit.* ; THOMPSON, JoLinda. *Implementing web-scale discovery services: a practical guide for librarians*. Lanham, Md : Rowman & Littlefield, 2014 ; VARNUM, Kenneth J. *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016. SMET, Egbert de et DHAMDHERE, Sangeeta N. (dir.). *E-discovery tools and applications in modern libraries*. Hershey, PA : Information Science Reference, 2016.

En outre, à l'exception de trois projets qui ont été très bien documentés<sup>18</sup>, les documents en langue française rendus publics sur la mise en place d'outils de découverte dans les bibliothèques restent relativement rares et isolés<sup>19</sup>.

C'est pourquoi nous avons complété la lecture de la bibliographie par une série d'entretiens avec des professionnels des bibliothèques de l'enseignement supérieur. À chaque fois que cela a été possible, nous avons échangé avec des professionnels qui étaient soit administrateurs directs de l'outil soit chargés de projet de mise en place de l'outil. Au terme de ce mémoire, nous avons pu mener 26 entretiens parmi lesquels nous pouvons distinguer notamment : 2 bibliothèques avec des outils libres, 5 bibliothèques concernées par des outils consortiaux et 2 bibliothèques avec des outils mixtes.

La liste des entretiens menés est détaillée dans la section « Sources » de ce travail. Les grilles sur lesquels ils se sont appuyés se trouvent intégrées en annexe (annexes 2 et 3). Ces grilles abordent différents aspects de la gestion de ces outils, des raisons de leur installation jusqu'aux informations concernant leur utilisation. Pour des raisons de confidentialité, dans le corps du mémoire, le nom des bibliothèques a été anonymisé avec un code numérique de type « BIB01 » (voir l'annexe 1 pour les détails).

Ces entretiens ont été complétés par des échanges avec des professionnels ayant une expertise spécifique dans le domaine des métadonnées, à l'ABES (pour le projet BACON) et à l'INIST (pour le projet ISTEEX).

En complément, nous avons également souhaité interroger les fournisseurs de services de découverte à index général centralisé, pour connaître leur point de vue sur les problématiques d'échanges et de flux de données. Il nous a été possible d'échanger avec deux d'entre eux : Ebsco (pour Ebsco Discovery Service) et Proquest (pour Summon). Le questionnaire sur lequel ces entretiens se sont appuyés se trouve aussi intégré en annexe (annexe 4).

Dans un premier temps, nous présenterons le contexte dans lequel sont apparus ces outils de découverte, en interrogeant plus précisément le modèle économique dont ils relèvent. Nous examinerons ensuite les différents aspects de l'implémentation de ces outils par les bibliothèques, avant de nous pencher sur la réception de ces outils et leur utilisation par le public et les professionnels. Enfin nous évoquerons des alternatives possibles aux outils de découverte ainsi que quelques perspectives d'évolution.

---

<sup>18</sup> Notamment à l'université de Liège, à la Bibliothèque Clermont Université et à l'université de Bordeaux (dans ces deux derniers cas, il s'agit d'interfaces libres). Pour Liège voir : RENAVILLE, François et RICHELLE, Laurence. Déploiement de la solution discovery Primo à l'Université de Liège. *Cahiers de la Documentation = Bladen voor Documentatie*. Décembre 2013, Vol. 67, n° 4, p. 41-53 ; RENAVILLE, François. *D'un Primo en local à un Primo dans le cloud: deux déploiements en deux ans d'une solution discovery à l'Université de Liège* [en ligne]. Novembre 2015. [Consulté le 14 juin 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/188168> ; NABIAS, Laurent, JESTIN, André et RENAVILLE, François. *Primo dans le Cloud ?* [en ligne]. Juin 2015. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/183397>. Pour Clermont et Bordeaux, voir *infra*, la partie IV, 1.

<sup>19</sup> HELLY, Perrine. Catalogue+ : arrêt du service. Op.cit. ; ABES. *Outils de découverte : signaler grâce à BACON*. Op cit. ; CURMIN, Frédéric et DOUILLY, Laurence. *Une autre interface publique en lien avec Koha Summon à Telecom Bretagne* [en ligne]. Vitry, 16 juin 2015. [Consulté le 10 juillet 2016]. Disponible à l'adresse : [http://kohaf.fr.org/sites/default/files/interface\\_summon.pdf](http://kohaf.fr.org/sites/default/files/interface_summon.pdf).

# I. CONTEXTE ET ENJEUX

---

## 1. A LA RECHERCHE D'UN OUTIL CENTRALISATEUR

Pour les bibliothèques, les outils de découverte sont apparus dans un paysage logiciel déjà complexe et très éclaté, marqué par une multiplication d'outils de signalement et de gestion des ressources numériques, avec lesquels ils doivent s'articuler.

### A. La multiplication d'outils de signalement et de gestion de ressources numériques

Historiquement, le premier élément de ce paysage logiciel est le SIGB. Or, comme l'a signalé Sandrine Berthier, si celui-ci reste le référent pour le signalement de la documentation sur support physique (cataloguée manuellement), il prend plus difficilement en compte la gestion des ressources numériques<sup>20</sup>. Cela est lié, d'après Delphine Coudrin<sup>21</sup>, au fait que les contenus des ressources numériques connaissent des mises à jour continues qui rendent leur traitement manuel très laborieux en raison de la grande quantité d'ajouts, de suppressions ou de modifications que peut connaître chaque titre ou chaque bouquet. Si le catalogage manuel de ces ressources de manière exhaustive dans le SIGB se révèle peu envisageable, le SIGB peut néanmoins se faire le réceptacle d'imports automatisés, notamment dans le cas des livres numériques lorsque l'éditeur en fournit une liste dans un format standard<sup>22</sup>.

Compte tenu de cette situation, plusieurs outils se sont développés indépendamment du SIGB pour gérer et signaler ces ressources numériques : les bases de connaissance (sur lesquelles s'appuient les listes alphabétiques de ressources et les résolveurs de liens) et les ERMS.

Les bases de connaissance sont le plus souvent des produits commerciaux qui contiennent des métadonnées avec une granularité de niveau titre de revues ou de livres numériques qui décrivent en détail la composition des bouquets de ressources numériques. Elles servent à générer des listes alphabétiques de ressources et permettent au résolveur de liens de construire les liens d'accès correspondants.

Une liste alphabétique de ressources (ou liste A-Z) est une interface qui « permet de consulter rapidement et très simplement la liste des revues, ebooks et bases de données souscrites par l'établissement. Ces outils sont fournis par différents éditeurs de logiciels ou agences d'abonnement et s'appuient sur une base de connaissance. La mise à jour des listes de titres et états de collection est ainsi automatique<sup>23</sup>. »

---

<sup>20</sup> BERTHIER, Sandrine. *Op. cit.*

<sup>21</sup> COUDRIN, Delphine. Les dispositifs de signalement en université. *Op. cit.*, p.112.

<sup>22</sup> COUPERIN. *Vade-mecum sur les ebooks* [en ligne]. [s. d.]. [Consulté le 21 avril 2016]. Disponible à l'adresse : <http://www.couperin.org/groupes-de-travail-et-projets-deap/ebook/vade-mecum-sur-les-ebooks/item/1227-signalement-valorisation-mediation>.

<sup>23</sup> COUDRIN, Delphine et HATT, Guillaume. Intégrer, signaler les ressources électroniques : 20 questions clés. Dans : *Intégrer des ressources numériques dans les collections*. Villeurbanne, France : Presses de l'enssib, 2014, p. 102.

Le résolveur de liens agit, quant à lui, au niveau des accès. Il s'agit d'un outil qui « permet à l'utilisateur en consultant une référence bibliographique sur une base de données de se voir proposer un lien direct d'accès au texte intégral de sa référence, disponible sur le site d'un des fournisseurs de la bibliothèque. Dans le cas où la bibliothèque n'a pas souscrit l'abonnement nécessaire, le résolveur permet également de proposer à l'utilisateur de relancer sa recherche sur le catalogue (revues imprimées) ou sur d'autres interfaces (autres catalogues, Google Scholar, *Open access...*)<sup>24</sup> ». Sandrine Berthier précise pour sa part que le résolveur doit permettre de générer « des liens directs et fonctionnels vers le périodique ou même l'article recherché et non vers le site du fournisseur du bouquet qui le contient où le lecteur doit réitérer sa recherche parfois dans une autre langue<sup>25</sup> ». Les résolveurs s'appuient le plus souvent sur les OpenURL qui permettent « la création de liens qui tiennent compte du contexte, c'est-à-dire des liens flexibles et capables de prendre en compte les affiliations institutionnelles de l'utilisateur et les licences de cette institution »<sup>26</sup>.

Concernant les ERMS, il s'agit d'outils à destination des professionnels, qui à l'instar du SIGB, permettent d'unifier toute la chaîne de traitement des ressources numériques dans un même système<sup>27</sup> : ils sont constitués notamment d'une base de connaissance, d'un module de gestion interne qui permet de garder traces des acquisitions et des licences, et d'un module de statistiques.

## B. Les moteurs de recherche fédérée

Les différentes solutions énumérées ci-dessus travaillent toutes avec une granularité de niveau titre de la revue ou du livre numérique, parfois même au niveau de la collection, du bouquet ou de la base de données. Si une liste A-Z permet ainsi d'interroger l'ensemble des titres de revues ou de livres numériques disponibles pour une bibliothèque donnée, la recherche précise à un niveau de granularité plus fin (article ou chapitre) a longtemps nécessité de passer par l'interrogation indépendante de chaque ressource concernée. Cela obligeait l'utilisateur à répéter sa recherche sur chaque plateforme, en reformulant à chaque fois sa requête selon des syntaxes différentes, et à appréhender de multiples interfaces d'interrogation.

Les moteurs de recherche fédérée, apparus en 2006, ont eu vocation à simplifier cette interrogation, en permettant de transmettre une même requête à une multitude de sources, afin de ne pas obliger l'utilisateur à recommencer à chaque fois la démarche pour chaque ressource interrogée. Comme l'a expliqué Marshall Breeding, les moteurs de recherche fédérée fonctionnent de la manière suivante<sup>28</sup> : l'utilisateur sélectionne les cibles de sa requête en indiquant les bases ou les disciplines sur lesquelles celles-ci doit porter ; la requête est reformulée dans différentes syntaxes

<sup>24</sup> *Ibid.*, p. 102.

<sup>25</sup> BERTHIER, Sandrine. *Op. cit.*, p. 47.

<sup>26</sup> COUPERIN. *OpenURL et les bases de connaissances : une vue d'ensemble* [en ligne]. [s. d.]. [Consulté le 26 novembre 2016]. Disponible à l'adresse : <http://www.couperin.org/groupes-de-travail-et-projets-deap/acces-aux-ressources-cat/bases-de-connaissances/kbart/item/892-openurl-et-les-bases-de-connaissances-une-vue-densemble>. Voir aussi à ce sujet MACHEFERT, Sylvain. *L'OpenURL dans les institutions françaises, une chance pour la valorisation des ressources électroniques ?* [en ligne]. [S. l.] : Institut national des techniques de la documentation du CNAM, 6 novembre 2007. [Consulté le 2 décembre 2016]. Disponible à l'adresse : [https://memic.ccsd.cnrs.fr/mem\\_00000613/document](https://memic.ccsd.cnrs.fr/mem_00000613/document).

<sup>27</sup> COUDRIN, Delphine. Les dispositifs de signalement en université. *Op. cit.*, p. 118.

<sup>28</sup> BREEDING, Marshall. Discovery product functionality. *Library Technology Reports*. 2015, Vol. 50, n° 1, p. 13.

selon les bases concernées ; elle est transmise ensuite aux ressources cibles grâce à différents connecteurs (protocoles Z39.50, SRU/SRW, API ou chemin XML/XPATH) ; les résultats sont renvoyés par les ressources cibles et affichés par lots à mesure qu'ils arrivent, regroupés par source ou discipline en fonction des choix de l'utilisateur.

Les moteurs de recherche fédérée permettaient donc une interrogation en une fois de plusieurs ressources au niveau de granularité de l'article, à travers une grande variété de plateformes. Ils comportaient cependant un certain nombre de faiblesses, qui ont été pointées à maintes reprises par la littérature professionnelle<sup>29</sup> : complexité d'interrogation, difficulté à développer ou à maintenir des connecteurs spécifiques pour chaque base, temps de réponse longs et asynchrones, absence de dédoublement, absence de classement par pertinence dans l'affichage des résultats et absence d'intégration avec les données du catalogue local.

Comme le signale la *National Information Standards Organization* (NISO), c'est en grande partie pour simplifier ce processus de recherche au niveau de l'article à travers différentes ressources, insuffisamment satisfait par la recherche fédérée, que les outils de découverte se sont développés : « dans une grande mesure, les services de découverte à index général centralisé sont apparus pour simplifier le processus de recherche, notamment les démarches complexes pour identifier et utiliser de multiples ressources différentes, chacune avec sa propre interface et ses propres spécificités de recherche<sup>30</sup> ».

## 2. LES GRANDES CATEGORIES D'OUTILS DE DECOUVERTE

Comme nous l'avions indiqué en introduction, les outils de découverte ne constituent pas un ensemble homogène mais peuvent se décliner en plusieurs catégories d'outils différents.

### A. Les interfaces de découverte

Les interfaces de découverte (ou OPAC de nouvelle génération) sont historiquement les plus anciennes : elles ont commencé à se développer à partir de 2006.

D'après Marshall Breeding, ces interfaces se caractérisent par une barre de recherche unique « orientée utilisateur », un moteur d'indexation qui indexe les

---

<sup>29</sup> BERTHIER, Sandrine. *Op. cit.*, p. 50. LUTHER, Judy et KELLY, Maureen C. The Next Generation of Discovery. *Library Journal*. Mars 2011, Vol. 136, n° 5, p. 67. VAUGHAN, Jason. Chapter 1: Web Scale Discovery What and Why? *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 6. ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. Herramientas de descubrimiento: ¿una ventanilla única? *Revista Española de Documentación Científica*. Mars 2015, Vol. 38, n° 1, p. 3. THOMSETT-SCOTT, Beth et REESE, Patricia E. Academic Libraries and Discovery Tools: A Survey of the Literature. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 6.

<sup>30</sup> « On a very large extent, index-based discovery services emerged to simplify the research process relative to the complex process of identifying and using many different individual content products, each with its own interface and search procedure. » National Information Standards Organization et Open Discovery Initiative Working Group. *Open discovery initiative: promoting transparency in discovery: a recommended practice of the National Information Standards Organization*. Baltimore, MD : National Information Standards Organization (U.S.), 2014. Disponible à l'adresse : [http://www.niso.org/apps/group\\_public/download.php/14820/rp-19-2014\\_ODI.pdf](http://www.niso.org/apps/group_public/download.php/14820/rp-19-2014_ODI.pdf), p.10.

données provenant de différentes sources, des résultats présentés par ordre de pertinence, une navigation par facettes et des fonctionnalités sociales<sup>31</sup>.

Ces interfaces sont avant tout des briques logicielles, elles ne contiennent pas elles-mêmes nativement leurs propres métadonnées. Leurs index sont alimentés le plus souvent par les notices extraites du SIGB, auxquelles peuvent s'ajouter éventuellement celles des archives institutionnelles et des listes A-Z de livres numériques ou de revues. Marshall Breeding signale que ces index peuvent avoir plusieurs millions de notices, mais qu'ils travaillent en général avec une granularité de niveau titre de l'ouvrage ou de la revue<sup>32</sup>.

Ces interfaces de découverte peuvent être des produits propriétaires ou *open source*. Parmi les produits propriétaires, on peut notamment signaler les suivants : Primo (d'Ex Libris), Enterprise (de SirsiDynix), BiblioCore (de BiblioCommons), Aquabrowser (de Proquest), Encore (Innovative interfaces). Ces interfaces de découverte sont produites par des sociétés productrices de logiciels de bibliothèque. Elles sont le plus souvent installées en local mais certaines, comme Primo, proposent aussi un mode d'installation en *cloud* (sur des serveurs délocalisés, accessibles en ligne).

Parmi les produits *open source*, Marshall Breeding distingue en particulier les produits suivants, conçus par des bibliothèques universitaires aux États-Unis : Blacklight (développé par l'université de Virginie), Vufind (développé par l'université Villanova), eXtensible Catalog (développé par l'université de Rochester), Franklin (développé par l'université de Pennsylvanie). Dans ce domaine, les deux produits majeurs sont Vufind et Blacklight. Ils s'appuient sur un code PHP pour Vufind et Ruby on Rails pour Blacklight, et utilisent des moteurs de recherche semblables, reposant sur une technologie Apache Lucene ou SOLR. Ces interfaces *open source* sont installées en local.

## B. Les services de découverte à index général centralisé

Les services de découverte à index général centralisé sont composés d'une interface de découverte telle que décrite précédemment et d'un index général centralisé, alimenté par des données de niveau universitaire : données d'éditeurs commerciaux, d'agrégateurs, de bases de références bibliographiques A&I (abstract et indexation), d'entrepôts en *open access*, auxquelles peuvent s'ajouter des ressources locales (SIGB, archive institutionnelle, bibliothèque numérique<sup>33</sup>, etc.). Il s'agit d'index de masse dont l'alimentation repose en grande partie sur des accords commerciaux, de nature privée, avec les fournisseurs de données<sup>34</sup>. Ces index ont fait leur apparition dans les années 2009-2010.

Comme le signale Marshall Breeding, il est important de distinguer ces index généraux centralisés des moteurs d'indexation décrits précédemment : le niveau de granularité dominant d'un index général centralisé n'est plus celui du titre d'ouvrage ou de revue mais celui du titre de l'article ou du chapitre, ce qui nous

<sup>31</sup> BREEDING, Marshall. Discovery product functionality. *Op. cit.*, p. 9-10. Voir également BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 2.

<sup>32</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p.17.

<sup>33</sup> L'expression « bibliothèque numérique » désignera dans ce travail les données relatives aux documents numérisés par une bibliothèque.

<sup>34</sup> Voir *infra* la partie I, 3, B « Les flux de données ».



place à une échelle beaucoup plus importante en termes de quantité de données, les index généraux centralisés pouvant contenir des milliards de références<sup>35</sup>.

Les services de découverte à index général centralisé sont donc, autant qu'une interface d'interrogation, un ensemble de données requêtables mises à disposition par le fournisseur. De ce fait, ils se trouvent dans une position charnière entre la brique logicielle et la ressource électronique.

À ce jour aucun service de découverte à index général centralisé n'a été créé sous licence *open source* sur le modèle de l'*open access*. Le marché de ces services de découverte à index général centralisé est intégralement composé de produits issus de fournisseurs commerciaux. Il s'agit de Primo Central Index d'Ex Libris, Ebsco Discovery Service (EDS) d'Ebsco, Summon de Proquest et Worldcat discovery service d'OCLC (auparavant Worldcat Local). Ces produits sont le fruit de sociétés éditrices de logiciels de bibliothèque ou de sociétés productrices de bases de données et/ou gestionnaires d'abonnements.

Si l'interface de découverte peut dans certains cas être installée en local, l'interrogation des données des index généraux centralisés se fait exclusivement en *cloud*. Ces outils mettent à disposition de la bibliothèque une interface d'administration ou base de connaissance (à travers laquelle les bibliothèques déclarent les ressources qu'elles souhaitent voir proposées à leurs utilisateurs), mais peuvent proposer aussi d'effectuer des recherches en dehors des ressources activées par la bibliothèque. Lorsque l'outil de découverte est fourni par le même fournisseur que le résolveur de liens, l'ERMS ou la liste A-Z, la bibliothèque pourra disposer d'une plateforme d'administration unique pour ces différents outils.

### C. Des outils mixtes : une interface et un index d'origines différentes

Un index général centralisé peut être articulé avec une interface de découverte qui ne lui est pas nativement liée, ce qui donne lieu à des produits hybrides qu'on désignera ici comme outils mixtes.

Ces outils mixtes peuvent être le fruit d'accords commerciaux entre des fournisseurs d'interface et des fournisseurs d'index qui choisissent de présenter une offre commune. L'outil Encore Duet articule ainsi l'interface produite par Innovative et l'index EDS, de même que l'outil Enterprise-EDS combine l'interface Enterprise de SirsiDynix et l'index EDS.

D'autre part, certaines bibliothèques ont mis en place leur propre outil mixte en combinant un index centralisé avec une interface libre : la bibliothèque de l'université de Chicago a ainsi implémenté Vufind avec EDS, celle de l'université Villanova, Vufind avec Summon, celle de l'université de Columbia, Blacklight avec Summon, et celle de l'université d'Alberta, Blacklight avec EDS<sup>36</sup>.

Dans le cas des outils mixtes, l'interface est installée en local, tandis que les données de l'index sont mises à disposition en *cloud* par le fournisseur commercial.

<sup>35</sup> *Ibid.* p. 17.

<sup>36</sup> POPOWICH, Sam. Agile open-source discovery : Blacklight with Ebsco Discovery Service. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 73-82 ; BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Library Technology Reports*. Mai 2015, Vol. 51, n° 4, p. 22-25 ; BREEDING, Marshall et National Information Standards Organization. *Op. cit.*

### 3. QUEL MODELE ECONOMIQUE POUR LES SERVICES DE DECOUVERTE A INDEX CENTRALISE?

#### A. Un marché fortement concurrentiel, évolutif et anglo-saxon

##### *Concentrations et convergences*

Le marché actuel des services de découverte à index général centralisé (et plus largement d'index centralisés pour les outils mixtes) est composé par un très petit nombre d'acteurs et présente un caractère fortement concurrentiel. Jusqu'à récemment ce marché comptait quatre fournisseurs majeurs (Ex Libris, Proquest, Ebsco et OCLC), dont trois seulement sont présents pour l'heure dans le paysage français (Ex Libris, Proquest, Ebsco). Cette tendance oligopolistique s'est accentuée en 2015 avec la fusion des sociétés Ex Libris et Proquest : désormais ils ne sont plus que trois fournisseurs, dont deux seulement présents en France.

Cette fusion récente soulève de nombreuses incertitudes : si la permanence des deux produits concernés a été ratifiée, on peut cependant s'interroger sur les modalités concrètes de ce maintien et sur le positionnement à long terme des deux produits au sein d'une même société. De nombreuses bibliothèques espèrent que de la fusion commerciale des deux sociétés découlera un gain qualitatif pour les deux produits, en particulier au niveau des métadonnées des index et des bases de connaissance dont la fusion a été annoncée, et peut être, à terme, sur les fonctionnalités et l'interopérabilité logicielle (pour les bibliothèques BIB26, BIB23, BIB19, BIB10, BIB18, BIB12, BIB01). Les bibliothèques espèrent également que de la fusion des supports techniques découlera un gain de savoir-faire et des échanges de pratiques.

Outre leur nombre extrêmement réduit, il faut noter que les fournisseurs qui composent ce marché font cohabiter au sein de leurs sociétés la production de services de découverte avec d'autres activités, notamment la production de logiciels de bibliothèque ou de contenus électroniques (bases de données).

Au niveau logiciel, la convergence au sein d'une même société de ces différentes branches d'activité peut être considérée *a priori* comme la garantie de l'interopérabilité. Une société qui produit un outil de découverte en même temps qu'un SIGB cherchera à rendre le premier compatible avec le second afin d'en assurer l'intégration avec une suite d'outils, dans une logique cohérente de *conceptual design*<sup>37</sup>. Il en va de même pour les bases de connaissance, résolveurs de liens et ERMS. Néanmoins, la très forte intégration de ces outils peut conditionner la stratégie d'une bibliothèque dans le choix de telle ou telle solution logicielle. Comme l'a pointé Marshall Breeding, la suite intégrée d'outils d'un seul fournisseur peut impliquer des avantages financiers et de maintenance, mais réduire la marge d'autonomie d'une bibliothèque en matière de paramétrages et de customisation de ses propres outils<sup>38</sup>.

Au niveau des contenus, une société qui produit ses propres contenus en plus d'un outil de découverte, aura bien sûr tout intérêt à rendre les premiers disponibles et visibles à travers le second. Si les fournisseurs affirment veiller à la

<sup>37</sup> BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Op. cit.*, p. 24-25.

<sup>38</sup> *Ibid.*, p. 24-25.

neutralité de leurs outils vis-à-vis de leurs contenus<sup>39</sup>, cette cohabitation a été cependant signalée par la NISO comme un facteur de risque pour la neutralité des outils : « une inquiétude particulière est liée à la possibilité d'un manque de neutralité, lorsque les services de découverte appartiennent à la même société mère que d'autres produits ou services »<sup>40</sup>.

En outre, dans ce cas de figure, une société aura tout intérêt à faire en sorte que les contenus qu'elle produit et qu'elle valorise à travers son propre outil ne soient pas mis à disposition d'outils concurrents. Ce problème a été signalé par de nombreux articles de la presse professionnelle<sup>41</sup>. On rappellera, en guise d'exemple, la controverse qui a opposé en 2014 les sociétés Ebsco et Ex Libris. Cette dernière a en effet avancé qu'Ebsco ne mettait pas les données des bases qu'elle produit ou diffuse à disposition d'autres outils concurrents, comme l'index Primo Central. La société Ebsco a argumenté pour sa part que l'intégration de ses données à des index concurrents, produits par d'autres fournisseurs, aurait entraîné une dégradation de la qualité de celles-ci<sup>42</sup>. Ainsi, comme le résumaient Lorena Ávila-García, Virginia Ortiz-Repiso et David Rodríguez-Mateos, la convergence d'activités jointe à une très grande concentration du marché peut conduire à délimiter le contenu de l'index général centralisé par des enjeux commerciaux au détriment des besoins documentaires des bibliothèques<sup>43</sup>.

### *Des produits évolutifs*

Ce contexte commercial fortement concurrentiel favorise, comme le note Marshall Breeding, des améliorations fréquentes qui suivent l'évolution des technologies de l'information<sup>44</sup>. En guise d'exemple, au cours même de la rédaction de ce travail, plusieurs bibliothèques ont connu ou allaient connaître, à très court terme, un changement de version de leurs outils (BIB25, BIB23, BIB18, BIB19, BIB26).

Ces produits adaptent progressivement leurs interfaces pour qu'elles soient consultables sur des dispositifs mobiles, soit grâce à des feuilles de style *responsive* soit grâce à des interfaces mobiles spécifiques, avec un résultat globalement satisfaisant pour de nombreuses bibliothèques (BIB23, BIB18, BIB19, BIB12, BIB22, BIB07, BIB14, BIB10, BIB20, BIB11, BIB17, BIB13).

Ces outils mettent par ailleurs à disposition des API, des *widgets* et des *plugins* qui permettent d'adapter les interfaces. Les *widgets* permettent d'intégrer

<sup>39</sup> *Primo - Open Discovery Initiative Conformance* [en ligne]. [s. d.]. [Consulté le 28 octobre 2016]. Disponible à l'adresse : <http://www.exlibrisgroup.com/?catid={8B84715C-234A-44A9-B3C1-DCDD2E18D049}#{FD99D14F-9A5C-43FB-BC5A-E716A661087C}> ; *ProQuest Summon / Web-scale discovery service increases the value of your library through an unprecedented research experience - The Summon® Service* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://www.proquest.com/products-services/The-Summon-Service.html>. Entretien téléphonique avec Mathieu Ponthenier (Ebsco). 16 septembre 2016.

<sup>40</sup> « Special concerns surround the possibility of bias when discovery services are owned by the same corporate parent as content products or services » National Information Standards Organization et Open Discovery Initiative Working Group. p. 2.

<sup>41</sup> ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.*, p. 9. BREEDING, Marshall. Discovery product functionality. *Library Technology Reports*. 2015, Vol. 50, n° 1, p. 14-15. BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Op. cit.*, p. 24. THOMPSON, JoLinda. *Op. cit.*, p. 37-38. HOEPPNER, Athena. The Ins and Outs of Evaluating Web-Scale Discovery Services. *Computers in Libraries*. Avril 2012, Vol. 32, n° 3, p. 9.

<sup>42</sup> BREEDING, Marshall. Discovery product functionality. *Op. cit.*, p. 14-15.

<sup>43</sup> ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.*, p. 9.

<sup>44</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 5.

facilement une barre de recherche de l'outil de découverte dans un site web donné. Les *plugins* mis à disposition par les outils peuvent servir à réaliser l'intégration des données d'un index de découverte à l'environnement numérique d'une bibliothèque<sup>45</sup>. Les API permettent d'échanger des données directement entre logiciels totalement distincts : elles reçoivent la requête d'un autre logiciel et lui répondent en lui renvoyant les données demandées. Ces API sont intéressantes pour les bibliothèques car elles leur donnent la possibilité de construire leurs propres interfaces sur mesure, adaptées à leurs besoins. Ainsi, lorsqu'un index de découverte met à disposition une API, celle-ci peut être interrogée par une autre couche logicielle afin d'afficher les résultats renvoyés par l'API selon une configuration spécifique à la bibliothèque qui pourra être différente de la configuration standard de l'outil de découverte (BIB21, BIB26). C'est également la présence d'une API qui permet l'articulation d'un index de découverte avec une interface non nativement liée dans le cadre d'outils mixtes. D'après Marshall Breeding, l'exploitation de ces API constitue un domaine particulièrement prometteur pour le développement d'outils mixtes avec des interfaces libres qui interrogent et retraitent les données d'index centralisés commerciaux<sup>46</sup>.

### *Des produits très liés au monde anglo-saxon*

Ces outils sont fortement marqués par le monde anglo-saxon dont ils sont issus.

Au niveau des formats de notices utilisés pour intégrer les données locales des bibliothèques, la gestion des notices ne se fait pas nécessairement en priorité en UNIMARC mais peut se faire dans d'autres formats MARC (MARC21, USMARC) qui nécessitent parfois des adaptations importantes de la part des bibliothèques (BIB13).

Par ailleurs, la plupart des interfaces publiques ainsi que des interfaces d'administrateur sont produites nativement en anglais. Les bibliothèques ont une marge de manœuvre plus ou moins étendue pour la traduction des libellés. Selon les cas, elles pourront effectuer elles-mêmes les traductions sur l'interface d'administration ou adresser leurs demandes au support technique. En règle générale, les nouvelles fonctionnalités sont d'abord disponibles en anglais et ensuite transposées en version française.

La traduction des interfaces peut représenter un travail important de la mise en œuvre, en particulier pour les premières installations d'outils en France (BIB12, BIB25) et, pour certaines installations récentes, des bibliothèques ont indiqué que leur travail de traduction n'était pas achevé au moment de l'entretien (BIB07, BIB14). Si dans un grand nombre de cas, la traduction des interfaces a été jugée satisfaisante (BIB20, BIB18, BIB11, BIB17, BIB15, BIB24), à plusieurs reprises des bibliothèques ont signalé avoir rencontré des problèmes à ce niveau avec des libellés non traduits ou incohérents, notamment pour ce qui concerne l'affichage des données issues de l'index centralisé (BIB22, BIB02, BIB21, BIB10, BIB25). Selon les modalités de maintenance choisies, il peut être nécessaire pour la bibliothèque de se concerter avec d'autres utilisateurs francophones relativement

<sup>45</sup> Voir à cet égard le retour d'expérience des bibliothèques du Swiss Education Group : PEEL, Lisa. KOHA/EDS application at César Ritz Colleges. Dans : *Symposium Koha 2016* [en ligne]. Université Jean Moulin Lyon 3, 28 juin 2016. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <https://youtu.be/-T5D48lilqM>.

<sup>46</sup> BREEDING, Marshall. Discovery product functionality. *Op. cit.*, p.28.

éloignés pour la traduction des intitulés, ce qui n'est pas optimal en termes de *workflow* (BIB22). Pour pallier ces dysfonctionnements, certaines bibliothèques utilisent des surcouches logicielles pour adapter les libellés à leurs besoins (BIB26, BIB21).

Les établissements qui disposent d'outils mixtes avec une interface propriétaire peuvent également rencontrer des difficultés de traduction. Les bibliothèques ont signalé que des libellés issus des données du catalogue pouvaient être correctement traduits sur l'interface de découverte, mais que la traduction pouvait être inopérante sur la partie qui interroge l'index centralisé (BIB13, BIB08). Dans ce cas, même quand des traductions sont disponibles pour l'index centralisé dans l'interface native de l'outil de découverte, c'est l'intégration de ces traductions dans le cadre de l'outil mixte qui s'avère difficile à réaliser.

Enfin, si ces outils permettent de prendre en compte correctement des caractères non latins dans la recherche et l'affichage des résultats (BIB15, BIB10, BIB20, BIB23), la gestion des diacritiques et des apostrophes a pu néanmoins poser des problèmes dans un certain nombre de cas (BIB23, BIB25, BIB24, BIB17, BIB08).

## B. Les flux des données

Les données qui alimentent l'index général centralisé (hors données locales des bibliothèques) proviennent de différentes sources : éditeurs ou agrégateurs, bases d'abstract et d'indexation (A&I), ressources en *open access*. Pour chaque fournisseur de données, l'alimentation de cet index implique un circuit de données particulier depuis la donnée source jusqu'à l'index centralisé.

### *La contractualisation*

Pour les contenus commerciaux, non disponibles en *open access*, les données sont mises à disposition dans les index des outils de découverte à travers des contrats de société à société entre les fournisseurs de données et les fournisseurs d'outils. Comme le résume la NISO, « les producteurs de ces services de découverte à index général centralisé passent des accords avec les fournisseurs de contenus à destination des bibliothèques afin de pouvoir recueillir les métadonnées bibliographiques et/ou le texte intégral de leurs ressources »<sup>47</sup>.

C'est le fournisseur de données, en tant que propriétaire de celles-ci qui choisit de quelle façon et avec quel fournisseur d'outils contractualiser ses données. Ainsi, la fourniture des données dépend avant tout de la volonté du fournisseur de données : « l'étendue du contenu indexé dans les services de découverte à index centralisé est intrinsèquement limitée par le périmètre dans lequel ces sociétés sont disposées à réaliser ces accords commerciaux d'échanges de données »<sup>48</sup>. Ces accords privés de société à société sont le plus souvent marqués par la confidentialité et le secret commercial. Comme le résume Marshall Breeding, « le caractère privé de ces accords d'échange de données entre

<sup>47</sup> « The creators of these index-based discovery services make arrangements with library-oriented content providers to receive citation and/or full-text metadata of their resources. » National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 11.

<sup>48</sup> « The universe of discoverable content through these index-based products is inherently limited by the extent to which commercial entities are willing to make these private agreements for data exchange ». BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p.14.

fournisseurs de contenus commerciaux et fournisseurs de services de découverte a attiré l'attention sur le besoin d'une plus grande transparence concernant les données utilisées pour alimenter ces index et la façon dont celles-ci sont exploitées dans le service de découverte.»<sup>49</sup> Il est donc tout à fait courant qu'une bibliothèque abonnée à une ressource n'ait pas accès aux données de cette ressource via son outil de découverte car il n'existe pas d'accord commercial entre le fournisseur de données pour cette ressource et le fournisseur d'outils<sup>50</sup>.

Ces données sont mises à disposition dans les outils de découverte selon des modalités différentes d'un fournisseur d'outil à un autre et d'une ressource à une autre. Interrogés sur leurs pratiques contractuelles, les fournisseurs d'outils ont donné des réponses variables : selon le fournisseur d'outil, ces contrats peuvent être plus ou moins systématiques et formalisés. Ainsi, Ebsco a déclaré n'intégrer aucune donnée sans contractualiser au préalable le versement avec les fournisseurs de données, y compris lorsqu'il s'agit de ressources en *open access*. Cela est considéré comme une garantie à la fois pour le fournisseur de données et pour les fournisseurs d'outils, ainsi que pour l'utilisateur final. Pour Ebsco, la qualité minimale des données versées, le format et le protocole de versement, la fréquence des mises à jour et le traitement des données dans l'index central après versement font l'objet de contractualisation, avec des périmètres variables selon les contrats. Pour Proquest, dans le cadre de Summon, la contractualisation de la fourniture de données se fait au cas par cas, non de manière standard, et ces différents éléments ne sont pas nécessairement contractualisés à chaque fois de manière explicite. Dans ce cas, cela relève plutôt d'un échange de bons procédés, sur le principe du donnant-donnant, les fournisseurs de données ayant eux-mêmes intérêt à voir leurs données indexées dans l'outil de découverte. Il s'agit donc davantage d'une démarche qui peut varier selon les cas et qui dépend de chaque fournisseur de données.

On peut toutefois se demander dans quelle mesure le fournisseur de données peut choisir de manière indépendante un outil, lorsque ses données sont elles-mêmes diffusées à travers des plateformes gérées par la même société qui produit l'outil. En effet, on peut considérer avec JoLinda Thompson et Marshall Breeding que ce double positionnement des fournisseurs d'outils-fournisseurs de données peut être un frein à la mise à disposition des données à des outils concurrents et peut conduire de fait à une situation d'exclusivité<sup>51</sup>. Ce problème d'exclusivité a également été soulevé par la NISO : « les inquiétudes concernent aussi les accords exclusifs ou d'autres rapports commerciaux entre un service de découverte et un fournisseur de contenu qui pourraient introduire une entorse à la neutralité »<sup>52</sup>. Interrogés sur l'existence d'accords exclusifs avec certains fournisseurs de données qui empêcheraient ces derniers de mettre leurs données à disposition d'outils concurrents, les fournisseurs ont mis en avant leur politique de

<sup>49</sup> « The private nature of data exchange agreements between commercial content providers and discovery service creators evoked an interest in more transparency regarding the data used to populate these indexes and how the data are exercised within the discovery services ». *Ibid.*, p. 14.

<sup>50</sup> THOMPSON, JoLinda. *Op. cit.*, p. 132-133. BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 28.

<sup>51</sup> THOMPSON, JoLinda. *Op. cit.*, p. 37-38. . BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 27-28.

<sup>52</sup> « Concerns also arise through exclusive arrangements or other business relationships made by a discovery service with a content provider that might introduce bias ». National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 2.

transparence et se sont défendus de tout accord d'exclusivité avec des fournisseurs de données. Ainsi, Proquest, dans le cadre de Summon, a indiqué n'avoir aucun accord d'exclusivité avec des fournisseurs de données. Pour sa part, Ebsco a indiqué que c'est le fournisseur de données qui choisit de verser ses données à tel ou tel fournisseur d'outil : il peut y avoir en effet des fournisseurs de données qui ont opté pour verser leurs données seulement à Ebsco. D'après Ebsco, s'il y a des exclusivités dans les accords commerciaux avec les fournisseurs de données, cela relève entièrement du choix du fournisseur de données et en aucun cas d'une clause imposée *a priori* par le fournisseur d'outil.

Concernant les droits d'accès aux données versées, l'accès au texte intégral sera bien sûr limité en règle générale aux utilisateurs reconnus, mais dans de nombreux cas l'accès aux métadonnées reste possible pour tout utilisateur. De nombreux fournisseurs de données ne limitent pas l'accès aux métadonnées uniquement aux établissements abonnés à la ressource. C'est bien évidemment le cas des ressources en *open access* pour lesquelles les métadonnées et le texte intégral sont librement consultables en ligne. C'est également le cas pour beaucoup de ressources sous droits pour lesquelles tout utilisateur de l'outil de découverte, même extérieur à une institution, pourra consulter les métadonnées, sans avoir accès au texte intégral. Mais il arrive que les métadonnées elles-mêmes fassent l'objet de contractualisation avec des restrictions d'accès<sup>53</sup>. Interrogés sur ce point, les fournisseurs ont donné des réponses variables. Proquest a indiqué que, dans le cadre de Summon, toutes les métadonnées sont consultables en mode non authentifié. Mais pour Ebsco qui intègre des métadonnées de bases de références bibliographiques, d'abstracts et indexation (A&I), il y a bien des restrictions d'accès. Dans ce cas, les contrats stipulent que la mise à disposition des métadonnées de la ressource dans l'outil de découverte sera limitée aux clients de la ressource : seuls les utilisateurs authentifiés pourront consulter ces métadonnées.

### *Le cas des bases d'A&I*

La valeur ajoutée de ces bases d'A&I réside dans les métadonnées elles-mêmes : si elles sont rendues publiques telles quelles, c'est tout l'intérêt de la base qui est remis en question. Ces bases d'A&I posent donc un problème spécifique aux outils de découverte car elles doivent continuer à prouver leur valeur ajoutée pour préserver leur intérêt économique. C'est ce qui explique la présence pour elles de ces clauses de restrictions à l'ouverture des données (*non-disclosure agreements*)<sup>54</sup>. Cela a de multiples conséquences à la fois sur l'accès (consultation des notices bibliographiques seules en mode non authentifié) et sur le dédoublement (comment dédoublement pour les clients et pour les non-clients). Il s'agit là d'un système de métadonnées protégées au sein de l'index centralisé, limitées dans leur utilisation aux seuls clients de ces plateformes.

Ce positionnement particulier des bases de données d'A&I a été relevé comme un point de vigilance par la NISO :

« Les services de découverte doivent limiter l'accès au contenu sous droits, tel que les résumés et les mots clés spécifiques, à des utilisateurs authentifiés rattachés à des établissements abonnés à la ressource. Au regard

<sup>53</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p.19.

<sup>54</sup> National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p.22.

de ces éléments, parmi beaucoup d'autres, les ressources d'A&I doivent faire l'objet d'une attention spécifique au sein de l'ensemble des fournisseurs de contenus qui peuvent potentiellement intéresser les services de découverte<sup>55</sup> ».

Cette spécificité a également été signalée par Marshall Breeding :

« La couverture de ressources d'A&I ou d'autres thesaurus spécialisés joue un rôle beaucoup plus complexe et controversé dans le domaine des services de découverte. [...] Les fournisseurs de ces produits sont souvent préoccupés par le fait que leur contenu sous droits puisse devenir disponible pour des utilisateurs non abonnés et, au-delà, parce que l'idée même d'un index de découverte peut affaiblir l'intérêt pour leurs ressources<sup>56</sup> ».

Le positionnement des bases d'A&I en tant que fournisseurs de données pour les outils de découverte n'est donc pas facile à trouver car ces derniers viennent bouleverser leur équilibre économique.

La *National Federation of Abstracting and Information Services* (NFAIS) qui rassemble les différents fournisseurs de bases d'A&I s'est fait bien sûr le relais de leurs préoccupations à ce sujet. Elle signale le risque qu'à travers les outils de découverte, les bases d'A&I puissent perdre leur pertinence sur le marché et leur valeur ajoutée si leurs données sont purement et simplement intégrées aux index des outils de découverte<sup>57</sup>. La NFAIS plaide pour l'inclusion de contenu contractuel seulement (respect des accords passés par les fournisseurs de données, en fonction de son désir de voir ses données exposées ou pas dans un outil de découverte) et pour la limitation de l'accès au contenu aux utilisateurs autorisés<sup>58</sup>.

La NFAIS relève que la spécificité des bases d'A&I est de contenir des termes d'indexation basés sur des thesaurus bien précis ainsi que des abstracts qui peuvent ne pas être disponibles par ailleurs et pour lesquels l'intégration dans un outil de découverte entraîne le risque d'une déperdition d'information. J. Thompson note également que parfois les métadonnées disponibles dans les bases A&I sont restructurées par la normalisation qu'impose leur intégration dans l'index centralisé. Ainsi, précise-t-elle, certaines bases d'A&I ne fournissent pas leurs données aux outils de découverte par peur que leurs articles ne soient pas exploitables de manière correcte, comme la *Bibliography of Asian studies*<sup>59</sup>.

On aboutit ainsi à la situation paradoxale décrite par la NISO où les multiples données d'une même unité documentaire peuvent être potentiellement fournies par différents fournisseurs de données : pour un article, le texte intégral pourra être disponible via l'éditeur principal, les métadonnées bibliographiques via une base de

<sup>55</sup> « Discovery services must limit access to proprietary content, such as abstracts and specialized vocabularies to authenticated users affiliated with mutually subscribing institutions. Given these factors, among many others, A&I resources must be treated with special consideration within the body of content providers that potentially contribute to discovery services ». *Ibid.*, p. 10.

<sup>56</sup> « The coverage of indexing and abstracting services or other subject indexes plays a much more complex and controversial role in the discovery services arena. [...] The providers of these products are often concerned that their proprietary content may become available to nonsubscribers and that the general idea of index-based discovery may weaken interest in their products. » BREEDING, Marshall. *Discovery product functionality*. *Op. cit.*, p. 14.

<sup>57</sup> National Federation of Abstracting and Information Services. *Recommended Practices: Discovery Services* [en ligne]. Philadelphia, PA : National Federation of Abstracting and Information Services, 2013. [Consulté le 20 mars 2016]. Disponible à l'adresse : [https://nfais.memberclicks.net/assets/docs/BestPractices/recommended\\_practices\\_final\\_aug\\_2013.pdf](https://nfais.memberclicks.net/assets/docs/BestPractices/recommended_practices_final_aug_2013.pdf), p. 3.

<sup>58</sup> *Ibid.*, p. 6-7.

<sup>59</sup> THOMPSON, JoLinda. *Op. cit.*, p. 39.



données d'agrégateur et les abstracts et le vocabulaire contrôlé via encore un autre fournisseur d'A&I, selon des modalités d'accès et des contrats différents<sup>60</sup>. Pour Marshall Breeding, l'absence d'intégration d'un grand nombre de bases d'A&I aux index centralisés, due à ces spécificités, reste une des lacunes les plus importantes à combler pour les outils de découverte<sup>61</sup>.

### **L'open access**

L'indexation des ressources en *open access* peut permettre d'éviter des négociations complexes avec un éditeur commercial. Néanmoins, les ressources en *open access* pourront aussi, selon les fournisseurs d'outils, faire l'objet d'une contractualisation pour la fourniture des données, comme le fait Ebsco.

Cette contractualisation peut être d'autant plus importante que les ressources en *open access* sont de nature extrêmement variée : archives disciplinaires, archives institutionnelles, revues en *open access* ou ressources en *open access* incluses dans des packages commerciaux auxquels la bibliothèque peut ne pas être abonnée.

Les ressources en *open access* sont le plus souvent moissonnables via le protocole OAI-PMH. Or, comme le rappelle François Renaville, ces ressources peuvent faire l'objet de moissonnages de manière directe ou indirecte. Ainsi, même lorsque l'entrepôt qui est la source initiale d'une donnée n'est pas moissonné directement, la donnée peut se retrouver dans l'index centralisé lorsque cet entrepôt initial a été moissonné par un entrepôt intermédiaire qui, lui, est bien intégré dans l'index centralisé<sup>62</sup>. Dans le flux de données de l'*open access*, il peut donc parfois être difficile de cibler ou d'identifier la source initiale d'une donnée en particulier.

## **C. Le modèle SaaS**

### **Des services en cloud**

Si certaines interfaces de découverte peuvent être installées en local comme par exemple Primo ou Encore pour l'outil mixte Encore-Duet, les services de découverte à index général centralisé et les données des index centralisés sont commercialisés sur le modèle de *software as a service* (abrégé *SaaS*). Ce modèle correspond à un des grands modèles de fourniture de services informatiques en *cloud*<sup>63</sup>. Sur le mode *SaaS*, les clients n'achètent pas une licence ou un logiciel mais s'abonnent à des services distants : les sociétés hébergent et maintiennent sur leurs serveurs les données des index centralisés ainsi que des interfaces, tandis que les bibliothèques et les usagers accèdent à l'outil à travers un navigateur web<sup>64</sup>. Ce modèle repose ainsi sur l'externalisation de l'infrastructure et des données auprès d'un fournisseur qui dessert plusieurs clients en même temps, c'est pourquoi ce

<sup>60</sup> National Information Standards Organization et Open Discovery Initiative Working Group, *Op. cit.*, p. 2.

<sup>61</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 28.

<sup>62</sup> RENAVILLE, François. Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 242.

<sup>63</sup> Sur les distinctions détaillées entre le modèle SaaS et les modèles *infrastructure as a service* (IaaS) et *platform as a service* (PaaS), voir SINGH, Saurabh, JEONG, Young-Sik et PARK, Jong Hyuk. A survey on cloud computing security: Issues, threats, and solutions. *Journal of Network and Computer Applications* [en ligne]. Novembre 2016, Vol. 75, p. 201.

<sup>64</sup> ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.* p. 5.

type de service est souvent qualifié de *multi-tenant* (une plateforme partagée entre différents clients). En dehors des outils de découverte, il peut concerner plusieurs produits destinés aux bibliothèques, tels que les SIGB ou d'autres systèmes de gestion pour les bibliothèques<sup>65</sup>.

La fourniture de service en *cloud* sur le modèle *SaaS* peut présenter plusieurs atouts pour les bibliothèques. D'abord, la connexion simultanée de nombreux clients à travers le monde à un même service peut être considérée comme une garantie de qualité car elle oblige le fournisseur à assurer une continuité de service et à avoir une infrastructure de sécurité qui le préserve des pannes (BIB23, BIB18). Comme le signalait Marshall Breeding, « la plupart d'offres en *SaaS* concernent plusieurs institutions ou particuliers qui partagent la même instance d'un logiciel. Cette approche *multi-tenant* permet au fournisseur de fournir la maintenance pour tous ses clients à travers une seule plateforme à grande échelle, sur laquelle toutes les mises à jour et les améliorations peuvent être effectuées en une seule fois, pour tous »<sup>66</sup>. Qui plus est, en *cloud*, les mises à jour du système sont effectuées régulièrement par le fournisseur ce qui affranchit la bibliothèque de cette tâche d'infrastructure informatique (BIB18), atout qui est également relevé par Marshall Breeding : « puisque le fournisseur de service prend en charge les améliorations, les mises à niveau et les correctifs, les clients sont dispensés de cette lourde responsabilité qu'il leur faudrait assurer avec un logiciel installé en local »<sup>67</sup>. Bien au contraire, lorsque des interfaces sont installées en local, il pourra y avoir des écarts de mise à jour du système plus importants entre la version en *cloud* et la version en local (BIB26, BIB06).

Cependant, les services proposés selon le modèle *SaaS* peuvent comporter, comme tous les services reposant sur le *cloud*, plusieurs risques de sécurité inhérents à ce dernier. L'Agence Européenne chargée de la sécurité des réseaux et de l'information (ENISA) a ainsi identifié huit risques de sécurité majeurs pour les services en *cloud*<sup>68</sup> :

- la perte de gouvernance des clients : le client cède en grande partie au fournisseur le contrôle et la maîtrise de son infrastructure ;
- le verrouillage technique : l'absence d'outils, de procédures ou de standards peut limiter la possibilité du client de migrer vers un autre fournisseur ou vers une instance locale, ce qui est une source de dépendance vis-à-vis d'un fournisseur en particulier ;
- des failles d'isolation des plateformes ou ressources *multi-tenant* : le risque de faille dans les mécanismes qui servent à séparer le stockage des données des différents clients ;
- des risques pour la certification : la migration vers le *cloud* peut compromettre un travail de certification si le fournisseur n'est pas en

<sup>65</sup> BREEDING, Marshall. A Cloudy Forecast for Libraries. *Computers in Libraries*. Septembre 2011, Vol. 31, n° 7, p. 34.

<sup>66</sup> « Most software-as-a-service offerings involve many organizations or individuals sharing a single instance of the software. This multitenant approach allows the provider to support all its customers on a single highly scalable platform, where all updates and enhancements can be applied once and for all. » *Ibid.*, p. 33.

<sup>67</sup> « Since the service provider takes care of enhancements, upgrades, and patches, users are relieved from this burdensome responsibility as it would apply to local software. » *Ibid.*, p. 33.

<sup>68</sup> European Union Agency for Network and Information Security. *Cloud Computing Risk Assessment* [en ligne]. [S. l.] : European Union Agency for Network and Information Security, novembre 2009. [Consulté le 4 novembre 2016]. Disponible à l'adresse : <https://www.enisa.europa.eu/publications/cloud-computing-risk-assessment>. p. 9-10. Voir également à ce sujet SINGH, Saurabh, JEONG, Young-Sik et PARK, Jong Hyuk. *Op. cit.*

mesure de prouver sa propre certification ou si le fournisseur n'autorise pas l'audit nécessaire pour la certification du client ;

- une sécurité insuffisante des interfaces d'administration : les interfaces d'administration sont accessibles sur internet et permettent des accès étendus à des paramètres, ce qui représente un risque important pour la sécurité, notamment quand elles sont accessibles à distance ;
- des risques pour la protection des données : le *cloud* implique de nombreux risques au niveau de la protection des données des clients et il peut être difficile pour le client de vérifier effectivement la politique de conservation des données du fournisseur et d'être sûr que les données sont utilisées de manière conforme à la loi ;
- des risques pour la suppression des données : en cas de demande de suppression de données de la part d'un client, celui-ci peut difficilement vérifier leur suppression effective et celle-ci peut être impossible à réaliser lorsque les données d'un client partagent un même espace de stockage avec celles d'autres clients ;
- des risques accrus en cas d'intrusion : en cas d'une intrusion dans le système, du fait du partage de données de plusieurs clients, les conséquences d'une intrusion sont décuplées.

### *La tarification*

Comme le signalaient Lorena Ávila-García, Virginia Ortiz-Repiso et David Rodríguez-Mateos, chaque société a son propre modèle de tarification pour les services de découverte<sup>69</sup>. Les entretiens menés auprès des bibliothèques ont fait apparaître la grande variabilité de cette tarification. Il semble en effet très difficile de déterminer le prix d'un outil de découverte, dans la mesure où celui-ci est rarement vendu seul mais inclus dans des packages à géométrie très variable. Ainsi, l'outil pourra être compris dans une proposition commerciale qui comportera aussi

- une suite logicielle complète : SIGB, résolveur de liens, ERMS, moteur de recherche fédérée ;
- une base de connaissance avec un résolveur de liens et une liste A-Z ;
- un abonnement à une ressource électronique : base de données, bouquets de revues ou de livres numériques ;
- une offre couplée mixte : une interface de découverte d'un fournisseur et un index centralisé d'un autre fournisseur.

De ce fait, il est très difficile d'établir des comparaisons tarifaires car il y a rarement deux packages identiques. Il faut noter aussi à ce propos que des connecteurs peuvent être nécessaires pour faire fonctionner ensemble les différents éléments de ces packages, par exemple pour articuler les données de la base de connaissance avec l'outil de découverte (BIB07, BIB14, BIB25). Or ces connecteurs peuvent ne pas être compris dans le prix global du package proposé et leur coût pourra être aussi onéreux que celui du package complet.

De même, il est difficile de déterminer clairement les critères de tarification. Ces derniers sont variables d'une société à une autre et peuvent être liés à un

<sup>69</sup> ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David *Op. cit.*, p. 8.

nombre d'utilisateurs, à la taille de l'institution, au nombre de notices locales moissonnées dans l'entrepôt<sup>70</sup>. Mais plusieurs autres paramètres peuvent entrer en ligne de compte pour la tarification. Ainsi, lorsque l'interface de découverte peut être installée en local et en *cloud*, les prix pourront différer selon le type d'installation. Marshall Breeding signale à ce sujet que « le coût d'abonnement annuel à un produit fourni en *SaaS* est en général plus élevé que les frais de maintenance liées à une application hébergée en local »<sup>71</sup>. Les moissonnages de bases extérieures prévus dès l'origine par la bibliothèque peuvent être compris dans le prix global (BIB04, BIB12) ou dans le cas contraire des frais pourront s'appliquer (BIB24). L'intégration de fonctionnalités particulières, comme des outils de bibliométrie ou de recommandations, pourra être facturée de manière distincte ou intégrée par défaut (BIB20, BIB02, BIB04, BIB23). Le fait que la bibliothèque soit déjà cliente par ailleurs d'un fournisseur n'entraînera pas nécessairement de remise tarifaire et inversement des remises commerciales peuvent être accordées aux nouveaux clients, surtout s'il s'agit des premières installations (BIB21, BIB17, BIB12). Pour les consortiums, le nombre de catalogues à intégrer, le nombre d'entrepôts institutionnels, le nombre d'instances du résolveur de liens et de l'outil de découverte peuvent également avoir un impact sur la tarification (BIB04, BIB15, BIB10).

## D. La maintenance

En ce qui concerne la maintenance, en mode *SaaS*, celle-ci n'est pas explicitement distincte de l'abonnement et ses modalités d'exécution peuvent être très variables. C'est en particulier pour les outils mixtes, fruit d'un accord conjoncturel entre deux sociétés que la question de la maintenance se pose de manière problématique sur le long terme (BIB13).

C'est pourquoi il a semblé important d'analyser plus précisément les modalités de maintenance. Pour cela, nous avons demandé aux bibliothèques de commenter la prise en charge de leur outil de découverte au regard des quatre items suivants : l'écoute des besoins, la réactivité, les développements à la demande et la qualité des solutions fournies. Les réponses que nous avons reçues sur ces quatre points ont été très hétérogènes, du très élogieux au très sévère, avec une grande variabilité au sein d'un même item et pour un même fournisseur. On est donc très loin d'un consensus en la matière.

Cela peut s'expliquer en partie par le fait que pour un même outil, il peut y avoir des modalités de prise en charge différentes. Notamment lorsqu'un même outil peut être installé en *cloud* ou en local, il peut exister plusieurs formules de maintenance possibles : selon les besoins de la bibliothèque, celle-ci pourra choisir de déléguer une grande partie de la gestion de la maintenance au fournisseur ou bien se réserver une plus grande latitude d'intervention sur ses instances.

Cela peut s'expliquer aussi par le fait que la composition du support et les conditions de prise en charge peuvent varier pour un même outil et une même bibliothèque entre la phase d'implémentation et la phase de production courante : la bibliothèque n'aura pas nécessairement affaire aux mêmes équipes d'une phase à

<sup>70</sup> *Ibid.*, p.8.

<sup>71</sup> « The annual subscription cost for a SaaS product is generally higher than the maintenance fees associated with a locally hosted application ». BREEDING, Marshall. Introduction and Concepts. *Library Technology Reports*. Juin 2015, Vol. 51, n° 4, p. 11.

une autre (BIB07, BIB14, BIB25, BIB05, BIB24). Si plusieurs bibliothèques dans ce cas ont jugé correct l'accompagnement en phase de mise en œuvre, elles ont été plus réservées sur le support en phase de fonctionnement courant.

### *L'écoute et la réactivité*

Pour ce qui est de l'écoute des besoins, quatorze bibliothèques, tous fournisseurs confondus, ont apprécié positivement – bonne, très bonne, voire excellente – l'écoute de leurs besoins de la part de leurs fournisseurs (BIB12, BIB01, BIB19, BIB18, BIB23, BIB04, BIB07, BIB14, BIB20, BIB15, BIB05, BIB17, BIB11, BIB13). Trois bibliothèques n'ont pas commenté cet item (BIB26, BIB16, BIB08). Néanmoins, l'écoute des besoins a été évaluée de manière plus mitigée – moyenne, mauvaise, voire très mauvaise – par sept bibliothèques, tous fournisseurs confondus (BIB22, BIB06, BIB10, BIB25, BIB02, BIB24, BIB21).

Concernant la réactivité, onze bibliothèques, tous fournisseurs confondus, ont apprécié de manière positive la réactivité de leur support (BIB01, BIB19, BIB18, BIB04, BIB20, BIB02, BIB05, BIB17, BIB13, BIB08, BIB14). Au contraire, dans douze cas, tous fournisseurs confondus également, les bibliothèques se sont montrées plus critiques, sévères voire très sévères au sujet de la réactivité de leur support (BIB12, BIB23, BIB22, BIB06, BIB26, BIB07, BIB10, BIB25, BIB15, BIB24, BIB21, BIB11). Une bibliothèque n'a pas commenté cet item (BIB16). C'est donc là un point qui, encore plus que le précédent, est loin de faire consensus.

Ces réponses amènent plusieurs remarques. D'abord, le constat unanime des bibliothèques est l'absence de supports techniques en France pour les outils de découverte, ce qui rend nécessaire de passer par des supports internationaux, localisés à l'étranger, y compris dans des pays situés en dehors de l'Union Européenne. Cela peut impliquer, outre des décalages horaires conséquents, une lenteur structurelle pour la maintenance et la résolution des problèmes. Ainsi, de nombreuses bibliothèques ont signalé que si la prise en charge d'une question par le support peut être rapide, il peut cependant y avoir des délais très longs pour le diagnostic du problème et la mise en place d'une solution effective (BIB12, BIB23, BIB26, BIB06, BIB24, BIB11, BIB21, BIB10, BIB05, BIB02, BIB19, BIB22, BIB07, BIB20, BIB25, BIB13). Ces délais peuvent s'étaler de deux mois pour une réponse simple (simple demande d'information) à plus d'un an et demi pour des solutions impliquant des développements de fonctionnalités. Certaines opérations, notamment l'intégration des données d'un SIGB peuvent prendre à elles seules plusieurs mois<sup>72</sup> (BIB11, BIB02, BIB24). À plusieurs reprises, les bibliothèques ont signalé qu'elles trouvaient les réponses avant la réponse du fournisseur soit en mettant en place une solution de contournement, soit en s'adressant à d'autres bibliothèques (BIB24, BIB21, BIB26). Cet éclatement géographique des supports et ces délais de mise en œuvre des solutions techniques se ressentent sur la communication avec les fournisseurs.

La communication avec les supports internationaux se fait généralement en anglais (BIB01, BIB19, BIB18, BIB23, BIB24, BIB10, BIB11, BIB21, BIB05, BIB02, BIB22, BIB07, BIB15, BIB13). La rédaction du message de dysfonctionnement directement en anglais peut faciliter la prise en charge par le support (BIB18, BIB19, BIB01, BIB23) mais cela peut aussi compliquer la

<sup>72</sup> Voir *infra* I, 4, C « Les besoins en termes de ressources humaines ».

communication lorsque les problèmes à traiter sont d'un niveau très technique (BIB13, BIB14). Dans certains cas, les bibliothèques ont à leur disposition une base de connaissances des tickets précédemment traités, ce qui est particulièrement utile (BIB26, BIB18, BIB01, BIB23, BIB19, BIB12, BIB04). Mais très souvent les échanges se font par courriel, ce qui ne facilite pas un aussi bon suivi des incidents. La délocalisation des supports et le prisme anglo-saxon des produits peut ralentir la compréhension d'un problème par le fournisseur, notamment lorsque celle-ci nécessite d'appréhender des spécificités locales françaises (BIB13, BIB20, BIB15). Plusieurs bibliothèques ont par ailleurs indiqué qu'elles n'étaient pas toujours prévenues suffisamment en amont de l'intervention des supports techniques sur leurs instances, découvrant ces interventions en cours de route ou en étant mises devant le fait accompli (BIB15, BIB11, BIB22, BIB25).

En outre, les bibliothèques ont signalé que lorsque plusieurs outils sont en jeu, par exemple le SIGB et l'outil de découverte, des difficultés de communication entre les différents supports peuvent surgir, y compris au sein d'un même fournisseur (BIB23, BIB22, BIB12, BIB24, BIB21). En effet, il n'y a pas toujours d'interfaçage entre les supports des différents produits d'une même société, ce qui conduit parfois les bibliothèques à devoir signaler un même dysfonctionnement à plusieurs supports différents d'un même fournisseur. Ces problèmes se complexifient dans le cadre d'outils mixtes ou d'outils consortiaux, qui font rentrer en jeu des supports de fournisseurs différents et où la compétence de chacun n'est pas clairement délimitée (BIB08, BIB13, BIB15, BIB04). Dans ces cas, comme le signale Marshall Breeding, le partage des responsabilités entre deux fournisseurs n'est pas facile à réaliser ni à maintenir sur le long terme<sup>73</sup>.

Pour plusieurs bibliothèques, le fonctionnement avec des supports distants, les difficultés de communication et les délais, joints au manque éventuel d'outils de suivi, ont des conséquences en matière de gestion de projet. En effet, il peut être difficile, dans ces conditions, d'établir et de suivre un planning prévisionnel, que ce soit pour l'implémentation ou pour la mise à jour (BIB11, BIB21, BIB24, BIB13, BIB17, BIB25). Cette situation est d'autant plus bloquante pour les bibliothèques que les outils sont moins paramétrables et plus verrouillés (BIB02, BIB06, BIB21).

### *Les développements et les solutions*

Concernant les développements spécifiques à la demande, là aussi les réponses apportées par les bibliothèques ont été très hétérogènes. Huit bibliothèques, tous fournisseurs confondus, ont apprécié positivement cet item au niveau de leur fournisseur, (BIB01, BIB04, BIB20, BIB15, BIB05, BIB24, BIB17, BIB23). Six autres bibliothèques se sont déclarées non concernées par cette question car elles n'avaient pas demandé de développements spécifiques (BIB08, BIB14, BIB07, BIB18, BIB26, BIB16). Enfin, dix bibliothèques, tous fournisseurs confondus, ont évalué de manière plus mitigée à très sévère les développements à la demande (BIB12, BIB19, BIB22, BIB06, BIB10, BIB25, BIB02, BIB21, BIB11, BIB13).

Au sujet de la qualité des réponses et des solutions fournies, les bibliothèques se sont là aussi montrées très partagées. Onze bibliothèques, tous fournisseurs confondus, ont signalé leur satisfaction sur ce point (BIB12, BIB01, BIB19,

<sup>73</sup> BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Op. cit.*, p. 24.

BIB18, BIB23, BIB04, BIB20, BIB02, BIB05, BIB08, BIB17). Trois bibliothèques ne se sont pas prononcées (BIB26, BIB14, BIB16). Dix bibliothèques, tous fournisseurs confondus, ont fait part d'un avis beaucoup plus mitigé et parfois sévère sur cette question (BIB22, BIB06, BIB07, BIB10, BIB25, BIB15, BIB21, BIB11, BIB13, BIB24). Comme pour les points précédents, il n'y a donc pas là non plus de consensus et il peut y avoir un grand écart entre l'écoute du fournisseur et la capacité de celui-ci à mettre en œuvre une solution concrète.

Ces réponses amènent là aussi plusieurs remarques. Les attentes vis-à-vis des développements peuvent varier selon que la bibliothèque souhaite utiliser le produit en version standard ou l'adapter davantage : dès que la bibliothèque souhaite aller au-delà de la version standard du produit, les interactions avec le support deviennent plus complexes. Cela peut être lié à différents cas de figure. Certains logiciels peuvent proposer des options de paramétrage avancées, qui donnent une grande souplesse aux administrateurs, mais ne sont pas pour autant faits pour être adaptés en dehors de leur suite logicielle. Dans ce cas les demandes d'amélioration pourront être impulsées grâce à des clubs utilisateurs<sup>74</sup>. Dans les autres cas, il s'agit plutôt de produits standard et les options de paramétrages sont nettement moins étendues. Dans cette configuration, lorsque des bibliothèques veulent néanmoins aller au-delà de la version standard, elles doivent passer par le support et c'est là que les demandes peuvent se heurter à l'absence de réponse ou à des délais plus importants de prise en charge. Ainsi, dans plusieurs cas de figure, que ce soit parce que l'outil n'est pas conçu pour être utilisé en dehors d'une suite logicielle donnée ou parce que le support prend difficilement en charge ces demandes d'amélioration, la mise en place de développements a pu être jugée difficile ou pas entièrement satisfaisante par plusieurs bibliothèques (BIB11, BIB21, BIB02, BIB24, BIB13, BIB04, BIB06, BIB12).

Par ailleurs, certaines bibliothèques ont indiqué que les solutions mises en place restent des solutions fragiles, propres à un contexte local donné, et qu'elles risquent de rentrer en conflit ou d'être inopérantes dans un autre contexte, voire à la prochaine version de l'outil (BIB21, BIB11, BIB02, BIB24). En effet, les solutions mises en place peuvent ressembler davantage à des « bricolages » qu'à des développements pérennes alors même que chaque amélioration demandée dans un outil fourni en *cloud* pourrait être réutilisable pour d'autres clients. Pour un même besoin, dans un même outil, auprès d'un même fournisseur, deux bibliothèques se verront proposer des solutions de contournement différentes, sans que pour autant une nouvelle fonctionnalité soit effectivement développée dans l'outil, ce qui permettrait de satisfaire tous les utilisateurs. Si elles permettent de remédier ponctuellement à un problème, ces solutions risquent de se révéler fragiles techniquement à long terme. Les bibliothèques ont également regretté que ces solutions soient mises en place sans qu'une explication de la cause du problème soit donnée par le support, ce qui limite fortement leur capacité d'intervention en cas de nouveau dysfonctionnement (BIB11, BIB24, BIB02, BIB21).

<sup>74</sup> Voir *infra* I, 4, C, « Des clubs utilisateurs inégalement développés ».

## 4. LES SPECIFICITES DU PAYSAGE FRANÇAIS

### A. Outils propriétaires, outils libres, outils mutualisés

La plupart des bibliothèques d'enseignement supérieur que nous avons interrogées ont installé un outil propriétaire (24 sur 26). Parmi elles, deux avaient installé un outil mixte.

Ces implantations s'échelonnent de 2010 à 2016 de la manière suivante :

Année	Nombre de bibliothèques
2010	3
2011	2
2012	2
2013	2
2014	4
2015	7
2016	5

Près de la moitié de ces installations sont très récentes (2015-2016) ce qui veut dire que nous avons disposé, dans ces cas, d'informations très fraîches sur la mise en place, mais d'informations un peu plus parcellaires quant aux mesures d'impact.

Alors que les interfaces libres sont utilisées par plusieurs bibliothèques universitaires importantes aux Etats-Unis (Cornell, Stanford, Michigan, Chicago), elles restent relativement peu répandues en France. Dans ce travail, deux bibliothèques interrogées étaient concernées par des interfaces libres : une effectivement mise en œuvre et une autre simplement en test.

Le contexte de mutualisation grandissante des structures d'enseignement supérieur, initié par les SICD pour les structures documentaires, prolongé avec les PRES et les COMUE, explique qu'on ait affaire à 5 bibliothèques concernées par des outils consortiaux. La mise en place de ces derniers s'accompagne le plus souvent d'une forte complexité à la fois technique, financière et politique. D'un point de vue technique, l'outil de découverte consortial devra s'articuler avec une multitude d'outils et de ressources des bibliothèques partenaires : il ne remplace pas dans l'immédiat les outils existants, mais se superpose à eux. Du point de vue des moyens, la pérennité des financements et des services supports pour ces outils consortiaux peut être sujette à caution sur le long terme. En effet, un certain nombre d'outils consortiaux ont été mis en place par des COMUE liées à des IDEX, dont la consolidation définitive est encore en cours<sup>75</sup>. La pérennité de l'outil reste donc soumise dans ces cas au maintien du financement de l'IDEX. Du point de vue politique, le projet d'outil de découverte s'insère dans un projet documentaire au niveau de la COMUE ou de l'IDEX : ce changement d'échelle implique des enjeux et des contraintes qui dépassent largement celles d'un établissement lambda.

<sup>75</sup> PIA 1: Initiatives d'excellence (Idex) - ESR: enseignementsup-recherche.gouv.fr [en ligne]. [s. d.]. [Consulté le 7 octobre 2016]. Disponible à l'adresse : <http://www.enseignementsup-recherche.gouv.fr/cid101570/pia-1-initiatives-d-excellence-idex.html>.



Par ailleurs, de nombreuses attentes en matière d'outils mutualisés sont cristallisées par le projet de Système de Gestion de Bibliothèque mutualisé (SGBm). Coordonné par l'ABES, porté par 9 sites pilotes et touchant 46 établissements signataires du groupement de commandes, ce projet a pour objectif de « fournir un système de gestion locale mutualisé aux bibliothèques des établissements membres du réseau SUDOC ; de fournir le cas échéant, un outil de découverte local associé au système de gestion et de garantir l'articulation entre le SGBm et les systèmes nationaux de signalement administrés par l'ABES »<sup>76</sup> notamment le SUDOC. En 2015, un premier appel d'offres a permis de sélectionner cinq candidats appelés à intégrer un dialogue compétitif. Ces candidats ont été auditionnés par les différentes commissions d'experts, parmi lesquelles on comptait une commission de découverte. Ce dialogue compétitif de seize mois a débouché sur la rédaction d'un nouveau CCTP qui a servi de base à un accord-cadre. Le 28 octobre 2016, l'avis d'attribution de cet accord-cadre a rendu publics les noms des quatre candidats retenus<sup>77</sup>. Ces quatre fournisseurs pourront donc équiper les sites pilotes et les établissements signataires avec des solutions informatiques qui pourront comprendre une brique de découverte. Cela dit, même après l'attribution de l'accord-cadre, l'ABES a signalé qu'il reste impossible pour l'heure de communiquer des informations techniques précises sur les offres.

## B. En amont de l'installation

### *Les raisons d'installer un outil de découverte*

En ouverture de chaque entretien, nous avons demandé aux bibliothèques, d'indiquer les raisons qui les avaient poussées à installer un outil de découverte, afin de connaître quelles étaient leurs attentes spécifiques à ce sujet : améliorer la visibilité et l'utilisation des ressources numériques, offrir un point d'accès unique aux collections, la couverture de l'index, l'ergonomie, la vitesse, le coût, ou d'autres raisons à expliciter.

La principale raison mise en avant par toutes les bibliothèques interrogées de manière unanime est le souhait d'améliorer la visibilité et l'utilisation des ressources numériques. Cela peut s'expliquer, entre autres, par la place croissante que prennent les acquisitions de ressources numériques dans les budgets de bibliothèques universitaires depuis ces dernières années et par la politique de transfert d'abonnements du papier vers l'électronique<sup>78</sup>. Dans de nombreux cas, cela se traduit par l'importance accordée à la couverture de l'index, retenue comme critère pour le choix de l'outil (BIB01, BIB19, BIB18, BIB23, BIB04, BIB06, BIB07, BIB14, BIB15, BIB02, BIB05, BIB24, BIB21, BIB08). Dans certains cas cette couverture a même été un critère déterminant pour le choix de l'outil (BIB15, BIB24, BIB05).

<sup>76</sup> Des nouvelles du projet SGBm. Dans : *Fil ABES* [en ligne]. 22 avril 2015. [Consulté le 4 novembre 2016]. Disponible à l'adresse : <https://fil.abes.fr/2015/04/22/des-nouvelles-du-projet-sgbm-3/>.

<sup>77</sup> Il s'agit de Decalog – Data management, Ex-Libris, Linagora – Biblibre, OCLC. Projet SGB mutualisé : attribution de l'accord cadre. Dans : *Fil ABES* [en ligne]. 28 octobre 2016. [Consulté le 2 novembre 2016]. Disponible à l'adresse : <https://fil.abes.fr/2016/10/28/projet-sgb-mutualise-attribution-de-laccord-cadre/>.

<sup>78</sup> Voir notamment ADBU. *Enquête ADBU 2002-2014 : évolution des budgets d'acquisition des bibliothèques universitaires* [en ligne]. [S. l.] : ADBU, [s. d.]. [Consulté le 24 octobre 2016]. Disponible à l'adresse : [http://adbu.fr/wp-content/uploads/2014/03/Enqu%C3%A0te\\_ADBU\\_2014.pdf](http://adbu.fr/wp-content/uploads/2014/03/Enqu%C3%A0te_ADBU_2014.pdf).

La deuxième raison, mise en avant par vingt bibliothèques, est la volonté d'offrir un point d'entrée unique aux collections (BIB12, BIB01, BIB19, BIB18, BIB23, BIB22, BIB04, BIB06, BIB26, BIB07, BIB14, BIB16, BIB25, BIB15, BIB02, BIB05, BIB24, BIB17, BIB13, BIB08). Celui-ci est considéré comme une avancée importante, notamment pour des bibliothèques qui avaient pu expérimenter une recherche fédérée et en constater les limites (BIB12, BIB06, BIB23, BIB19, BIB01). Pour les instances consortiales, ce point d'entrée unique peut être particulièrement précieux afin de pouvoir proposer un portail unifié pour toutes les institutions partenaires (BIB04, BIB15).

L'ergonomie a été signalée comme une raison d'installer l'outil par dix bibliothèques (BIB01, BIB19, BIB22, BIB04, BIB20, BIB15, BIB05, BIB17, BIB08, BIB03). Mais il faut noter que certaines bibliothèques qui n'ont pas désigné explicitement cet élément ont toutefois signalé que l'outil de découverte présentait des améliorations importantes en matière d'ergonomie par rapport à leur interface précédente (BIB11, BIB25, BIB13, BIB16).

Dans une moindre mesure, pour neuf bibliothèques, c'est la vitesse de l'outil qui a pu constituer une raison de choix (BIB12, BIB19, BIB01, BIB23, BIB22, BIB04, BIB15, BIB17, BIB21). Là encore, pour les bibliothèques qui avaient expérimenté un moteur de recherche fédérée, l'outil de découverte représente une avancée majeure sur ce plan-là.

Enfin pour six bibliothèques (BIB22, BIB06, BIB20, BIB25, BIB17, BIB13) le coût de l'outil figure parmi les raisons de l'installer (quoique jamais seul). Deux bibliothèques ont soulevé l'opportunité et le bon rapport qualité/prix d'une offre commerciale donnée (BIB13, BIB25).

Dans une moindre mesure, les bibliothèques ont indiqué d'autres raisons, telles que le besoin d'intégration avec les systèmes locaux, comme des suites logicielles ou des services d'authentification (BIB23, BIB18, BIB08, BIB01, BIB19). D'autres bibliothèques ont signalé le souhait d'avoir une cohérence des outils d'administration entre d'un côté l'outil de découverte, et de l'autre soit le résolveur de liens et la liste A-Z soit un ERMS, afin de simplifier les interfaces d'administration pour les gestionnaires de ressources numériques (BIB14, BIB20, BIB10).

Certaines bibliothèques ont mis en avant aussi des raisons plus spécifiques. Ainsi, l'ouverture technique d'une solution, la présence de *web services* et d'une API étaient des éléments importants pour garantir la possibilité de développements ultérieurs et d'adaptations locales (BIB21). De même, des contextes locaux spécifiques, comme le projet de réinformatisation d'un réseau, ont pu être des éléments déterminants dans l'adoption d'un outil de découverte (BIB10).

### *Les tests avant la mise en œuvre*

Il nous a semblé intéressant de demander aux bibliothèques si elles avaient pu tester un ou plusieurs outils avant de décider de leur mise en œuvre. De nombreuses bibliothèques interrogées ont signalé qu'elles n'ont pas eu l'occasion d'effectuer des tests approfondis avant de choisir leur outil (BIB19, BIB18, BIB23, BIB22, BIB07, BIB10, BIB15, BIB05, BIB17). Dans plusieurs cas c'est un travail de benchmarking ou une présentation par les commerciaux qui tiennent lieu de test (BIB15, BIB05, BIB17).

Mais dans plusieurs autres cas, les bibliothèques ont indiqué avoir disposé d'une instance de test leur permettant de vérifier certains aspects des outils avant la mise en œuvre (BIB21, BIB11, BIB24, BIB02, BIB12, BIB01, BIB06, BIB14, BIB25, BIB13, BIB08).

Néanmoins, la portée de ces tests peut être limitée lorsque la bibliothèque n'a pas accès à toute l'interface d'administration complète (BIB24) ou lorsqu'elle ne peut pas faire tester l'interface par le public ou des utilisateurs autres que les administrateurs (BIB02).

En outre, les tests avant installation ne concernent pas l'intégration des données du SIGB et il reste donc très difficile pour une bibliothèque de vérifier quel sera effectivement le résultat de cette intégration dans l'outil de découverte (BIB11, BIB14, BIB21). Comme le signalaient Stephen Bollinger et Kate Siltan, « tous les fournisseurs promettent l'interopérabilité, mais tester réellement les instances de production de bibliothèques avec la même infrastructure informatique devient très vite difficile »<sup>79</sup>. C'est pourquoi ces difficultés d'intégration du catalogue apparaissent le plus souvent après le choix de l'outil, lorsque le projet est déjà lancé. De son côté, Craig Milberg a également signalé les réticences des fournisseurs à mettre en place un test complet, du fait de la grande complexité des opérations d'intégration des données du catalogue (qui impliquent une migration des données locales) :

« Nous avons demandé à tous les fournisseurs la possibilité de mener des tests sur leurs systèmes. Tous les fournisseurs ont refusé dans la mesure où ils ont légitimement avancé que pour mener un test significatif, tout le travail de fond de *mapping*, d'importation des données et de configuration devait être abouti, et cela avant même un engagement commercial. Ainsi, nous avons dû étudier les installations d'autres bibliothèques. Il s'agissait la plupart du temps d'une simple exploration visuelle, car nous pouvions étudier les fonctionnalités jusqu'à une certaine limite, mais nous étions incapables d'évaluer le fonctionnement de l'intégration de ressources en texte intégral dans la mesure où nous ne disposions pas d'accès à leurs bases de données en texte intégral. Nous étions aussi incapables de juger le fonctionnement de l'algorithme de pertinence, dans la mesure où nous ne connaissions pas en profondeur leurs collections.<sup>80</sup> »

Dans tous les cas concernés, le travail de test a été un travail conséquent en termes ressources mobilisées. Comme le signalait Joseph Deodato, « pour certaines bibliothèques, la dépense de temps et d'efforts impliqués dans la configuration

<sup>79</sup> « All of the vendors appear to promise interoperability, but verifying actual production instances of libraries using the same technology stack quickly becomes difficult ». BOLLINGER, Stephen et SILTON, Kate. When You Rent Your Front Door: A Mid-Size Library's Experiences Relying on Discovery Vendors. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 28.

<sup>80</sup> « We have requested the option from all vendors to carry out a trial of their systems. All the vendors refused to do so as they legitimately pointed out that to have meaningful trial, all of the labor-intensive mapping, data importation, and configuring had to happen at their end prior to a financial commitment. Thus we were forced to rely on examining other institutions' installations. This was particularly vexing, as we could examine functionality to some extent; however we were unable to evaluate how well full-text resource integration worked, as we lacked access to their full text databases. We were also unable to gauge how well relevancy worked, as we weren't intimate with their collections. » MILBERG, Craig I. A Tale of Two Systems: A Case Study on the Implementation of Two Discovery Systems at Davidson College. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 147-148.

d'un outil de découverte peut rendre la perspective de mener des tests prohibitive »<sup>81</sup>.

## C. Les besoins en termes de ressources humaines

### *Une forte mobilisation des équipes*

Nous avons demandé aux bibliothèques d'estimer les moyens humains mobilisés par l'outil de découverte. Le constat général est que les outils de découverte exigent un travail important de configuration, de paramétrage et de gestion courante qui réclame des ressources humaines qualifiées : leur impact sur le temps de travail ne doit donc pas être sous-estimé (BIB22, BIB02, BIB13, BIB01, BIB19, BIB21).

Il est difficile d'isoler les besoins en ressources humaines d'un outil de découverte par rapport à son environnement global, sans prendre en compte la gestion de ressources numériques, du résolveur de liens, du catalogue et du site web. On peut tout de même tenter de distinguer ce que l'outil exige en ressources humaines selon qu'on est en phase d'implémentation ou en phase de fonctionnement courant.

En phase de mise en œuvre, les bibliothèques ont indiqué que les groupes de projet pouvaient mobiliser plusieurs personnes, notamment : le responsable du SIGB, de la documentation numérique, un informaticien ou développeur, l'administrateur du site web, le responsable du catalogue ou du signalement, l'administrateur ou chargé de projet pour l'outil de découverte. Les groupes de projet se composent le plus souvent d'environ cinq personnes, mais peuvent être plus importants lorsque le projet concerne une réinformatisation complète, une refonte de site web ou un outil consorsial. Selon les difficultés rencontrées, ce mode projet peut se prolonger pendant plusieurs mois (de deux mois à un an). À partir du moment où elle comprend l'intégration du catalogue, la phase d'implémentation dépasse le plus souvent six mois.

En mode courant, d'après les estimations effectuées par les bibliothèques interrogées, la gestion de l'outil de découverte mobilise entre 0,2 et 3 ETP. Là encore il est difficile de distinguer la part de temps de travail consacrée spécifiquement à l'outil de manière isolée, dans la mesure où il est nécessaire d'avoir une très forte intégration de toutes les applications informatiques de la bibliothèque pour que l'ensemble fonctionne bien. Ce qui se dégage néanmoins, c'est qu'il n'y a pas d'outil, y compris en mode courant, y compris dans l'installation la plus standard possible, qui puisse fonctionner en autonomie sans mobilisation des équipes. Celle-ci peut varier en fonction du degré d'adaptation souhaité par la bibliothèque, selon que les données du SIGB sont intégrées ou pas, selon qu'il s'agit d'un outil sans trop de modifications ou au contraire très personnalisé (plus la bibliothèque s'éloigne de la version standard, plus la demande en termes de ressources humaines est importante), mais elle n'est jamais nulle. Dans certains cas, si l'on compte la gestion et la mise à jour des bases de connaissance, l'administration fonctionnelle, les développements web et l'articulation avec les différents systèmes de signalement, cela peut aller jusqu'à 2 ou 3 personnes qui travaillent sur l'outil en continu, avec un temps de travail

<sup>81</sup> « For some libraries, the expenditure of time and effort involved in configuring a web-scale discovery service can make the prospect of conducting trials prohibitive. » DEODATO, Joseph. Evaluating Web-Scale Discovery: A Step-by-Step Guide. *Information Technology & Libraries*. Juin 2015, Vol. 34, n° 2, p. 30.

variable (BIB01, BIB21, BIB23, BIB26). Ce sont donc là des besoins importants qui sont loin d'être sans conséquences sur l'organisation d'une bibliothèque universitaire en termes de ressources humaines. On peut à ce propos rappeler ce commentaire relevé par Marshall Breeding dans une enquête conduite auprès des bibliothèques avec des outils de découverte en 2014 : « comprendre, gérer et signaler les erreurs au niveau des métadonnées est presque un temps complet en soi ! »<sup>82</sup>

### *Des clubs utilisateurs inégalement développés*

Le cas échéant, les bibliothèques peuvent s'appuyer sur des clubs utilisateurs existants pour faire évoluer leurs outils ou pour construire leurs formations, comme dans le cas de Sirsi Dynix (AUSIDEF) ou d'Ex Libris (ACEF). Ainsi, pour ce dernier, le club utilisateur fait passer des enquêtes annuelles afin de recueillir les besoins des clients, qui peuvent voter pour l'ajout de nouvelles fonctionnalités lors de deux congrès annuels. Ces différentes demandes sont pondérées par le fournisseur qui arbitre entre elles et leur attribue un ordre de priorité. En cas de demande conjointe, les clients peuvent ainsi se regrouper pour impulser tel ou tel développement en fonction de leurs besoins.

Pour Proquest et pour Ebsco il n'existe pas à ce jour de club utilisateur comparable. Dans le cas d'Ebsco, c'est une liste de diffusion française, créée à l'initiative des professionnels de bibliothèque, qui permet les échanges entre utilisateurs. Dans le cas de Proquest, il existe une liste de diffusion anglophone. La fusion de Proquest et d'Ex Libris pourra entraîner une reconfiguration du club utilisateurs avec l'intégration des utilisateurs de Proquest au sein de l'ACEF. Au printemps 2016, une liste de diffusion commune aux administrateurs d'outils des BU françaises (tous fournisseurs confondus) a également été créée, à l'initiative des professionnels concernés.

---

<sup>82</sup> « Understanding, managing and reporting the errors in the metadata is almost a full time job in itself! ». BREEDING, Marshall. Chapter 2: Major Discovery Product Profiles. *Library Technology Reports*. Février 2014, Vol. 50, n° 1, p. 44.



## II. LES SERVICES DE DECOUVERTE A L'EPREUVE DE LEUR IMPLEMENTATION

---

### 1. APPREHENDER L'INDEX GENERAL CENTRALISE

L'index général centralisé rassemble, pour un service de découverte, des données issues d'une multitude de sources (éditeurs, agrégateurs, entrepôts en *open access* etc.). Cet index général centralisé constitue véritablement l'ossature de l'outil et représente, pour les bibliothèques qui mettent en place un service de découverte, le premier élément qu'elles doivent appréhender.

#### A. L'index général centralisé et ses inconnues

Dans le cadre de son travail sur l'*Open Discovery Initiative*<sup>83</sup>, la NISO a fait remarquer que les bibliothèques ont besoin de comprendre clairement la nature et l'étendue du contenu qui est signalé dans l'index de découverte : elles ont besoin d'informations détaillées et exactes sur les ressources couvertes, le niveau d'indexation (du texte intégral, des références bibliographiques seules ou les deux), la provenance des données (éditeur, agrégateur, bases de données de références bibliographiques A&I, *open access*) ainsi que les modalités d'accès (données ouvertes à tous les utilisateurs ou soumises à authentification)<sup>84</sup>. La NFAIS a souligné, quant à elle, que l'absence de ces données peut nuire à la bonne utilisation de l'outil lorsque la profondeur, l'étendue et la source des contenus intégrés ne sont pas suffisamment définis et détaillés dans la description du produit ou lorsque les utilisateurs manquent d'informations à jour sur les ressources indexées<sup>85</sup>.

Or la situation actuelle semble très éloignée du souhait de la NISO. Marshall Breeding avait déjà signalé que tous les fournisseurs de services de découverte à index centralisé ne mentionnaient pas publiquement le contenu détaillé de leur index<sup>86</sup>. Le constat général fait par les bibliothèques, c'est qu'elles manquent d'informations essentielles sur le contenu détaillé de l'index de l'outil de découverte : calcul de la couverture (période couverte, notamment pour les périodiques), niveau d'indexation (texte intégral, mots clés, métadonnées de base) et provenance des données (BIB21, BIB23, BIB12, BIB26, BIB18, BIB01, BIB10, BIB20, BIB25, BIB07, BIB14, BIB11, BIB05, BIB08, BIB17, BIB22, BIB13, BIB02, BIB06). Si elles indiquent être informées des ajouts importants de ressources ou des grandes soustractions, les bibliothèques pointent néanmoins un manque de transparence sur les informations concernant les données négociées avec les éditeurs qui constituent la matière de l'index (BIB21, BIB23, BIB12, BIB26, BIB18, BIB01, BIB10, BIB20, BIB25, BIB07, BIB14, BIB05, BIB08, BIB17, BIB22, BIB11, BIB13, BIB02). En effet, les négociations juridiques qui précèdent et déterminent l'ajout d'une ressource restent sous le sceau de la confidentialité commerciale.

---

<sup>83</sup> Ce point sera développé en détail *infra* IV, 3, A, « L'Open Discovery Initiative ».

<sup>84</sup> National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 2.

<sup>85</sup> National Federation of Abstracting and Information Services. *Op. cit.*, p. 3.

<sup>86</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 5.

Les bibliothèques signalent par ailleurs que les listes de ressources intégrées dans un index donné indiquent souvent un niveau de couverture à travers des pourcentages, mais sans suffisamment de précisions sur la façon dont ces pourcentages sont calculés (BIB21, BIB23, BIB12, BIB18, BIB20, BIB25, BIB07, BIB14, BIB15, BIB05, BIB08, BIB17, BIB22, BIB13, BIB02). Dans ces conditions, il est très difficile pour une bibliothèque d'effectuer des comparaisons, d'apprécier la couverture réelle d'une ressource et de mettre en place des outils d'évaluation. Signe de cette opacité, au moment de l'entretien, certaines bibliothèques ont indiqué ne disposer d'aucun document de leur fournisseur détaillant les ressources indexées qui soit daté postérieurement à 2014 (BIB10, BIB20, BIB25, BIB07, BIB14).

Interrogés sur les moyens par lesquels ils calculent le taux de couverture d'une ressource dans l'index centralisé lorsqu'ils n'ont pas d'accord contractuel avec le fournisseur de données, les fournisseurs d'outils ont donné des réponses variables : Ebsco s'est défendu d'intégrer les données via des tiers et de calculer des taux de couverture considérés comme peu exploitables et peu pertinents, tandis que Proquest a indiqué que, dans le cadre de Summon, il existe bien une possibilité de chercher en dehors de la collection de la bibliothèque, pour étendre la recherche à tout ce qui est indexé dans Summon. Mais dans ce cas-là, le fournisseur a signalé qu'il était difficile d'évaluer ce qui sera indexé pour les ressources avec lesquelles il n'y a pas d'accord avec le fournisseur de données.

Les bibliothèques pointent également le manque d'information sur la fréquence des mises à jour des données dans l'index (BIB21, BIB12, BIB02, BIB26, BIB04, BIB06, BIB10, BIB20, BIB25, BIB07, BIB11, BIB15, BIB05, BIB13, BIB08, BIB17, BIB24, BIB22). La plupart du temps elles disposent d'une indication de fréquence qui est simplement d'origine contractuelle (de type quotidienne, hebdomadaire, mensuelle, bimestrielle, trimestrielle, semestrielle, annuelle etc.) et qui peut être très variable d'une ressource à une autre.

Il peut donc y avoir des décalages entre les ressources effectivement disponibles pour la bibliothèque, directement sur une plateforme éditeur par exemple, et les ressources référencées dans l'outil de découverte. À de nombreuses reprises, les bibliothèques ont indiqué avoir remarqué des ressources qui auraient pu être signalées dans l'index et qui ne l'étaient pas, par exemple : articles ou numéros les plus récents d'une revue, version la plus récente d'un livre numérique, licences nationales (BIB06, BIB23, BIB24, BIB22, BIB02, BIB10, BIB19). Ce problème de synchronisation des données a été pointé également par Marshall Breeding comme un problème récurrent pour la mise à jour entre les données du fournisseur de données et celles de l'index<sup>87</sup>.

Concernant l'*open access*, on peut signaler avec François Renaville que l'intégration de ces ressources pose des problèmes spécifiques quant aux domaines couverts : il est parfois difficile pour la bibliothèque de distinguer ce qui est contenu par une ressource en *open access*<sup>88</sup>. La bibliothèque peut désactiver un entrepôt mais celui-ci aura pu entretemps avoir été moissonné par une source tierce, ce qui peut être difficile à gérer pour la bibliothèque (BIB26).

---

<sup>87</sup> *Ibid.*, p. 27.

<sup>88</sup> RENAVILLE, François. *Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? Op. cit.*



Toutefois, des bibliothèques ont pu constater une évolution vers plus de transparence : si initialement il était très difficile d'évaluer la couverture, à présent la situation tend à devenir plus transparente et les informations sur la couverture s'améliorent progressivement (BIB18).

## B. La granularité de l'indexation

La plus-value des outils de découverte est d'apporter une granularité plus fine que les catalogues traditionnels, au niveau de l'article ou du chapitre de livre, tandis que la granularité de niveau titre de la revue ou de l'ouvrage reste plutôt gérée par d'autres outils (SIGB ou base de connaissance). Pour un certain nombre de bibliothèques cette plus-value se conjugue de manière harmonieuse avec les titres de livres numériques et de revues qui sont importés soit de la base de connaissance soit du catalogue (BIB26, BIB19, BIB11, BIB15).

Cependant, plusieurs problèmes de gestion de la granularité au sein de l'outil de découverte peuvent apparaître, en particulier quand la connexion entre la base de connaissance et l'outil de découverte n'est pas entièrement opérationnelle<sup>89</sup> (BIB21, BIB15, BIB13, BIB08, BIB07, BIB14, BIB25). Cela peut provoquer des difficultés pour trouver les ouvrages ou les revues par leur titre, et c'est notamment gênant pour les livres numériques lorsqu'aucune autre indexation de niveau titre n'est disponible dans l'index centralisé.

Dans le cas d'outils mixtes, la connexion entre la base de connaissance et l'outil de découverte peut être également laborieuse, engendrant là aussi des difficultés pour trouver les ebooks ou les revues numériques par leur titre (BIB13, BIB08). La bibliothèque maintient alors parallèlement un accès direct à la liste A-Z de revues et de livres numériques pour permettre la recherche au niveau du titre et conseille aux utilisateurs une interrogation parallèle des deux outils (BIB13).

Parfois les bibliothèques signalaient déjà, avant même l'installation de leur outil de découverte, leurs livres numériques dans leur SIGB, ce qui leur permet de contourner ce problème (BIB20, BIB10, BIB14). Mais dans d'autres cas, pour pallier cet inconvénient, des bibliothèques choisissent de cataloguer ou d'importer des titres de revues ou de livres numériques dans le SIGB, pour que ces titres soient ensuite intégrés dans l'outil de découverte depuis le SIGB (BIB07, BIB25, BIB24, BIB02). Cette solution de contournement peut néanmoins entraîner des soucis au niveau des accès, tels que des conflits de rebonds possibles entre plusieurs liens ou des URL d'accès différentes entre le chapitre et le titre (BIB07, BIB25, BIB02).

Par ailleurs, le lien de l'article ou du chapitre vers la notice du numéro ou de l'ouvrage n'est globalement pas géré dans ces outils. Selon les outils, la bibliothèque pourra disposer d'un lien de l'article vers la notice globale de la revue, avec éventuellement un encart au niveau de la revue pour chercher à l'intérieur de celle-ci (BIB11, BIB24, BIB05). Mais il ne s'agit pas d'un lien pour remonter du chapitre vers l'ouvrage complet, ni de l'article vers le numéro précis dont il est extrait.

Certaines bibliothèques se demandent dans quelle mesure la coexistence de deux niveaux de granularité pour les livres numériques est susceptible de confondre l'utilisateur (BIB20). Pour d'autres, ce problème de gestion de

---

<sup>89</sup> Voir *infra* II, 2, B « L'intégration avec la(les) base(s) de connaissance ».

granularité est essentiellement un réflexe de bibliothécaire et la coexistence de deux niveaux de granularité pour un même livre numérique n'est pas spécialement gênante (BIB02, BIB14).

### C. Des variations selon les éditeurs, les langues et les disciplines

La couverture des données indexées est très variable en fonction des éditeurs, des bases et des disciplines.

Les éditeurs importants sont plutôt bien couverts (comme Elsevier ou Springer) mais ce n'est pas le cas pour toutes les ressources. Si les bibliothèques comprennent le non référencement de bases très spécialisées, le manque de ressources considérées comme incontournables pour une discipline donnée est un peu plus difficile à gérer. Si ce manque n'est pas toujours imputable à l'outil lui-même mais au fournisseur de données, il est néanmoins dommage que les fournisseurs d'outils mettent en avant leur caractère global et unificateur alors que les données restent partielles (BIB22).

Dans son étude pour la NISO, Breeding avait pointé le besoin d'améliorer la représentation des ressources dans les langues autres que l'anglais, plus spécialement dans les langues non-européennes et non-latines<sup>90</sup>. Le domaine francophone ne fait pas exception. Certes un certain nombre de ressources sont d'ores et déjà disponibles dans différents outils : *Cairn*, *Gallica*, *OpenEdition*, *Persée*, HAL, DUMAS (BIB23, BIB15, BIB12, BIB21, BIB13, BIB08, BIB01, BIB22). Néanmoins, il y a encore une grande marge de progression, notamment quant aux ressources juridiques. La demande conjointe de plusieurs bibliothèques pour une même ressource peut pousser le fournisseur à entamer des négociations, mais la progression reste au final assez lente (BIB24, BIB15, BIB20). Dans certains cas, les bibliothèques sont amenées à se tourner vers les éditeurs eux-mêmes afin qu'ils incitent les fournisseurs d'outils à mettre en ligne leurs données (BIB20). Les données des ressources francophones sont en outre jugées plus faibles en qualité que celle des ressources anglo-saxonnes, ce qui peut être gênant quand il manque des données essentielles, comme le lien direct au niveau de l'article (BIB19, BIB01, BIB24). Des améliorations restent encore possible dans le signalement des thèses françaises en ligne (BIB21).

Au niveau des disciplines, la couverture peut être également très disparate. Une étude de BIB10 estimait que, pour ses ebooks, le taux de couverture par domaines disciplinaires rapporté aux souscriptions était de 92 % en STM, 52 % en LSH et 7 % en droit-économie-gestion, ce qui révèle d'importantes variations.

Une des disciplines les plus touchées par ces faiblesses est le droit. Les bibliothèques concernées ont pointé l'absence des données des bases juridiques : *Dalloz*, *Lamy*, *Jurisclasseur* (BIB26, BIB12, BIB14, BIB07, BIB01, BIB15, BIB17, BIB11, BIB22, BIB24, BIB10, BIB16, BIB02). En effet dans ce domaine, les éditeurs sont particulièrement réservés quant à la mise à disposition de leurs contenus sur des outils commerciaux. Cela n'est pas imputable à l'outil en soi, mais pose particulièrement problème pour des établissements avec une composante juridique importante, alors que l'outil de découverte a tendance à se présenter comme un guichet unique : ainsi, la BIB16 n'a pas souhaité pendant longtemps

---

<sup>90</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 28-29.

intégrer le *widget* de recherche de l'outil de découverte sur son site en raison du manque de ressources juridiques.

Selon les outils, les bibliothèques ont pu signaler aussi l'absence des *Techniques de l'ingénieur* (BIB22), ainsi que des livres numériques *EMC* dans le domaine médical (BIB11, BIB15, BIB22). Ces bibliothèques ont toutefois indiqué qu'en ce qui concerne les ressources en santé, une base comme *PubMed* était bien intégrée.

La psychologie représente un cas un peu à part, une base de données bibliographiques payante comme *Psycinfo* ayant choisi de verser ses données à un seul outil de découverte (BIB21).

Les bibliothèques constatent également des manques sur des bases généralistes de presse ou d'usuels, notamment *Europresse*, *l'Encyclopaedia Universalis* ou les dictionnaires (BIB12, BIB22, BIB11, BIB24)

Pour remédier à ces difficultés les bibliothèques mettent en place des stratégies de contournement afin de signaler quand même les ressources peu ou pas représentées dans l'outil de découverte :

- Signalement sur la page web de la bibliothèque (BIB26).
- Signalement via un *widget* sur la page de résultats (BIB15, BIB24, BIB05, BIB02).
- Signalement via des suggestions de base de données (BIB25, BIB14, BIB20, BIB07).
- Signalement dans la liste A-Z, éventuellement grâce à BACON<sup>91</sup> (BIB12, BIB01, BIB19). Ce signalement n'est toutefois pas entièrement satisfaisant pour les cas où il n'y a pas d'URL univoque au niveau de l'article et où en accédant à la ressource l'utilisateur devra refaire sa recherche.

En ce qui concerne l'*open access*, les grands acteurs du domaine commencent à avoir une meilleure couverture dans les index de découverte<sup>92</sup> : les archives disciplinaires ou institutionnelles, peuvent être, selon les index, moissonnées directement ou indirectement via OAISTER<sup>93</sup>. Etudiant la couverture des ressources en *open access* des entrepôts institutionnels dans un index en particulier (en l'occurrence Primo Central Index), François Renaville a montré leur évolution croissante, stabilisée entre 2013 et 2015<sup>94</sup>. Néanmoins, le niveau d'indexation dans l'outil peut varier d'un entrepôt à un autre et a pu évoluer dans le temps : en 2013, l'outil de découverte pouvait indexer tout type de ressource en *open access* du moment que celle-ci répondait à certains prérequis techniques (métadonnées DC ou MODS et moissonnage OAI-PMH), mais depuis 2015 l'outil de découverte indexe seulement la partie qui est intégralement en *open access* (avec texte intégral).

---

<sup>91</sup> Voir *infra*. IV, 3, C, « BACON ».

<sup>92</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 29.

<sup>93</sup> Développé par l'université du Michigan puis racheté par l'OCLC, le catalogue OAISTER référence des notices de ressources en *open access* moissonnées par le protocole OAI-PMH : il compte environ 30 millions de notices, fournies par plus de 1500 organismes participants. *OAIster* [en ligne]. [s. d.]. [Consulté le 26 novembre 2016]. Disponible à l'adresse : <http://www.oclc.org/fr-CA/oaister.html/>.

<sup>94</sup> RENAVILLE, François. Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? *Op. cit.*, p. 240-241.

## D. Adapter l'index général centralisé à la bibliothèque

Pour faire en sorte que les données interrogeables dans l'index général centralisé correspondent à leurs abonnements effectifs, les bibliothèques déclarent les ressources auxquelles elles veulent circonscrire les résultats, à travers l'interface d'administration de l'outil de découverte. Cette sélection se fait dans une base de connaissance avec une granularité de niveau titre de bouquet ou base de données. Ce paramétrage initial peut être important en matière de temps de travail car il doit s'effectuer ressource par ressource (BIB07, BIB14, BIB10, BIB17).

Selon les outils et la taille de la ressource activée, entre le moment où la ressource est déclarée dans la base de connaissance de l'outil de découverte et le moment où les données sont effectivement disponibles dans l'index, il peut s'écouler un délai moyen variable, de 24h à 8 jours (BIB23, BIB12, BIB26, BIB19, BIB22, BIB20, BIB07, BIB14, BIB25, BIB17, BIB10, BIB08). Les bibliothèques ont signalé que ce délai de prise en compte de leurs paramétrages leur semblait globalement acceptable mais que, dans certains cas, il pouvait se prolonger sans que les causes en soient clairement explicitées (BIB10, BIB23).

Si les bibliothèques peuvent normalement activer dans leur outil de découverte les ressources de leur choix, il arrive qu'elles rencontrent des problèmes pour les identifier correctement dans la base de connaissance. Ainsi, il peut arriver qu'une bibliothèque soit abonnée à une seule partie d'une base de données plus large mais qu'elle dispose seulement de la possibilité de déclarer cette base dans son ensemble sans pouvoir cibler spécifiquement la partie qui la concerne (BIB08). De même, plusieurs bibliothèques ont indiqué des difficultés pour trouver ou faire ajouter les bouquets correspondant aux offres négociées au sein de l'ESR français (licences nationales et bouquets Couperin) qui peuvent être manquantes ou incomplètes dans la base de connaissance de l'outil de découverte (BIB23, BIB25, BIB10, BIB08, BIB24). Plusieurs bibliothèques ont indiqué que des progrès seraient néanmoins en cours dans ce domaine et on peut espérer qu'à l'avenir les bouquets les plus importants de ces offres soient effectivement disponibles de manière plus complète (BIB24, BIB17, BIB10).

Par ailleurs, il peut arriver que des ressources soient activées par défaut dans l'index de l'outil de découverte, même lorsque la bibliothèque ne les a pas cochées dans sa base de connaissance (BIB21, BIB05, BIB17). Dans ce cas, il s'agit le plus souvent de ressources en *open access*.

## 2. L'INTEGRATION AVEC LES SYSTEMES LOCAUX

*Sometimes we need to customize things that aren't easily customizable.*

*Sometimes we need to try to get data out of systems that just don't want to give it up.*

*Sometimes we just need to make things work together that weren't meant to work together.*<sup>95</sup>

(Jason Thomale)

Pour pouvoir accomplir pleinement sa mission de « *discovery and delivery* », l'index centralisé doit s'articuler correctement avec de nombreux systèmes locaux. Comme l'ont signalé de multiples écrits professionnels<sup>96</sup>, cette bonne intégration de l'outil de découverte avec l'ensemble des éléments qui composent l'infrastructure et l'environnement informatique d'une bibliothèque – SIGB, résolveur de liens, base de connaissance, proxy, systèmes d'authentification – est cruciale pour le bon fonctionnement de l'outil.

### A. L'intégration avec le SIGB

Le premier de ces systèmes locaux est sans doute le SIGB. L'intégration des données de celui-ci dans l'outil représente un travail long et considérable pour les bibliothèques. Au moment où nous avons effectué les entretiens, toutes les bibliothèques avaient procédé à l'intégration des données de leur catalogue à l'exception de la BIB20, qui a effectué cette intégration ultérieurement.

#### *L'intégration des notices bibliographiques*

La récupération des données bibliographiques du SIGB peut se faire selon deux procédés :

- un moissonnage OAI-PMH. NB : tous les SIGB ne supportent pas nécessairement un serveur fournisseur OAI-PMH et celui-ci doit être paramétré en amont.
- un import via (S)FTP des notices, en format MARC, MARCXML ou ISO2709, le plus souvent compressé.

Dans la configuration la plus simple, quand il s'agit d'un outil de découverte d'une seule BU et que l'outil de découverte prévoit l'intégration native avec un SIGB (outils produits par le même fournisseur et conçus pour travailler ensemble), l'articulation pourra être fluide. Mais les procédures peuvent se compliquer

---

<sup>95</sup> « Parfois on a besoin d'adapter des choses qui ne sont pas facilement adaptables. Parfois on a besoin d'extraire des données de systèmes qui ne veulent tout simplement pas les lâcher. Parfois on a juste besoin de faire travailler ensemble des choses qui ne sont pas censées travailler ensemble. » THOMALE, Jason. *You Gotta Keep 'em Separated: The Case for « Bento Box » Discovery Interfaces* [en ligne]. Portland, OR, 5 décembre 2014. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://code4lib.org/conference/2015/thomale>. Présentation filmée disponible sur <https://www.youtube.com/watch?v=i853dLWavZA>.

<sup>96</sup> WEBSTER, Peter. *The Web-Scale Discovery Environment and Changing Library Services and Processes*. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 646-661 ; BOLLINGER, Stephen et SILTON, Kate. *Op. cit.* ; BREEDING, Marshall et National Information Standards Organization (U.S.). *Op.cit.* ; BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Op. cit.* ; THOMPSON, JoLinda. *Op. cit.*

considérablement lorsqu'il s'agit de logiciels de fournisseurs différents pour lesquels il n'est pas prévu d'intégration native avec l'outil de découverte, et à plus forte raison en contexte consorsial, qui implique des bibliothèques de plusieurs établissements, avec des configurations techniques, des tutelles, des DSI et des supports de maintenance différents.

Le plus souvent, l'intégration des données du SIGB dans l'outil de découverte implique d'établir un *mapping* (ou tables de correspondance ou questionnaire ou règles de normalisation) entre les champs du SIGB (en format MARC ou dérivés) et les champs de l'outil de découverte (dans un format propriétaire spécifique à l'outil de découverte). Ce *mapping* va servir à la conversion des données MARC au format de l'outil de découverte, mais également à l'affichage des notices et à la génération des facettes. Selon les cas, les bibliothèques auront plus ou moins de marge de manœuvre dans la configuration des *mappings*. Il existe des systèmes très paramétrables, où la bibliothèque pourra configurer elle-même les *mappings* et garder la main sur un certain nombre d'affichages ainsi que sur les index (BIB04, BIB18, BIB19, BIB01, BIB23, BIB26). Dans d'autres cas, la bibliothèque doit envoyer au fournisseur un tableau renseigné avec les correspondances souhaitées. C'est alors le fournisseur qui peut les paramétrer et éventuellement les retoucher. Le fonctionnement de cette « moulinette » reste le plus souvent opaque pour les bibliothèques qui n'ont que peu de visibilité sur la façon dont leurs données sont retravaillées au sein de l'outil de découverte (BIB02, BIB05, BIB07, BIB10, BIB11, BIB14, BIB15, BIB21, BIB24, BIB25).

La définition des besoins d'une bibliothèque en matière de *mapping* constitue une étape essentielle dans la mise en place d'un outil de découverte. Pour cela il semble important de faire au préalable un relevé détaillé des zones que l'on souhaite indexer, afficher ou masquer, des champs à usage local, et des zones avec des informations spécifiques (provenance, fonds particuliers...). Ce travail de relevé est indispensable pour que les données du SIGB puissent être indexées et affichées dans l'outil de découverte d'une façon qui convienne aux besoins de la bibliothèque. La BIB25 a ainsi demandé à faire ressortir une collection spécifique, correspondant à un fonds étiqueté de manière particulière dans le SIGB, pour que ce dernier reste identifiable dans l'outil de découverte. Comme les données du catalogue sont mélangées à des données de sources très hétérogènes dans l'outil de découverte, il est important qu'elles soient bien structurées dans le format MARC. Il convient donc de vérifier la qualité du catalogue et la présence des bonnes informations dans les bonnes zones, notamment pour les informations stratégiques pour la bibliothèque. Par ailleurs, il peut être utile de vérifier la présence de notices correspondant à des ressources numériques dans le SIGB, en particulier des livres numériques, pour éviter de créer des doublons avec les données de l'index de découverte (BIB18, BIB07, BIB05, BIB10).

Les modèles de *mapping* sont livrés par le fournisseur. Ces modèles de *mapping* se présentent de manière différente selon les fournisseurs, les possibilités de paramétrage et selon la structure des données propre à chaque outil. Il pourra s'agir par exemple d'un *mapping* notice bibliographique / exemplaire (BIB24, BIB21, BIB11, BIB15, BIB02), ou encore d'un *mapping* type de document / notice (BIB07, BIB14, BIB10, BIB25). En règle générale, les champs cibles sont verrouillés et ne sont pas modifiables, mais les bibliothèques pourront choisir les champs sources (les données que l'on va faire correspondre à ces champs cibles). Concrètement, cela signifie donc qu'en matière de *mappings*, la bibliothèque ne

pourra pas introduire de nouveaux champs ou de nouvelles zones non prévues par le fournisseur sauf si l'outil permet de créer des champs locaux spécifiques (BIB18). Ces *mappings* peuvent par ailleurs être moins fins que le format MARC d'origine (BIB24, BIB10). Cela se reflète notamment au niveau de la gestion des types de documents : des bibliothèques qui auraient souhaité distinguer au niveau du *mapping* les thèses imprimées des thèses électroniques, ou bien les mémoires des thèses, ou bien les livres numériques et les livres imprimés, ou encore les revues électroniques des périodiques imprimés, n'ont pas la possibilité de le faire car cela signifierait l'ajout d'un nouveau champ cible au niveau des types de document, modification qui ne peut pas être faite dans l'outil de découverte (BIB07, BIB18, BIB02, BIB25). Cela peut impacter également d'autres zones MARC, comme les champs de notes (zones 300) qui peuvent se retrouver concaténés dans le même champ de l'outil de découverte (BIB24). En raison de ces contraintes de paramétrage, les bibliothèques voient parfois leurs données présentées de manière moins cohérente ou moins riche dans l'outil de découverte que dans le SIGB (BIB18, BIB05, BIB25, BIB02).

En contexte consortial, les outils peuvent permettre de paramétrer autant de *mappings* que de SIGB intégrés (BIB15, BIB04). Néanmoins, la multiplication des *mappings* créés implique la multiplication des *mappings* à vérifier et à maintenir. Il peut donc être intéressant d'utiliser un même *mapping* pour deux instances différentes d'un même SIGB utilisé dans deux établissements différents d'un même consortium (BIB04). Mais cette possibilité entraîne d'autres contraintes : une bibliothèque qui partage un *mapping* avec une autre institution pourra voir des données affichées qui ne sont pas pertinentes pour son public car ce champ est « partagé » dans le *mapping* avec un autre établissement et donc nécessairement présent (BIB06). Selon les outils il est parfois possible de créer des exceptions à un *mapping*, mais la multiplication des exceptions peut être au final tout aussi difficile à gérer que la création d'un nouveau *mapping*.

Lorsque les *mappings* établis doivent passer par le support technique, le paramétrage par le fournisseur peut prendre plusieurs mois avant que les configurations soient prises en compte (BIB15, BIB21, BIB11, BIB05, BIB02). Cela s'explique par un détour via les supports internationaux qui allonge le délai de mise en œuvre. Une fois le *mapping* paramétré, les demandes de modifications ultérieures nécessiteront aussi de passer par le fournisseur. Modifier un *mapping* implique de refaire par la suite toute l'indexation des données du catalogue dans l'outil de découverte. Cela a pu néanmoins être fait dans certains cas pour rectifier le *mapping* des livres numériques ou des thèses (BIB07, BIB25). Dans ces cas, il a été nécessaire de passer par le support et de refaire une indexation totale des données du SIGB.

Lorsque les *mappings* sont configurés, l'export ou le moissonnage peuvent se mettre en place. Ils peuvent concerner :

- soit la totalité des données bibliographiques du SIGB à chaque fois : on exporte toutes les notices et on écrase (on remplace) les notices présentes jusqu'alors dans l'outil de découverte. Cela implique des temps d'export et d'indexation qui peuvent être longs mais cela peut être une solution si la bibliothèque n'a pas les moyens de gérer un export différentiel (BIB11) ;
- soit un export différentiel : on exporte uniquement les notices créées, modifiées et supprimées depuis l'import précédent. Cette façon de procéder permet de raccourcir des temps de traitement (import et

indexation) mais implique de pouvoir détecter chacun de ces différents cas, ce qui peut parfois être difficile selon les SIGB. De manière générale, toutes les bibliothèques qui peuvent le faire s'orientent vers un export différentiel (c'est le cas de toutes les autres bibliothèques concernées).

Il faut noter que parfois certains traitements complémentaires peuvent être nécessaires sur les notices exportées, par exemple lorsque les *mappings* ne permettent pas de gérer de façon souhaitée certaines zones très spécifiques (BIB11).

L'encodage des caractères (notamment les diacritiques et les guillemets) demeure aussi un point de vigilance lors de ces procédures : le SIGB et l'outil de découverte ne gèreront pas nécessairement le même encodage ce qui peut provoquer des affichages insolites et troubler la lisibilité (BIB24, BIB13, BIB25, BIB17). Ces problèmes peuvent être difficiles à résoudre par la bibliothèque du fait de l'opacité et du verrouillage technique des outils (BIB24).

Dans la mesure du possible, les exports sont automatisés, de préférence tous les jours, afin de réduire les hiatus entre le catalogue et l'outil de découverte. Parfois les bibliothèques choisissent des cadences différentes pour les exports : export quotidien pour les notices créées et modifiées, export hebdomadaire pour les notices supprimées (BIB24). D'autres doublent les exports différentiels quotidiens par un export complet mensuel (BIB02). Mais l'automatisation de l'export dépend souvent des possibilités techniques et des compétences informatiques spécifiques pour mettre en place un script d'export. Dans certains cas, lorsque l'automatisation n'est pas possible, les exports se font manuellement une fois par semaine, ce qui implique un décalage d'autant entre les données bibliographiques du SIGB et celles de l'outil de découverte (BIB11, BIB15).

Outre ces difficultés, il faut noter aussi que l'import des données n'est pas synonyme d'indexation. Selon les outils, entre l'import des données dans l'outil et leur indexation effective il peut y avoir un délai, le plus souvent d'entre 24 et 72 heures, mais qui peut parfois être plus important (BIB02, BIB07, BIB21, BIB25). Les bibliothèques n'ont pas toujours les moyens de vérifier via l'interface d'administration que cette indexation est effectivement faite ni de distinguer la part de leur catalogue indexée dans l'outil de découverte (BIB02, BIB07, BIB21). Des lenteurs peuvent être dues aussi au fait que dans le *cloud* il peut y avoir une file d'attente pour l'indexation. Ainsi, plusieurs BU ont constaté des délais accrus (de 5 à 10 jours) pour l'indexation de leurs données du SIGB, sans que ces lenteurs, rallongées par périodes, ne trouvent d'explication (BIB10, BIB07, BIB25). Or cela peut avoir des répercussions lorsqu'il y a eu des changements importants dans la structure des données du SIGB, tels que des fusions de bibliothèques. Dans deux cas, suite à ces modifications, des bibliothèques qui n'existaient plus dans le SIGB continuaient néanmoins d'exister dans l'outil de découverte, ce qui posait des problèmes pour la recherche documentaire (BIB10, BIB25).

### ***La récupération des données d'exemplaires en temps réel***

L'intégration en temps réel des données d'exemplaires (en particulier disponibilité et localisation) constitue également une étape importante et longue dans la mise en place d'un outil de découverte.



Cette intégration peut se faire via l'exploitation de *web services* ou d'API pour communiquer avec le SIGB (BIB24, BIB21, BIB15, BIB01, BIB11, BIB17, BIB02). Dans le cas de SIGB qui s'intègrent nativement avec un outil de découverte de la même suite, l'intégration pourra être fluide et transparente pour le lecteur (BIB23, BIB26, BIB01, BIB04, BIB19, BIB08).

Dans certains cas lorsque le SIGB est d'un autre fournisseur que celui de l'outil de découverte, et que le SIGB ne fournit pas de *web service* ou que celui-ci ne peut pas être complètement exploité par l'outil de découverte, le processus peut devenir plus complexe et long à mettre en place. Parfois, l'intégration des données d'exemplaires requiert de mettre en place un *web service* spécifique développé par la bibliothèque (BIB10).

Une autre solution peut être d'exploiter le catalogue public par une technique de *web scrapping* lorsque cela est possible (BIB07, BIB14, BIB15). En récupérant la page de code HTML du catalogue public en temps réel pour une notice précise, on extrait les informations concernant les exemplaires.

Dans d'autres cas, c'est une interrogation du SIGB via le protocole Z39.50 qui permet de récupérer les données d'exemplaire (BIB25, BIB15). La plupart du temps il s'agit de SIGB qui fonctionnent avec une logique classique de client / serveur et non pas en mode *full web*. Le paramétrage et l'ajustement des affichages des données des exemplaires via Z39.50 ont pu représenter dans ce cas un travail long et complexe (BIB25). Par ailleurs, si le SIGB est limité en nombre de sessions ou si le nombre de connexions simultanées est réduit, le serveur peut être surchargé et ne plus répondre. Dans ce cas, la bibliothèque peut élaborer une solution de contournement, comme par exemple : si le SIGB ne répond pas au bout de 10 secondes, faire en sorte que l'outil de découverte affiche les informations de base du catalogue et un message spécifique indiquant de consulter le catalogue (BIB15).

Lorsque ni l'utilisation de *web services* ni l'interrogation via Z39.50 ne sont possibles, il peut être envisageable de mettre en place un import des données d'exemplaires du SIGB plusieurs fois par jour, par exemple toutes les demi-heures (BIB13). C'est là une solution alternative qui permet de rester assez fidèle au « temps réel ». Mais il faut pour cela que le SIGB concerné soit capable de produire l'export des données d'exemplaires dans un temps relativement court, ce qui n'est pas le cas de tous les SIGB qui ont parfois des temps d'export plus longs (BIB04).

Dans certains cas, quand il n'y a pas de compatibilité avec le *web service* d'un SIGB et pas de solution alternative, il ne peut pas y avoir d'intégration des données d'exemplaires (BIB06, BIB04). Une solution pérenne dans ce cas impliquerait des développements des deux côtés (SIGB et outil de découverte) qui ne sont pas complètement assumés ni par l'un ni par l'autre.

Même quand l'intégration est possible, il faut néanmoins avoir présent à l'esprit que le *web service* part de l'outil de découverte vers le SIGB, ce qui peut avoir des répercussions plus ou moins importantes en fonction de l'écart chronologique entre les notices présentes dans le SIGB et celles importées dans l'outil de découverte. Lorsque des notices sont présentes dans le SIGB mais pas encore importées dans l'outil de découverte, *a fortiori* les données d'exemplaires correspondantes ne pourront pas être affichées dans l'outil de découverte. Inversement, lorsqu'une notice est toujours présente dans l'outil de découverte, mais supprimée dans le SIGB, les informations d'exemplaires ne pourront pas être

affichées non plus dans l'outil de découverte. Par ailleurs, lorsqu'il y a beaucoup d'exemplaires pour une même notice, le calcul des disponibilités peut être plus difficile (BIB26).

Des problèmes d'affichage et/ou de traduction peuvent surgir également sur les données d'exemplaires du SIGB : libellés en anglais, affichage des diacritiques, affichage des formats de dates ou affichage d'un même exemplaire sur plusieurs lignes (BIB05, BIB07, BIB17, BIB11). Par ailleurs, selon les données que l'outil garde en cache, ce n'est pas toujours le premier exemplaire disponible qui sera affiché dans l'outil de découverte, mais ce peut être un exemplaire non disponible, tandis que les autres sont masqués (BIB07).

En conséquence, pour des portails consortiaux, selon le SIGB source, il pourra y avoir, tout comme pour les données bibliographiques, une très grande hétérogénéité de l'affichage des données d'exemplaires : certaines seront intégrées en temps réel, tandis que pour d'autres il s'agira simplement d'un lien vers le catalogue (BIB04, BIB15).

### *Intégration avec des services*

L'intégration avec les services du SIGB peut se faire de manière plus ou moins fluide selon que le SIGB et l'outil de découverte sont nativement conçus pour travailler ensemble.

Dans le cas d'intégration native, l'utilisateur pourra accéder à son compte depuis l'interface de découverte pour profiter des différents services distants (prolongations, réservations, suggestions, éventuellement demandes de PEB ou réservation de salles) ainsi que des listes de références bibliographiques (BIB26, BIB23, BIB12, BIB18, BIB01, BIB08, BIB13). Pour plusieurs des bibliothèques dans ce cas, l'intégration avec les services a été un point crucial du choix de l'outil et un gain en matière de services offerts (BIB23, BIB08, BIB13).

Dans tous les autres cas, quand il n'y a pas d'intégration native, les bibliothèques ajoutent le plus souvent à l'interface de découverte un lien vers le compte utilisateur du SIGB ou vers l'ENT ou site web de la BU. Dans cette configuration, l'utilisateur quittera la page de l'outil de découverte en cliquant sur ce lien, pour se rendre sur celle de l'OPAC ou de l'ENT.

Il est parfois possible d'intégrer un lien vers le compte utilisateur dans chaque notice via un bouton « réserver » présent sur chaque notice du catalogue mais cela ne fonctionne pas nécessairement pour tous les SIGB ou n'est pas toujours contextualisé en fonction des besoins de la bibliothèque (BIB24, BIB05, BIB11, BIB02).

Les services de découverte proposent aussi des comptes qui correspondent à des paniers de gestion de références bibliographiques, alertes et historiques de recherche. Si ces paniers permettent de gérer les notices issues de l'index centralisé, ils ne permettront pas nécessairement de gérer celles provenant du catalogue (BIB13). Parfois les bibliothèques choisissent de ne pas afficher ce panier de gestion de références bibliographiques afin de ne pas induire en erreur l'utilisateur par rapport au compte utilisateur classique de l'OPAC (BIB07, BIB10, BIB05, BIB02).

Certaines bibliothèques ont ajouté aussi à leur outil des liens vers d'autres services : lien vers un service de questions/réponses, *chat* ou compte Refworks, forum ou foire aux questions (BIB20, BIB25, BIB14).

Excepté les cas d'intégration native, l'interopérabilité difficile avec des comptes utilisateurs existants sur les OPAC ou ENT constituent un point faible des services de découverte à index général centralisé, par rapport aux OPAC traditionnels qui prenaient correctement en compte ces fonctionnalités.

## **B. L'intégration avec la(les) base(s) de connaissance**

Pour permettre d'accéder efficacement aux ressources, l'outil de découverte doit s'appuyer sur une base de connaissance. Le fait que l'outil de découverte et la base de connaissance soient produits par un même fournisseur semble une condition nécessaire, mais non suffisante, pour assurer le bon fonctionnement de l'outil. En cas de fournisseurs différents, les bibliothèques auront plusieurs bases de connaissance : une pour l'outil de découverte, où s'effectuera la déclaration des ressources dans l'index centralisé, et une différente pour le résolveur de liens et la liste A-Z.

### *L'intégration d'une liste alphabétique de ressources*

Pour signaler leurs revues ou ebooks avec une granularité de niveau titre, les bibliothèques peuvent importer dans leur outil de découverte le contenu de leur liste alphabétique de ressources, extraite de leur base de connaissance.

En cas d'intégration native entre les deux outils, l'import de cette liste dans l'outil de découverte ne pose pas de problème majeur. De nombreuses bibliothèques procèdent ainsi à un import journalier de leur liste A-Z à travers un *mapping* spécifique (BIB26, BIB19, BIB18, BIB04, BIB01, BIB06). Dans ce cas, les données de la base de connaissance deviennent, après import, des données parmi d'autres de l'outil de découverte. Lorsque le signalement est correct dans la liste A-Z, ces bibliothèques ne connaissent pas de problème de gestion de la granularité de niveau titre de revues et d'ebooks. Les bibliothèques pourront également signaler dans leurs bases de connaissance des ressources qui ne se trouvent pas dans l'index, notamment des ressources francophones, pour alimenter l'index depuis la base de connaissance (BIB01).

Mais pour d'autres bibliothèques, même quand il s'agit de deux outils d'un même fournisseur, il n'y a pas d'intégration complète de la base de connaissance dans l'outil de découverte. Il s'agit juste d'une interrogation simultanée de l'outil de découverte et de la base de connaissance : la recherche dans l'outil de découverte se répercute de manière simultanée dans l'index centralisé et dans la liste A-Z avec un affichage synchronisé des réponses, mais ce n'est pas une recherche entièrement unifiée (BIB11, BIB24, BIB21, BIB05, BIB02). Dans ce cas, lorsqu'une chaîne recherchée correspond exactement à un titre dans la base de connaissance, cela génère un encart en haut des résultats avec le titre de la revue ou du livre concerné. Le fait qu'il y ait une fonction d'autocomplétion qui s'appuie sur les données de la base de connaissance peut faciliter cette correspondance exacte et la recherche de titres de revues. Mais même en cas de correspondance exacte, cette fonctionnalité peut être difficile à utiliser, car l'encart est paramétré pour fonctionner avec un titre seul, ce qui peut poser problème lorsque la correspondance exacte concerne plus d'un titre dans la liste A-Z. Dans ce cas, l'encart n'affichera aucun résultat, même quand le titre avait été proposé grâce à l'autocomplétion (BIB02, BIB21). On peut signaler aussi que ce système d'encart ne prend en compte que les titres de revues et pas les actes de colloque ou d'autres documents dans la base de connaissance (BIB02, BIB21). Pour une interrogation

identique de niveau titre, l'interface publique de la liste A-Z pourra donc fournir plus de réponses que l'outil de découverte.

Ainsi, même quand ce sont des outils d'un même fournisseur, il arrive qu'il n'y ait pas toujours une véritable intégration de la base de connaissance dans l'outil de découverte. Le constat fait par de nombreuses bibliothèques est que la communication entre les outils d'un même fournisseur peut être problématique et que l'articulation entre ces différentes briques demeure fragile (BIB21, BIB02, BIB07, BIB14, BIB07, BIB25). À cet égard, des bibliothèques souhaiteraient voir à terme les notices de la base de connaissance intégrer véritablement l'index centralisé, afin de générer cette granularité de niveau titre, développement qui leur a été annoncé par le fournisseur concerné (BIB02, BIB13). En attendant, lorsque les bases de connaissance fournissent une API, les bibliothèques peuvent mettre en place des développements spécifiques pour interroger la liste A-Z via une surcouche logicielle et retraiter les réponses afin de mieux gérer ce niveau de granularité (BIB21).

Il arrive également que, même pour des outils d'un même fournisseur, un connecteur spécifique s'avère nécessaire pour pouvoir intégrer les données de la liste alphabétique dans l'interrogation depuis l'outil de découverte, afin de disposer d'une granularité de niveau titre pour les revues et les livres numériques (BIB25, BIB07, BIB14, BIB20). En l'absence de ce connecteur, même si l'information est présente dans la base de connaissance, la granularité de niveau titre ne sera pas gérée dans l'outil de découverte. L'information sera bien disponible d'un côté dans la base de connaissance mais pas dans l'index centralisé de l'outil de découverte. Ce problème d'interaction entre deux outils d'un même fournisseur, peut se traduire pour les bibliothèques par la mise en place de circuits de traitement de données supplémentaires, tels que l'import des livres numériques dans le catalogue (BIB07).

Ces problèmes d'interopérabilité se complexifient davantage lorsqu'ils impliquent des fournisseurs différents : soit parce qu'il s'agit d'outils mixtes (avec un fournisseur pour l'interface et un autre pour l'index centralisé), soit parce que la bibliothèque a deux fournisseurs différents pour deux bases de connaissance distinctes (une pour l'outil de découverte et une autre pour le résolveur de liens et liste A-Z).

Pour les outils mixtes, s'il peut y avoir une bonne interaction avec la liste A-Z dans la version initiale de l'outil, les évolutions successives peuvent être à terme plus difficiles à gérer lorsqu'un des outils connaît des modifications ou des mises à jour. Par exemple, lorsqu'une société fait évoluer son interface d'interrogation de liste A-Z, les fonctionnalités et la structure de la nouvelle version pourront être plus difficilement prises en compte dans l'outil mixte, alors qu'elles sont correctement gérées dans l'interface native (BIB08, BIB13).

Lorsque l'outil de découverte et la base de connaissance qui sert à générer la liste A-Z sont produits par des fournisseurs différents, les risques de problèmes d'articulation entre les deux outils se multiplient. En effet, d'un outil à l'autre, les ressources activées pourront avoir des périmètres différents, sans véritablement pouvoir obtenir une correspondance adéquate des deux côtés, ce qui implique une interaction plus difficile entre les données de la liste A-Z et l'outil de découverte (BIB10, BIB04, BIB15). Ce problème a été également rencontré par Shaundra

Walker et Iyanna Sims du réseau des bibliothèques HBCU<sup>97</sup> et par Stephen Bollinger et Kate Siltan de la Bluford Library : « En plus d'activer toutes ces ressources, nous devons les confronter avec celles [de l'autre] base de connaissance. Dans certains cas, le nombre total de titres, les périodes couvertes ou d'autres informations ne correspondaient pas entre les deux bases de connaissance<sup>98</sup>. »

Pour remédier à ce problème, il peut être possible d'écrire des scripts spécifiques pour importer la base de connaissance d'un fournisseur dans l'outil d'un autre fournisseur (BIB15). Ces scripts peuvent toutefois être difficiles à mettre en œuvre dans la mesure où un grand nombre de bouquets de la base de connaissance sont différents ou absents de ceux de l'outil de découverte et qu'il n'y a pas de coïncidence entre les deux (BIB10).

### *Signalement et gestion des accès*

Pour que l'outil puisse remplir correctement sa fonction, il est essentiel qu'il ait une bonne intégration avec le résolveur de liens afin d'assurer un signalement exact et une gestion fluide des accès. Comme le note Peter Webster « étant donné qu'au final les utilisateurs ne distinguent aucun élément de la chaîne d'accès à l'information, la réussite de tout le processus dépend du caractère direct et de la fluidité avec laquelle les utilisateurs sont connectés à l'information dont ils ont besoin<sup>99</sup> ». Cette intégration est si importante que plusieurs bibliothèques ont déclaré avoir choisi leur outil de découverte en fonction du résolveur de liens ou réciproquement (BIB21, BIB17, BIB14, BIB04, BIB06).

Le résolveur de liens travaille sur une base de connaissance afin de construire les URL qui permettent l'accès à une ressource correctement reconnu. Cette base de connaissance pourra être la même que celle de l'outil de découverte, en cas de fournisseur identique, ou une autre, en cas de fournisseurs différents. Quand plusieurs plateformes offrent l'accès à une même ressource, le résolveur pourra offrir le choix de la plateforme, par une page intermédiaire entre l'outil de découverte et la ressource.

Lorsque les outils sont nativement conçus par un même fournisseur et qu'ils s'appuient sur une base de connaissance commune, l'articulation entre l'outil de découverte et le résolveur pourra être globalement fluide, sans problèmes majeurs de gestion des accès (BIB12, BIB23, BIB22, BIB19, BIB01, BIB04, BIB17, BIB07, BIB15).

Cependant, même dans le cas d'outils produits par un même fournisseur, plusieurs bibliothèques ont signalé l'existence de conflits entre le résolveur de liens et l'outil de découverte, ce qui peut être source de confusion pour l'utilisateur (BIB26, BIB23, BIB24, BIB21, BIB11, BIB20, BIB02, BIB06). À plusieurs occasions les bibliothèques ont pointé des difficultés d'intégration et de contextualisation du résolveur de liens dans l'outil de découverte : par exemple,

---

<sup>97</sup> WALKER, Shaundra et SIMS, Iyanna L. Implementing a Discovery Tool at Two HBCUs. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 204.

<sup>98</sup> « In addition to activating all of these resources, they had to be checked against [the other] knowledge base. In some cases, title counts, dates of coverage, and other information did not match up in both knowledge bases ». BOLLINGER, Stephen et SILTON, Kate. *Op. cit.*, p. 30.

<sup>99</sup> « Since users do not ultimately distinguish any element of the supply chain, the success of the whole process hinges on how directly and seamlessly users are linked to needed information » WEBSTER, Peter, *Op. cit.*, p. 654.

même si une ressource est effectivement accessible, le résolveur n'est pas en mesure de proposer le lien correct depuis l'outil de découverte ou affiche à la fois deux informations contradictoires : absence du texte intégral, mais présence d'un lien d'accès (BIB26, BIB23, BIB02). Ces conflits peuvent s'accroître en contexte consortial où il peut y avoir beaucoup de chemins d'accès pour une même ressource : ressource propre à la bibliothèque, ressource consortiale, licences nationales (BIB06).

L'articulation entre le résolveur de liens et l'outil de découverte peut se voir complexifiée par la coexistence de plusieurs systèmes de gestion de liens différents. Les liens peuvent être rentrés « en dur » dans les notices de l'outil de découverte ou construits par le résolveur de liens. La coexistence de plusieurs types de liens différents (liens propres à la bibliothèque, liens directs, OpenURL) peut entraîner des difficultés pour le fonctionnement et la lisibilité de l'outil ainsi que pour sa maintenance (BIB11, BIB21, BIB24, BIB02). Un changement d'URL chez l'éditeur peut se traduire par exemple par des liens cassés sans que la bibliothèque soit en mesure de trouver la cause et de résoudre le problème car il est difficile de déterminer s'il provient de la ressource, du paramétrage du résolveur ou de la façon dont cela s'affiche dans l'outil de découverte (BIB02, BIB24, BIB11). La gestion de l'OpenURL peut également être source de difficultés, ce qui se traduit parfois par le fait que des utilisateurs doivent s'authentifier pour consulter une ressource en *open access* (BIB21, BIB13, BIB17).

Pour que l'outil puisse pleinement remplir sa mission de *discovery and delivery*, il faut que la construction du lien par le résolveur puisse se faire de manière univoque pour chaque item présent dans l'index. Cela n'est pas toujours possible pour des bases qui ne fournissent pas de lien univoque et pérenne pour chacun de leurs articles comme *Factiva* ou *Europresse* (BIB14, BIB01, BIB24, BIB16). Ce problème a été pointé, entre autres, par la NFAIS<sup>100</sup>. L'outil de découverte et le résolveur ne pourront pas permettre à une recherche d'aboutir jusqu'à l'article en l'absence de ce lien vers l'article : dans ce cas on arrivera sur la page d'accueil de la ressource, où il faudra refaire la recherche. Il faut également que les métadonnées permettant de construire le lien d'accès (par exemple pour un article l'ISSN et le titre du périodique) soient correctement renseignées dans la notice de l'index pour que le résolveur de liens puisse y donner accès (BIB08, BIB26, BIB23). Il est nécessaire aussi que le lien soit correctement construit à chaque niveau de granularité concerné : dans l'outil de découverte, l'URL d'accès au niveau du chapitre peut ne pas être la même qu'au niveau du titre du livre numérique ce qui peut entraîner des difficultés d'accès (BIB07).

Pour les bibliothèques qui ont un outil de découverte et un résolveur de liens de producteurs différents, on retrouvera, au niveau de la gestion des accès, les mêmes problèmes liés aux périmètres différents des deux bases de connaissance. La cohabitation des deux outils n'étant pas fluide, il pourra y avoir des conflits entre le résolveur de liens et l'outil de découverte, lorsque celui-ci prend le pas sur le lien construit par le résolveur en proposant un lien direct vers la ressource. Or ce lien direct ignorera les étapes nécessaires pour que l'accès soit correctement reconnu et autorisé (par exemple le passage par un proxy), ce qui peut se traduire par un échec d'accès à la ressource (BIB10). C'est un problème semblable qu'ont rencontré les professionnels de la Bluford Library : « une grande part de cet échec peut être attribuée à ce problème d'infrastructure informatique : étant donné qu'on

<sup>100</sup> National Federation of Abstracting and Information Services. *Op. cit.*, p. 3.

utilisait deux bases de connaissance de fournisseurs différents, il y avait beaucoup de cas où il n'y avait pas de correspondance entre les métadonnées, ce qui empêchait de générer des liens fonctionnels »<sup>101</sup>. Ainsi, du fait de l'absence de concordance des deux bases de connaissance, alors que la ressource est bien signalée dans le résolveur, l'accès peut ne pas être effectif depuis l'outil de découverte.

Pour des installations consortiales, le résolveur de liens doit en outre permettre de cibler les différents accès aux ressources de chaque établissement. Dans l'idéal, les informations d'accès qui s'affichent dans l'outil de découverte doivent être contextualisées, c'est-à-dire rattachées à un établissement, afin de ne pas proposer à l'utilisateur des ressources qui ne seront ensuite pas accessibles pour lui. Cela peut être fait au moyen de vues spécifiques propre à chaque instance (BIB04, BIB15, BIB06) ou au niveau de la vue globale par une identification visuelle des liens avec les logos des institutions de rattachement (BIB15). Cette identification claire est d'autant plus importante qu'au sein d'un consortium, il peut y avoir plusieurs instances de résolveurs de liens qui cohabitent : il est donc crucial de pouvoir distinguer les ressources spécifiques de chaque établissement. Lorsque cette contextualisation n'est pas réalisée, cela peut être perturbant pour l'utilisateur : soit parce qu'en l'absence de résultats sur l'instance locale le résolveur conduit vers l'instance consortiale (BIB06), soit parce que l'utilisateur ne peut pas filtrer les ressources en fonction de son établissement d'origine (BIB10).

### C. L'intégration avec les systèmes d'authentification

La façon dont l'outil de découverte s'articule avec les différents systèmes d'authentification est également importante pour le bon fonctionnement de l'outil, au point que certaines bibliothèques choisissent leur outil en fonction de leur possibilité d'interfaçage avec le système d'authentification (BIB08, BIB23).

Le système d'authentification est souvent géré par les bibliothèques universitaires via des systèmes de CAS<sup>102</sup> ou Shibboleth<sup>103</sup> articulés avec un système de reverse-proxy (ou serveur mandataire inverse) (BIB23, BIB14, BIB07, BIB26, BIB17, BIB05, BIB15, BIB18, BIB04, BIB11). La plupart du temps, cette authentification ne s'effectue donc pas dans l'outil lui-même, mais elle s'appuie sur un annuaire de l'établissement. Pour les portails consortiaux, les utilisateurs peuvent choisir leur profil en fonction de leur établissement et se connecter via l'annuaire (BIB15). L'absence de cet annuaire peut être un frein au bon fonctionnement du système d'authentification, et pourra empêcher les établissements de bénéficier de l'authentification (BIB04). En contexte consorcial, la reconnaissance des différents proxys des établissements partenaires peut également être complexe à mettre en place (BIB06).

---

<sup>101</sup> « Much of this failure could be attributed to the technology stack problem; by using two different vendors' knowledge bases, there were many instances where the metadata just did not match up to create a resolvable link » BOLLINGER, Stephen et SILTON, Kate. *Op. cit.*, p. 31.

<sup>102</sup> Central Authentication Service (CAS) : système d'authentification unique *open source* développé par l'université de Yale. CAS / Apereo [en ligne]. [s. d.]. [Consulté le 26 novembre 2016]. Disponible à l'adresse : <https://www.apereo.org/projects/cas>.

<sup>103</sup> Shibboleth est un système de fédération d'identités *open source* développé par le consortium éponyme. Shibboleth [en ligne]. [s. d.]. [Consulté le 26 novembre 2016]. Disponible à l'adresse : <http://shibboleth.net/>.

Pour les comptes utilisateurs, il peut arriver que les usagers d'une bibliothèque s'authentifient chez le fournisseur. Dans ce cas, il est important de veiller à la façon dont les données personnelles transitent, notamment à leur chiffrement (cryptage) car l'absence de celui-ci peut être problématique au regard de la réglementation de la Commission Nationale Informatique et Libertés sur la sécurité des données personnelles (BIB22). Concernant la politique de confidentialité des données personnelles, les fournisseurs d'outils mettent à disposition des serveurs en Europe (BIB07) car il s'agit là très souvent d'un prérequis des CCTP.

Le niveau auquel on demande aux utilisateurs de s'authentifier pose également question. Traditionnellement la consultation des OPAC est libre et n'impose pas aux utilisateurs de s'authentifier. Mais si l'on considère l'outil de découverte comme une base de données donnant accès à des ressources payantes, la question de l'authentification à l'entrée, avant même toute consultation, se pose plus légitimement : certains services de découverte demandent à leurs utilisateurs de s'authentifier pour pouvoir effectuer des recherches dans l'outil. Les bibliothèques ont tout de même la possibilité de mettre en place la consultation en mode invité, afin de permettre aux utilisateurs non authentifiés de pouvoir quand même effectuer des recherches dans l'outil, sans avoir accès au texte intégral (BIB11, BIB21, BIB17, BIB05, BIB24, BIB02). Cette question du niveau d'authentification rejoint celle du niveau de proxification des liens : on peut en effet se demander s'il convient de proxifier la recherche elle-même ou seulement l'accès au texte intégral (BIB02, BIB21).

En lien avec la question de ces différents niveaux d'authentification possible, on retrouve celle de l'affichage des résultats provenant des bases d'A&I (dont la valeur ajoutée réside justement dans le référencement bibliographique seul). Dans certains cas ces notices ne s'affichent pas du tout en mode non authentifié (BIB17, BIB14, BIB02, BIB13) : pour une même recherche on pourra avoir un nombre de résultats différents selon qu'on est authentifié ou pas. Dans d'autres cas, ces notices s'affichent avec le titre masqué par un message du type « connectez-vous pour accéder à la ressource » (BIB21, BIB08, BIB24). Cet affichage peut être le fruit d'un choix plus ou moins contraint, car le fait de masquer ces notices peut entraîner d'autres problèmes au niveau de la gestion du résolveur de liens (BIB24). Ces problèmes d'affichage de résultats différents selon qu'on est authentifié ou pas ont été pointés également par la NISO : « Certains résultats peuvent être affichés pour des utilisateurs authentifiés, et d'autres non, mais la définition de ces distinctions, si elle existe, est parfois vague. »<sup>104</sup>

Ainsi, de nombreuses bibliothèques incitent leurs utilisateurs à s'authentifier dès que possible sur l'outil (BIB17, BIB24, BIB08, BIB11, BIB01, BIB02) en proxifiant au maximum les liens d'accès, y compris aux ressources en *open access* (BIB17, BIB24, BIB11, BIB21, BIB02). Il s'agit avant tout de prévenir l'effet de frustration qui découlerait d'un accès tronqué, mais également de pouvoir générer facilement des statistiques en exploitant les logs (ou traces enregistrées lors de consultation) du proxy<sup>105</sup> ou encore de remédier aux faiblesses de gestion de l'OpenURL par les résolveurs de liens.

---

<sup>104</sup> « Some elements might be displayed to authenticated users, and some not, but definitions of these distinctions are sometimes vague, if they are described at all ». National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 3.

<sup>105</sup> Voir *infra* III, 1, C, « Mesures de l'impact sur la consultation de la documentation numérique ».



## D. L'articulation avec d'autres bases

### *Intégration via OAI-PMH*

Plusieurs bibliothèques ont mis en place le moissonnage dans leur outil de découverte d'une ou plusieurs bases locales via OAI-PMH (BIB26, BIB18, BIB01, BIB04, BIB06, BIB10, BIB24, BIB15, BIB05, BIB19). Il peut s'agir d'un entrepôt de publications institutionnelles, de thèses et mémoires ou de bibliothèques numériques patrimoniales ou thématiques.

Comme pour l'intégration du catalogue, cette démarche implique un travail d'adaptation des données qui passe par un *mapping* des champs à moissonner. Des problèmes spécifiques peuvent se poser pour la gestion de certains champs au format Dublin Core ou pour l'automatisation du moissonnage (BIB18, BIB01, BIB10).

Les délais de mise en œuvre du moissonnage peuvent être de plusieurs mois (BIB15, BIB24, BIB10). Comme pour les données locales, il peut y avoir d'abord un moissonnage initial global, puis un moissonnage incrémental. La fréquence du moissonnage pourra répondre ensuite à la fréquence souhaitée par la bibliothèque : par exemple quotidienne ou hebdomadaire. Parfois des décalages ou des interruptions de moissonnage peuvent se présenter comme pour d'autres données locales (BIB05).

Néanmoins, malgré ces difficultés de mise en route, une fois mis en place, le fonctionnement de ce moissonnage est apprécié positivement par les bibliothèques concernées et répond de manière satisfaisante à leurs besoins.

### *Complémentarité avec un moteur de recherche fédérée, des widgets, des systèmes de recommandation, des ERMS*

Malgré sa vocation unificatrice, dans plusieurs cas, l'outil de découverte n'a pas entièrement remplacé les moteurs de recherche fédérée. Certaines bibliothèques maintiennent un outil de recherche fédérée en complément, par exemple pour générer une liste de bases de données afin de permettre une recherche par nom de la base de données (BIB12, BIB23, BIB19).

Dans d'autres cas, ce sont des *widgets* affichés sur la page de résultats qui permettent de rendre visibles des bases non intégrées dans l'index (BIB24, BIB02, BIB05). Ces *widgets* ne sont pas contextuels et s'affichent quelle que soit la recherche effectuée. En cliquant sur le *widget* on quitte l'outil de découverte et il faut refaire la recherche dans la base concernée.

Certaines bibliothèques ont pu paramétrer un système de recommandations en fonction de mots clés afin de renvoyer vers ces bases non moissonnées par l'outil (BIB25, BIB07, BIB14, BIB20). On pourra ainsi faire remonter le nom d'une base de données, en l'associant à un terme de recherche : par exemple, en associant le résultat de la base de données « Dalloz » à une recherche « Dalloz », « droit » ou « jurisprudence ». Si ce système est limité à une seule suggestion pour chaque mot clé, il est intéressant de noter qu'il peut être détourné pour afficher des informations générales sur la bibliothèque, comme par exemple afficher une notice de la bibliothèque qui corresponde au mot « horaires » ou « BU ».

Enfin, on peut signaler également la possibilité de combiner la recherche dans l'outil de découverte avec l'interrogation de fichiers extraits d'un ERMS

(BIB02). Dans ce cas, la bibliothèque a paramétré l'extraction automatique d'un fichier de l'ERMS qui génère dans la page de résultats de l'outil de découverte un encart pour la ressource recherchée. Cet affichage permet de mettre en valeur la ressource et de faciliter l'accès à une base de données dans l'outil de découverte. C'est un *widget* de l'outil de découverte qui interroge en temps réel ce fichier. C'est là une solution de contournement pour compenser les absences de certains éditeurs dans l'index centralisé et, au-delà, pour proposer un accès direct aux bases de données (qu'elles soient intégrées dans l'index ou pas).

### 3. UNE RECHERCHE UNIFIEE ?

#### A. Une recherche unifiée, des métadonnées distinctes

##### *Des métadonnées très hétérogènes*

Outre l'intégration avec les systèmes locaux, de nombreux articles de la presse professionnelle soulignent que la qualité des métadonnées est cruciale pour permettre à l'outil de fonctionner efficacement<sup>106</sup>. Or la qualité des métadonnées peut être difficile à évaluer. Lorena Ávila-García, Virginia Ortiz-Repiso et David Rodríguez-Mateos notent que les fournisseurs donnent très peu d'informations sur les traitements qu'ils font subir aux métadonnées et n'indiquent pas avec suffisamment de précision s'il s'agit des données d'origine ou si ces données sont créées par leurs propres systèmes de découverte<sup>107</sup>. Plusieurs bibliothèques ont ainsi pointé l'absence de visibilité et de possibilités de rectification sur la façon dont les données sont intégrées et structurées dans l'index centralisé, plus particulièrement en ce qui concerne les *mappings* et les traitements effectués après versement (BIB23, BIB26, BIB15, BIB21, BIB24, BIB07, BIB19).

Interrogés sur ces processus, les fournisseurs ont déclaré travailler sur des formats internationaux : il s'agit notamment du MARC, MARC21, MARCXML, UNIMARC, Dublin Core pour EDS ; et du MARC21, du KBART et du Dublin Core pour Summon. Ils ont indiqué que le versement des données peut se faire de multiples manières : transfert FTP, *web service*, moissonnage OAI ou envoi de

<sup>106</sup> DEMPSEY, Lorcan. Thirteen ways of looking at libraries, discovery and the catalogue: scale, workflow, attention. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 179-202; ELLERO, Nadine. Integration or Disintegration: Where Is Discovery Headed? *Journal of Library Metadata* [en ligne]. Octobre 2013, Vol. 13, n° 4, p. 311-329. ; E-Data Quality Working Group OCLC. *Success strategies for electronic content and access: a cross-industry white paper*. Dublin, OH : OCLC, 2014. [Consulté le 20 mars 2016]. Disponible à l'adresse : <http://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>; Groupe de Travail du Réseau des Bibliothèques de l'Université De Moncton. *Les outils de découverte : Rapport du groupe de travail du Réseau des bibliothèques de l'Université de Moncton. Version abrégée destinée au public* [en ligne]. [S. l.] : [s. n.], mars 2011. [Consulté le 21 février 2016]. Disponible à l'adresse : [http://www.umoncton.ca/umcm-bibliotheque-champlain/files/umcm-bibliotheque-champlain/wf/wf/pdf/Outils\\_decouverte\\_Rapport\\_final\\_abrege\\_mars\\_2011.pdf](http://www.umoncton.ca/umcm-bibliotheque-champlain/files/umcm-bibliotheque-champlain/wf/wf/pdf/Outils_decouverte_Rapport_final_abrege_mars_2011.pdf); HAN, Myung-Ja K. et WEATHERS, William Fletcher. Know Thy Metadata: Metadata Challenges in Discovery Services. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 267-276 ; ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.*; KABASHI, Arta, PETERSON, Christine, PRATER, Tim, et al. *Discovery services : a white paper for the Texas State Library and Archives' Commission* [en ligne]. [S. l.] : [s. n.], 2014. [Consulté le 20 mars 2016]. Disponible à l'adresse : [https://www.tsl.texas.gov/sites/default/files/public/tslac/lot/TSLAC\\_WP\\_discovery\\_\\_final\\_TSLAC\\_20140912.pdf](https://www.tsl.texas.gov/sites/default/files/public/tslac/lot/TSLAC_WP_discovery__final_TSLAC_20140912.pdf); HUDSON, Cole et HUKILL, Graham. One-to-Many: Building a Single-Search Interface for Disparate Resources. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 141-153 ; ZHU, Julie et KELLEY, Jalyn. Collaborating to Reduce Content Gaps in Discovery: What Publishers, Discovery Service Providers, and Libraries Can Do to Close the Gaps. *Science & Technology Libraries* [en ligne]. Octobre 2015, Vol. 34, n° 4, p. 315-328.

<sup>107</sup> ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.*, p.7.

fichier Excel. Les deux fournisseurs interrogés ont déclaré que certaines données sont obligatoires selon le type de contenu concerné (ces champs obligatoires ne seront pas les mêmes selon qu'il s'agit d'un article ou d'un livre par exemple) et Proquest a précisé en particulier que la présence d'identifiants comme le DOI ou le PMID est bien évidemment souhaitée. Les deux fournisseurs interrogés ont déclaré effectuer un contrôle sur les données versées pour vérifier la conformité du fichier. Dans le cadre de Summon, Proquest a indiqué toutefois que certains fournisseurs de données n'étaient malheureusement pas toujours compatibles avec l'OpenURL. En outre, les fournisseurs ont également déclaré enrichir et normaliser les données de différentes manières : *mappings* propres à chaque type de contenu, format et fournisseur de données (Ebsco) ; attribution d'un titre autoritaire basé sur la Bibliothèque du Congrès (Proquest) ; enrichissement des sujets (les deux).

Malgré ces traitements, la qualité des données intégrées dans l'index de découverte, du fait de leur grande variété de sources, peut se révéler très hétérogène, comme le signale Peter Webster : « Les métadonnées utilisées dans les services de découverte sont issues de différentes sources et ont été élaborées avec des visées différentes. Les différents ensembles de métadonnées, décrivant des livres, des articles de revues, ou des ressources d'entrepôts en ligne, ne sont pas centralisés avec cohésion dans un seul et même entrepôt »<sup>108</sup>.

Les commentaires des bibliothèques interrogées sur la qualité des métadonnées reflètent cette hétérogénéité : pour un certain nombre de bibliothèques, la qualité est satisfaisante pour extraire l'information nécessaire pour l'utilisateur (BIB14, BIB17, BIB06, BIB18). Pour d'autres, tout en reconnaissant que la qualité des métadonnées pourrait être améliorée, cela relève essentiellement d'une préoccupation de bibliothécaires mais ne constitue pas un problème bloquant pour les utilisateurs (BIB02, BIB25). D'autres encore préfèrent privilégier le besoin d'exhaustivité à la qualité (BIB01), ou la rapidité des réponses à la complétude (BIB19). Mais pour de nombreuses bibliothèques, la qualité des métadonnées a pu être jugée comme inégale, voire relativement pauvre, quel que soit le fournisseur concerné, que ce soit pour les articles ou pour les livres numériques (BIB26, BIB12, BIB24, BIB19, BIB08, BIB10, BIB11, BIB23, BIB21). On peut citer à ce propos un commentaire relevé par Marshall Breeding dans son enquête de 2014 : « Nous étions préparés à l'idée que le service de découverte allait rendre visibles des incohérences dans nos données, ce qui est normal [...]. Cependant, nous n'étions pas préparés à la somme d'incohérences qui allaient apparaître en raison des erreurs ou d'approches non standardisées que les fournisseurs avaient adoptées avec leurs propres données. »<sup>109</sup> Plusieurs bibliothèques espèrent qu'une amélioration de la qualité de l'index suivra la fusion commerciale des fournisseurs Proquest et Ex Libris (BIB12, BIB10, BIB19).

La faible qualité des métadonnées peut avoir plusieurs causes. Comme le signalent Julie Zhu et Jalyn Kelley dans une étude sur la qualité des données des

<sup>108</sup> « The metadata used in discovery services is maintained by different sources, and was created for different purposes. The separate bodies of metadata, which describe books, journal articles, and online repository materials, are not cohesively centralized in one repository. » WEBSTER, Peter. *Op. cit.*, p. 655.

<sup>109</sup> « We were prepared for the fact that [the discovery service] would expose inconsistencies in our data which is fine [...]. However we were not prepared for the amount of inconsistencies that would occur as a result of errors / nonstandard approaches that providers have taken with their own data. » BREEDING, Marshall. Chapter 2: Major Discovery Product Profiles. *Op. cit.*, p. 44.

outils de découverte, la chaîne de fourniture de métadonnées peut être imparfaite à chacune de ses étapes, sans que cela soit attribuable à un seul outil<sup>110</sup>.

Au niveau du fournisseur de données, Julie Zhu et Jalyn Kelley signalent que les données peuvent être manquantes ou incomplètes pour différentes raisons : commerciales (volonté de commercialiser ces données par ailleurs), techniques (manque de moyens pour indexer correctement les articles), ou stratégiques (manque de vision globale de l'utilité de ce type de données et de la conformité aux standards). À ce propos, les bibliothèques interrogées ont constaté des cas où c'est bien la donnée à la source, chez l'éditeur ou l'entrepôt *open access*, qui est pauvre ou erronée (BIB13, BIB08, BIB18, BIB24). Elles ont pointé aussi des différences de qualité selon la langue source (BIB19, BIB05). Interrogés sur ce point les fournisseurs d'outils ont déclaré effectuer des contrôles systématiques sur les versements des données (avec un certain nombre de champs obligatoires par type de contenu). Ce contrôle se traduit d'abord par une vérification automatique (pour tester la conformité du fichier versé) et, en cas d'erreur détectée, par une vérification et une correction manuelle.

Au niveau des fournisseurs d'outils, Julie Zhu et Jalyn Kelley signalent que les données peuvent ne pas être indexées de la façon la plus riche possible pour différentes raisons : des aléas techniques (défaillances des transferts FTP ou du moissonnage OAI, *bugs*, pertes de données), le choix des fournisseurs de n'indexer qu'une partie des données fournies pour gagner du temps ou alléger l'index, des écarts temporels entre la publication du contenu et l'indexation dans l'outil, notamment pour les actes de colloque et les livres numériques qui peuvent être plus longs à traiter.

La façon dont les index centralisés intègrent les données sources plus riches peut éventuellement conduire à une déperdition d'information avec une indexation qui pourra être moins précise que ce que la bibliothèque aurait souhaité, par exemple par la perte des mots clés français ou encore par l'attribution dans l'index de type de documents ou de langues erronés (BIB23, BIB08, BIB01). En particulier pour les types de documents, les différents *mappings* utilisés par les index centralisés pour intégrer les données des différentes sources peuvent parfois permettre moins de nuance ou de précision que la typologie des données d'origine<sup>111</sup>.

Le mode de récupération des métadonnées peut avoir également un effet sur leur qualité : lorsque les métadonnées proviennent d'une base diffusée par la même société que celle qui diffuse l'outil de découverte, il y aura plus de possibilités d'exploiter les données fournies (BIB15, BIB11). Inversement, lorsque les données ne sont pas récupérées directement auprès du diffuseur de données mais intégrées depuis un tiers ou une API, il pourra y avoir des pertes d'informations et les données pourront être intégrées plus difficilement dans l'interface de découverte (BIB23, BIB04, BIB19). En conséquence, pour un même titre, on pourra disposer d'informations variables selon la source et la provenance de la notice.

D'après Julie Zhu et Jalyn Kelley, en bout de chaîne, une déperdition d'information peut aussi avoir lieu au niveau des bibliothèques elles-mêmes en cas

<sup>110</sup> ZHU, Julie et KELLEY, Jalyn. *Op. cit.*, p. 325.

<sup>111</sup> Par exemple pour le traitement des données de ressources en *open access* voir GAINES, David J. et FAGAN, Jody Condit. OAIster on EBSCO Discovery Service, FirstSearch, and OAIster.worldcat.org. *The Charleston Advisor* [en ligne]. Avril 2015, Vol. 16, n° 4, p. 9-10. RENAVILLE, François. Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? *Op. cit.*, p. 242.

de paramétrage inadapté des ressources déclarées dans la base de connaissance de l'outil de découverte ou du résolveur de liens.

### *Le dédoublement*

Si l'index général centralisé est censé dédoubler, les procédés utilisés pour ces opérations ne sont pas toujours clairement visibles pour les bibliothèques. Marshall Breeding signale que les fournisseurs n'ont pas un fonctionnement homogène à ce sujet : certains fusionnent les notices dans une stratégie d'enregistrement unique, d'autres gardent des enregistrements différents pour chaque fournisseur de données. Ces différences dans la structure de la base de données peuvent avoir des implications sur la taille et le fonctionnement de l'index<sup>112</sup>. Certaines bibliothèques ont indiqué que dans leur cas il s'agit moins d'un dédoublement, que d'une frbrisation ou d'un regroupement de plusieurs notices sous une même notice chapeau (BIB01, BIB19).

Interrogés sur leurs pratiques de dédoublement, les fournisseurs ont déclaré que celui-ci est possible à partir du moment où le système est assuré d'avoir plusieurs fois une référence identique, mais ils ont assuré avoir des procédures différentes pour détecter et traiter les doublons.

Pour Ebsco, en ce qui concerne les données de l'index général centralisé, les doublons se repèrent à des données absolument identiques (sans aucun écart, ne serait-ce que typographique) dans les quatre champs suivants : titre, auteur, date, ISSN ou ISBN. L'inversion du nom ou du prénom, une ponctuation différente, ou encore la présence d'un identifiant avec ou sans tiret empêchera donc le dédoublement. Ebsco reconnaît que ce système n'est pas parfait, mais à partir du moment où ces éléments coïncident, l'outil privilégie la référence la plus complète afin d'afficher la référence la plus riche. En effet Ebsco a affirmé ne pas réaliser de fusions de notices, en combinant certains champs d'une notice avec certains champs d'une autre : il n'y a donc pas de mélange de métadonnées de différentes notices dans une notice tierce qui les agrégerait toutes. Pour Ebsco, cela est exclu car les métadonnées appartiennent aux fournisseurs de données et ce faisant, le fournisseur d'outil serait en train de manipuler des métadonnées dont il ne serait pas propriétaire. Dans ce cas, le producteur d'outils de découverte n'a donc pas le droit de mixer ou de fusionner les données des fournisseurs de données.

Pour sa part, Proquest a indiqué que, dans le cadre de Summon, le dédoublement automatique est réalisé par des scripts qui vont effectivement fusionner les notices si besoin. Ce dédoublement s'appuie sur les champs suivants : titre, auteur, date de publication, pages. Sur ces critères, le script cherche des chaînes des caractères identiques. Lorsque certains critères sont concordants mais un champ diffère (par exemple une date), il y a une vérification humaine, grâce à un identifiant, comme le DOI par exemple, pour être sûr de conserver la bonne information et si besoin effectuer des corrections dans l'index centralisé.

Pour les bibliothèques, les problèmes de dédoublement peuvent se poser à différents niveaux. Il peut s'agir de doublons présents dans l'index. Plusieurs bibliothèques ont constaté des doublons dans l'index de découverte, notamment pour les articles et les chapitres de livres numériques (BIB26, BIB22, BIB07, BIB25, BIB05, BIB24).

---

<sup>112</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.* p. 5.

Par ailleurs, lorsque le fournisseur procède par notices chapeau, le procédé de regroupement est rarement contextualisé en fonction des ressources déclarées par la bibliothèque. Il peut donc y avoir plusieurs doublons d'une même notice regroupés au niveau de l'index central sous une même notice chapeau, mais cette architecture persistera même dans le cas où une bibliothèque n'aura activé qu'une seule source de la notice et pas les autres : au bout de cette notice chapeau, il n'y aura qu'une seule notice, ce qui rend l'étape de regroupement inutile (BIB23, BIB01, BIB19).

Il peut s'agir aussi de doublons entre l'index de découverte et le SIGB. Cela peut notamment avoir lieu pour les notices de livres numériques lorsque celles-ci sont issues du catalogue (BIB07, BIB10, BIB15). Selon les outils et leur niveau de paramétrabilité les bibliothèques pourront établir des règles plus ou moins fines de dédoublement des notices de leur SIGB en fonction de différents champs clés comme l'identifiant, le titre, l'année... (BIB04, BIB26). Mais l'écart de qualité entre les données de l'index et celle d'un catalogue peut être important : lorsque des notices riches, provenant du SUDOC, se retrouvent en doublon avec des notices de l'index, le dédoublement risquerait de les appauvrir (BIB04).

En contexte consorsial, les soucis de dédoublement peuvent être liés à la fonction de « métacatalogue » de l'outil de découverte lorsque celui-ci assure la fusion de multiples catalogues préexistants (BIB15). Comme le titre est un élément du dédoublement entre les différentes notices des SIGB, pour deux notices avec un titre signifiant identique, il suffira de la présence d'une simple mention "texte imprimé" dans l'une et de son absence dans l'autre, pour empêcher le regroupement.

Les bibliothèques ont aussi signalé des doublons entre l'index et les données de leur entrepôt institutionnel : par le jeu des moissonnages réciproques entre différents entrepôts en *open access*, une notice peut se trouver dans un entrepôt institutionnel moissonné directement ainsi que dans un autre entrepôt intégré dans l'index centralisé (BIB26). Ainsi, il arrive qu'une bibliothèque choisisse de ne pas activer la base d'archives ouvertes HAL dans l'index de découverte par crainte de ces doublons (BIB11). Ce problème des moissonnages multiples des entrepôts en *open access* et des risques de doublons qu'ils entraînent a également été soulevé par François Renaville dans son étude sur le signalement de l'*open access* dans un outil de découverte<sup>113</sup>.

Ces risques de doublons ont pu pousser les bibliothèques à repenser leurs circuits de signalement de leurs ressources numériques. Pour éviter les doublons, il arrive qu'une bibliothèque décide de ne plus cataloguer de ressources numériques dans leur SIGB mais de les signaler uniquement dans leur base de connaissance, à travers leur liste A-Z (BIB18).

Cela dit, certaines bibliothèques estiment que les inconvénients liés aux doublons restent maîtrisables, dans la mesure où il est possible de n'activer dans l'outil que les ressources absentes du SIGB (BIB02), ou qu'ils ne constituent pas un frein majeur pour les utilisateurs (BIB06, BIB13).

---

<sup>113</sup> RENAVILLE, François. Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? *Op. cit.*, p. 242-243.

### *La gestion des autorités*

En ce qui concerne les autorités, de nombreuses bibliothèques ont signalé que la gestion des autorités dans l'outil de découverte n'est pas complètement satisfaisante, ce qui se traduit par des doublons et des facettes manquantes ou peu exploitables (BIB12, BIB23, BIB07, BIB14, BIB15, BIB05, BIB21, BIB10). Ce contrôle encore imparfait des autorités concerne de manière générale l'ensemble des auteurs, tous fournisseurs confondus. Les données de l'index provenant de sources diverses ne sont pas (ou pas assez) normalisées : les données d'autorités de l'index restent peu ou prou aussi hétérogènes que celles des données sources. C'est donc là aussi un aspect sur lequel les outils de découverte ont une grande marge de progression.

Interrogés sur ce point, les fournisseurs d'EDS et de Summon ont déclaré qu'ils effectuent des contrôles sur les autorités. Ces contrôles sont automatisés en priorité et complétés éventuellement par une vérification ou une correction manuelle. Ces corrections peuvent s'appuyer sur des référentiels existants, comme celui de la Bibliothèque du Congrès par exemple pour Summon. Les fournisseurs d'outils regrettent que tous les fournisseurs de données n'utilisent pas un même standard international pour les autorités, ce qui permettrait d'harmoniser les données à la source. Il peut en effet y avoir beaucoup de présentations différentes des noms d'auteur selon les fournisseurs de données et l'exploitation qui pourra en être faite ne sera pas complètement satisfaisante en l'absence d'un même standard appliqué par tous.

Néanmoins des progrès sont en cours en ce sens car l'usage de référentiels autorités se développe dans les outils de découverte, notamment pour ce qui a trait au référencement des autorités dans le web de données. Actuellement les URI ne sont pas gérées dans les outils de découverte, mais des évolutions sont en cours dans ce domaine. Ainsi, la BIB04 travaille à générer des URI dans son outil de découverte afin de pouvoir exploiter le référentiel IdRef mis en place par l'ABES.

En ce qui concerne les autorités-matière, la qualité de l'indexation est mise en avant par certains fournisseurs d'outils. Néanmoins la mise en correspondance ou la normalisation des thésaurus existants impliquerait un travail conséquent et il n'est pas sûr que ce soit une priorité pour les fournisseurs (BIB02, BIB24, BIB11). Interrogés sur ce point, les deux fournisseurs ont signalé que pour les sujets, ils effectuent un travail d'enrichissement et de réindexation en fonction de normes internationales (en anglais), avec des mots clés spécifiques. Mais il semble peu envisageable de réaliser ce travail pour toutes les sources, toutes les langues, toutes les bases.

Malgré cela, pour plusieurs bibliothèques, ces problèmes liés aux autorités ne constituent pas des inconvénients majeurs, soit parce qu'elles-mêmes n'utilisent pas les autorités dans leur SIGB soit parce qu'elles estiment qu'ils ne représentent pas des écueils pour l'utilisateur (BIB22, BIB20, BIB25 BIB13, BIB02, BIB01, BIB11, BIB10).

## **B. Les facettes**

La navigation par facettes, permettant de filtrer les résultats *a posteriori*, au lieu d'imposer le choix d'un index *a priori* à l'utilisateur, est présentée comme une des avancées des interfaces de découverte.

Or comme le faisait remarquer Nadine Ellero ainsi que Lorena Ávila-García, Virginia Ortiz-Repiso et David Rodríguez-Mateos, le fonctionnement des facettes est profondément lié à la qualité des métadonnées indexées<sup>114</sup> : plus les données sont riches, homogènes et contrôlées, plus les facettes ont des chances d'être efficaces.

Les principaux points problématiques soulevés par les bibliothèques concernant les facettes, quels qu'elles soient, sont :

- Leur choix limité (parmi les facettes proposées par les fournisseurs).
- Leur application partielle (aux métadonnées de l'index ou aux métadonnées du SIGB).
- Leur fonctionnement inégal dû au faible niveau de normalisation des métadonnées.
- Une possibilité de paramétrage variable, notamment pour ce qui est des *mappings* et des traductions des intitulés.

On examinera ici les facettes les plus généralement présentes dans les outils de découverte et qui ont suscité le plus de commentaires lors des entretiens.

### ***Bibliothèque***

La facette « bibliothèque » ne fonctionne, dans la totalité des cas, que sur les données locales : données provenant du catalogue, de listes A-Z ou de bibliothèques numériques. Néanmoins, elle est fortement utilisée, spécialement dans des interfaces uniques où il n'y a pas d'autre accès aux données du SIGB.

Les données de la facette « bibliothèque » proviennent du SIGB ou des bases locales via le(s) *mapping(s)*. En conséquence, elles sont le reflet de ce qui existe dans le SIGB et les bases locales. En cas de fusion de bibliothèques dans le SIGB, il est nécessaire de prévenir le fournisseur pour que le nouvel affichage soit effectif dans l'outil (BIB25). Cela peut être également problématique lorsque la bibliothèque souhaite avoir un affichage des bibliothèques dans l'outil de découverte qui soit différent de celui du SIGB : par exemple une facette regroupant dans l'outil de découverte plusieurs bibliothèques existantes dans le SIGB (BIB23).

Dans tous les cas, la facette « bibliothèque » ne fonctionne pas sur les données de l'index centralisé, même lorsqu'il s'agit de données correspondant à des abonnements qui sont détenus par la bibliothèque : concrètement, la facette bibliothèque ne porte pas sur les données de l'index centralisé et quand on l'utilise on fait disparaître ces résultats.

Inversement, la facette « bibliothèque » disparaît quand on limite les résultats à la documentation électronique, sauf si des livres numériques sont signalés dans le catalogue avec un exemplaire fictif ou si les imports de la base de connaissance intègrent une mention de la bibliothèque (par exemple en contexte consortial avec contextualisation du résolveur). Ainsi, lorsque l'outil de découverte est utilisé pour explorer la documentation numérique uniquement et qu'il n'y a pas de distinction entre les différents accès à signaler selon les bibliothèques, la facette « bibliothèque » n'est plus opérationnelle et est le plus souvent masquée.

---

<sup>114</sup> ELLERO, Nadine. *Op. cit.* p. 316. ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.*, p. 10.



Pour les consortiums, ces caractéristiques de la facette « bibliothèque » prennent une importance particulière, en raison du nombre important des bibliothèques participantes et de la complexité des droits d'accès différents aux ressources numériques que l'outil doit gérer. Il peut en effet devenir problématique d'avoir des filtres de bibliothèques qui ne fonctionnent que sur les données locales, alors qu'un utilisateur d'une bibliothèque peut trouver un article dans l'index centralisé, croire dès lors qu'il y a accès, et ne pas pouvoir y accéder au final car il n'aura pas la possibilité de filtrer sur la documentation électronique propre à son établissement (BIB10, BIB15, BIB04). Les bibliothèques mettent en place alors d'autres moyens de prévenir l'utilisateur, en passant par exemple par des vues spécifiques ou une signalisation bien différenciée des liens d'accès.

### *Type de document*

La facette « type de document » peut être considérée comme capitale dans un outil de découverte, étant donné la grande variété de sources et de supports brassés. L'information du type de document est le plus souvent obligatoire à la fois pour les données de l'index et pour les données locales mappées. Ce fait explique qu'elle soit une facette fortement utilisée, voire la plus utilisée (BIB18, BIB01).

Sur ce point il faut prendre en compte la diversité de ce que l'on peut comprendre par « type de document » dans le vocabulaire de l'information scientifique et technique : il peut s'agir d'une notion de support, d'une notion de contenu, d'une notion de circulation (règles de prêt), ou encore de source de la notice ou de « contenant ». Donc si l'information est le plus souvent obligatoire, cela ne garantit en rien l'homogénéité des données en question.

Plusieurs bibliothèques ont constaté que cette facette ne fonctionne pas de manière entièrement satisfaisante en raison du manque d'alignement des données de l'index, des problèmes de traduction des intitulés et de l'impossibilité d'agir sur les *mappings* comme elles le souhaiteraient (BIB12, BIB21, BIB11, BIB15, BIB13, BIB08, BIB02).

Les données qui proviennent des sources locales sont intégrées via les *mappings* mais la bibliothèque n'a pas forcément la main sur le résultat final de cette normalisation des types de documents et peut difficilement rectifier les données en cas de besoin (BIB21, BIB11, BIB13, BIB12, BIB23, BIB02).

Etant donné l'impossibilité de créer de nouveaux types de documents via le *mapping*, les bibliothèques sont parfois amenées à sacrifier une partie de leurs nuances de types de documents (mémoires/thèses, thèses imprimées/électroniques, livres imprimés/livres numériques) dans les filtres proposés en facette (BIB18, BIB07, BIB05, BIB02).

Cela peut aboutir à des résultats parfois surprenants dans la facette « types de documents ». Il pourra y avoir par exemple deux entrées différentes et absolument identiques « ressources électroniques » qui ne sont pas fusionnées (BIB15). Des livres numériques signalés dans le catalogue pourront apparaître ainsi dans l'outil de découverte en tant que « livres » ou « ressources électroniques », à côté d'une autre catégorie « ebooks » renvoyant à des chapitres de livres numériques provenant de l'index centralisé (BIB12, BIB02). Cela peut être source de confusion, car en filtrant sur « ebooks », l'utilisateur exclura les notices des livres numériques issues du catalogue et rassemblées sous la rubrique « livres ». On retrouve ce problème dans d'autres cas, par exemple pour des titres de revues électroniques qui s'affichent dans la catégorie « magazines » tandis que la

catégorie « revues universitaires » rassemble des articles issus de revues universitaires et non les revues elles-mêmes (BIB02). Le premier intitulé correspond à une indication de contenu, tandis que le deuxième est plutôt une indication de source, les deux apparaissant sous la même rubrique « type de document ». De même, les thèses et mémoires en ligne pourront figurer dans la facette « type de document » comme « ressources électroniques » au lieu de « thèses et mémoires » (BIB02).

Pour ces raisons, la facette « type de document » est parfois considérée comme difficilement exploitable (BIB11, BIB21, BIB02).

### *Sujet*

La facette « sujet » fait apparaître la grande diversité de données présente dans l'index centralisé.

Il faut noter d'abord que les sujets ne sont pas toujours présents dans les notices : de nombreuses ressources n'ont pas de sujet et ne pourront pas en avoir (BIB01). C'est donc une facette qui fonctionnera nécessairement de manière partielle.

D'autre part, la plupart des bibliothèques s'accordent sur le fait que la facette sujet rassemble des données très hétérogènes, s'appuyant sur des langages d'indexation divers (liste d'autorités-matière comme RAMEAU, LC ou LCSH, thésaurus de bases de données, mots libres), et souffre d'importants écarts selon les différents niveaux de granularité, l'indexation sujet des articles étant dans son ensemble peu fine (BIB26, BIB12, BIB01, BIB10, BIB07, BIB14, BIB20, BIB25, BIB21, BIB11, BIB24, BIB05, BIB13, BIB17, BIB08). En ce sens, William Breitbach avait signalé un problème général au niveau de la recherche par mots sujet : « comme les métadonnées spécifiques des sujets varient selon les disciplines, les mots recherchés peuvent perdre de leur sens<sup>115</sup> ». Cette remarque reste valable pour l'utilisation des sujets dans les facettes.

Les écarts de couverture linguistique ont aussi une influence sur le fonctionnement de la facette « sujet ». Dans les outils de découverte, les sujets disponibles sont le plus souvent en anglais, sans rebonds possibles de l'anglais vers le français ou réciproquement, ce qui limite l'usage de la facette « sujet » (BIB26, BIB23, BIB10, BIB20, BIB21, BIB11, BIB17, BIB13, BIB08, BIB02). Les utilisateurs ont donc intérêt à faire des recherches directement en anglais. À ce propos, on peut signaler le manque d'un algorithme de traduction pour les sujets qui permettrait de proposer des mots sujets français ou de rebondir sur des mots sujets français depuis des mots sujets anglais (BIB10). Marshall Breeding signale également que la question de la diversité linguistique dans la recherche par sujet constitue un domaine avec une forte marge de progression possible, notamment en ce qui concerne les techniques de recherche multilingues<sup>116</sup>.

Par ailleurs, comme l'a pointé Jody Condit Fagan : « les outils de découverte combinent entre eux différents thesaurus, mélangent des collections de revues et ne peuvent offrir des filtres efficaces que lorsque les métadonnées sont communes à

---

<sup>115</sup> « Since subject specific metadata varies between disciplines the semantic meaning of the search terms may be lost » BREITBACH, William. *Op. cit.*, p. 639.

<sup>116</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p.28.

plusieurs ressources »<sup>117</sup>. La très grande spécificité d'un certain type de sujets (indexation d'articles en anglais provenant de bases de données spécifiques) et l'absence d'alignement limitent l'étendue du filtre à une ressource en particulier. Ainsi, en choisissant un sujet, l'utilisateur sera en train de choisir un type de document et/ou une source (articles indexés dans cette base de données spécifique seulement) sans qu'il en soit clairement averti, car ce sujet ne sera utilisé que pour un type de ressource spécifique (BIB07, BIB05, BIB18).

C'est pourquoi le fonctionnement de la facette « sujet » peut être difficile à saisir pour l'utilisateur, qui ne peut disposer d'une information claire sur ce qui est effectivement interrogé au moyen de ce filtre (BIB14, BIB17, BIB16).

D'après plusieurs bibliothèques, ces problèmes ne constituent cependant pas des écueils majeurs pour les utilisateurs (BIB25, BIB24, BIB02). Certaines bibliothèques ont toutefois travaillé à l'amélioration de l'affichage de la facette « sujet » afin de faire ressortir le genre, la couverture géographique et la période historique en exploitant les vedettes RAMEAU récupérées via les *mappings* dans les notices du catalogue (BIB25).

### *Auteur*

La facette « auteur » est l'une de celles qui posent le plus de problèmes aux bibliothèques, en raison du manque d'alignement des données dans l'index centralisé. La disponibilité de cette facette et l'étendue de son d'application sont variables selon les outils. Certaines bibliothèques disposent bien d'une facette « auteur » qui s'applique à l'ensemble des données présentes dans l'outil de découverte : à la fois les données importées du SIGB et celles correspondant aux ressources activées dans l'index centralisé (BIB04, BIB01, BIB12, BIB26). Mais, pour plusieurs autres bibliothèques, la facette « auteur » ne peut fonctionner que sur les données provenant du catalogue (BIB07, BIB14, BIB25, BIB13). Dans beaucoup de cas la facette « auteur » n'est pas du tout disponible dans l'outil (BIB02, BIB11, BIB21, BIB24, BIB17, BIB20, BIB08). Dans d'autres, ce sont les bibliothèques elles-mêmes qui ont fait le choix de ne pas l'afficher afin d'éviter la confusion pour les utilisateurs (BIB10, BIB23).

Pour plusieurs bibliothèques, le manque d'une facette « auteur » constitue une faiblesse importante au niveau de la recherche documentaire par rapport à des fonctionnalités bien rendues par un catalogue classique (BIB11, BIB21, BIB24, BIB08). Cela peut se révéler gênant en cas d'interface unifiée où la recherche sur les données du catalogue passe aussi par l'outil de découverte. Quelques bibliothèques ont demandé des développements sur la facette « auteur » mais cela n'a pas abouti à une solution pleinement satisfaisante : avec les développements proposés, la facette ne pouvait porter que sur un nombre de résultats limités et ne constituait pas une solution viable (BIB11, BIB21). Les aspects relatifs aux autorités restent donc un domaine avec une grande marge d'amélioration possible y compris au niveau des facettes.

---

<sup>117</sup> « Discovery tools combine thesauri together, blend collections of journals, and can offer effective limiters only where metadata is shared among resources ». FAGAN, Jody Condit. *Discovery Tools and Information Literacy. Journal of Web Librarianship* [en ligne]. Juillet 2011, Vol. 5, n° 3, p. 172.

### *Editeur, date de publication*

Tout comme la facette « auteur », la facette « éditeur » est souvent problématique en raison du manque d'alignement des noms d'éditeur et d'une faible normalisation.

Cette facette n'est pas toujours disponible : certaines bibliothèques n'ont pas du tout la possibilité de l'activer (BIB10, BIB20, BIB25, BIB07, BIB14).

Dans certains cas, cette facette n'est disponible que de manière partielle pour les données provenant du catalogue ou des bases locales, archives ouvertes, bibliothèques numériques (BIB26, BIB23) mais pas pour les données de l'index.

Pour d'autres bibliothèques, la facette « éditeur » est bien disponible pour toutes les données (locales et index) (BIB02, BIB11, BIB21, BIB24, BIB17, BIB05). Néanmoins, du fait du manque de normalisation, certaines l'ont estimée peu exploitable (BIB11, BIB21) tandis que d'autres ont fait le choix de ne pas l'activer en raison du manque d'harmonisation des données concernées (BIB23, BIB12, BIB13). Tout comme l'auteur, l'éditeur est une donnée importante pour la recherche documentaire, et l'absence ou la faiblesse de cette facette peut compliquer la recherche d'une référence précise dans le cadre d'une interface unifiée.

La date de publication est une facette qui peut être très utilisée (BIB18, BIB01). Dans certains cas, comme pour l'éditeur, cette facette ne fonctionne que sur les documents provenant du catalogue (BIB17). Mais parfois, même quand elle est disponible, certaines bibliothèques ont choisi de ne pas l'afficher en raison du manque de fiabilité et d'harmonisation dans la présentation des métadonnées de l'index centralisé. En effet, la coexistence au sein de l'index centralisé de plusieurs formats de dates différents avec des niveaux de précision variables (par exemple : année seule vs année – mois – jour) peut rendre piégeuse l'application de ce filtre (BIB23, BIB13). D'autres bibliothèques l'affichent tout de même, mais par contrainte, pour permettre de filtrer par date les résultats issus de notices du catalogue dont la fiabilité est plus sûre (BIB12).

### *Langue*

La facette « langue » peut être également une facette très utilisée (BIB01).

Les ressources signalées dans l'outil de découverte peuvent correspondre à de très nombreuses langues, mais les ressources anglo-saxonnes y sont très largement prédominantes. Or ces mentions de langue ne sont pas homogènes, une même langue pourra être indiquée de manière différente dans la facette d'un même outil (*french, français*). Par ailleurs, ce qui est indexé comme langue peut être variable d'une source de données à une autre : ce ne sera pas nécessairement la langue de l'article mais cela peut correspondre à une langue contenue dans un article (par exemple dans une citation) et la mention de la langue peut ne pas être présente sur toutes les données. L'utilisation d'un filtre sur la facette langue avec des données erronées peut conduire à exclusion des ressources francophones pour lesquelles leur signalement est absent ou incorrect (BIB01). C'est pourquoi les bibliothèques choisissent parfois de ne pas afficher cette facette (BIB23).

## C. L'algorithme de pertinence

Comme l'explique la NFAIS, « un algorithme de pertinence est un programme utilisé par un moteur de recherche pour classer les résultats d'une recherche et les afficher dans un ordre décroissant de pertinence selon un ensemble de facteurs prédéfinis. Si chaque moteur de recherche peut pondérer à sa manière les différents facteurs de pertinence, ils considèrent tous en général les mêmes facteurs en priorité. »<sup>118</sup>. Cet algorithme de pertinence est essentiel au fonctionnement de l'outil de découverte, en raison de la fenêtre de recherche unique (qui favorise les recherches larges), de la diversité des sources intégrées et du grand nombre d'enregistrements dans l'index.

Les algorithmes de pertinence reposent le plus souvent sur des calculs impliquant le nombre de mots, la proximité des mots et le nombre de mots de la notice. Leur fonctionnement a été étudié par Till Kinstler<sup>119</sup> : il s'agit d'estimer la meilleure correspondance en confrontant les chaînes de caractères de la requête avec celles de chaque notice de l'index, pour leur attribuer un taux de similarité qui permettra de trier les résultats en fonction d'un ordre de similarité décroissant. Till Kinstler explique que le calcul de similarité repose sur un modèle spatial vectoriel qui traduit les écarts de similarité en distances entre la requête initiale et les notices de l'index : les vecteurs sont pondérés en fonction des paramétrages propres à chaque outil et le calcul des distances doit fournir le meilleur facteur de pertinence. Ce modèle est appelé *term frequency - inverse document frequency* (abrégié *TF-IDF*). Il s'appuie tout d'abord sur la fréquence d'occurrence des termes recherchés (*term frequency*). Or celle-ci, utilisée toute seule, fait remonter les mots les plus courants de la langue, à savoir les mots vides (ceux qui ne comportent pas d'informations utiles pour la description documentaire). C'est pourquoi ce modèle utilise en contrepartie la fréquence inverse des mots dans le document (*inverse document frequency*) pour identifier les mots vides et réduire leur pondération afin de compenser leur effet sur le calcul de pertinence. La combinaison de *TF* et *IDF* est utilisée pour calculer le poids final d'un terme dans un document de sorte que les mots avec beaucoup d'occurrences dans un document en particulier, mais avec peu d'occurrences dans beaucoup de documents, ont une forte pondération *TF-IDF*. Le modèle spatial vectoriel prend en compte la longueur du document concerné en normalisant les vecteurs en fonction de la longueur du document pour faire en sorte qu'il soit applicable aussi bien à des notices longues ou courtes. Cela peut néanmoins poser problème pour les titres accrocheurs mais pas forcément pertinents, car les mots du titre pourront être fortement pondérés dans les paramètres d'un outil.

Etant donné le grand nombre de résultats ramenés par l'index centralisé, le bon paramétrage de cet algorithme de pertinence est fondamental pour la lisibilité de l'outil. Or, les bibliothèques qui mettent en place un outil de découverte n'ont pas la main sur l'algorithme de pertinence et le fonctionnement de celui-ci demeure opaque pour un certain nombre d'entre elles, y compris parfois après plusieurs années de fonctionnement (BIB12, BIB20, BIB11, BIB05, BIB10, BIB25, BIB02, BIB24, BIB13, BIB21). Cette opacité a fait l'objet de nombreux commentaires critiques,

<sup>118</sup> « A ranking algorithm is the program used by a search engine to rank search results and display them in order of decreasing relevance based upon a set of pre-defined factors. While each search engine may assign different values to the various ranking factors, they usually consider the same factors overall ». National Federation of Abstracting and Information Services. *Op. cit.*, p. 11.

<sup>119</sup> KINSTLER, Till. Making search work for the library user. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 22-24.

relevés par Marshall Breeding dans une enquête auprès des utilisateurs d'outils de découverte en 2015, comme par exemple :

« Si les bibliothécaires n'ont pas l'entière maîtrise de l'algorithme de pertinence utilisé pour présenter les résultats, la bibliothèque a cédé à son fournisseur un élément clé dont elle dispose pour servir ses utilisateurs finaux. Cela est à nos yeux risqué et ne devrait pas être considéré comme acceptable. Il est crucial pour l'avenir des bibliothèques d'être en mesure de déterminer comment leurs ressources sont présentées à leurs utilisateurs finaux. »<sup>120</sup>

On examinera ici les paramétrages de l'algorithme de pertinence d'abord pour les données locales (notamment celles du SIGB) et ensuite pour les données natives de l'index.

### *Données locales*

De nombreuses bibliothèques qui ont intégré leur SIGB ont souhaité que les résultats du catalogue soient placés assez haut dans l'outil de découverte, tout comme éventuellement leurs bases d'archives ouvertes ou bibliothèques numériques locales (BIB06, BIB25, BIB17, BIB02, BIB12, BIB19, BIB18, BIB26, BIB23, BIB10, BIB11, BIB05, BIB24, BIB15, BIB08). Cela répond à la volonté des professionnels de prévenir des effets de bruit et aussi à une crainte de voir les notices du catalogue ou des ressources locales reléguées très loin dans la liste des résultats. Dans un certain nombre de cas, la bonne visibilité des données locales est une condition *sine qua non* avant de faire de l'outil de découverte l'interface documentaire principale (BIB24, BIB17, BIB02).

Du fait de nombreuses demandes de bibliothèques en ce sens, la mise en avant des données locales est souvent proposée comme fonctionnalité de base par l'outil lui-même (BIB25, BIB11, BIB07, BIB15, BIB17, BIB14). Mais cela ne se vérifie pas dans tous les cas et, pour mettre en avant leurs ressources locales, certaines bibliothèques ont demandé à leur fournisseur d'effectuer les modifications nécessaires (BIB10, BIB05).

D'autres ont pu modifier elles-mêmes certains critères de calcul de la pertinence des ressources locales (BIB02, BIB24, BIB12, BIB23, BIB26, BIB19, BIB15). Dans certains cas, les bibliothèques ont pu attribuer elles-mêmes un niveau de pondération global au catalogue à travers l'interface d'administration de l'outil (BIB24, BIB02). Pour d'autres, la modification a pu être effectuée en attribuant des scores à chaque critère de pertinence (mots clés titres, auteur, sujet, texte intégral, date de publication...) qui servira au calcul de la pondération des résultats selon les besoins de la bibliothèque pour ses données locales (BIB12, BIB23, BIB26, BIB19, BIB18, BIB01). Selon le niveau de paramétrabilité des outils, il peut être également possible de modifier le pourcentage de notices locales par rapport aux notices de l'index centralisé affichées dans les résultats : des bibliothèques ont ainsi choisi de proposer sur leur première page de résultats un nombre donné de notices provenant de leurs ressources locales (BIB18, BIB26,

---

<sup>120</sup> « If librarians don't have full control of the relevance ranking algorithm being utilized to present results, the library has surrendered to their vendor an important differentiator for their library to utilize in serving end-users. This to me, is dangerous and should not be found acceptable. It is key to librarians' future to be able to determine how their assets are presented to end-users. » BREEDING, Marshall. Discovery product functionality. *Op. cit.*, p. 24.

BIB01, BIB24), ou d'afficher toujours un résultat local en une position donnée (BIB08).

En contexte consortial, lorsqu'un outil propose plusieurs vues par établissement, chaque vue pourra avoir des critères de pertinence différents correspondant aux souhaits de la bibliothèque (BIB04). En revanche, introduire des pondérations dans la vue consortiale peut être un exercice délicat, moins d'un point de vue technique que politique, car cela reviendrait à donner la priorité à une source de données et pas à une autre au sein d'un outil commun et partagé entre plusieurs établissements.

Cela dit, même quand l'algorithme de pertinence est modifiable par la bibliothèque, cette modification peut s'avérer délicate, étant donné le manque de visibilité sur le fonctionnement effectif de l'algorithme de pertinence. S'agissant de développements qui sont au cœur de l'investissement industriel de ces sociétés, les bibliothèques ont pu constater que la documentation fournie à ce sujet reste très vague (BIB10, BIB11, BIB21, BIB13). Il peut donc être difficile pour la bibliothèque de connaître les modifications qu'elle doit effectuer dans la mesure où elle n'a que très peu d'informations sur le fonctionnement de l'algorithme (BIB12, BIB23, BIB24). Ainsi, les professionnels ont peu de moyens de prévoir l'impact de leurs modifications sur les pondérations de leurs données locales, ce qui entraîne parfois des effets de bord non souhaités (BIB24, BIB13, BIB21). Cela explique que même quand des paramétrages sont possibles, les bibliothèques ne les tentent pas car ils comportent un risque d'altérer des systèmes qui par ailleurs jusque-là fonctionnent correctement (BIB13).

D'autre part, certaines pondérations sont globales et réalisées pour tous les clients par défaut (BIB18, BIB24). Or lorsqu'il s'agit de collections numériques locales autres que le SIGB (par exemple des bibliothèques numériques moissonnées via OAI-PMH), le paramétrage par défaut peut faire remonter des résultats qui ne sont pas nécessairement ceux que la bibliothèque estime être les plus pertinents pour son public cible (BIB06), ou ne pas être entièrement efficace sur des notices ayant une qualité de métadonnées plus faible (BIB05).

Lorsque les outils de découverte mettent à disposition une API, celle-ci peut être exploitée à travers une surcouche logicielle. Il pourrait être envisageable de recalculer des facteurs de pertinence de base autres que ceux de la version standard d'un outil grâce à ce travail depuis l'API. Cependant, cela demanderait un travail de développement pour créer son propre algorithme de pertinence à partir de l'API.

Il peut arriver que la bibliothèque ne puisse pas adapter l'algorithme de pertinence d'une manière satisfaisante, ce qui entraîne le risque de voir les ressources locales dévalorisées, confondues dans les résultats issus de l'index centralisé (BIB22). Mais dans de nombreux cas, la pertinence est plutôt ressentie comme satisfaisante quant aux ressources locales (BIB25, BIB26, BIB24, BIB06, BIB15).

### *Données de l'index*

En ce qui concerne le fonctionnement de l'algorithme de pertinence sur les données de l'index, là aussi la caractéristique générale est celle de l'opacité du fonctionnement. Ce point a également été soulevé dans le rapport de la NFAIS<sup>121</sup>.

---

<sup>121</sup> National Federation of Abstracting and Information Services. *Op. cit.*, p.11.

En raison de la quantité des métadonnées présentes, les bases importantes (gros entrepôts en *open access*, bases de données de références bibliographiques, bases de grands éditeurs) peuvent être surreprésentées dans les résultats, surtout dans les cas d'une recherche large (BIB02, BIB23). Cela reste néanmoins difficile à quantifier et à vérifier par les professionnels administrateurs de l'outil en raison de l'opacité qui entoure le fonctionnement de l'algorithme. De même la qualité très variable des données de l'index, entre les très complètes et les très sommaires, font que certaines ressources faiblement décrites émergent difficilement (BIB05).

En règle générale, les bibliothèques n'ont pas assez d'éléments d'évaluation pour juger de la pertinence des données remontées par l'outil de découverte lorsqu'il s'agit de l'index, c'est-à-dire qu'elles ont peu d'éléments concrets pour en vérifier la neutralité. Lors de l'enquête conduite par Marshall Breeding en 2015, plusieurs bibliothèques interrogées questionnaient fortement cette neutralité<sup>122</sup>. En l'absence d'information sur le fonctionnement de l'algorithme et sur la couverture de l'index qui permettraient d'évaluer la neutralité de l'outil, ce problème demeure un des soucis majeurs soulevés par Marshall Breeding ainsi que par la NISO : « Un autre problème est lié à la possibilité que le fonctionnement d'un service de découverte favorise certaines bases de données ou d'autres ressources en se basant sur des relations commerciales plutôt que sur les besoins des utilisateurs ou les choix de la bibliothèque »<sup>123</sup>. Interrogés sur ce point, les deux fournisseurs d'EDS et de Summon, se sont strictement défendus de privilégier un contenu par rapport à un autre dans les résultats affichés grâce à l'algorithme de pertinence. Pour eux, cela est absolument exclu, car cela remettrait en cause l'objectivité de leurs produits.

De manière plus radicale, Jason Thomale, William Hicks et Mark Phillips ont questionné les principes mêmes sur lesquels reposent les algorithmes de pertinence des outils de découverte. Selon eux, ces outils ne font qu'imiter les caractères superficiels d'une recherche web, mais ne peuvent pas traiter efficacement les problèmes fondamentaux d'une recherche de la même manière qu'un moteur du web<sup>124</sup>. D'après ces auteurs, si les outils de découverte imitent le fonctionnement de la toile, leurs algorithmes de pertinence ne peuvent pas fonctionner aussi bien que ceux des moteurs de recherche du web, car ce qui rend ces derniers efficaces c'est la structure même des données du web. Or celle-ci est différente de la structure des ressources documentaires. Ces auteurs soutiennent par exemple que pour les bibliothèques, dès qu'il s'agit d'un contenu académique, le type de document est très important et variable selon la discipline. Or l'outil de découverte ne peut pas choisir les types de documents les plus importants pour une recherche donnée s'il ne sait pas quels doivent être les types de documents à prendre en compte en priorité. Si les utilisateurs ne filtrent pas les données via les facettes, la présentation des résultats dans une longue liste à l'image de celle des moteurs de recherche du web pourra être entièrement non pertinente au niveau du type de document dominant.

Cela revient à questionner la contextualisation de la recherche et de l'affichage des résultats : l'outil de découverte s'appuie sur une liste de résultats

---

<sup>122</sup> BREEDING, Marshall. Discovery product functionality. *Op. cit.*, p. 21-25.

<sup>123</sup> « Another topic concerns factors related to whether or not a discovery service functions with a bias toward certain databases or content products based on business relationships rather than user needs or library preferences. » National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 3.

<sup>124</sup> THOMALE, Jason, HICKS, William et PHILLIPS, Mark. The Bento Box Design Pattern. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 126.



unique, mais cet affichage unifié devient de plus en plus problématique à mesure que le contenu disponible s'accroît de manière exponentielle.

## D. Des interfaces multiples

### *Interface unique ou interface principale ?*

Un certain nombre de bibliothèques ont pu mettre en place une interface unique intégrant à la fois les données bibliographiques et d'exemplaires du SIGB, ainsi que le compte utilisateur de ce dernier, dans l'interface de découverte (BIB12, BIB01, BIB18, BIB26, BIB23, BIB08, BIB13).

Néanmoins, force est de constater que cette interface unique est loin d'être la solution la plus générale.

Au moment où les entretiens ont été réalisés, il y avait tout d'abord un cas où la bibliothèque n'avait pas encore procédé à l'intégration des données de son SIGB, ce qui faisait nécessairement coexister ce dernier avec l'OPAC (BIB20).

Par ailleurs, dans beaucoup d'autres cas, si l'interface d'interrogation de l'outil de découverte est mise en avant sur le site web, devenant ainsi l'interface documentaire principale de la bibliothèque, l'OPAC classique continue d'exister en parallèle, notamment pour permettre l'accès au compte utilisateur (BIB06, BIB21, BIB11, BIB25, BIB02, BIB05, BIB17). Dans ce cas la bonne intégration des données du catalogue semble tout de même une étape importante pour pouvoir mettre en avant l'outil de découverte comme interface principale (BIB02, BIB05). De même, pour les instances consortiales l'interface de l'outil de découverte coexiste le plus souvent avec les portails des établissements partenaires (BIB15, BIB04, BIB10).

Mais il arrive aussi que ce soit l'OPAC traditionnel qui continue à être mis en avant par la bibliothèque et à être la principale interface, parfois en raison du caractère trop récent de l'installation de l'outil ou de la forte identité professionnelle de l'OPAC<sup>125</sup>(BIB10, BIB07, BIB14, BIB24). En outre, lorsque les bibliothèques disposent de postes dédiés à la consultation de l'OPAC (verrouillés à la navigation extérieure), la décision de mettre en avant l'outil de découverte comme interface unique ou principale entraîne une réflexion sur le devenir de ces postes, car il est difficile d'envisager la consultation d'un outil de découverte sans autoriser l'accès aux ressources numériques qui y sont foncièrement liées (BIB14).

### *Une seule recherche, plusieurs affichages possibles*

L'intérêt de l'outil de découverte est de proposer à travers une seule recherche une multitude de types de résultats, ce qui se traduit, la plupart du temps, par une liste de résultats unique, affichage qui a été adopté par de très nombreuses bibliothèques (BIB02, BIB01, BIB04, BIB05, BIB06, BIB08, BIB10, BIB11, BIB12, BIB14, BIB15, BIB16, BIB17, BIB18, BIB22, BIB24, BIB25).

Néanmoins, dans certains cas, il peut être intéressant de pouvoir générer des affichages de résultats distincts. Pour les installations consortiales en particulier, les outils de découverte peuvent permettre de générer des vues propres à chaque établissement partenaire qui dispose ainsi d'une interface ciblée sur ses propres

---

<sup>125</sup> Ce fort attachement sera développé ultérieurement : cf. *infra* III, 2, « Du côté des utilisateurs professionnels ».

ressources (BIB15, BIB04). L'intérêt de ce dispositif est de pouvoir passer de manière fluide de la vue d'un établissement à la vue globale sans obliger l'utilisateur à refaire sa recherche.

Mais, même en contexte non consorcial, certaines bibliothèques ont décidé de présenter les résultats de manière séparée. Dans certains cas les bibliothèques ont fait évoluer leur liste unique de résultats vers des listes de résultats différenciées. Ainsi, des bibliothèques ont mis en place des onglets distincts, avec possibilité de passer de l'un à l'autre en gardant les termes de recherche, afin de satisfaire les demandes des étudiants les plus avancés et des enseignants-chercheurs qui souhaitaient un accès direct aux articles. Dans ce cas de figure-là, le découpage se fait selon le niveau de granularité : d'un côté les articles et les chapitres, de l'autre les documents de niveau titre, livres imprimés / numériques, revues imprimées / numériques (BIB19, BIB23). Ce choix de présentation pourrait néanmoins être remis en question en cas de nouvelle version de l'outil, par exemple pour ce qui est de l'affichage sur des dispositifs mobiles. Il est intéressant de noter que cette distinction a été mise en place par des bibliothèques qui ont une intégration fluide de leur SIGB et de l'outil de découverte et pour lesquelles l'interface unique ne pose pas de problèmes techniques majeurs.

Il est possible également de séparer les résultats non pas en fonction du niveau de granularité mais en fonction de la source de la notice (catalogue local vs index centralisé). Cela peut être proposé comme une fonctionnalité dans les outils mixtes : on aura ainsi une recherche unique avec l'affichage des résultats dans deux onglets différents et la possibilité de passer de l'un à l'autre, tout en gardant les termes de la recherche initiale (BIB13).

On peut envisager aussi de séparer les résultats par type de document ou de support, en les affichant dans la même page mais dans des sections bien différenciées, ce qu'on a surnommé le mode « *bento* ». Ce terme a été utilisé pour la première fois par les bibliothécaires de la North Carolina State University<sup>126</sup>, et adopté ensuite par les professionnels des bibliothèques pour désigner un type d'interface utilisateur dans laquelle, à l'image de la boîte à repas compartimentée, les résultats issus de différents systèmes sont regroupés par catégorie et s'affichent à l'écran dans des encadrés séparés. En bibliothèque, ces interfaces ne sont pas issues d'une fonctionnalité commerciale, mais sont le fruit de développements locaux, ce qui implique une possibilité d'adaptation accrue et une plus grande marge de manœuvre de la bibliothèque sur les codes sources qu'elle utilise. Outre la North Carolina State University, plusieurs bibliothèques universitaires des États-Unis ont fait ce choix de présentation des résultats dans leur outil de découverte (Stanford, Columbia, Cornell, Michigan entre autres).

Plusieurs articles professionnels se sont penchés sur les atouts et faiblesses du mode « *bento* »<sup>127</sup>. Dans cette configuration, la liste unique de résultats est remplacée par plusieurs sections de contenus plus homogènes et contextualisés.

<sup>126</sup> LOWN, Cory, SIERRA, Tito et BOYER, Josh. How Users Search the Library from a Single Search Box. *College & Research Libraries*. Mai 2013, Vol. 74, n° 3, p. 227-241.

<sup>127</sup> Voir notamment LOWN, Cory, SIERRA, Tito et BOYER, Josh. *Op. cit.* ; THOMALE, Jason, HICKS, William et PHILLIPS, Mark. *Op. cit.*; THOMALE, Jason. *You Gotta Keep 'em Separated: The Case for « Bento Box » Discovery Interfaces* [en ligne]. Portland, OR, 5 décembre 2014. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://code4lib.org/conference/2015/thomale>. Présentation filmée disponible sur <https://www.youtube.com/watch?v=i853dLWavZA> ; SINGLEY, Emily. To Bento or Not to Bento – Displaying search results. Dans : *usable libraries* [en ligne]. 4 janvier 2016. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://emilysingley.net/to-bento-or-not-to-bento-displaying-search-results/>.

Cet affichage permettrait notamment de lutter contre le bruit et la surabondance d'information en regroupant les résultats à l'intérieur de chaque catégorie, profitant du fait que les ressources de bibliothèque sont déjà caractérisées par des types de documents dominants. Le mode « *bento* » serait aussi plus à même d'optimiser le fonctionnement de l'algorithme de pertinence au niveau de chaque système cible : l'utilisateur pourra avoir un aperçu des premiers résultats par type de document, ce qui fera ressortir des éléments dans chacune de ces sections qui n'auraient sans doute pas pu s'afficher dans les premiers éléments de la liste globale. Cela permettrait ainsi de contrebalancer l'algorithme de pertinence pour mieux mettre en avant les résultats propres à chaque catégorie. On peut néanmoins pointer que le mode « *bento* » a l'air de fonctionner de manière semblable à une recherche fédérée (plusieurs requêtes sont envoyées en même temps et les résultats sont récupérés dans un affichage dynamique).

La présentation en mode « *bento* » a été choisie par deux des bibliothèques que nous avons interrogées et réalisée grâce à des surcouches logicielles qui retraitent les données issues des API des outils de découverte, afin de les présenter dans des affichages séparés et contextuels (BIB21, BIB26). Pour une même recherche, on a ainsi plusieurs affichages distincts des résultats, avec des regroupements propres à la bibliothèque (en rayon / en ligne, livres et livres numériques / articles / revues etc.). Ce mode de présentation de résultats permet de répondre à un besoin de contextualisation, notamment pour les usagers des premiers cycles, tout en laissant l'affichage de l'OPAC classique intact, lorsque celui-ci coexiste avec l'outil de découverte. Ce mode de présentation en *bento* est éventuellement envisagé par d'autres bibliothèques afin d'éviter que les utilisateurs soient perdus dans la très grande masse des résultats issus de l'index centralisé (BIB07).



### III. LES OUTILS DE DECOUVERTE ET LEUR UTILISATION

---

#### 1. DU COTE DU PUBLIC

Le bilan de l'utilisation semble bien plus difficile à saisir que celui de la mise en œuvre. Les entretiens que nous avons menés ont fait apparaître le besoin de consolider les outils de mesure d'activité des outils de découverte ainsi que les outils de mesure de la consultation de la documentation numérique. Dans de nombreux cas, du fait du caractère très récent des installations, les données d'évaluation dont disposent les bibliothèques sont encore lacunaires ou insuffisantes pour en tirer des informations exploitables. Nous avons tenté de faire la synthèse de ces données tout en les sachant approximatives, en ayant recours à la littérature professionnelle existante dans le domaine pour compléter notre réflexion.

##### A. Quel public cible ?

Les bibliothèques qui mettent en place un outil de découverte n'ont pas toutes pour objectif de desservir le même public cible.

Pour un certain nombre de professionnels interrogés, l'outil de découverte est adapté aux recherches très larges plutôt qu'aux recherches précises, ce qui le rend davantage susceptible de répondre aux besoins des étudiants (BIB04, BIB18, BIB11, BIB02, BIB12, BIB16). Il permettrait de débroussailler un sujet, en fournissant une vision d'ensemble très rapide, mais serait moins performant pour une recherche fine pour laquelle ce sont plutôt les bases de données spécifiques qui restent pertinentes. Pour les chercheurs et les étudiants avancés, notamment en STM, l'outil pourra difficilement se substituer à des bases de données spécifiques ou à la plateforme à laquelle ces publics sont habitués : plateforme d'un éditeur, bases de données, liste A-Z des revues ou Google Scholar (BIB12, BIB04, BIB18, BIB11, BIB02, BIB16). Plusieurs articles professionnels abondent en ce sens, indiquant que les outils peuvent satisfaire davantage un public d'étudiants de premier cycle qu'un public avancé, et insistent sur la nécessité de maintenir des outils alternatifs (telles que les bases de données spécialisées) pour les chercheurs<sup>128</sup>. Néanmoins ces écrits portent essentiellement sur les pratiques documentaires du monde anglo-saxon, dans lequel les étudiants sont amenés à utiliser des articles dès la première année<sup>129</sup>, alors qu'en France la recherche d'articles intervient souvent plus tard dans les cursus.

---

<sup>128</sup> GUTHRIE, Ana et MCCOY, Rhonda. A Glimpse at Discovery Tools within the HBCU Library Landscape. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 177-191; KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. Discovering the Discovery Tool: The Introduction and Impact on Research and Instruction at Seminole State College of Florida. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 158-176 ; FALCIANI-WHITE, Nancy. Understanding Information Seeking Behavior of Faculty and Students: A Review of the Literature. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 1-21 ; ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. *Op. cit.* FAWLEY, Nancy et KRYSAK, Nikki. Information Literacy Opportunities within the Discovery Tool Environment. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 87-94. KEYSER, Piet de, *Op. cit.*, p. 58.

<sup>129</sup> FALCIANI-WHITE, Nancy. *Op. cit.*, p. 5.

Pour d'autres, étant donné la prédominance des articles dans les index centralisés et leur versant clairement anglophone, les outils de découverte correspondent moins aux besoins de recherche documentaire des premiers cycles qu'à ceux de la recherche (BIB21, BIB19, BIB26). Les étudiants de premier cycle ont à chercher des références en français, en général des livres, souvent des manuels, ce qui n'est pas le type de document dominant dans l'index centralisé. Dans ce cas, l'outil serait plutôt tourné vers les doctorants et les chercheurs, en raison de la grande spécificité des ressources indexées (BIB21, BIB26), pouvant parfois être perçu comme un outil servant essentiellement à la recherche de documentation électronique (BIB24, BIB19, BIB20).

Enfin, les bibliothèques peuvent aussi viser les deux usages : un usage « pédagogique » pour les premiers cycles cohabitant avec un usage de recherche (BIB07, BIB14).

## B. Outils de découverte et *information literacy*

Comme le signalaient Judy Luther et Maureen Kelly, la grande masse d'informations rendues disponibles dans les index centralisés des outils de découverte peut représenter un défi pour les compétences informationnelles d'un utilisateur : « Lorsque le contenu abonde, trouver le contenu adéquat devient un défi »<sup>130</sup>.

L'*information literacy* ou maîtrise de l'information désigne précisément la capacité d'une personne à savoir reconnaître et exprimer son besoin d'information ainsi qu'à trouver, évaluer et utiliser correctement les sources d'information correspondant à ce besoin. Proclamée par l'UNESCO comme « un droit humain fondamental »<sup>131</sup> en 2005, la maîtrise de l'information constitue un champ d'action dans lequel, comme le signalait Sylvie Chevillotte, les bibliothèques universitaires sont très investies et jouent un rôle essentiel<sup>132</sup>. L'Association of College & Research Libraries (ACRL) a, pour sa part, défini en détail un certain nombre de critères et de standards correspondant à l'*information literacy*, à l'usage des bibliothèques d'enseignement supérieur et de recherche<sup>133</sup>. Ces critères concernent notamment la capacité à :

- déterminer la nature et l'étendue de son besoin d'information ;
- accéder à cette information de manière efficace et efficiente ;
- évaluer l'information et les sources de manière critique, intégrer l'information dans son bagage intellectuel ;
- utiliser l'information efficacement pour accomplir un travail spécifique ;

<sup>130</sup> « When content is abundant, finding the right content becomes the challenge » LUTHER, Judy et KELLY, Maureen. *Op. cit.*, p. 67.

<sup>131</sup> *Proclamation d'Alexandrie sur la maîtrise de l'information et l'apprentissage tout au long de la vie* [en ligne]. [s. d.]. [Consulté le 21 octobre 2016]. Disponible à l'adresse : [http://portal.unesco.org/ci/fr/ev.php-URL\\_ID=20891&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/fr/ev.php-URL_ID=20891&URL_DO=DO_TOPIC&URL_SECTION=201.html).

<sup>132</sup> CHEVILLOTTE, Sylvie. Bibliothèques et Information Literacy. *Bulletin des bibliothèques de France*. Janvier 2005, n° 2, p. 42-48.

<sup>133</sup> ASSOCIATION OF COLLEGE & RESEARCH LIBRARIES (ACRL). Information Literacy Competency Standards for Higher Education. Dans : *Association of College & Research Libraries (ACRL)* [en ligne]. 1 septembre 2006. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://www.ala.org/acrl/standards/informationliteracycompetency>.

- comprendre les enjeux socio-économiques et juridiques de l'utilisation de l'information, y accéder et utiliser les sources d'information dans le respect de la légalité.

Chacun de ces critères se décompose ensuite dans une série d'items qui doivent servir à évaluer l'acquisition de chacune de ces compétences par un étudiant.

Les services de découverte à index général centralisé visant spécifiquement le monde académique, il nous a paru intéressant de les confronter avec cette notion d'*information literacy*, afin de comprendre quelles pouvaient être leurs implications en bibliothèque universitaire dans ce domaine.

### *Des fonctionnalités prometteuses pour l'information literacy*

Les outils de découverte proposent un certain nombre de fonctionnalités intéressantes pour l'*information literacy*. Il s'agit tout d'abord d'outils clairement « orientés utilisateurs<sup>134</sup> » qui permettent à ces derniers de s'approprier facilement l'interface proposée grâce à une barre de recherche unique « à la Google », sans leur imposer des syntaxes complexes ou des filtres *a priori*. Cette adaptation des outils de bibliothèques aux pratiques actuelles du web est saluée comme un gain d'utilisabilité et d'efficacité par plusieurs articles professionnels<sup>135</sup> ainsi que par plusieurs bibliothèques (BIB18, BIB19, BIB13, BIB04, BIB23). Pour un certain nombre d'entre elles, le fait de pouvoir se référer à une interface principale permet de simplifier l'utilisation des ressources en proposant un portail centralisateur (BIB06, BIB13, BIB24, BIB15, BIB04).

Avec une interface qui se veut facilement utilisable, l'outil de découverte peut permettre de faire en sorte que l'utilisateur se tourne en priorité vers les ressources de la bibliothèque plutôt que vers un moteur de recherche du web<sup>136</sup>. Il faut noter à ce propos que l'index centralisé recouvre, en dehors de certaines ressources en *open access* pouvant être déclarées par défaut, des ressources activées par la bibliothèque qui en reste l'administrateur final. Il s'agit donc, dans l'ensemble, de ressources sélectionnées par les bibliothécaires et auxquelles ces derniers proposaient souvent déjà des accès via d'autres chemins (BIB17). Elles conservent par là leur caution de qualité intellectuelle et académique.

En outre, de nombreux outils de découverte présentent plusieurs fonctionnalités qui peuvent favoriser une recherche d'information efficace. La fonctionnalité d'autocomplétion est ainsi présente dans de nombreux outils (BIB26, BIB10, BIB07, BIB20, BIB14, BIB25, BIB11, BIB05, BIB02, BIB13, BIB17, BIB16, BIB24). Cette autocomplétion peut s'appuyer sur des titres effectivement présents dans l'index ou dans la liste A-Z. Elle peut aussi se baser sur une liste de recherches déjà effectuées, mais son effet risque alors d'être limité,

<sup>134</sup> DECOSTER, Sara et RENAVILLE, François. *Défis et opportunités des outils de découverte : côté professionnel, côté usager* [en ligne]. Mars 2016. [Consulté le 14 juin 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/194784>. CHRISTENSEN, Anne. Next-generation catalogues: what do users think? Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 1-14.

<sup>135</sup> DEMPSEY, Lorcan. *Op. cit.* HANRATH, Scott et KOTTMAN, Miloche. Use and Usability of a Discovery Tool in an Academic Library. *Journal of Web Librarianship* [en ligne]. Janvier 2015, Vol. 9, n° 1, p. 1-21. ; COMEAUX, David J. Usability Testing of a Web-Scale Discovery System at an Academic Library. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 69-86.

<sup>136</sup> GUTHRIE, Ana et MCCOY, Rhonda. *Op. cit.*

voire trompeur pour un utilisateur qui s'attendrait à voir correspondre ces termes proposés à des résultats précis. Il est parfois possible de mettre en place un service d'autocomplétion amélioré, en exploitant l'API d'un outil de découverte pour obtenir des données enrichies (BIB26).

Le système de suggestions de corrections orthographiques (en anglais *did you mean*) est également intéressant du point de vue de l'utilisabilité et peut favoriser l'*information literacy* en évitant des phénomènes de silence dus à des saisies erronées. Ce système de suggestions est présent dans tous les outils mais il faut néanmoins noter qu'il reste plus adapté aux recherches en langue anglaise qu'en langue française (BIB14, BIB12, BIB23, BIB15) et difficile à évaluer (BIB07, BIB20, BIB10). Certaines bibliothèques n'ont d'ailleurs pas du tout activé cette fonctionnalité (BIB21, BIB13). En outre, même dans un contexte anglophone, l'effet de ces suggestions orthographiques reste mitigé<sup>137</sup>.

Les outils de découverte peuvent proposer également des services de recommandations d'articles. Ces derniers peuvent favoriser l'*information literacy*, dans la mesure où ils permettent d'alerter les utilisateurs sur la parution d'articles sur des sujets proches, qu'ils n'auraient pas recherchés explicitement. Toutefois, ces recommandations proviennent d'une base globale de recherches déjà faites, qui n'est pas spécifiquement ciblée en fonction d'une bibliothèque. Les effets de cette fonctionnalité restent difficiles à estimer pour une bibliothèque, car pour pouvoir l'évaluer correctement il serait nécessaire de connaître en profondeur un domaine scientifique donné (BIB04). Comme le signale Marshall Breeding, le bilan de cette fonctionnalité reste donc pour le moment incertain, avec une grande marge d'amélioration possible<sup>138</sup>. Il en va de même pour les encarts issus d'encyclopédies en ligne qui peuvent être affichés en haut d'une page de résultats pour définir un terme recherché (BIB20, BIB24, BIB15, BIB13). Si *a priori* c'est là une fonctionnalité intéressante pour l'*information literacy*, des bibliothèques ont indiqué qu'il était dommage que les encyclopédies interrogées ne soient pas le plus souvent des encyclopédies francophones mais anglophones, ce qui tend à limiter leur impact (BIB24, BIB15).

Tous les outils proposent des fonctionnalités permettant de citer facilement une référence, compatibles avec des logiciels de gestion de références bibliographiques. Ces outils de citation permettent aux étudiants, d'après Dianne Cmor et Xin Li, de se concentrer davantage sur le sens et la valeur de la citation, plutôt que sur sa forme bibliographique<sup>139</sup>. En complément, certains outils proposent aussi des fonctionnalités bibliométriques en s'appuyant sur les données produites par Altmetric ou PlumX<sup>140</sup> (BIB20, BIB02, BIB05). Celles-ci permettent de connaître le nombre de citations d'un article par d'autres et peuvent être utiles à un public de chercheurs ou d'étudiants avancés, en mettant à leur disposition un outil d'évaluation de l'information académique<sup>141</sup>. Ces fonctionnalités sont parfois

<sup>137</sup> HANRATH, Scott et KOTTMAN, Miloche. *Op. cit.* p. 16

<sup>138</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p.31.

<sup>139</sup> CMOR, Dianne et LI, Xin. Beyond boolean, towards thinking: discovery systems and information literacy. *Library Management* [en ligne]. Octobre 2012, Vol. 33, n° 8/9, p. 455.

<sup>140</sup> Altmetric. Dans : *Altmetric* [en ligne]. [s. d.]. [Consulté le 12 décembre 2016]. Disponible à l'adresse : <https://www.altmetric.com/>. *Plum Analytics Altmetrics - Plum Analytics* [en ligne]. [s. d.]. [Consulté le 12 décembre 2016]. Disponible à l'adresse : <http://plumanalytics.com/>.

<sup>141</sup> Pour le calcul du nombre de citations, Altmetric déclare s'appuyer essentiellement sur Scopus, tandis que PlumX indique utiliser Scopus mais aussi Crossref, Pubmed Central, SSRN, Repec et d'autres. *Citation Metrics - Plum Analytics* [en ligne]. [s. d.]. [Consulté le 12 décembre 2016]. Disponible à l'adresse :



inclues par défaut dans l'outil, sans que cela soit une demande expresse de la bibliothèque.

Par ailleurs, on peut estimer avec Tammera Race que certaines fonctionnalités des outils de découverte permettent de favoriser la sérendipité, comme par exemple la recherche dans un grand ensemble de métadonnées indexées avec du texte intégral, un accès facilité aux ressources, la possibilité de parcourir des résultats et de filtrer par les facettes, l'aide en ligne, les recommandations, les recherches sauvegardées ou les dossiers de références<sup>142</sup>. Toutefois les difficultés préalablement soulevées par certaines de ces fonctionnalités (comme les facettes) peuvent conduire à nuancer ce propos.

### *Un gain incontesté : la découverte d'articles*

Mais au-delà de ces fonctionnalités dont certaines sont encore à consolider, un grand nombre des bibliothèques interrogées s'accordent sur le fait que la valeur ajoutée des outils de découverte en matière *d'information literacy* est de donner les moyens à un utilisateur lambda d'effectuer des recherches simplifiées à un niveau de granularité bien précis : celui de l'article (BIB01, BIB18, BIB06, BIB07, BIB13, BIB04, BIB20, BIB11, BIB19, BIB02). Pour ces bibliothèques, les index centralisés permettent de donner de la visibilité à des références d'articles qu'il fallait au préalable rechercher dans des bases de données spécifiques ou des archives institutionnelles. Ils mettent à disposition des références de niveau académique, issues de revues à comité de lecture ou de travaux universitaires, ce qui permet à l'utilisateur débutant de se confronter d'emblée à cette littérature (BIB19). Pour les chercheurs, la possibilité simplifiée de recherche d'articles peut représenter un saut qualitatif notable par rapport aux outils précédents, conduisant à une adoption extrêmement rapide de l'outil (BIB06, BIB19, BIB23).

La grande étendue de la couverture des ressources dans les index centralisés permet également, d'après Dianne Cmor et Xin Li de faire en sorte que cette découverte d'articles s'effectue de manière interdisciplinaire, en dépassant les limites d'une recherche dans des ressources natives, limitées à des éditeurs ou à des thématiques particulières. Pour ces deux auteurs, cela est particulièrement utile dans la phase initiale d'une recherche, par exemple pour établir la carte heuristique d'un sujet et trouver l'angle d'attaque le plus pertinent : «Etant donné le caractère interdisciplinaire de l'outil et les différents types de contenus indexés, une recherche simple et large, comme par exemple sur le clonage, montrera immédiatement les différents angles d'attaque du sujet à un étudiant de première année : les aspects médicaux, religieux, technologiques, éthiques etc. »<sup>143</sup>

---

<http://plumanalytics.com/learn/about-metrics/citation-metrics/>. Our sources. Dans : *Altmetric* [en ligne]. 9 juillet 2015. [Consulté le 12 décembre 2016]. Disponible à l'adresse : <https://www.altmetric.com/about-our-data/our-sources/>.

<sup>142</sup> RACE, Tammera. Resource discovery : supporting serendipity. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*: Hershey, PA : IGI Global, 2012, p. 139-152. CHRISTENSEN, Anne. *Op. cit.*

<sup>143</sup> « Due to the interdisciplinary nature of the tool and the various content types, a simple and broad search on cloning, for example, will immediately show many angles of the topic to a first year student – medicine, religion, technology, ethics, etc. ». CMOR, Dianne et LI, Xin, *Op. cit.*, p. 455.

### *Des points de vigilance*

Sans nier ces apports potentiels à l'*information literacy*, il faut noter que les services de découverte peuvent entraîner une gêne pour la recherche d'information efficace, notamment en raison du bruit généré par les grandes quantités de données présentes dans l'index centralisé ou par les paramétrages parfois laborieux de l'algorithme de pertinence (BIB04, BIB12, BIB15, BIB14, BIB25, BIB16, BIB22, BIB08). La crainte de générer du bruit a ainsi pu pousser des bibliothèques à ne pas activer certaines ressources généralistes en *open access* qui auraient élargi beaucoup le spectre des résultats, comme *Hathi Trust*, DOAJ, BASE (BIB11, BIB12, BIB23). Cet effet de bruit a été signalé à de nombreuses occasions par la littérature professionnelle, en particulier en ce qui concerne la recherche d'une référence précise, pour laquelle l'outil de découverte peut être moins performant qu'un catalogue traditionnel<sup>144</sup>. Comme le résumait Marshall Breeding :

« Un défi supplémentaire concerne la capacité des services de découverte à trouver des références précises. En particulier quand on recherche une ressource dont le titre contient des mots très courants de type *Nature* ou *Time*, les algorithmes de classement par pertinence ne répondent pas toujours avec les résultats attendus. Les différents services de découverte ont amélioré la recherche de références précises, mais ce point continue à susciter des critiques quant à leur fonctionnement. »<sup>145</sup>

La présence d'abondantes références issues de ressources électroniques, numériquement dominantes par rapport aux données d'un catalogue local, peut elle aussi expliquer que les résultats ne soient pas toujours perçus comme pertinents et suscitent la confusion des utilisateurs<sup>146</sup>. Des bibliothèques ont signalé que dans certaines disciplines très attachées à la documentation imprimée, notamment les LSH, le nombre parfois important de résultats issus de ressources numériques peut embrouiller le processus de recherche pour les utilisateurs (BIB22, BIB18). Il est donc essentiel d'affiner minutieusement les ressources activées dans l'index centralisé afin d'exclure des références qui pollueraient trop les résultats (BIB12, BIB23, BIB14).

Selon plusieurs articles professionnels<sup>147</sup>, cet effet de bruit serait lié à un problème plus profond qui est celui de la décontextualisation des données présentées dans l'outil de découverte. Or ce besoin de contextualisation en fonction d'un vocabulaire, de codes et de références propres à une discipline

<sup>144</sup> HANRATH, Scott et KOTTMAN, Miloche. *Op. cit.*, p. 11. KABASHI, Arta, PETERSON, Christine, PRATER, Tim, et al. *Op. cit.*, p. 7. BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 30. FAHEY, Sue, GORDON, Shannon et ROSE, Crystal. Seeing Double at Memorial University: Two WorldCat Local Usability Studies. Partnership: *The Canadian Journal of Library and Information Practice and Research*; Vol 6, No 2 (2011) [en ligne]. Disponible à l'adresse : <https://journal.lib.uoguelph.ca/index.php/perj/article/view/1552>, p. 10 et suivantes. KEYSER, Piet de, *Op. cit.*, p. 58.

<sup>145</sup> « One additional challenge lies in the ability of discovery services to find known items. Especially when searching for resources with one-word titles or common words, such as *Nature* or *Time*, relevancy-based retrieval may not always return the expected results. Each of the discovery services has improved its handling of known-item searching, but this continues as a point of criticism of performance. » BREEDING, Marshall et NATIONAL INFORMATION STANDARDS ORGANIZATION (U.S.). *Op. cit.*, p. 30.

<sup>146</sup> ASHER, Andrew, DUKE, Lynda et WILSON, Suzanne. Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources. *College & Research Libraries*. Septembre 2013, Vol. 74, n° 5, p. 474.

<sup>147</sup> BREITBACH, William. *Op. cit.* ; FAGAN, Jody Condit. *Op. cit.* ; THOMSETT-SCOTT, Beth et REESE, Patricia E. *Op. cit.* ; SWANSON, Troy A. et GREEN, Jeremy. Why We Are Not Google: Lessons from a Library Web site Usability Study. *The Journal of Academic Librarianship* [en ligne]. Mai 2011, Vol. 37, n° 3, p. 222-229.

scientifique serait justement l'apanage des travaux universitaires et ferait pleinement partie de *l'information literacy* de niveau académique. Comme le rappelle l'ACRL, dans le monde universitaire de l'information :

« L'autorité est un objet construit et contextuel : les ressources d'information reflètent l'expertise et la fiabilité de leurs auteurs et sont évaluées en fonction d'un besoin d'information donné et du contexte dans lequel cette information sera utilisée. L'autorité est un objet construit dans la mesure où différentes communautés reconnaîtront différents types d'autorité. Elle est contextuelle dans la mesure où le besoin d'information peut aider à déterminer le niveau d'autorité requise ».<sup>148</sup>

C'est également ce besoin de contextualisation qui constitue d'après William Breitbach la ligne de démarcation entre le besoin d'information généraliste et le besoin d'information de niveau universitaire :

« Les travaux universitaires impliquent d'établir des rapports et des connexions entre différents éléments, d'avoir une vue globale du contexte et des interrelations ; le besoin d'information rapide est largement satisfait par des informations ponctuelles ou factuelles sans nécessité de comprendre leur contexte ou leurs interrelations. Les travaux universitaires sont évalués en fonction du niveau, de l'étendue et de la profondeur des éléments qu'ils intègrent dans un ensemble global ; le besoin d'information rapide est largement évalué par sa capacité à fournir la « bonne » réponse ou au contraire à tourner au « feu de brousse » informationnel. »<sup>149</sup>

Or, pour William Breitbach, les index centralisés, en fusionnant les données de différentes sources empêchent cette contextualisation : « le fait de fusionner les données de nombreuses disciplines dans des systèmes qui ne soulignent pas la spécificité de leurs contenus et de leurs métadonnées, peut confondre et perdre les utilisateurs puisque le contenu et la signification des métadonnées seront différents selon les disciplines »<sup>150</sup>. Jody Condit Fagan est du même avis : « Même un chercheur débutant peut concevoir le besoin d'étudier un sujet de la façon dont les experts du domaine le font et les outils de découverte n'offrent aucune potentialité en ce sens<sup>151</sup> ». Quelques années auparavant, Edward Swanson and Daniel Joudrey avaient formulé une critique semblable à destination des moteurs de recherche du web : « baisser le niveau des systèmes de description et des mots sujets qui ont prouvé leur efficacité, pour répondre aux besoins des recherches d'informations occasionnelles, dessert ceux qui ont un besoin fondamental d'accéder à des

<sup>148</sup> « Authority is constructed and contextual : information resources reflect their creators' expertise and credibility, and are evaluated based on the information need and the context in which the information will be used. Authority is constructed in that various communities may recognize different types of authority. It is contextual in that the information need may help to determine the level of authority required. » ACRL. Framework for Information Literacy for Higher Education. Dans : *Association of College & Research Libraries (ACRL)* [en ligne]. 9 février 2015. [Consulté le 21 octobre 2016]. Disponible à l'adresse : <http://www.ala.org/acrl/standards/ilframework>.

<sup>149</sup> « Scholarship requires linkages, connections, contexts and overviews of relationships; quick information seeking is largely satisfied by discrete information or facts without the need to establish the contexts and relationships surrounding them. Scholarship is judged by the range, extent, and depth of elements its integrates into a whole ; quick information seeking is largely judged by whether it provides a "right" answer or puts out an immediate informational "brush fire". » BREITBACH, William. *Op. cit.*, p. 641.

<sup>150</sup> « Merging numerous disciplines together in systems that do not highlight specific content and metadata may throw users off the "scent trail" because the content and meaning of metadata will differ between disciplines ». *Ibid.*, p. 639.

<sup>151</sup> « Even a beginning searcher can conceptualize the need to look at a topic the way experts in the field do, and there are no affordances in a discovery tool to apply this perspective ». FAGAN, Jody Condit. *Op. cit.*, p. 172.

ressources structurées dans des catégories conceptuelles, avec des interrelations bibliographiques et thématiques »<sup>152</sup>. De nombreux articles professionnels ont signalé, en complément, que cette décontextualisation peut toucher également les types de documents, que les usagers ont parfois du mal à identifier malgré les icônes qui les représentent<sup>153</sup>.

Dans les disciplines littéraires et les sciences humaines, les usagers peuvent aussi être perturbés par le mélange au sein de l'index centralisé de sources primaires et secondaires. Or la capacité à distinguer des types de documents et les sources primaires et secondaires est un des critères qui constituent l'*information literacy* pour l'ACRL<sup>154</sup>. D'après Jody Condit Fagan, le fonctionnement des outils de découverte ne permet pas d'effectuer correctement ces distinctions<sup>155</sup>.

Par ailleurs, il pourrait être contreproductif de penser que l'outil de découverte, en tant que point d'accès unique ou principal aux ressources, puisse satisfaire tous les besoins en matière d'*information literacy* de manière équivalente. De nombreux articles professionnels signalent que les outils de découverte ne peuvent pas répondre par eux-mêmes à tous les besoins de recherche d'information : s'en contenter risquerait d'appauvrir la recherche documentaire, de développer des pratiques utilitaristes de l'information, au détriment de l'*information literacy*<sup>156</sup>.

Dans ce sens, plusieurs articles signalent que l'affichage des résultats dans l'outil de découverte, reposant sur un algorithme de pertinence à l'image de celui des moteurs de recherche du web, peut s'avérer lui aussi contreproductif pour l'*information literacy*, car il peut renforcer l'attraction pour les résultats les plus immédiats et les plus visibles<sup>157</sup>. À cet égard, plusieurs bibliothèques ont signalé que sous l'influence des moteurs de recherche web, les utilisateurs ont tendance à s'arrêter à la première page de résultats et/ou aux références qui ont une image de couverture (BIB21, BIB07, BIB22). Cet état de fait a pu pousser les bibliothèques à faire remonter les références issues de leur catalogue dans les résultats, afin de les rendre visibles dès la première page.

Enfin, plusieurs articles professionnels signalent que les outils de découverte à index général centralisé peuvent conduire à masquer la complexité du processus

<sup>152</sup> « Dumbing-down effective systems of description and subject access to meet the needs of casual information seekers does a disservice to those who genuinely need access to information resources that have been organized into conceptual categories, which reveal bibliographic and subject relationships. » SWANSON, Edward et JOUDREY, Daniel N. Book Review: The Changing Nature of the Catalog and Its Integration with Other Discovery Tools: Final Report, and Rethinking How We Provide Bibliographic Services for the University of California: Final Report. *Library Resources & Technical Services* [en ligne]. Octobre 2006, Vol. 50, n° 4, p. 296.

<sup>153</sup> THOMSETT-SCOTT, Beth et REESE, Patricia E. *Op. cit.* ; KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. *Op. cit.* ; DJENNO, Mireille, INSUA, Glenda, GREGORY, Gwen M., et al. Discovering Usability: Comparing Two Discovery Systems at One Academic Library. *Journal of Web Librarianship* [en ligne]. Juillet 2014, Vol. 8, n° 3, p. 267 COMEAUX, David J. *Op. cit.* WALKER, Shaundra et SIMS, Iyanna L. *Op. cit.*, p. 204.

<sup>154</sup> ASSOCIATION OF COLLEGE & RESEARCH LIBRARIES (ACRL). Information Literacy Competency Standards for Higher Education. *Op. cit.*

<sup>155</sup> FAGAN, Jody Condit. *Op. cit.*, p. 172-173.

<sup>156</sup> BREITBACH, William. *Op. cit.* ; KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. *Op. cit.* ; FAGAN, Jody Condit. *Op. cit.* ; FAWLEY, Nancy et KRYSAK, Nikki. *Op. cit.* ; ASHER, Andrew, DUKE, Lynda et WILSON, Suzanne. Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources. *College & Research Libraries*. Septembre 2013, Vol. 74, n° 5, p. 464-488.

<sup>157</sup> DJENNO, Mireille, INSUA, Glenda, GREGORY, Gwen M., et al. *Op. cit.*, Andrew, DUKE, Lynda et WILSON, Suzanne. *Op. cit.* ; THOMSETT-SCOTT, Beth et REESE, Patricia E. *Op. cit.*

de recherche d'information en raison de l'immédiateté des ressources proposées<sup>158</sup>. Des bibliothèques ont ainsi indiqué que la disponibilité immédiate de l'accès au texte intégral peut inciter à télécharger un résultat en texte intégral juste parce que celui-ci est disponible (BIB21, BIB24) : on peut supposer que ces effets de hasard seraient moins importants dans une base de données spécifique dont l'usage requiert un apprentissage préalable plus important. Les prérequis et les différentes étapes d'une recherche documentaire se trouveraient gommées par cet outil transverse et multi-ressources. Or la construction raisonnée d'une recherche documentaire fait également partie des compétences informationnelles de l'*information literacy*. À ce propos, Ana Guthrie et Rhonda McCoy vont même jusqu'à parler, de manière hyperbolique, d'une « mort des véritables compétences bibliographiques »<sup>159</sup>.

### *Un besoin d'accompagnement renforcé*

Les outils de découverte se veulent intuitifs et faciles à manier par un néophyte, ce qui pourrait laisser croire que les bibliothèques seraient moins impliquées dans la formation des usagers. Or de nombreux articles professionnels insistent sur le besoin de maintenir des formations pour accompagner les étudiants dans leur utilisation d'un outil de découverte<sup>160</sup>. Les bibliothèques sont bien conscientes de ce besoin d'accompagnement et elles sont très nombreuses à avoir intégré l'outil de découverte à leur parcours de formation, à leurs journées portes ouvertes ou à leurs séances de rendez-vous avec un bibliothécaire (BIB18, BIB24, BIB21, BIB11, BIB26, BIB01, BIB19, BIB16, BIB12, BIB06, BIB02, BIB13, BIB08). Y compris pour des installations très récentes, les bibliothèques avaient aussi indiqué, au moment de l'entretien, avoir prévu de les intégrer à leurs parcours de formation (BIB14, BIB07, BIB15, BIB04). Inversement, lorsque l'outil n'est pas montré en formation, les bibliothèques ont constaté que cela pouvait nuire à la connaissance de l'outil et à son appropriation par le public (BIB22, BIB10).

Les formations peuvent notamment servir à familiariser les étudiants avec les différents types de ressources présentées (BIB18, BIB19) ainsi qu'avec leur caractère anglophone (BIB17, BIB13) et les problématiques d'accès (BIB15, BIB13). Elles peuvent indiquer la nécessité d'interroger plusieurs outils de manière complémentaire pour obtenir le bon niveau de granularité (BIB13).

De manière plus générale, d'après Sara Decoster, la mise en place d'un service de découverte doit conduire à modifier les séances de formation à la recherche documentaire déjà existantes dans la bibliothèque : « un tel changement radical doit nécessairement se refléter dans les formations à la recherche documentaire »<sup>161</sup>. Elle préconise, entre autres, l'utilisation des cartes heuristiques pour aider les étudiants à balayer leur champ de recherche. Pour elle, les outils de découverte impliquent de travailler davantage la stratégie documentaire dans

<sup>158</sup> BREITBACH, William. *Op. cit.* ; GUTHRIE, Ana et MCCOY, Rhonda. *Op. cit.* FAGAN, Jody Condit, *Op. cit.* FAWLEY, Nancy et KRYSAK, Nikki. *Op. cit.* ; COMEAUX, David J. *Op. cit.* CMOR, Dianne et LI, Xin, *Op. cit.*

<sup>159</sup> « Death of true bibliographic skills » GUTHRIE, Ana et MCCOY, Rhonda. *Op. cit.*, p. 185.

<sup>160</sup> GUTHRIE, Ana et MCCOY, Rhonda. *Op. cit.* ; KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. *Op. cit.* ; DECOSTER, Sara et RENAVILLE, François. *Op. cit.* ASHER, Andrew, DUKE, Lynda et WILSON, Suzanne. *Op. cit.* ; HANRATH, Scott et KOTTMAN, Miloche. *Op. cit.*

<sup>161</sup> DECOSTER, Sara. *Relever le défi des outils « discovery » : les formations orientées vers l'utilisateur* [en ligne]. Mai 2015. [Consulté le 12 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/182048>. p. 1.

l'optique d'une résolution de problèmes, d'inciter les étudiants à garder un œil critique sur les métadonnées et de mettre l'accent sur l'évaluation de l'information. Ces besoins de recentrer les formations sur l'évaluation de l'information, en même temps que sur sa pertinence et sur le caractère itératif de la recherche sont également soulevés par plusieurs articles de la presse professionnelle<sup>162</sup>.

Pour l'instant, un certain nombre de bibliothèques déclarent n'avoir pas suffisamment de données pour évaluer l'effet de ces formations (BIB13, BIB26, BIB04). Par ailleurs beaucoup d'installations récentes n'avaient pas bénéficié, au moment où l'entretien a été réalisé, d'une année universitaire complète au niveau de la formation (BIB15, BIB04, BIB14, BIB07, BIB20).

### C. Une évaluation de l'utilisation encore balbutiante

Que ce soit au niveau de l'évaluation de l'utilisation de l'outil lui-même ou de son effet sur la consultation des ressources numériques, la mesure et l'exploitation de ces données est encore en élaboration. Les outils de mesure sont eux-mêmes en cours de consolidation et il est difficile de tirer des enseignements concluants sur ce point.

#### *Mesures de l'utilisation de l'outil de découverte*

Les outils fournissent différents modules statistiques pour permettre d'en mesurer l'utilisation. Comme le signale JoLinda Thompson, ces modules fournissent en général des rapports mensuels sur le nombre de recherches, de téléchargements, de consultations, avec éventuellement des informations complémentaires sur les facettes et les filtres ou d'autres fonctionnalités le cas échéant<sup>163</sup>. Néanmoins ces outils statistiques ont besoin de consolidation, comme l'a signalé le groupe de travail de l'université de Moncton, la NFAIS et la NISO<sup>164</sup>.

Pour l'heure, peu d'établissements disposent d'informations statistiques fiables sur l'utilisation de leur service de découverte. La plupart déclarent n'avoir pas encore recueilli de données ou ne pas avoir eu l'occasion de les exploiter suffisamment (BIB26, BIB06, BIB04, BIB14, BIB20, BIB07, BIB17, BIB05, BIB15). Plusieurs regrettent l'absence d'un outil statistique consolidé pour le service de découverte (BIB21, BIB24) : il arrive que l'administrateur ne puisse pas facilement faire la différence entre la consultation d'une base et celle de l'outil de découverte lui-même. Le mode de comptage, peu explicité, de l'outil statistique de l'outil de découverte a pu ainsi provoquer la perplexité : une bibliothèque a pu décompter jusqu'à sept millions de requêtes dans son outil, mais sans avoir la possibilité de comprendre en profondeur les modalités de ce calcul et la signification de ce chiffre (BIB24). Dans certains cas, les installations sont trop récentes, ou ont connu des changements importants (passage à une interface principale), pour pouvoir donner lieu à une évaluation quantitative pertinente (BIB14, BIB07, BIB05, BIB17, BIB02). Il arrive aussi que la bibliothèque n'ait pas du tout d'outil statistique pour évaluer les usages, par exemple sur ce qui

<sup>162</sup> ASHER, Andrew, DUKE, Lynda et WILSON, Suzanne. *Op. cit.* ; HANRATH, Scott et KOTTMAN, Miloche. *Op. cit.* CMOR, Dianne et LI, Xin, *Op. cit.*, p. 455.

<sup>163</sup> THOMPSON, JoLinda. *Op. cit.*, p. 60-61.

<sup>164</sup> National Federation of Abstracting and Information Services. *Op. cit.* p. 3-4. Groupe de Travail du Réseau des Bibliothèques de l'université de Moncton. *Op. cit.* p. 19. National Information Standards Organization Et Open Discovery Initiative Working Group. *Op. cit.* Voir infra IV, 3, A « Des bonnes pratiques à l'Open Discovery Initiative ».

concerne les interrogations de l'index centralisé dans un outil mixte (BIB13) ou qu'elle envisage de se servir d'autres outils, comme Piwik, pour mesurer cette utilisation (BIB05).

Pour les installations consortiales en particulier, ce qui rend également la mesure des usages difficile à appréhender, c'est l'impossibilité de la comparer avec une autre mesure. Dans ces cas, l'outil de découverte vient se greffer sur les outils de chaque établissement partenaire. Or la situation de départ de ces derniers, avant la mise en place d'un outil de découverte, peut être très variable d'un établissement à un autre, ce qui rend les comparaisons difficiles et non pertinentes (BIB04, BIB15). Lorsqu'il existe des vues spécifiques au sein d'un outil consorsial, il est difficile d'établir des comparaisons entre les différentes statistiques d'utilisation de chaque vue, ce qui impliquerait un travail d'analyse en détail du nombre de connexions et de requêtes.

Malgré cela, quelques bibliothèques ont exploité les outils statistiques disponibles, notamment le nombre de recherches effectuées et des facettes utilisées. Là aussi les données sont contrastées.

Dans un certain nombre de cas les mesures signalent une forte utilisation de l'outil et une appropriation par le public (BIB18, BIB01, BIB19, BIB23, BIB25). Dans le cas de la BIB18, ces mesures ont fait apparaître la persistance de l'utilisation de la recherche avancée durant la première année suivant la mise en place de l'outil, ce qui semblerait *a priori* contraire à l'esprit d'un outil de découverte et de sa barre de recherche unique. Cela pouvait signifier de la part de l'utilisateur la volonté de limiter un nombre de résultats trop importants. Néanmoins, la bibliothèque a constaté que cette tendance s'est progressivement inversée et désormais les recherches avancées plafonnent tandis que les recherches simples connaissent une croissance exponentielle. Pour la BIB01, le grand nombre de recherches effectuées (plus de 500 000 par an) dans l'outil de découverte laisse également supposer que le public s'est pleinement approprié l'outil.

Lorsque l'outil de découverte coexiste avec un catalogue public, la mesure de l'utilisation n'est pas simple non plus. L'outil de découverte peut faire l'objet d'une consultation plus importante que celle de l'OPAC traditionnel, même si parfois celui-ci reste encore consulté pour la documentation papier (BIB25, BIB02).

Dans d'autres cas, ce sont les statistiques de consultation des deux outils (OPAC et outil de découverte) qui augmentent conjointement, et cela alors même que l'outil de découverte est mis en avant et que le catalogue public n'est plus montré en formation (BIB11). Dans ce cas, cela est probablement dû au fait que pour accéder aux services de réservations et de renouvellement des prêts, le passage par le catalogue public reste indispensable, en l'absence d'un compte utilisateur qui propose ces fonctionnalités sur l'outil de découverte.

Inversement, il peut arriver qu'un catalogue public classique fasse concurrence à l'outil de découverte et que la comparaison des mesures d'utilisation soit au détriment de ce dernier, loin derrière le catalogue traditionnel (BIB22, BIB10, BIB22). Ainsi, la BIB10 comptabilisait en 2015 plus de 10 fois plus de consultations sur l'OPAC que sur l'outil de découverte, quoique les chiffres de consultation de ce dernier aient eu tendance à augmenter régulièrement depuis 2012. Pour la BIB21, qui dispose actuellement d'un affichage en mode « *bento* », les résultats de l'OPAC étaient de même toujours plus consultés que ceux de l'outil de découverte.

### *Mesures de l'impact sur la consultation de la documentation numérique*

L'augmentation de la consultation de la documentation numérique qui suivrait la mise en place d'un outil de découverte est souvent, pour les fournisseurs, un argument commercial en faveur de leurs produits<sup>165</sup>. Cet argument semble cependant difficile à vérifier en tant que tel, dans la mesure où un tel impact ne peut être évalué à partir de l'outil de découverte seul. Pour évaluer l'impact de l'outil de découverte sur la consultation de ressources numériques, il serait nécessaire de pouvoir croiser plusieurs paramètres : les données de consultation de l'outil de découverte avec celles fournies par d'autres plateformes ou outils de gestion de ressources numériques (plateformes natives, résolveurs de liens, proxy)<sup>166</sup>.

D'une part, du côté de l'outil de découverte, l'absence d'outils de mesure d'utilisation consolidés ainsi que l'impossibilité de déterminer à un instant  $t$  les ressources qui sont effectivement couvertes par l'index centralisé rendent très difficile d'estimer l'impact de la mise en place de cet outil sur la consultation des ressources numériques.

D'autre part, les outils de mesure de la consultation des ressources numériques, sont eux-mêmes en cours de consolidation. Il s'agit notamment du code de bonnes pratiques COUNTER et du projet ezPAARSE / ezMESURE.

Lancé en 2002, le code de bonnes pratiques COUNTER qui en est aujourd'hui à sa 4<sup>e</sup> version, a pour but de « faciliter l'enregistrement, l'échange et l'interprétation des données d'utilisation en ligne ; il établit pour ce faire des normes et des protocoles ouverts et internationaux permettant aux fournisseurs d'informations de produire des statistiques d'utilisation, qui sont cohérentes, crédibles et compatibles<sup>167</sup>. » En effet, comme le rappelle Julien Logre, avant la mise en place de COUNTER, ces mesures étaient dépendantes des méthodes, des outils et de la volonté de chaque fournisseur<sup>168</sup>. Par rapport à cette situation, le code COUNTER représente une première avancée en matière de standardisation des mesures, car il permet de construire des dénominateurs communs. Ainsi, par exemple pour les revues, les rapports produits selon le code COUNTER peuvent fournir, par mois et par revue, le nombre de requêtes réussies d'articles en texte

---

<sup>165</sup> EBSCO Discovery Service™ increases use of school library resources and prepares students for college research [en ligne]. [s. d.]. [Consulté le 21 octobre 2016]. Disponible à l'adresse : <https://discovery.ebsco.com/pulse/article/ebsco-discovery-service-increases-use-of-school-library-resources-and-prepa> ; Primo Discovery and Delivery [en ligne]. [s. d.]. [Consulté le 21 octobre 2016]. Disponible à l'adresse : <http://www.exlibrisgroup.com/category/PrimoOverview> ; ProQuest Summon | Web-scale discovery service increases the value of your library through an unprecedented research experience - The Summon® Service [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://www.proquest.com/products-services/The-Summon-Service.html>.

<sup>166</sup> WAY, Doug. The Impact of Web-scale Discovery on the Use of a Library Collection. *Serials Review* [en ligne]. Décembre 2010, Vol. 36, n° 4, p. 214-220 ; BREEDING, Marshall et NATIONAL INFORMATION STANDARDS ORGANIZATION (U.S.). *Op. cit.* ; THOMPSON, JoLinda. *Op. cit.*, p.118.

<sup>167</sup> COUPERIN. *Statistiques d'utilisation - COUNTER* [en ligne]. [s. d.]. [Consulté le 15 avril 2016]. Disponible à l'adresse : <http://www.couperin.org/groupes-de-travail-et-projets-deap/statistiques-dusage/couperin>.

<sup>168</sup> LOGRE, Julien. *Autour de COUNTER: l'utilisation des Ressources électroniques onéreuses spécialisées dans l'enseignement supérieur et la recherche en France mesures, indicateurs, pilotage* [en ligne]. S.l. : s.n., 2013. [Consulté le 15 avril 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/60371-autour-de-counter-l-utilisation-des-ressources-electroniques-onereuses-specialisees-dans-l-enseignement-superieur-et-la-recherche-en-france-mesures-indicateurs-pilotage.pdf>, p. 26-27.



intégral, le nombre de requêtes réussies d'articles en texte intégral d'une archive, et le nombre de refus de connexion.

Cependant, comme le notent Thomas Jouneau et Thomas Porquet, beaucoup d'éditeurs ne respectent pas toujours le code COUNTER, « soit en n'offrant aucune donnée, ou aucune garantie sur le monde de calcul, soit en proposant des modes de comptage (« crédits », « pages vues » etc.) ne pouvant être rapprochés des éléments couverts par COUNTER »<sup>169</sup>. Mais même dans les cas où l'éditeur fournit des données respectant le code de bonnes pratiques COUNTER, celles-ci ont besoin d'être affinées, car elles sont fournies de manière globale pour un établissement et une plateforme, et non pas en fonction d'un bouquet précis. Or, comme l'expliquait Dominique Rouger, cela peut être problématique dans la mesure où il peut y avoir des articles accessibles sur les plateformes d'éditeurs hors abonnement (par exemple des articles en *open access* ou *online first*<sup>170</sup>) et que les données COUNTER ne correspondront pas au périmètre des abonnements effectivement souscrits par la bibliothèque<sup>171</sup>. Par ailleurs, la notion de requête utilisée dans le code COUNTER peut elle-même prêter à confusion car il n'y a pas d'équivalence d'unité entre le décompte d'une requête et celui d'un article distinct : « les requêtes sont bien distinctes, mais pas obligatoirement les articles<sup>172</sup> ». Un même article ayant pu être téléchargé plusieurs fois, l'écart entre le nombre de requêtes et le nombre d'articles distincts effectivement téléchargés peut être considérable. Enfin, les rapports fournis selon le code COUNTER ne peuvent pas offrir « de visibilité sur l'usage des ressources par différentes catégories d'utilisateurs (discipline de spécialité, statut, niveau d'études...) »<sup>173</sup>.

Porté par l'université de Lorraine, l'INIST et le consortium Couperin, le projet ezPAARSE a pour objectif de remédier à certaines insuffisances des données issues des rapports COUNTER : il doit permettre d'obtenir des mesures plus fines de la consultation des ressources numériques en complétant ou en consolidant les données COUNTER, dans les cas où ces dernières ne sont pas fournies ou pas assez détaillées<sup>174</sup>. Ce projet consiste à analyser, exploiter les traces (ou logs) produites lors des consultations des ressources numériques et enregistrées par un proxy, grâce à une instance du logiciel libre ezPAARSE, hébergée par la plateforme AnalogIST<sup>175</sup>. Les lignes composant les fichiers de logs étant formatées de manière standard<sup>176</sup>, le travail d'étude préalable de leur

<sup>169</sup> JOUINEAU, Thomas et PORQUET, Thomas. Evaluer l'usage des ressources numériques : quel enjeux, quelles solutions? Dans : *Evaluer la bibliothèque par les mesures d'impact*. Villeurbanne, France : ENSSIB, 2016, p. 141.

<sup>170</sup> *Online first* : articles approuvés pour publication, disponibles en ligne sur la plateforme d'un éditeur avant leur publication effective dans un numéro de revue imprimée.

<sup>171</sup> ROUGER, Dominique. *Regard critique sur les statistiques d' "utilisation" de la documentation numérique*. ENSSIB, 13 mars 2012, p. 18.

<sup>172</sup> ROUGER, Dominique. « Don't let me be Miss Understood » ou les bibliothécaires lisent-ils le Counter dans le chiffre ? Dans : *L'information scientifique et technique dans l'univers numérique*. Paris : ADBS, 2010, p. 117.

<sup>173</sup> JOUINEAU, Thomas et PORQUET, Thomas. Evaluer l'usage des ressources numériques : quels enjeux, quelles solutions? *Op. cit.*, p. 141.

<sup>174</sup> JOUINEAU, Thomas et PORQUET, Thomas. *Exploitation des traces de passage vers la documentation numérique : ezPAARSE AnalogIST* [en ligne]. Enssib, 14 janvier 2016. [Consulté le 31 octobre 2016]. Disponible à l'adresse : [http://www.enssib.fr/sites/www/files/documents/manifestations/2016/ezpaarse\\_\\_\\_analogist\\_-\\_enssib\\_-\\_14\\_01\\_2016.pdf](http://www.enssib.fr/sites/www/files/documents/manifestations/2016/ezpaarse___analogist_-_enssib_-_14_01_2016.pdf).

<sup>175</sup> *start [analogist]* [en ligne]. [s. d.]. [Consulté le 6 avril 2016]. Disponible à l'adresse : <http://analogist.couperin.org/>.

<sup>176</sup> JOUINEAU, Thomas et PORQUET, Thomas. Evaluer l'usage des ressources numériques : quels enjeux, quelles solutions? *Op. cit.*, p. 142.

structure, afin de construire l'outil d'analyse des logs (ou parseur), a pu être réparti et mutualisé entre plusieurs bibliothèques d'enseignement supérieur<sup>177</sup>. Après le nettoyage des lignes non significatives, l'analyse du fichier de logs par le logiciel ezPAARSE présente des informations plus détaillées que celles des rapports COUNTER : elle fournit notamment l'identifiant de la revue et de l'article, afin de pouvoir décompter le nombre d'articles distincts<sup>178</sup>. Lorsque l'établissement dispose d'un annuaire unifié permettant de détailler les catégories d'utilisateurs et d'une authentification préalable requise pour l'accès aux ressources, ces informations pourront éventuellement être exploitées dans les logs afin de disposer de données plus fines sur les consultations<sup>179</sup>.

Si le logiciel ezPAARSE a désormais été mis à disposition de la communauté sous licence libre, malheureusement, pour un certain nombre de plateformes, les logs n'ont pas été analysables sur ce modèle<sup>180</sup>. Par ailleurs, ezPAARSE « trouve son plein intérêt dans le cas d'une proxification complète des accès »<sup>181</sup> : lorsqu'une bibliothèque ne proxifie qu'une partie de ses accès, de nombreuses données ne laisseront donc pas de traces sur le proxy. Depuis 2016, le projet ezMESURE a pris le relais du projet ezPAARSE. Son objectif est de « disposer d'un entrepôt national de données homogènes et indépendantes sur les accès aux ressources numériques payantes par les établissements de l'ESR ». Cet outil doit rassembler les données de consultation locales produites par les instances d'ezPAARSE installées dans les établissements et proposer une interface de visualisation des données (avec notamment une cartographie). Il s'agit là aussi d'un projet en cours de réalisation.

Des deux paramètres qui seraient nécessaires pour pouvoir correctement évaluer l'impact du service de découverte sur la consultation des ressources numériques (à savoir les mesures de l'utilisation de l'outil de découverte et celles de la consultation de la documentation numérique), aucun ne repose donc sur un outil entièrement consolidé. Sur ce point, le bilan semble encore une fois très difficile à établir.

D'abord, de nombreuses bibliothèques signalent ne pas avoir assez de données ou le recul suffisant pour en tirer les enseignements quant à l'effet sur la consultation de la documentation numérique (BIB20, BIB07, BIB14, BIB06, BIB15, BIB08, BIB17).

Dans plusieurs autres cas, il est impossible pour les bibliothèques de mesurer l'impact de l'outil dans ce domaine parce que le périmètre de leur offre de documentation numérique a pu varier d'une année sur l'autre (BIB18, BIB01, BIB12). En l'absence d'un dénominateur commun, la comparaison d'une année sur l'autre devient impossible à réaliser. À ce propos, on peut signaler qu'il est dommage que les services de découverte ne mettent pas à disposition un outil de mesure du type COUNTER (BIB24). Pour les outils consortiaux en particulier,

<sup>177</sup> ezPAARSE / AnalogIST - Suivi des plateformes / Trello [en ligne]. [s. d.]. [Consulté le 20 avril 2016]. Disponible à l'adresse : <https://trello.com/b/wEaLnz8d/ezpaarse-analogist-suivi-des-plateformes>.

<sup>178</sup> JOUNEAU, Thomas et PORQUET, Thomas. Évaluer l'usage des ressources numériques : quels enjeux, quelles solutions? *Op. cit.*, p. 143.

<sup>179</sup> *Ibid.*, p. 145 et suivantes.

<sup>180</sup> *platforms:start [analogist]* [en ligne]. [s. d.]. [Consulté le 31 octobre 2016]. Disponible à l'adresse : [http://analogist.couperin.org/platforms/start#plateformes\\_editeursanalyse\\_et\\_developpement\\_des\\_parseurs](http://analogist.couperin.org/platforms/start#plateformes_editeursanalyse_et_developpement_des_parseurs).

<sup>181</sup> JOUNEAU, Thomas et PORQUET, Thomas. Évaluer l'usage des ressources numériques : quels enjeux, quelles solutions? *Op. cit.*, p. 140-149.

lorsque les bibliothèques partenaires ne disposaient pas au préalable de ressources numériques ou d'outils de signalement de celles-ci, il est difficile de déterminer de manière précise l'influence de l'outil (BIB04).

Malgré ces difficultés, plusieurs bibliothèques ont signalé que la mise en place de l'outil de découverte a été suivie d'un effet sensiblement positif sur les chiffres de consultation de ressources numériques fournis par les éditeurs, dont certaines fortes augmentations (BIB18, BIB23, BIB19, BIB01, BIB04, BIB21, BIB25, BIB24, BIB02). En guise d'exemple, pour les revues électroniques, en nombre constant de titres, la BIB01 a indiqué avoir expérimenté une progression de 250% sur les téléchargements après la mise en place de l'outil, sur une période de 5 ans, en se basant sur des rapports COUNTER.

Cela dit, il reste difficile, même en cas d'augmentation constatée, d'établir une relation de cause à effet entre la mise en place de l'outil de découverte et la consultation de la documentation numérique car celle-ci peut avoir de tout autres causes : l'évolution générale des usages, la banalisation et la disponibilité immédiate du texte intégral, les progrès de l'*open access* ou encore l'augmentation des ressources disponibles en ligne dans toutes les disciplines. Ainsi, même en cas d'augmentation constatée, la bibliothèque n'a pas les moyens d'attribuer cette augmentation à la mise en place de l'outil de découverte (BIB18, BIB23, BIB21, BIB24, BIB02). Cette difficulté avait été soulevée par Scott Garrison, Anne Prestamo et Juan Carlos Rodriguez :

« On peut vraisemblablement déduire qu'une augmentation significative des téléchargements en texte intégral pendant la période qui suit immédiatement la mise en place de l'outil de découverte, est liée à cette dernière. Les rapports COUNTER fournis par les plateformes d'éditeurs en texte intégral peuvent signaler de telles augmentations d'utilisation, mais il n'y a pas de moyen de connaître l'origine des recherches qui ont conduit à cette augmentation<sup>182</sup>. »

Il faut noter également que l'augmentation constatée des téléchargements depuis la mise en place de l'outil de découverte peut aller de pair avec une baisse du nombre de recherches dans les bases de données spécifiques, notamment dans celles de références bibliographiques seules (BIB23, BIB18, BIB25, BIB26). Cette tendance à la baisse de la consultation des bases de données de références bibliographiques suite à la mise en place d'un outil de découverte, a été signalée à maintes reprises par la littérature professionnelle<sup>183</sup>. Mais il semble là aussi difficile d'établir une relation de cause à effet, puisque plusieurs autres paramètres peuvent entrer en ligne de compte dans cette situation (l'évolution des usages et l'attente vis-à-vis du texte intégral). Cette baisse de la consultation des bases de données spécifiques ne se vérifie d'ailleurs pas toujours : pour la BIB02, la consultation des bases de données spécifiques reste stable suite à la mise en place

<sup>182</sup> « It is plausible to conclude that a significant increase in full text downloads in the period immediately following discovery implementation is attributable to that implementation. COUNTER reports from full text vendors can indicate usage increases, but there is no way to know the origin of searches that lead to the increased use ». GARRISON, Scott, PRESTAMO, Anne et RODRIGUEZ, Juan Carlos. *Putting Library Discovery Where Users Are*. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p.400.

<sup>183</sup> THOMPSON, JoLinda. *Op. cit.* p. 120. WAY, Doug. *Op. cit.* p. 218. RENAUVILLE, François, RICHELLE, Laurence et THIRION, Paul. « *Where are my Marc records?* » - *Librarians' perception of discovery tools* [en ligne]. Berlin, 8 septembre 2013. [Consulté le 3 avril 2016]. KORNBLAU, Amy I., STRUDWICK, Jane et MILLER, William. *How Web-Scale Discovery Changes the Conversation: The Questions Librarians Should Ask Themselves*. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 38.

de l'outil, ce qui laisse supposer qu'une catégorie d'utilisateurs continue d'utiliser directement les bases de données sans passer par l'outil de découverte.

Dans d'autres cas, les bibliothèques signalent ne pas avoir constaté, après la mise en place d'un outil de découverte, une augmentation sensible de la consultation de la documentation numérique (BIB22, BIB11, BIB13, BIB10, BIB17). Pour certaines bibliothèques, celle-ci semble bien moindre que celle qu'avait impliquée la mise en place d'un moteur de recherche fédérée ou du résolveur de liens (BIB22, BIB17). Cela peut s'expliquer par le fait que les utilisateurs continuent d'accéder de manière directe aux ressources spécifiques ou par un contexte local spécifique (BIB11, BIB10).

À terme, au-delà des mesures fournies par l'outil de découverte lui-même, plusieurs bibliothèques ont indiqué leur volonté de disposer d'outils de mesure de la consultation de la documentation numérique plus performants, notamment grâce à l'exploitation des logs (c'est-à-dire les traces), afin d'avoir des statistiques plus fines sur les ressources numériques effectivement consultées, même si, actuellement, au niveau des logs analysés par ezPAARSE, les bibliothèques ne disposent pas de l'URL d'origine et ne peuvent donc pas identifier la source du clic (BIB26, BIB19, BIB23, BIB06, BIB24, BIB01, BIB21, BIB17). Plusieurs bibliothèques travaillent de manière plus poussée sur les logs de leur proxy pour croiser les informations de consultation avec les informations des profils utilisateurs (BIB01, BIB24, BIB21, BIB17, BIB26). Dans ces cas, le passage obligé des utilisateurs par le proxy permet d'identifier les articles téléchargés et d'obtenir des informations sur le profil des utilisateurs (niveau et discipline)<sup>184</sup>.

Il reste donc à développer des outils de mesure plus fins pour mieux comprendre le comportement des utilisateurs vis-à-vis des résultats disponibles dans l'outil de découverte. Mais il est important de garder à l'esprit que quel que soit l'outil de mesure mis en place et son degré de finesse, les informations fournies ne concerneront jamais l'usage effectif de la ressource une fois entre les mains (ou sur le bureau) de l'utilisateur. Comme le signalait Dominique Rouger au sujet de COUNTER :

« Ce qui est enregistré dans les fichiers traces, c'est la trace d'une action mais en aucun cas celle d'une intention et encore moins d'une utilisation. Les traces prises en compte pour COUNTER indiquent que le serveur a bien envoyé le document demandé (par l'activation d'un lien hypertexte) mais, du fait du protocole http, ne disent rien de la réception du document sur le poste demandeur, et encore moins de ce qu'il en est fait une fois arrivé<sup>185</sup>. »

L'usage réel fait par l'utilisateur de la ressource finale reste donc une inconnue, dont la variabilité peut être encore plus grande du fait de de l'accès facilité par l'outil de découverte. C'est pourquoi il est important d'avoir aussi des retours qualitatifs des utilisateurs.

<sup>184</sup> Voir à ce sujet le retour d'expérience de l'université de Lorraine JOUNEAU, Thomas et PORQUET, Thomas. *Evaluer l'usage des ressources numériques : quel enjeux, quelles solutions?* *Op. cit.*

<sup>185</sup> ROUGER, Dominique. « Don't let me be Miss Understood » ou les bibliothécaires lisent-ils le Counter dans le chiffre ? Dans : *L'information scientifique et technique dans l'univers numérique*. Paris : ADBS, 2010, p. 123-124.

*Enquêtes et tests sur les outils de découverte*

De même que pour les données quantitatives, les données concernant l'évaluation qualitative de l'outil sont là aussi à consolider. Au moment où ont été réalisés les entretiens, de nombreuses bibliothèques ont signalé ne pas encore avoir pu conduire de tests ou d'enquêtes auprès des utilisateurs au sujet des outils de découverte, souvent du fait d'installations très récentes ou en raison de plusieurs autres projets concomitants (BIB12, BIB04, BIB06, BIB07, BIB14 BIB10, BIB15, BIB02, BIB05, BIB17, BIB11, BIB13).

Ainsi, au moment des entretiens, peu de bibliothèques ont déclaré avoir effectué une enquête spécifique sur l'outil de découverte (BIB18, BIB22, BIB25), même si plusieurs ont indiqué avoir des enquêtes en cours ou prévues à court terme (BIB20, BIB24, BIB15, BIB21). Dans le cadre de ce travail, nous avons pu disposer des résultats des enquêtes menées par les bibliothèques BIB18, BIB22, BIB25.

Au niveau de la BIB18, deux enquêtes successives ont montré une appropriation progressive de l'outil. Une première enquête, conduite immédiatement après la mise en place de l'outil, a permis de faire apparaître des remarques très sévères sur l'outil de découverte aussi bien de la part des étudiants que des enseignants-chercheurs, avec toutefois des raisons différentes selon les disciplines. Dans les disciplines de STM, les enseignants-chercheurs gardaient leur prédilection pour les bases de données spécifiques, tandis qu'à l'opposé, les enseignants-chercheurs de LSH restaient nettement focalisés sur l'imprimé et boudaient la documentation électronique. Un an plus tard une autre enquête menée auprès de la communauté universitaire sur le nouveau site web ainsi que sur l'outil de découverte a permis de constater une plus grande acceptation de l'outil. Dans le cas de la BIB18, le fait que l'outil de découverte corresponde à une interface unique, l'OPAC traditionnel n'étant plus accessible, est considéré par la bibliothèque comme un facteur crucial pour l'adoption progressive de l'outil. De même le renouvellement progressif de la population étudiante (qui n'a pas connu l'ancien OPAC) a pu pousser lui aussi vers l'adoption de l'outil de découverte.

La situation est différente pour la BIB22, où l'outil de découverte coexistait avec l'OPAC traditionnel et où l'enquête de publics, effectuée un an après la mise en place de l'outil, portait sur les deux interfaces à la fois. Cette enquête a révélé dans l'ensemble une appropriation difficile de l'outil de découverte et une préférence globale pour l'OPAC traditionnel. Sur de nombreux points évalués, pour lesquels l'outil de découverte aurait dû répondre de manière plus efficace (notamment la pertinence, la vitesse et l'ergonomie), la comparaison entre les deux outils se faisait systématiquement en faveur de l'OPAC. En termes de préférences d'utilisation, le public dans son ensemble a déclaré utiliser davantage l'OPAC traditionnel (56%) que l'outil de découverte (16%), avec toutefois des variantes par catégorie d'utilisateurs. Les enseignants-chercheurs constituaient le seul groupe à considérer l'outil de découverte comme leur outil de recherche prioritaire, tandis que c'était sans conteste l'OPAC traditionnel qui suscitait l'approbation des étudiants, au détriment de l'outil de découverte, jugé nettement moins pertinent. Cette enquête a permis également à la bibliothèque de prendre conscience que l'outil de découverte, contrairement au catalogue public, restait largement méconnu d'une grande partie de son public. C'est en effet lors de cette enquête que de nombreux usagers de la BIB22 ont été amenés à utiliser pour la première fois l'outil de découverte.

La BIB25 a pu, quant à elle, mener des tests d'utilisabilité auprès du public et du personnel. Pour les usagers, ces tests ont pris la forme d'entretiens à la bibliothèque, pendant lesquels ils se voyaient proposer dix actions à réaliser (adaptables en fonction du déroulement de l'entretien). Ils ont concerné en tout quatorze étudiants et un lecteur extérieur. Le bilan a fait ressortir une satisfaction de ces usagers quant à l'ergonomie de la plateforme (en particulier le déroulé vertical de la page), l'affichage des notices détaillées, les fonctionnalités bibliométriques et le panier de références bibliographiques, qui étaient tous des éléments très appréciés. En revanche, les étudiants interrogés ont été plutôt gênés par les facettes, dont les intitulés et fonctionnement restaient pour eux partiellement obscurs, et par l'impossibilité de filtrer par certains types de documents (distinguer les livres imprimés des livres numériques). L'affichage des informations d'exemplaires pouvait susciter aussi des difficultés pour le repérage de la cote. De même, l'articulation de l'outil de découverte avec l'OPAC traditionnel est apparue comme une source de confusion possible, les étudiants ne comprenant pas pourquoi ils étaient renvoyés vers l'interface de l'OPAC quand ils cliquaient sur un exemplaire.

### *D'autres retours qualitatifs*

D'autre part, plusieurs bibliothèques ont signalé disposer d'informations grâce à d'autres enquêtes qui, même si elles ne portent pas directement sur l'outil, permettent de recueillir des informations à ce sujet : enquêtes Libqual (BIB01, BIB24, BIB02), enquêtes sur le site web (BIB23, BIB02), enquêtes sur les pratiques des publics (BIB26, BIB19, BIB16, BIB10, BIB08, BIB15).

Des enquêtes Libqual ont parfois pu permettre à des bibliothèques d'obtenir des informations sur l'usage de l'outil de découverte. Dans le cas de la BIB24 et de la BIB02 les commentaires obtenus sur l'outil de découverte par des enquêtes Libqual étaient plutôt sporadiques et les usagers se sont peu prononcés à ce sujet. Mais dans le cas de la BIB01, une première enquête Libqual conduite juste après la mise en place de l'outil, a d'abord recueilli des commentaires très sévères sur le changement brusque d'interface en cours d'année universitaire et sur des fonctionnalités perdues avec le nouvel outil (la liste de bases de données précédemment générée par l'outil de recherche fédérée). Toutefois, pour la même bibliothèque, l'enquête Libqual suivante a permis de constater, quatre ans plus tard, une évolution nettement positive et une appropriation totale de l'outil par le public.

D'autres bibliothèques ont pu obtenir des informations sur l'utilisation de leur outil de découverte grâce à des enquêtes portant sur leur site web. Ainsi, pour la BIB23 c'est suite aux commentaires recueillis lors d'une enquête sur le site web que le choix d'une recherche distincte en fonction de la granularité a été mis en place, afin de mieux répondre aux besoins des enseignants-chercheurs de séparer la recherche d'articles des autres documents.

Plus largement, les enquêtes de public peuvent également permettre à une bibliothèque de recueillir des informations sur la réception de l'outil de découverte. Ainsi, pour la BIB19, dont l'interface de découverte distingue la recherche par niveaux de granularité, une enquête auprès des enseignants-chercheurs a révélé une grande satisfaction et une appréciation très positive de l'outil par cette population, notamment pour la recherche d'articles et les ressources en ligne. De même, pour la BIB15 où les enseignants-chercheurs ont été sollicités après la mise en place de l'outil, leurs commentaires ont été très positifs

quant aux fonctionnalités de recherche dans la documentation électronique et les articles. À la BIB26, des enquêtes de public ont pu montrer que l'outil de découverte restait plutôt utilisé par les enseignants-chercheurs ou doctorants en priorité, plus que par les étudiants. Le changement récent de la présentation de l'interface (désormais en mode « *bento* ») pourra éventuellement conduire à une évolution de ces pratiques.

Dans le cas de la BIB08, c'est une enquête générale de satisfaction, intervenue quatre mois après la mise en œuvre de l'outil, qui a permis de recueillir les avis des utilisateurs. Cette enquête a fait apparaître quelques commentaires critiques récurrents, en particulier sur le bruit généré dans les résultats et la difficulté à distinguer les ressources pour lesquelles le texte intégral était effectivement disponible et celles pour lesquelles l'index ne fournissait que la référence bibliographique. En dépit de ces inconvénients, l'outil de découverte obtenait plus de 70% de réponses satisfaites lors de cette enquête.

À l'opposé, pour la BIB10, où l'outil de découverte coexiste avec l'OPAC traditionnel, une enquête conduite 2 ans après la mise en place de l'outil a confirmé que celui-ci n'était toujours pas du tout l'outil prioritaire choisi par les usagers : sur plus de 9000 répondants, 42 % des personnes sondées déclaraient le connaître (contre 77 % pour l'OPAC traditionnel) et 18% indiquaient l'utiliser (contre 54 % pour l'OPAC).

Du côté du public, les données dont nous avons disposé, encore parcellaires à de nombreux égards, font donc apparaître pour l'heure des résultats très contrastés selon les bibliothèques quant à l'utilisation des outils de découverte.

## 2. DU COTE DES UTILISATEURS PROFESSIONNELS

### A. Un bilan mitigé

Du côté des utilisateurs professionnels, la réception des outils n'est pas non plus homogène. Dans un certain nombre de cas, la réception de l'outil par les professionnels a été plutôt positive ou tend progressivement vers l'acceptation (BIB25, BIB11, BIB13, BIB02, BIB18, BIB07, BIB06, BIB24). Cela s'explique notamment par l'étendue de ressources et de fonctionnalités offertes, l'unification des interfaces, la modernisation de celles-ci (en particulier par rapport à des OPAC vieillissants) et par l'adaptation aux modes de recherches actuels : rapidité, ergonomie, adaptation aux supports mobiles, qualités graphiques des outils.

Néanmoins, nombreuses sont les bibliothèques qui ont signalé avoir éprouvé des réticences des professionnels avant ou après la mise en place de l'outil de découverte (BIB22, BIB12, BIB18, BIB19, BIB11, BIB07, BIB14, BIB10, BIB17, BIB13, BIB02, BIB15, BIB25, BIB16). Comme l'ont signalé David Dahl et Patricia MacDonald<sup>186</sup>, ces réticences peuvent atteindre des niveaux très importants.

Cette résistance est très souvent liée aux difficultés à comprendre pleinement et à s'appropriier le fonctionnement de l'outil, en raison de la barre de recherche unique (qui prend le contrepied des recherches articulant plusieurs index avec des

<sup>186</sup> DAHL, David et MACDONALD, Patricia. Implementation and Acceptance of a Discovery Tool: Lessons Learned. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 366-387.

opérateurs booléens), de l'utilisation d'un format de notices différent du format MARC et d'un algorithme de pertinence sur lequel la bibliothèque n'a pas entièrement la main (BIB12, BIB18, BIB15, BIB16, BIB21, BIB11, BIB25, BIB02). Ces réticences peuvent également être liées à l'impossibilité d'effectuer une recherche sur des critères précis (comme les vedettes RAMEAU) disponibles sur les catalogues traditionnels ainsi qu'à l'incapacité de justifier une liste de résultats obtenus (BIB11). De même, les facettes partielles ou inadaptées aux fonds de la bibliothèque peuvent être déroutantes pour les professionnels, lorsque les intitulés ne sont pas pertinents sur des champs fondamentaux tel que le type de document ou lorsque le fonctionnement des filtres est modifié en cours de route par un changement de version (BIB02, BIB25). D'autres fonctionnalités plus accessoires (outils de citation ou de bibliométrie) ont pu elles aussi être considérées par les professionnels comme peu intuitives et sources de confusion (BIB25).

Les points de méfiance concernent également les données de l'index centralisé. L'impossibilité de connaître le périmètre et le niveau d'indexation des ressources indexées, ainsi que la grande fluctuation des données dans de brefs laps de temps et la gestion parfois difficile des doublons ou de la granularité peuvent freiner l'appropriation de l'outil par les professionnels (BIB23, BIB20, BIB13, BIB02, BIB11, BIB07). En particulier pour les formateurs, l'absence de visibilité sur les ressources effectivement indexées et l'articulation parfois conflictuelle avec le résolveur de liens, peuvent devenir très problématiques, lorsque les professionnels ne sont pas en mesure de transmettre une information claire et lisible sur les outils qu'ils doivent présenter (BIB23, BIB02, BIB11, BIB16). Il est difficile en effet de faire connaître l'outil lors de formations quand on ne comprend pas soi-même ce que l'on manipule. Ce point est relevé comme étant la source d'un problème majeur d'acceptation des outils de découverte par le personnel des bibliothèques, dans l'étude de David Dahl et Patricia MacDonald :

« Les équipes de la bibliothèque ont aussi exprimé leur confusion et leurs incertitudes quant aux ressources qu'elles explorent. [...] Elles n'ont pas de vision claire de ce qui est indexé exactement. Avec la complexité croissante des accords commerciaux entre fournisseurs et éditeurs, ceci continuera vraisemblablement à être un problème dans les années à venir. »<sup>187</sup>

Une autre crainte très partagée dans ce domaine est celle de voir les ressources locales, notamment celles du catalogue, « noyées » dans les résultats issus de ressources numériques, plus particulièrement lorsque l'outil devient l'interface unique ou principale de la bibliothèque (BIB18, BIB10, BIB22, BIB02, BIB24, BIB17, BIB15). Or le catalogue de bibliothèque garde une importance symbolique forte aux yeux de nombreux professionnels. Plusieurs bibliothèques ont conservé dans un premier temps l'accès au catalogue traditionnel pour faciliter la transition vers une interface unique ou principale (BIB24, BIB17, BIB02, BIB10). Mais dans certains cas, ce choix peut ralentir la mise en place de cette dernière, les professionnels continuant d'utiliser l'ancien OPAC pour la recherche de références imprimées ou n'intégrant pas l'outil dans les formations (BIB10, BIB22, BIB12, BIB02). Le caractère massivement anglophone des ressources de l'index centralisé peut expliquer en partie cette préférence. Enfin, la lisibilité des

<sup>187</sup> « Library staff also expressed confusion and uncertainty over what materials they are searching. [...] They are unclear on exactly what it indexes. With increasingly complex licensing agreements between vendors and publishers, this will most likely continue to be an issue in the future. » *Ibid.*, p. 377.



outils peut être elle aussi remise en cause, en particulier dans le cas d'installations consortiales où l'outil de découverte se rajoute à de nombreuses interfaces déjà existantes (BIB10, BIB06, BIB04, BIB15).

Au niveau de l'administration de l'outil, des professionnels ont signalé que le faible niveau d'adaptabilité des outils, du fait que de nombreux paramètres leur échappent, pouvait être un frein à l'appropriation (BIB07, BIB02, BIB06, BIB25). De même, si l'évolution rapide des outils et l'ajout de nouvelles fonctionnalités sont appréciés en termes de nouveauté, ils peuvent se révéler à double tranchant pour ce qui est de leur acceptation. Il peut être en effet difficile pour les professionnels d'avoir un discours stable et cohérent, sur des outils dont ils ne maîtrisent pas les technologies et qui évoluent à une grande vitesse, sur le mode du « *perpetual beta* » (version bêta perpétuelle) : « l'absence d'un produit abouti peut faire que les équipes de la bibliothèque doutent de leur capacité à fournir des services cohérents et stables aux utilisateurs »<sup>188</sup>.

Malgré ces difficultés, et y compris dans des cas où les bibliothèques ne disposent pas encore (ou pas assez) de données qualitatives concernant les usages par le public, on peut noter que, dans de nombreux cas, l'expérience générale de l'outil est appréciée de manière positive par les professionnels interrogés (BIB17, BIB13, BIB04, BIB06, BIB02, BIB25, BIB24, BIB16, BIB20, BIB19, BIB23, BIB01, BIB18). Ces derniers y voient notamment la possibilité d'offrir des services élargis et des interfaces modernisées pour les usagers.

## B. Accompagner l'évolution des pratiques professionnelles

Les outils de découverte peuvent donc venir bouleverser des pratiques professionnelles : une barre de recherche « à la Google » au lieu d'index prédéfinis, une recherche plein texte au détriment des vocabulaires contrôlés, l'abandon des opérateurs booléens, la perte de visibilité du catalogue, etc. Sur tous ces points, la logique de l'outil de découverte est différente de celle de la recherche documentaire traditionnelle et va à l'encontre de pratiques professionnelles bien implantées, ce qui peut susciter le malaise ou l'incompréhension des équipes.

En outre la mise en place de l'outil peut aussi se traduire par des changements de *workflow* au sein des équipes. Ainsi, à la BIB18, la mise en place de l'outil a entraîné l'arrêt du catalogage des ressources numériques dans le SIGB ce qui a signifié un changement de pratiques internes d'indexation et de classification ainsi que la modification des tâches d'une vingtaine de chargés de collection.

Le déploiement d'un outil de découverte exige donc la mise en place d'une pédagogie à destination des professionnels afin d'accompagner le changement d'outil et de créer les conditions de son acceptation (BIB12, BIB18, BIB02, BIB24, BIB07, BIB25). Il s'agit là d'un travail non négligeable qui peut nécessiter plusieurs dizaines d'heures de formation, en particulier à destination des équipes de formateurs qui seront ensuite aux premières lignes pour présenter l'outil aux usagers (BIB18, BIB02). Le passage à une interface principale avec l'abandon symbolique de l'OPAC traditionnel, ou à un changement lourd de version de l'outil de découverte, nécessite, lui aussi, un accompagnement renforcé (BIB24,

<sup>188</sup> « The lack of polish product can leave library staff feeling unsure about their ability to deliver consistent, stable services to users ». *Ibid.*, p. 377.

BIB02, BIB25). Le cas échéant, les bibliothèques peuvent s'appuyer sur les clubs utilisateurs existants pour préparer ces formations.

Il peut être également souhaitable d'investir le plus possible les équipes de la bibliothèque dans le projet de déploiement d'un outil de découverte (BIB18). Inversement, lorsque le personnel est peu sollicité ou n'est pas formé au nouvel outil, celui-ci peut rester un outil secondaire à ses yeux (BIB10).

L'effet de ces formations peut être estimé par des enquêtes auprès du personnel. Ainsi, à la BIB18, une première enquête après la mise en place de l'outil avait révélé de fortes réticences des équipes, notamment au sujet de la perte des formats MARC dans l'affichage des notices et de la fenêtre de recherche unique. Cette dernière cristallisait en elle-même de nombreuses réticences, avec 60% de répondants qui estimaient que le point d'entrée unique pour toutes les ressources de la bibliothèque ne constituait pas une avancée. Trois ans plus tard, une autre enquête auprès du personnel, partiellement renouvelé, révélait un degré d'acceptation beaucoup plus important de la part des équipes, avec plus de 70% d'acceptation du point d'entrée unique.

Enfin pour la BIB18, la mise en place de l'outil a également permis de développer des compétences professionnelles, notamment une plus grande maîtrise des concepts sur l'évolution des catalogues (les problématiques de frbérisation) et plus largement sur les ressources numériques. Dans ce cas, cela s'est traduit par une nette montée en compétences du personnel dans la maîtrise et la compréhension des outils d'informatique documentaire et de documentation numérique.

## IV. PERSPECTIVES

---

Dans le cadre de ce travail, nous avons souhaité aussi proposer des alternatives aux outils commerciaux et aux index centralisés, en présentant l'expérience de différentes bibliothèques ayant opté pour d'autres solutions. Nous signalerons ensuite les possibilités d'amélioration et d'évolution pour les outils à index général centralisé, notamment en ce qui concerne la qualité des données.

### 1. DES ALTERNATIVES AUX INTERFACES COMMERCIALES : LES OUTILS LIBRES

Les outils *open source* dans le domaine concernent uniquement des interfaces de découverte et non pas des outils à index général centralisé. En France, les bibliothèques ayant mis en place des interfaces de découverte libres restent relativement peu nombreuses. On signalera ici deux cas d'installations achevées : la Bibliothèque Clermont-Université (BCU) et le portail Babord+ des universités de Bordeaux. Dans les deux cas, il s'agit d'interfaces qui travaillent avec une granularité de niveau titre de l'ouvrage ou de la revue, et qui intègrent différentes ressources locales mais qui, contrairement à d'autres outils mis en place<sup>189</sup>, ne se combinent pas avec un index centralisé commercial.

Ces outils sont installés en local et ont pu être testés avant la mise en œuvre. Ils ont permis aux bibliothèques concernées de développer des fonctionnalités adaptées à leurs besoins, afin d'offrir une meilleure intégration des services en ligne.

#### A. La mise en place

##### *Blacklight à la Bibliothèque Clermont-Université*

La BCU a mis en place l'outil libre Blacklight<sup>190</sup> en 2013<sup>191</sup>. Cet outil est écrit en Ruby on Rails et travaille avec le moteur d'indexation SOLR<sup>192</sup>.

L'outil Blacklight de la BCU intègre plusieurs sources de données locales, dont les données du catalogue. Il faut noter à ce sujet que la version initiale de Blacklight ne connaissait pas le protocole OAI-PMH ni Z39.50. Son moteur

---

<sup>189</sup> Comme celles de la bibliothèque de l'université de Chicago, ou celle d'Alberta. Cf. I, 2, « Des outils mixtes »

<sup>190</sup> Home — Blacklight [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://projectblacklight.org/>.

<sup>191</sup> Voir les différents billets publiés dans le blog de la BCU à ce propos : BUSSEMEY, Raphaële. *Mise en place de Blacklight : pourquoi ?* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://alambic.hypotheses.org/4163> ; BUSSEMEY, Raphaële. *Mise en place de Blacklight : comment ?* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4172> ; BACAULT, Mathieu. *Blacklight, deuxième partie : les données d'exemplaires* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/2114> ; BACAULT, Mathieu. *Blacklight : déploiement avec Ruby On Rails* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/3456> ; BACAULT, Mathieu. *Blacklight, troisième partie : les services au lecteur* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/2785>.

<sup>192</sup> Ruby on Rails [en ligne]. [s. d.]. [Consulté le 27 novembre 2016]. Disponible à l'adresse : <http://rubyonrails.org/>. Apache Solr - [en ligne]. [s. d.]. [Consulté le 27 novembre 2016]. Disponible à l'adresse : <http://lucene.apache.org/solr/>.

d'indexation SOLR peut cependant intégrer des notices MARC en XML ou en ISO2709. Ayant été développé aux Etats-Unis, Blacklight utilise le format MARC21 : la BCU a donc réalisé en un premier temps des *mappings* entre l'UNIMARC et le MARC21, ce qui a impliqué des modifications non négligeables pour un certain nombre de zones, telles que les auteurs secondaires ou les titres complémentaires. Pour mettre en place l'import des notices du SIGB dans Blacklight au format ISO2709, la BCU a demandé à son fournisseur de SIGB de créer une procédure d'export quotidienne. Toutes les nuits, les données du catalogue sont ainsi importées dans Blacklight : il ne s'agit pas d'un import incrémental mais d'un import total qui écrase l'ensemble des données du catalogue présentes dans Blacklight. Cet import total est pour la bibliothèque plus facile à gérer qu'un import différentiel et ne pose pas de problèmes de lenteur : chaque nuit, l'import dans Blacklight et l'indexation des notices dans SOLR sont réalisés très rapidement.

Pour interroger la disponibilité et la localisation des exemplaires en temps réel, la BCU s'est appuyée sur l'API ZOOM<sup>193</sup> qui lui a permis de développer des services en Z39.50. Grâce à cette API, Blacklight peut interroger la base de données du SIGB par le protocole Z39.50 et le serveur Z39.50 du SIGB lui renvoie les données d'exemplaire en temps réel au format XML. C'est également une interrogation du SIGB en Z39.50 qui permet d'afficher les informations du compte utilisateur du SIGB sur l'interface de découverte. Il a fallu pour cela utiliser un connecteur spécifique développé par le fournisseur du SIGB, afin de pouvoir interroger les données du fichier lecteur et permettre à l'utilisateur d'effectuer des prolongations en ligne, des demandes en magasin et des réservations. Que ce soit pour les exemplaires ou pour le compte utilisateur, il faut noter toutefois que cette connexion en Z39.50 est très gourmande en sessions, ce qui peut induire certaines lenteurs, car le SIGB de la BCU n'a pas été conçu pour fonctionner avec une connexion web. Le point positif reste que l'utilisateur ne doit pas se rendre sur une autre interface pour consulter son compte personnel.

La BCU a aussi mis en place l'import dans Blacklight de sa liste de revues numériques, de ses livres numériques et de sa bibliothèque numérique. La liste de revues numériques est exportée une fois par mois manuellement depuis la base de connaissance commerciale de la BCU dans un fichier au format XML qui est ensuite converti en MARC21 puis en UNIMARC. En complément, les éléments de ce fichier se voient attribuer la mention « revues électroniques ». Une fois importé dans Blacklight et indexé par SOLR, ce fichier remplace celui du mois précédent. Pour ses listes de livres numériques, la BCU effectue aussi un import mensuel dans Blacklight d'un fichier en ISO2709 qui est indexé dans la foulée par SOLR, tandis que l'ancien fichier est supprimé : dans la mesure où il n'y a pas de livres numériques dans le catalogue, il n'y a pas de risques de générer des doublons. Blacklight disposant désormais d'un plugin moissonneur en OAI-PMH, la BCU a pu mettre en place en 2016 le moissonnage quotidien des métadonnées de sa bibliothèque numérique sous Omeka. Les notices de la bibliothèque numérique, disponibles au format Dublin Core, sont converties grâce à un *mapping* spécifique au format UNIMARC, pour permettre leur import et leur indexation par SOLR.

Le moteur d'indexation SOLR a, quant à lui, fait l'objet de nombreux paramétrages et ajustements successifs à la BCU. Il a fallu notamment faire en

<sup>193</sup> ZOOM - Home Page [en ligne]. [s. d.]. [Consulté le 27 novembre 2016]. Disponible à l'adresse : <http://zoom.z3950.org/>.

sorte qu'il prenne correctement en compte les caractéristiques de la langue française grâce à un fichier de configuration spécifique afin qu'il puisse analyser les structures de mots et corriger une faute d'orthographe si besoin. La pondération des résultats sur laquelle repose le calcul de l'algorithme de pertinence a également fait l'objet d'ajustements : SOLR a été paramétré pour prendre en compte la présence des mots recherchés dans le titre (poids le plus important), ensuite dans les mots sujets ou à défaut dans le reste de la notice (résumé, notes etc.). Le nombre de mots d'une recherche pris en compte par SOLR a dû aussi être adapté. En effet, dans sa configuration par défaut, SOLR prenait en compte seulement les trois premiers mots d'une recherche, ce qui peut ne pas être pertinent lorsqu'un utilisateur tape le titre complet d'un ouvrage. La BCU a donc modifié ce paramétrage pour que SOLR prenne désormais en compte six mots au lieu de trois. C'est aussi un paramétrage spécifique du moteur d'indexation SOLR qui permet de mieux exploiter les vedettes RAMEAU au niveau des facettes, en séparant l'affichage des subdivisions géographiques (\$y) et chronologiques (\$z) dans deux facettes distinctes (« région » et « période ») tout en laissant les vedettes complètes intactes au niveau de la notice.

Ces différentes améliorations répondent au souci de garder une interface de découverte qui reste lisible pour les étudiants : ainsi la BCU ne souhaite pas pour le moment intégrer dans Blacklight les données des archives ouvertes HAL, car cela risquerait de brouiller la présentation des résultats, en introduisant un niveau de granularité plus fine. À ce jour, la BCU a constaté que Blacklight a été très bien pris en main par les étudiants des premiers cycles dont le besoin documentaire est essentiellement centré sur les monographies. De même, il ressort des formations effectuées par la BCU que Blacklight répond beaucoup mieux aux attentes du public étudiant que l'ancien OPAC. Les chercheurs, quant à eux, ont indiqué regretter l'absence d'un formulaire de recherche avancée pour pouvoir effectuer des requêtes plus complexes. Cette possibilité de recherche avancée a donc été rajoutée à l'outil de la BCU afin de répondre à ce besoin des enseignants-chercheurs.

### **Babord+**

Le portail bordelais Babord+<sup>194</sup> a été mis en ligne en 2011 et il est le fruit d'une longue tradition de coopération documentaire inter-établissements à Bordeaux<sup>195</sup>. À l'origine du projet, les différents établissements partenaires disposaient déjà d'un SIGB commun et chacun avait sa propre liste alphabétique de ressources. Ils manquaient en revanche d'une interface de consultation des documents numérisés et d'un ERMS. Dans ce contexte, le portail Babord+ a été la première brique d'un projet de réinformatisation qui devait comprendre la brique de découverte, l'ERMS, le SIGB, une bibliothèque numérique, une archive ouverte institutionnelle et une plateforme de revues universitaires en ligne<sup>196</sup>.

<sup>194</sup> Recherche - Université de Bordeaux - Documentation [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : <https://babordplus.univ-bordeaux.fr/>.

<sup>195</sup> COUDRIN, Delphine. Les dispositifs de signalement en université. *Op. cit.* ; COUDRIN, Delphine. *Faire évoluer un SID : enjeux / retour d'expérience - Intervention DCB 25* [en ligne]. ENSSIB, 29 juin 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : [http://moodle.enssib.fr/pluginfile.php/20764/mod\\_resource/content/0/Intervention%20Delphine%20Coudrin.pdf](http://moodle.enssib.fr/pluginfile.php/20764/mod_resource/content/0/Intervention%20Delphine%20Coudrin.pdf). COUDRIN, Delphine. BABORD+ : Une interface de recherche documentaire innovante pour les universités bordelaises. *Arabesques*. 2012, n° 66, p. 20.

<sup>196</sup> COUDRIN, Delphine. *Faire évoluer un SID*. *Op. cit.*

Le portail Babord+, hébergé par l'université de Bordeaux, a été conçu et mis en œuvre à partir de différentes briques logicielles libres et, comme Blacklight, il s'appuie sur le moteur d'indexation SOLR.

L'index de Babord+ est alimenté par de multiples sources à travers différents protocoles : pour les données bibliographiques du SIGB, il s'agit d'imports en ISO2709, pour les métadonnées des listes A-Z et de l'ERMS il s'agit d'imports de fichiers Excel ou csv, pour les thèses ou bases extérieures, ce sont des moissonnages en OAI-PMH. Le portail Babord+ intègre aussi des ressources pédagogiques numériques, comme les vidéos de Canal U. Pour toutes ces ressources, Babord+ propose une barre de recherche unique et l'affichage des notices est enrichi de différents éléments provenant de bases extérieures, comme par exemple les indicateurs bibliométriques issus de la base Scopus et du code couleur Romeo indiquant la politique éditoriale des revues.

Si Babord+ est un portail consortial, chaque établissement membre dispose d'un profil spécifique qui lui permet d'adapter graphiquement l'interface et de paramétrer les ressources interrogées ainsi que les options de recherche par défaut<sup>197</sup>. Pour les utilisateurs de chaque établissement, l'interface Babord+ permet d'accéder aux services distants à travers un compte personnel : « L'utilisateur accède à un profil établissement via les environnements web ou les environnements numériques de travail (ENT) dans lesquels l'établissement a encapsulé l'interface. L'authentification des usagers, appuyée sur le protocole Shibboleth, permet de proposer des services personnalisés<sup>198</sup>. » Grâce à cette authentification, l'utilisateur a accès à la sauvegarde de notices, aux alertes, aux commentaires ainsi qu'à son compte personnel, sans devoir se rendre sur une autre interface pour consulter ce dernier.

Dès sa mise en ligne en 2011, le portail Babord+ a été lauréat d'un prix de l'innovation numérique en bibliothèque décerné par l'ENSSIB<sup>199</sup>. Il a également été signalé comme une interface documentaire innovante dans une récente étude de l'Enssib qui précise : « l'innovation de Babord+ tient à sa capacité d'agrégation des ressources y compris en *open access*, à la qualité de l'indexation et à la rapidité de la recherche<sup>200</sup> ».

À Bordeaux, comme à Clermont, le souci de préserver une interface de découverte adaptée aux étudiants et une granularité de niveau titre de revue ou d'ouvrage, exclut l'intégration des données des archives ouvertes HAL<sup>201</sup>. Ce parti-pris a réussi dans la mesure où l'interface Babord+ a été très rapidement adoptée par les usagers avec des retours très positifs dès les premiers mois<sup>202</sup>. Une enquête avec des focus groups menée après la mise en place du portail a toutefois permis d'introduire plus de nuances dans ce constat : si, pour les étudiants, Babord+ est très fortement utilisé, en particulier pour la recherche de documentation imprimée, les chercheurs sont restés beaucoup plus

<sup>197</sup> COUDRIN, Delphine. Les dispositifs de signalement en université. *Op. cit.*, p. 122.

<sup>198</sup> COUDRIN, Delphine. BABORD+ : Une interface de recherche documentaire innovante pour les universités bordelaises. *Op. cit.*, p. 20.

<sup>199</sup> *Babord + : un des trois projets primés lors de la biennale du numérique de l'enssib | Enssib* [en ligne]. [s. d.]. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <http://www.enssib.fr/brevues/2011/11/22/babord-un-des-trois-projets-primés-lors-de-la-biennale-du-numérique-de-l-enssib>.

<sup>200</sup> MULLER, Catherine. *Étude et repérages sur les interfaces documentaires innovantes*. Villeurbanne, France : EnssibLab, 2014. [Consulté le 21 février 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/64911-etude-et-reperages-sur-les-interfaces-documentaires-innovantes.pdf>, p. 9.

<sup>201</sup> COUDRIN, Delphine. Les dispositifs de signalement en université. *Op. cit.*, p. 125.

<sup>202</sup> COUDRIN, Delphine. BABORD+ : Une interface de recherche documentaire innovante pour les universités bordelaises. *Op. cit.*, p. 20.

circonspects devant cet outil, avec de fortes disparités d'usage selon les disciplines. Ainsi, les chercheurs en SHS utilisent davantage Babord+ que les chercheurs en STM pour lesquels Google Scholar, *PubMed* et la liste alphabétique de ressources demeurent des outils incontournables<sup>203</sup>.

## B. Les moyens mobilisés

La mise en place de ces interfaces a demandé dans les deux cas une forte mobilisation des équipes et un investissement important en termes de temps de travail et de compétences, que ce soit en interne ou avec un prestataire extérieur.

À la BCU, la mise en place de Blacklight a impliqué pour les administrateurs d'acquérir des compétences techniques considérables. Comme Blacklight ne comporte pas d'interface d'administration, les différentes opérations sont réalisées directement en ligne de commande sur le serveur. Il a donc fallu dans un premier temps apprendre le langage Ruby on Rails dans lequel Blacklight est écrit. À cet égard, la présence dans l'équipe de la BCU d'un développeur dédié à ce projet a été décisive pour mener à bien l'implémentation de l'outil. La mise en place des connecteurs entre Blacklight et le SIGB pour la récupération des données d'exemplaires en temps réel et l'intégration du compte utilisateur, ont été aussi des étapes particulièrement complexes d'un point de vue technique. De même, les différents paramétrages apportés au moteur d'indexation SOLR ont exigé de comprendre de manière approfondie son fonctionnement. Il s'agit donc là d'un projet qui a été très demandeur en matière de compétences et qui a fortement mobilisé les équipes sur le long terme. La phase de mise en œuvre, de la première installation à la mise en test publique, a ainsi duré douze mois et a elle-même été suivie d'une longue période d'évolution fonctionnelle, y compris jusqu'en 2016, avec l'intégration de la bibliothèque numérique. Au quotidien, Blacklight mobilise deux personnes, dont le développeur, ce qui reste important, même si cela ne correspond pas à deux temps complets à proprement parler.

Pour le portail Babord+, le caractère consorsial du projet et la taille des établissements concernés impliquaient d'eux-mêmes une importante mobilisation des équipes : l'équipe projet de Babord+ était ainsi composée de 14 personnes correspondant à 4,25 ETP<sup>204</sup>. Dans ce cas, la mise en place de l'outil n'a pas été assurée directement par les établissements eux-mêmes, mais est passée par un appel d'offres ouvert à la fois aux solutions libres et aux solutions propriétaires. Au préalable, afin de préparer la rédaction du CCTP de Babord+, l'équipe projet a fait appel à un consultant pour l'assistance à maîtrise d'ouvrage<sup>205</sup>. À l'issue de la mise en concurrence, le choix s'est porté sur une solution libre avec un société de prestations informatiques chargée de l'implémentation, des développements et de la maintenance du portail. Dans cette configuration, avec des équipes techniques externalisées, le coût annuel par rapport à une solution propriétaire reste faible, mais les coûts initiaux du projet Babord+ ont été, en revanche, très importants<sup>206</sup>. Désormais, les équipes en charge de Babord+ assurent directement la maintenance du portail, sans passer par un prestataire extérieur.

<sup>203</sup> COUDRIN, Delphine. *Faire évoluer un SID*, Op. cit.

<sup>204</sup> *Ibid.*

<sup>205</sup> *Ibid.* et MULLER, Catherine. *Étude et repérages sur les interfaces documentaires innovantes*. Op. cit.

<sup>206</sup> COUDRIN, Delphine. *Faire évoluer un SID*, Op. cit.

## 2. DES ALTERNATIVES AUX INDEX CENTRALISES

Pour signaler et donner accès au contenu de ressources numériques, notamment avec une granularité de niveau article ou chapitre, sans passer par un service de découverte à index général centralisé, plusieurs alternatives sont envisageables.

### A. Paramétrer un résolveur de liens avec Google Scholar

À de nombreux égards pour les bibliothèques, l'utilisation de Google Scholar peut représenter une alternative aux services de découverte à index général centralisé<sup>207</sup>. Paramétré avec un résolveur de liens, Google Scholar peut permettre l'accès à des données avec une granularité de niveau article. Cette solution a été mise en place, en France, par l'université de Bretagne Occidentale ou encore par celle de Lyon 1<sup>208</sup> et, à l'étranger, par les universités d'Utrecht et de Gand ainsi que par l'université Libre de Bruxelles<sup>209</sup>. Cette articulation de Google Scholar avec un résolveur de liens avait été par ailleurs un des scénarios envisagés par l'étude de l'ABES sur un éventuel outil de découverte national<sup>210</sup>.

Si l'emploi d'un résolveur de liens commercial reste bien évidemment payant pour les bibliothèques, les fonctionnalités de recherche et d'utilisation de liens offertes par Google Scholar sont actuellement sans frais. Toutefois, pour faire fonctionner le résolveur de liens avec Google Scholar, le fournisseur du résolveur ou la bibliothèque devront indiquer au moteur de recherche les données concernant les adresses IP de la bibliothèque et celles de sa base de connaissance<sup>211</sup>. Ainsi, à l'instar de bien d'autres activités de Google, telles que la numérisation de masse, l'apparente gratuité du service s'accompagne en contrepartie de la fourniture de données au moteur de recherche par des établissements publics.

S'il indexe des ressources de niveau universitaire et de recherche, Google Scholar fonctionne sur un modèle différent de celui des services de découverte à index centralisés commerciaux, car il est alimenté essentiellement par des processus automatiques : ce sont des robots moissonneurs qui collectent les ressources académiques comme ils collectent les sites web en général<sup>212</sup>. Pour Marshall Breeding, le fait que les robots de Google Scholar moissonnent des contenus sur des serveurs sécurisés en temps réel fluidifie la mise à jour des contenus, mieux que lors des transferts par lots utilisés par les services de découverte à index général centralisé<sup>213</sup>.

<sup>207</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 6-9.

<sup>208</sup> Accueil-Ubodoc. Dans : *Ubodoc* [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://ubodoc.univ-brest.fr/>. Accéder aux publications référencées dans Google Scholar - BU Lyon 1 - Université Lyon 1. Dans : *BU Lyon 1* [en ligne]. [s. d.]. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <http://portaildoc.univ-lyon1.fr/les-collections/bibliotheque-en-ligne/accéder-aux-publications-referencées-dans-google-scholar-840654.kjsp>

<sup>209</sup> KEYSER, Piet de, *Op. cit.*, p. 62-63.

<sup>210</sup> Voir : ABES. Vers un outil de découverte national. *Op. cit.* Et aussi : ABES. L'étude d'impact du SGB mutualisé et l'étude sur un dispositif de découverte. Dans : *SGB mutualisé* [en ligne]. 30 mars 2013. [Consulté le 30 novembre 2016]. Disponible à l'adresse : <https://sgbm.abes.fr/2013/03/30/letude-dimpact-du-sgb-mutualise-et-letude-sur-un-dispositif-de-decouverte/>.

<sup>211</sup> Google Scholar Support for Libraries [en ligne]. [s. d.]. [Consulté le 14 décembre 2016]. Disponible à l'adresse : <https://scholar.google.com/intl/en/scholar/libraries.html>

<sup>212</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 7.

<sup>213</sup> « This live harvesting facilitates a more current representation of recently published articles compared to the bulk transfers employed by the library-oriented index-based discovery services » *Ibid.*, p. 7.



La politique d'indexation de Google Scholar repose comme celle des outils de découverte propriétaires sur des accords commerciaux avec des éditeurs. Pour pouvoir indexer du contenu propriétaire, Google Scholar a passé des accords avec les fournisseurs de données : ainsi Google Scholar précise dans sa politique à destination des éditeurs que ces derniers gardent le contrôle sur l'accès à leurs articles mais que le résumé ou le contenu de la première page doit être accessible pour les internautes<sup>214</sup>. Ces accords interdisent notamment la livraison des données moissonnées par Google Scholar à des tiers. Ainsi, interrogé sur la possibilité de développer une API de Google Scholar pour permettre d'exploiter les données de ce dernier depuis d'autres plateformes, Anurag Acharya, co-développeur du moteur de recherche, a formellement répondu par la négative : « Je ne peux pas faire cela. Nos accords commerciaux avec les éditeurs l'excluent. Nous sommes autorisés à scanner tous les articles, mais non pas à distribuer ces informations en gros à des tiers »<sup>215</sup>. De manière plus générale, plusieurs articles ont signalé que les utilisateurs disposent de peu d'informations détaillées sur l'étendue, la couverture et le niveau d'indexation des ressources référencées dans Google Scholar<sup>216</sup>. À terme, comme l'a signalé Marshall Breeding, la marge de manœuvre des bibliothèques pour influencer les évolutions ultérieures de Google Scholar dans le sens d'une plus grande transparence concernant les ressources indexées et le fonctionnement général semble plutôt réduite : « Les bibliothèques ne seront vraisemblablement pas en mesure de jouer un rôle important dans la conception de l'avenir du service ou d'obtenir plus de transparence concernant son mode de fonctionnement détaillé ou le contenu couvert. »<sup>217</sup>

## B. Intégrer des API d'autres bases à un catalogue public

En 2015, le réseau des bibliothèques d'enseignement supérieur stéphanois BRISE-ES a intégré au catalogue public de son SIGB Koha, renommé pour l'occasion BRISE-ES+<sup>218</sup>, les API mises à disposition par deux bases bibliographiques : ISTE<sup>219</sup> et BASE<sup>220</sup>. Ces deux bases apparaissaient comme complémentaires en termes de contenus : la première référence les données des licences nationales et la deuxième correspond principalement à un moteur de recherche de données en *open access* maintenu par l'université de Bielefeld. Elles

<sup>214</sup> *Google Scholar - Publisher Support* [en ligne]. [s. d.]. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <https://scholar.google.com/intl/en/scholar/publishers.html#policies>.

<sup>215</sup> « I can't do that. Our indexing arrangements with publishers preclude it. We are allowed to scan all the articles, but not to distribute this information to others in bulk. » VAN NOORDEN, Richard. *Google Scholar pioneer on search engine's future*. *Nature* [en ligne]. Novembre 2014. [Consulté le 11 novembre 2016].

<sup>216</sup> ORDUÑA-MALEA, Enrique, AYLLÓN, Juan M., MARTÍN-MARTÍN, Alberto, et al. *Methods for estimating the size of Google Scholar*. *Scientometrics* [en ligne]. Septembre 2015, Vol. 104, n° 3, p. 931-949. BREEDING, Marshall et National Information Standards Organization. *Op. cit.* TAY, Aaron. *8 surprising things I learnt about Google Scholar* [en ligne]. [s. d.]. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <http://musingsaboutlibrarianship.blogspot.com/2014/06/8-surprising-things-i-learned-about.html>.

<sup>217</sup> « Libraries are not likely to have a large role in shaping the future of the service or gaining additional transparency regarding its details of operation or content covered. » BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 8.

<sup>218</sup> *BRISE-ES Catalogue* [en ligne]. [s. d.]. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <https://brisees-opac.univ-st-etienne.fr/>.

<sup>219</sup> *Plateforme ISTE / - ISTE* [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://www.istex.fr/plateforme-istex/>. A noter que dans le cas d'ISTE, l'API est encore en phase d'expérimentation et le réseau BRISE-ES en est un des testeurs, avec l'université de Lorraine et celle de Strasbourg.

<sup>220</sup> *BASE - Bielefeld Academic Search Engine / About BASE* [en ligne]. [s. d.]. [Consulté le 22 novembre 2016]. Disponible à l'adresse : <https://www.base-search.net/about/en/>.

allaient permettre d'apporter plusieurs millions de documents supplémentaires et un niveau de granularité plus fin que celui du catalogue, allant jusqu'à l'article ou au chapitre d'ouvrage, sans devoir pour autant gérer localement ces métadonnées.

Sur BRISE-ES+, l'interrogation du catalogue et de ces deux bases extérieures se fait dans une même interface, celle du catalogue public de Koha. L'interrogation de ces bases est présentée comme une extension possible du périmètre de recherche, qui par défaut est celui du catalogue public. Le choix de cette recherche « + » n'est pas coché d'office (il est ensuite conservé dans un cookie pour que l'utilisateur retrouve son choix précédent). L'idée étant de faire connaître (et de donner accès à) cette possibilité d'extension, mais sans l'imposer.

Une fois la recherche « + » lancée, les résultats du catalogue et ceux renvoyés par les API sont affichés au sein de la même interface, le catalogue public de Koha, mais de manière séparée : il s'agit bien de la page de résultats habituelle du catalogue public, mais dans laquelle vient s'insérer, au-dessus des réponses du catalogue, un bandeau spécifique avec des liens vers les résultats des bases extérieures ISTEEX ou BASE. C'est en cliquant sur l'un ou l'autre de ces liens que les réponses de chacune sont rendues visibles pour l'utilisateur, avec à chaque fois une liste de résultats distincte, proposant ses propres facettes et la possibilité de relancer la recherche dans la base concernée. De cette façon, les réponses renvoyées par les différentes sources de données ne sont pas mélangées dans une liste unique, mais sont affichées dans des espaces distincts et le passage de l'un à l'autre se fait grâce aux liens situés dans le bandeau. Les résultats du catalogue ne sont donc pas confondus dans une grande masse de références d'articles issues de ces bases bibliographiques, chaque type de résultat restant associé à sa propre source de données dans un espace bien spécifique.

L'utilisation que l'on peut faire d'une API est bien entendu contrainte par la liste de métadonnées qu'impose cette API, mais pour autant elle laisse une grande marge de manœuvre à la bibliothèque pour des adaptations locales ou des développements. Ainsi, par exemple, dans le cas de BRISE-ES+, le DOI d'un article est utilisé non pour l'accès au texte intégral (pour lequel c'est le lien vers ISTEEX qui est privilégié), mais pour permettre à l'utilisateur de bénéficier d'outils disponibles sur les plateformes d'éditeurs, comme par exemple « cité par ». Un développement a également été réalisé sur BRISE-ES+ pour permettre de récupérer toute une liste de résultats au format RIS (notamment pour les utilisateurs de logiciels de gestion de références bibliographiques). Enfin, alors que les données d'ISTEEX relèvent d'une granularité de niveau article ou chapitre, un développement a été effectué pour générer le sommaire d'un numéro de revue ou d'un ouvrage à partir d'une notice d'article ou de chapitre d'ISTEEX (en réinterrogeant la même API mais avec une question ad hoc). L'utilisateur dispose ainsi, dans l'espace de résultats d'ISTEEX, d'un lien de l'article ou du chapitre vers le sommaire complet du numéro de la revue ou de l'ouvrage (et réciproquement), ce qui permet d'exploiter de manière riche et cohérente différents niveaux de granularité. Il s'agit là d'un développement particulièrement intéressant qui n'a pas son pareil dans les outils de découverte commerciaux.

L'intégration de ces API dans le catalogue de Koha fait appel à des techniques "classiques" d'interface web (code HTML, CSS, JavaScript, jQuery, Ajax et ici, en raison des particularités d'installation du logiciel Koha dans le

réseau BRISE-ES, PHP)<sup>221</sup>. On peut noter cependant que l'interrogation conjointe du catalogue et des bases extérieures a nécessité d'effectuer des correspondances entre les différents index de chaque base et, au-delà, d'écrire un algorithme pour analyser finement et retraiter la syntaxe des requêtes, afin que celles-ci puissent être correctement prises en compte par chaque base<sup>222</sup>. Cette intégration, réalisée en interne, a donc nécessité de disposer au sein de l'équipe de compétences informatiques touchant ces différents domaines.

### C. Exploiter les données de Crossref ?

L'agence Crossref est une organisation à but non lucratif qui est chargée de l'attribution et de l'enregistrement des DOI. En 2016, la base de Crossref référençait plus de 80 millions d'enregistrements<sup>223</sup>. Un DOI étant l'identifiant unique qui permet de tracer et de citer une production intellectuelle, y compris avant sa publication effective chez un éditeur, la base Crossref a l'avantage de présenter un niveau de fraîcheur de l'information très élevé.

Crossref met à disposition des utilisateurs et des développeurs une API ouverte pour exploiter les métadonnées de la base. Cette API permet de construire des interfaces de recherche, de croiser des informations des éditeurs afin de permettre le développement d'applications de *data mining* et d'obtenir des rapports sur les métadonnées Crossref. Cette API est utilisée par différentes plateformes collaboratives d'édition de littérature grise en ligne (comme Authorea ou Paperhive) qui s'en servent pour permettre à leurs utilisateurs d'insérer facilement des citations dans leurs articles grâce à une recherche dans la base Crossref sur le nom de l'auteur, l'année, le DOI ou l'identifiant ORCID.

Cependant, l'API de Crossref ne semble pas pour autant faite pour supporter des recherches bibliographiques plus larges. Le fonctionnement de l'API est en effet limité en termes de temps de réponse et de nombre de résultats, ainsi que par les possibilités de recherche proposées (absence de l'opérateur « AND »). De ce fait l'usage massif de l'API de Crossref depuis une interface de découverte semble difficilement envisageable en tant que tel, mais elle reste plutôt opérationnelle pour des interrogations plus restreintes, qui peuvent permettre d'ajouter des fonctionnalités supplémentaires à un outil de découverte existant. Par exemple, il est possible d'interroger les données Crossref de manière ciblée pour la veille bibliographique afin de proposer les articles les plus récents dans un domaine donné grâce à une interrogation sur la date de création des articles et les mots du titre (BIB26).

### D. Des corpus et des réseaux : le dispositif CERCLES

En complément, les bibliothèques peuvent aussi s'appuyer sur le travail mené dans le réseau SUDOC pour obtenir des données de qualité des ressources

<sup>221</sup> ROUGER, Dominique. Intégration des services web de recherche dans des bases bibliographiques au sein d'un catalogue Koha. Dans : *Symposium Koha 2016* [en ligne]. Université Jean Moulin Lyon 3, 28 juin 2016. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <https://youtu.be/-T5D48liqM>. (Entre 1:40:35 et 2:19:00).

<sup>222</sup> L'API d'ISTEX est sensible à la casse et exige des règles de syntaxe bien spécifiques (comme par exemple les espaces et la ponctuation après les noms d'index). Par ailleurs, l'opérateur booléen par défaut de l'API d'ISTEX est « OR » (alors que dans le catalogue public c'est « AND ») et dans la version initiale de l'API cet opérateur ne pouvait pas être redéfini.

<sup>223</sup> *crossref.org* [en ligne]. [Consulté le 13 novembre 2016]. Disponible à l'adresse : <http://www.crossref.org/>.

numériques, avec une granularité de niveau titre de la revue ou de l'ouvrage. En effet, afin d'améliorer la qualité des données des ressources numériques dans le SUDOC, l'ABES a lancé en 2015 le projet CERCLES<sup>224</sup> (Correction et Enrichissement par le Réseau de Corpus de l'Enseignement Supérieur). Ce dispositif s'appuie sur les établissements du réseau SUDOC pour enrichir les notices de corpus numériques, en attribuant à chaque établissement participant un chantier qualité correspondant à un corpus donné. Les établissements qui s'y engagent deviennent responsables qualité des leurs corpus respectifs et ont un rôle de référent pour le réseau. De son côté, l'ABES coordonne les chantiers entre les établissements et apporte une aide technique et méthodologique en mettant à disposition des établissements des outils de suivi. L'ABES a également pour fonction de recueillir et diffuser les informations sur l'avancée des chantiers au sein du réseau SUDOC.

Ce dispositif, qui implique l'ABES et le réseau, doit aboutir à un meilleur signalement de la documentation électronique, en particulier pour des corpus acquis par des licences nationales. En novembre 2016, le projet CERCLES avait à son actif quatre chantiers terminés et onze chantiers en cours<sup>225</sup>. L'enrichissement qualitatif des notices se traduit par exemple au niveau des différents corpus par les opérations suivantes : création des autorités auteur (zones 7XX), ajout des liens vers les autorités-matière et auteurs (zones 6XX et 7XX), ajout des liens réciproques en zones de liens (zones 4XX), vérification ou ajout d'identifiants (ISSN, ISBN), vérification des URL d'accès, dédoublonnage, et traitement des zones de notes d'information interne au SUDOC en zone 309<sup>226</sup>. Les établissements qui participent au dispositif peuvent s'appuyer notamment sur les équipes du hub de métadonnées à l'ABES afin de faire le maximum de traitements automatisés (liens auteurs, insertion de la table de matières, ajout d'ISBN, ajout d'indices Dewey etc.) et de concentrer leurs efforts sur l'enrichissement intellectuel des données.

À ce jour, plusieurs retours d'expérience d'établissements participants ont été diffusés<sup>227</sup>. Ils ont relevé l'importance de ce positionnement des établissements comme référents pour le réseau ainsi que l'effet positif de l'émulation entre établissements. Ils ont signalé également que ces chantiers, quoique importants en termes de temps de travail, sont valorisants d'un point de vue professionnel, à la fois pour les catalogueurs et pour les chargés de chantier.

Les données améliorées fournies par le dispositif CERCLES peuvent ainsi représenter une alternative à une base de connaissance commerciale et un

<sup>224</sup>Le dispositif CERCLES [en ligne]. [s. d.]. [Consulté le 13 novembre 2016]. Disponible à l'adresse : [http://documentation.abes.fr/sudoc/manuels/controle\\_bibliographique/cercles/co/Le\\_dispositif\\_CERCLES.html](http://documentation.abes.fr/sudoc/manuels/controle_bibliographique/cercles/co/Le_dispositif_CERCLES.html).

<sup>225</sup>Récapitulatif des corpus dans CERCLES [en ligne]. [s. d.]. [Consulté le 13 novembre 2016]. Disponible à l'adresse : [http://documentation.abes.fr/sudoc/manuels/controle\\_bibliographique/cercles/co/Recapitulatif\\_des\\_corpus\\_dans\\_CERCLES.html](http://documentation.abes.fr/sudoc/manuels/controle_bibliographique/cercles/co/Recapitulatif_des_corpus_dans_CERCLES.html).

<sup>226</sup>CERCLES : 4 nouveaux corpus. Dans : *Punktokomo* ; [en ligne]. 17 juillet 2015. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2015/07/17/cercles-4-nouveaux-corpus/>.

<sup>227</sup>CERCLES : Retour d'expérience de La Sorbonne & l'École Française d'Athènes. Dans : *Punktokomo* ; [en ligne]. 15 janvier 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/01/15/cercles-retour-dexperience-de-la-sorbonne-lecole-francaise-dathenes/> ; CERCLES retour d'expérience du SCD de l'Université de Picardie Jules Verne sur le corpus CAIRN. Dans : *Punktokomo* ; [en ligne]. 7 septembre 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/09/07/cercles-retour-dexperience-du-scd-de-luniversite-de-picardie-jules-verne-sur-le-corpus-cairn/>. PASTOR MARTÍNEZ, Nuria. *Un cercle vertueux : E-books Licence Nationale Springer* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <http://www.abes.fr/Publications-Evenements/Journees-ABES/Journees-ABES-2016-10-11-mai-2016>.

complément aux outils de découverte pour les données avec une granularité de niveau titre. Il arrive que des bibliothèques qui ont un outil de découverte, s'investissent aussi dans ce travail d'amélioration de la qualité des métadonnées dans le SUDOC dans le cadre de CERCLES, tout en conservant des outils commerciaux. Ainsi, la BIB01 travaille à importer dans sa base de connaissance commerciale les métadonnées enrichies du SUDOC grâce à CERCLES, afin de pouvoir intégrer ce travail d'amélioration effectué par le réseau dans son outil de découverte. L'enjeu est bien à terme de disposer de plus en plus de métadonnées de qualité, produites et enrichies par un travail collectif de curation des données.

### 3. DES INDEX DE DECOUVERTE A LA CURATION DES DONNEES

#### A. Des bonnes pratiques à l'*Open Discovery Initiative*

Du côté des outils commerciaux, plusieurs initiatives, issues à la fois des professionnels des bibliothèques et des fournisseurs commerciaux, se sont succédées pour aboutir à l'*Open Discovery Initiative* qui plaide pour améliorer la qualité et la disponibilité des données pour les outils de découverte.

##### *Les recommandations de bonnes pratiques*

Dès 2010, à la suite de l'apparition sur le marché des premiers index centralisés, l'*International Coalition of Library Consortia* (ICOLC) avait alerté sur la nécessité pour les fournisseurs de données de rendre disponibles leurs données pour les outils de découverte, de la manière la plus large et la plus ouverte possible :

« Nous incitons les éditeurs à autoriser la diffusion de leur contenu à travers de nombreux fournisseurs adaptés à leurs thématiques. Nous incitons aussi les diffuseurs et les agrégateurs à permettre que leurs métadonnées soient incluses dans les briques de découverte en développement, sur un mode non-exclusif. Des plateformes à accès multiples permettront aux bibliothèques et aux consortiums de sélectionner les ressources et les outils de découverte qui seront adaptés et viables pour leurs membres. »<sup>228</sup>

Un peu plus tard, en 2011, le *Joint Information Systems Committee* (JISC) a lancé au Royaume-Uni une *discovery initiative* afin d'améliorer les pratiques d'échanges de métadonnées, en énonçant des principes fondamentaux pour favoriser l'ouverture des données pour les outils de découverte<sup>229</sup>. Cette initiative a cessé son activité après 2012.

En 2013, c'est la NFAIS qui a formulé des recommandations de bonnes pratiques<sup>230</sup> sur la manière dont les outils gèrent les données des fournisseurs de bases

<sup>228</sup> « We encourage publishers to allow their content to be made available through numerous vendors appropriate for their subject matter. We also encourage online providers and aggregators to allow their metadata to be included in emerging discovery layer services on a non-exclusive basis. Multiple access platforms will permit libraries and consortia to select content and discovery tools that are suitable and affordable for their constituents. » ICOLC. *Revised Statement on the Global Economic Crisis and Its Impact on Consortial Licenses* | ICOLC Website [en ligne]. 14 juin 2010. [Consulté le 30 septembre 2016]. Disponible à l'adresse : <http://icolc.net/statement/revised-statement-global-economic-crisis-and-its-impact-consortial-licenses>.

<sup>229</sup> *Discovery / Home* [en ligne]. [s. d.]. [Consulté le 20 mars 2016]. Disponible à l'adresse : <http://discovery.ac.uk/>.

<sup>230</sup> National Federation of Abstracting and Information Services. *Op. cit.*

d'A&I. Sans surprise, cette initiative provient d'un secteur commercial très touché par le développement des services de découverte à index général centralisé. La NFAIS a établi une liste des droits et obligations à l'intention des fournisseurs de données, de fournisseurs d'outils de découverte, des abonnés et des utilisateurs, afin de satisfaire les besoins de tous les acteurs. Ces recommandations doivent servir à maintenir l'équilibre entre toutes les parties prenantes du processus, afin que chacune comprenne comment les différents aspects du problème touchent les uns et les autres, notamment les enjeux économiques du marché de fourniture de données<sup>231</sup>.

Concernant les données indexées, pour la NFAIS, l'acceptation de ces bonnes pratiques devrait permettre de faciliter l'ouverture et la transparence de la chaîne de données au bénéfice des utilisateurs<sup>232</sup>. Toutefois, les recommandations de la NFAIS insistent sur l'importance pour le fournisseur de données de pouvoir limiter l'accès au contenu exposé dans l'outil de découverte aux seuls utilisateurs autorisés, ainsi que sur l'identification de la propriété du contenu dans l'outil de découverte<sup>233</sup>. Ainsi, pour la NFAIS, le fournisseur de données doit avoir une visibilité sur les contenus exposés dans l'outil de découverte et sur la façon dont cette exposition est réalisée<sup>234</sup>.

De même, d'après la NFAIS, les fournisseurs d'outils de découverte doivent fournir une description complète et détaillée du contenu de leur index<sup>235</sup>. La mise à jour des données et la rectification font également partie des points sur lesquels la NFAIS a attiré l'attention : la fourniture régulière et ponctuelle de l'information doit correspondre aux accords contractuels existants, entre les fournisseurs de contenus et les fournisseurs d'outils. Les fournisseurs de données ou de plateformes ont l'obligation de fournir ces mises à jour ponctuelles et régulières aux fournisseurs d'outils et les fournisseurs d'outils doivent intégrer ces mises à jour pour les bibliothèques<sup>236</sup>.

En ce qui concerne l'algorithme de pertinence, d'après la NFAIS, le fournisseur d'outil de découverte doit donner une description du fonctionnement de l'algorithme de pertinence qui doit être claire et explicite à la fois pour les bibliothèques et pour les fournisseurs de données. À ce propos, la recommandation de la NFAIS exclut toute description vague ou simplement commerciale : « Les abonnés, les utilisateurs et les fournisseurs de contenus ont le droit de disposer d'une description claire et détaillée de l'algorithme de pertinence utilisé par le service de découverte afin de pouvoir comprendre, les cas échéant, quel type d'information est priorisé. Le service de découverte a l'obligation de fournir une telle description. »<sup>237</sup>

<sup>231</sup> « It is the intent of NFAIS that these Recommended Practices will contribute to an awareness of the issues surrounding Discovery Service relationships and that it will create an understanding of how these issues affect each participant in the relationship ». *Ibid.*, p. 1.

<sup>232</sup> « Acceptance of these Recommended Practices will help facilitate full disclosure and transparency for the ultimate benefit of information seekers ». *Ibid.*, p. 1.

<sup>233</sup> « A Content Owner has the right to restrict access to only that content that is either purchased or licensed by an institution and a Discovery Service has the obligation to use reasonable efforts to prevent access by unauthorized users » *Ibid.* p. 6-7. Et « Content Owners have a right to a reasonable display of brand identification. » *Ibid.*, p. 7.

<sup>234</sup> « A Content Owner has the right and obligation to audit how their content is displayed in a Discovery Service » *Ibid.*, p. 8.

<sup>235</sup> « Discovery services have an obligation to provide full disclosure on the content of their index » *Ibid.*, p.7.

<sup>236</sup> « The regular delivery of and timely access to information must be in accord with the specific, predetermined agreements that exist between participants in a Discovery Service arrangement. It is the obligation of both the Content Owner and/or the Platform to provide such timely access to the Discovery Service provider. It is the right of the Content Owner, the Platform, the Subscribers and Users to expect that the Discovery Service will perform the updates in accordance with their predetermined agreements. » *Ibid.*, p. 8.

<sup>237</sup> « Subscribers, users, and the content owner have a right to a clear, high-level description of the ranking algorithm used by the Discovery Service in order to understand the type of information, if any, that is given a priority. The Discovery Service has an obligation to provide such a description ». *Ibid.*, p.7.

La NFAIS a également précisé plusieurs points concernant les statistiques, les formations pour les administrateurs et les données personnelles : pour la NFAIS, les fournisseurs d'outils doivent fournir des rapports d'utilisation standard, normalisés et réguliers ; ils doivent effectuer des formations pour les administrateurs et doivent avertir les fournisseurs de données et les administrateurs d'outils de tout changement majeur ; ils doivent veiller aussi à la confidentialité des données personnelles des utilisateurs.

Conjointement à la NFAIS, un groupe de travail de l'OCLC sur la qualité des données (*E-Data Quality Working Group*) a également formulé des recommandations dans un livre blanc sur les stratégies à adopter pour la diffusion des contenus numériques et l'accès aux données<sup>238</sup>. Dans ce document, le groupe de travail de l'OCLC formule des recommandations qui devraient s'appliquer à toute la chaîne de fourniture de données afin de garantir qu'elles soient exploitables et accessibles. Pour le groupe de l'OCLC, toutes les parties engagées dans la chaîne de fourniture de données doivent mettre en place les structures et les outils nécessaires pour permettre aux usagers de la bibliothèque d'accéder de manière efficace au contenu.

Le groupe de travail de l'OCLC reconnaît une responsabilité partagée entre fournisseurs de données, fournisseurs d'outils et bibliothèques pour améliorer la qualité des données<sup>239</sup>. Mais elle se concentre sur les recommandations à l'usage des bibliothèques pour traiter leurs données : elle recommande notamment l'utilisation d'identifiants électroniques standard (ISSN électronique, ISBN électronique, DOI), des formats de données normalisés (en particulier MARC et KBART), et l'utilisation d'un fichier de suivi des mises à jour des données.

### *L'Open Discovery Initiative*

En s'appuyant sur ces travaux précédents, et notamment sur les recommandations de la NFAIS, la NISO a lancé en 2014 un programme de travail sur l'*Open Discovery Initiative* (ODI) afin de promouvoir la transparence des outils de découverte. À l'issue de ce travail, la NISO a produit un document de bonnes pratiques à destination des fournisseurs de données et des fournisseurs d'outils<sup>240</sup>. À terme, l'objectif de l'ODI est de favoriser « un dialogue constructif entre les fournisseurs d'outils, les fournisseurs de données et les bibliothèques<sup>241</sup> ».

Pour la NISO, le constat de départ a été que le contenu des index centraux était rarement connu des bibliothèques avec le niveau de détail qu'elles souhaitaient : « l'impression d'un manque de transparence sur ces flux de données a provoqué le besoin de formuler clairement ces enjeux et de clarifier quelles métadonnées ou éléments de données sont rendus disponibles à quels acteurs et sous quelles conditions<sup>242</sup> ». Ce constat joint à d'autres interrogations quant à la neutralité, à

<sup>238</sup> E-Data Quality Working Group OCLC. *Success strategies for electronic content and access: a cross-industry white paper*. Dublin, OH : OCLC, 2014. [Consulté le 20 mars 2016]. Disponible à l'adresse : <http://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>.

<sup>239</sup> « We recognize that all of us, as participants in the content supply chain, have a shared interest in improving content discovery and access for library users through better quality bibliographic metadata and holdings data. We also recognize that we have a shared responsibility to improve the quality of the data exchanged and to implement more effective data exchange workflows. » *Ibid.*, p. 2.

<sup>240</sup> National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*

<sup>241</sup> « A need for more transparency and educational dialog between discovery providers, content providers, and libraries. » *Ibid.*, p. 14.

<sup>242</sup> « A perceived lack of transparency across these data flows prompted the need to develop definitions of the data points and to clarify what metadata or data elements are made available to which parties and under which conditions. » *Ibid.*, p. 3.

l'opacité et à l'exclusivité des accords commerciaux, rendaient indispensable, d'après l'ODI, de proposer des bonnes pratiques qui devraient à terme améliorer la transparence du circuit de données, réaffirmer la neutralité des outils et garantir l'accès au contenu de la façon la plus large et précise possible.

L'ODI s'intéresse à tous les types de données qui peuvent concerner les collections numériques des bibliothèques, sans distinction en fonction du modèle économique de l'acquisition ou du type de licence (commerciale ou *open access*).

### Bonnes pratiques pour les fournisseurs de données :

L'ODI apporte aux fournisseurs de données des indications techniques sur les données à fournir. Ces recommandations distinguent deux niveaux de données : *core metadata* (métadonnées fondamentales) et *enriched content* (contenu enrichi : texte intégral, résumé, indexation sujet). Pour l'ODI, les fournisseurs de données devraient, dans tous les cas, rendre disponibles pour les fournisseurs d'outils au minimum les métadonnées fondamentales et, à chaque fois que cela sera possible, le contenu enrichi.

Pour l'ODI, les champs nécessaires pour les métadonnées fondamentales sont<sup>243</sup> :

Champ	Définition
Titre de l'élément	Le titre principal de l'élément
Auteurs	L'(les) auteur(s) de l'élément. Les auteurs personnes physiques devront figurer dans l'ordre : Nom, Prénom
Editeur	Le nom de l'éditeur de la source
Volume	Le volume de la source, si applicable
Numéro	Le numéro de la source, si applicable
Pages	La pagination, si applicable
Date ou période couverte	La date de publication. Pour une publication en série, les intervalles de dates de la période couverte
Identifiant	Un ou plusieurs identifiants standards pour la version imprimée ou électronique de l'élément, par exemple : l'ISSN, le numéro d'identification OCLC, l'ISBN, le DOI etc. L'identifiant devra être précédé d'un libellé indiquant le type d'identifiant
Titre de la source	Le titre de la publication (source) d'où est extrait l'élément, par exemple : pour des articles de revues, le titre de la revue ; pour des chapitres de livre, le titre du livre
Identifiant de la source	L'identifiant standard de la source tel que défini plus haut, par exemple : l'ISSN, le numéro d'identification OCLC, l'ISBN, le DOI etc. L'identifiant devra être précédé d'un libellé indiquant le type d'identifiant
URL de l'élément	Soit un OpenURL soit un lien direct vers le texte intégral de l'élément

<sup>243</sup> *Ibid.*, p. 14 et suivantes.



Référent <i>Open access</i>	Conformément aux recommandations de la NISO <i>Open access</i> Metadata and Indicators (OAMI), si un élément est en <i>open access</i> , ce statut devra être indiqué avec la mention « free_to_read ». Dans le cas contraire il devra être laissé vide. Voir les recommandations à ce sujet sur <a href="http://www.niso.org/workrooms/oami/">www.niso.org/workrooms/oami/</a>
Indicateur de texte intégral	Champ oui/non pour indiquer si le fournisseur de données rend l'élément disponible en texte intégral (ou s'il s'agit de contenus non textuels, la version intégrale ou à grande résolution) afin qu'il soit indexé par le fournisseur d'outil de découverte. On peut espérer que cet indicateur soit à l'avenir exposé par l'outil de découverte aux bibliothèques dans la description de la couverture d'indexation du contenu d'un titre ou d'une collection
Type de document	Champ conseillé pour décrire le type de document d'après la liste normalisée MARC 21 06-Type of Record position <sup>244</sup>
Type de source	Champ conseillé pour décrire le type de source d'après la liste normalisée MARC 21 07-Bibliographic Level position <sup>245</sup>

Pour le contenu enrichi, l'ODI énumère les champs suivants :

Champ	Définition
Données d'indexation sujet	Un ou plusieurs mots clés (issus de vocabulaires libres ou contrôlés) qui décrivent le contenu de l'élément
Texte intégral ou transcription	Pour du contenu textuel, l'intégralité du contenu de l'élément. Pour du contenu audio ou vidéo, la transcription intégrale du contenu sonore de l'élément. N'est pas forcément pertinent pour tout le contenu indexé
Résumé / description	Un résumé du contenu ou, pour les éléments non textuels, une description de l'élément

L'ODI note néanmoins la très grande hétérogénéité que l'on peut avoir dans les données d'indexation sujet selon les fournisseurs de données et la source concernée.

Quant au format des métadonnées livrées et à la méthode de transfert, l'ODI recommande aux fournisseurs des données d'utiliser les standards existants et interopérables, comme la recommandation KBART, le protocole OAI-PMH ou ResourceSync.

Afin de favoriser la transparence sur les accords commerciaux, l'ODI préconise que les fournisseurs de données rendent disponibles pour les bibliothèques abonnées à

<sup>244</sup> Voir les détails sur : *MARC 21 Format for Bibliographic Data: lead: Leader (Network Development and MARC Standards Office, Library of Congress)* [en ligne]. [s. d.]. [Consulté le 10 septembre 2016]. Disponible à l'adresse : <https://www.loc.gov/marc/bibliographic/bdleader.html>.

<sup>245</sup> *Ibid.*

leurs ressources, les informations suivantes concernant leur niveau de participation dans les outils de découverte :

- la couverture des données fournies pour l'indexation (pour les périodiques l'intervalle des dates couvertes ; pour les autres ressources s'il s'agit de l'intégralité ou d'une partie) ;
- le niveau de données fournies pour l'indexation (métadonnées, indexation sujet, résumé, texte intégral) ;
- leur statut en tant que fournisseur de données (éditeur, agrégateur, base d'A&I, entrepôt en *open access*) ;
- le destinataire des données fournies : s'il s'agit de données rendues disponibles pour tous les outils de découverte ou seulement pour certains d'entre eux via des accords commerciaux. Dans ce cas, le fournisseur de données devra fournir la liste des outils de découverte concernés par ces accords.

Si l'ODI prend acte de l'existence d'accords commerciaux limitant l'accès à certains contenus uniquement aux abonnés (*non disclosure agreements*), elle signale que ces clauses ne devraient pas être utilisées pour empêcher la transparence et l'ouverture des données<sup>246</sup>. Cela ne permet pas de résoudre le problème des bases d'A&I dont la valeur ajoutée réside dans leurs métadonnées mêmes : l'absence de clause limitative quant à l'accès se traduirait-elle, dans ce cas, par un appauvrissement des données indexées dans l'outil de découverte ou par une indexation à plusieurs niveaux (une indexation de base accessible sans authentification et une indexation complète pour les usagers authentifiés) ? Le problème demeure d'actualité.

Pour les fournisseurs de données en *open access*, la NISO a formulé une recommandation complémentaire à l'ODI<sup>247</sup>. Pour améliorer l'accès aux données des ressources en *open access* la NISO préconise l'utilisation de deux nouveaux éléments de métadonnées XML : *free\_to\_read* et *license\_reference*. Le premier sert à indiquer que le contenu peut être consulté par tout utilisateur gratuitement et peut éventuellement être complété par des dates de début et de fin lorsque des barrières temporelles s'appliquent. Le second doit être utilisé pour renvoyer vers la licence ou la page qui explique les conditions d'accès à la ressource, de préférence via une URI, et peut également contenir des précisions de dates lorsque des licences ont changé avec le temps. La combinaison de ces deux éléments doit permettre, selon la NISO, de répondre à la plupart des cas de ressources en *open access*, en respectant les politiques des archives institutionnelles ainsi que celles des éditeurs. Marshall Breeding espère qu'une fois que cette recommandation sera adoptée par les éditeurs proposant des ressources en *open access*, les outils de découverte, pourront mieux identifier ces ressources et proposer un accès plus efficace dans tous les cas, indépendamment de l'existence d'un abonnement<sup>248</sup>.

<sup>246</sup> « We recognize that some discovery service providers and/or content providers may wish to incorporate non-disclosure agreements (NDAs) and similar provisions into the contracts and other agreements that they enter into with one another. But NDAs should not be used to avoid the disclosures and transparency recommended in Section 3.3.2, and elsewhere in this document. » National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 22.

<sup>247</sup> National Information Standards Organization. *Access and License Indicators*. Baltimore, MD : National Information Standards Organization (NISO), 2015. [Consulté le 3 octobre 2016]. Disponible à l'adresse : <http://www.niso.org/workrooms/ali/>, p. 5-6.

<sup>248</sup> BREEDING, Marshall et National Information Standards Organization. *Op. cit.*, p. 29.

### Bonnes pratiques pour les fournisseurs d'outils :

L'ODI recommande aux fournisseurs d'outils de donner les informations les plus complètes possibles aux bibliothèques. L'ODI reconnaît l'importance des métadonnées au niveau de l'élément qui constitue l'unité documentaire mais préconise, à chaque fois que cela pourra être pertinent, l'adoption d'une granularité de niveau titre de la source pour lequel elle recommande les éléments suivants<sup>249</sup> :

Champ	Définition
Titre de la publication	Le titre principal
Date ou période couverte	La date de publication. Pour une publication en série, les intervalles de dates de la période couverte
Identifiant	Un ou plusieurs identifiants standards pour la version imprimée ou électronique de l'élément, par exemple : l'ISSN, le numéro d'identification OCLC, l'ISBN, le DOI etc. L'identifiant devra être précédé d'un libellé indiquant le type d'identifiant
Type de document	Conseillé pour décrire le type de document d'après la liste normalisée MARC 21 06-Type of Record position <sup>250</sup>
Type de source	Conseillé pour décrire le type de source d'après la liste normalisée MARC 21 07-Bibliographic Level position <sup>251</sup>
Contenu indexé	Description du contenu indexé : métadonnées, indexation sujet, résumé/description, texte intégral

Pour l'ODI, l'application de ces recommandations devrait permettre d'améliorer la gestion de la granularité de niveau titre pour les revues et les livres numériques.

En complément l'ODI fait également un certain nombre de préconisations sur les aspects techniques suivants :

- l'encodage des caractères utilisé, notamment l'UTF8 ;
- le nommage des fichiers fournis de manière standardisée, par exemple : FournisseurOutil\_AAAA-MM-DD ;
- l'utilisation de schémas de métadonnées normalisés, notamment le MARC, MODS, METS, Dublin Core, KBART, EAD ;
- les extensions de fichier utilisées pour la fourniture de données, notamment csv, txt, mrc ou xml ;
- les méthodes de transfert : OAI PMH, ResourceSync, API, FTP.

<sup>249</sup> National Information Standards Organization et Open Discovery Initiative Working Group. *Op. cit.*, p. 23 et suivantes.

<sup>250</sup> MARC 21 Format for Bibliographic Data: lead: Leader (Network Development and MARC Standards Office, Library of Congress) [en ligne]. [s. d.]. [Consulté le 10 septembre 2016]. Disponible à l'adresse : <https://www.loc.gov/marc/bibliographic/bdleader.html>.

<sup>251</sup> *Ibid.*

L'ODI préconise en outre que les fournisseurs d'outils informent les bibliothèques et les fournisseurs de données de tout changement dans leur infrastructure qui puisse affecter la récupération et l'affichage des résultats, l'algorithme de pertinence ou l'ordre de présentation des liens. D'après l'ODI, les fournisseurs d'outils devraient clairement décrire leurs capacités techniques, leurs limites et leurs préférences quant aux formats de versement de données. Ils devraient aussi informer le fournisseur de données à chaque fois qu'une modification de leurs formats, de leurs schémas de métadonnées ou de leurs mécanismes de transferts est susceptible d'avoir un impact sur le fonctionnement de l'outil.

L'ODI énonce aussi des bonnes pratiques sur le fonctionnement global des outils et sur le circuit des métadonnées, afin de répondre à des points de préoccupation soulevés par les bibliothèques : les accords commerciaux, la neutralité, la fréquence de mise à jour des données indexées et les statistiques.

Parallèlement aux fournisseurs de données, les fournisseurs d'outils devraient mettre à disposition des bibliothèques un état des lieux annuel, expliquant leurs liens commerciaux avec les fournisseurs de données, y compris ceux avec lesquels ils ont des relations de dépendance et ceux avec lesquels ils ont négocié des accords commerciaux pour la fourniture de métadonnées, la fourniture de liens directs ou d'autres accords spécifiques.

Concernant la neutralité, pour l'ODI, les fournisseurs d'outils devraient offrir une preuve tangible de la neutralité de leurs algorithmes utilisés pour générer les résultats, le classement de pertinence et l'ordre des liens, que ce soit vis-à-vis de l'orientation politique ou idéologique de la ressource, du fournisseur de données (en particulier pour ses propres contenus ou ceux d'une société parente) ou tout autre facteur d'importance. Les outils de découverte devraient s'abstenir de privilégier, en fonction de leurs accords commerciaux, certains fournisseurs de contenus (particulièrement les leurs) dans l'algorithme de pertinence, la récupération et l'affichage des résultats, et l'ordre de présentation des liens si plusieurs sont disponibles (ce que résume l'expression *fair linking*). Les outils de découverte devraient permettre aux bibliothèques de choisir vers quelles plateformes elles souhaitent renvoyer les utilisateurs et dans quel ordre de priorité elles souhaitent les présenter aux usagers. La gestion des liens devrait être configurable par chaque bibliothèque.

Sur la fréquence des mises à jour des métadonnées dans l'index, d'après l'ODI, les bibliothèques devraient recevoir des mises à jour régulières de leurs métadonnées, au moins trimestrielles, ou à la demande.

L'ODI fournit également des recommandations pour les statistiques d'usage que les outils devraient notamment être en mesure de fournir de manière mensuelle : le nombre total de recherches dans l'outil de découverte, les recherches réparties par base de données et source, le nombre de clics sur des résultats et si possible par source, le nombre de clics d'accès via un OpenURL ou un lien direct, le nombre de visiteurs uniques, le nombre de clics d'accès au texte intégral.

L'ODI recommande qu'on intègre une URL *referrer* (URL d'origine) pour identifier les recherches provenant de l'outil de découverte. Etant donné l'attente des usagers vis-à-vis du texte intégral, l'ODI estime que la mesure du nombre de clics d'accès au texte intégral devrait permettre aux bibliothèques d'avoir une meilleure mesure d'usage pour évaluer la valeur d'un contenu plutôt que les recherches sur des plateforme natives.

En complément de ces bonnes pratiques, l'ODI a préconisé la création d'un groupe de travail pour superviser l'adoption de ces bonnes pratiques, fournir de

l'assistance aux fournisseurs de données et aux fournisseurs d'outils, mettre en place un forum de discussion sur le sujet pour toutes les parties prenantes (fournisseurs de données, fournisseurs d'outils, bibliothèques).

## B. Quid d'un index de découverte en *open access* ?

Si l'offre de services de découverte à index centralisé reste dominée par les fournisseurs commerciaux, l'existence d'interfaces de découverte libres et la croissance des entrepôts en *open access*, permettent de formuler l'hypothèse d'un outil de découverte libre fonctionnant sur un index en *open access* et interrogeable à travers une interface de découverte libre. Un index centralisé de ce type pourrait permettre aux bibliothèques de disposer d'informations plus transparentes sur les flux de données et sur les caractéristiques de l'indexation.

Dans son rapport pour la NISO, Marshall Breeding a exploré la faisabilité de création d'un tel index centralisé en *open access*<sup>252</sup>. Cela exigerait, d'après cet auteur, une infrastructure capable de gérer l'indexation d'un ou deux milliards de références, y compris une grande partie en texte intégral, et des moyens considérables :

- une plateforme technologique de qualité à grande échelle ;
- un large programme de relations avec les fournisseurs de données pour négocier et appliquer des accords relatifs au versement de données dans l'index central ;
- le développement de logiciels pour les interfaces, l'indexation, l'algorithme de pertinence et d'autres aspects techniques.

Cette perspective se heurterait, pour Marshall Breeding, à de nombreux obstacles, dont le premier serait celui des ressources et des moyens disponibles : les besoins en financement pour mettre sur pied et maintenir un index de découverte en *open access* seraient très importants. La création d'un index de découverte en *open access* nécessiterait un financement, du personnel et des ressources techniques au moins au niveau d'un fournisseur d'outil de découverte commercial : à ce propos, Marshall Breeding estime que pour un fournisseur commercial comme Ebsco il y a plus de 300 personnes qui travaillent seulement sur l'outil de découverte<sup>253</sup>. D'importantes ressources semblent donc nécessaires pour créer et maintenir un outil de découverte à index centralisé. S'il reconnaît que cet outil de découverte pourrait être cogéré par des bibliothèques à travers un (ou des) consortium(s) au bénéfice des bibliothèques participantes, Marshall Breeding se demande dans quelle mesure la mutualisation et le financement collaboratif entre les bibliothèques intéressées pourraient suffire à développer cet index. Il s'interroge de même sur la capacité à faire baisser le seuil de dépenses minimales et la complexité des processus économiques qui jusqu'ici ont limité le développement du marché des services de découverte à index général centralisé à un petit nombre de fournisseurs commerciaux. Toutefois, il avance qu'avec un tel outil, la communauté de bibliothèques à travers le monde serait moins dispersée et plus à même de négocier, à terme, l'accès au contenu propriétaire.

Le deuxième ensemble de problèmes concerne, pour Marshall Breeding, la propriété intellectuelle et le modèle économique touchant l'index de découverte en *open access* et les contenus sous droits. En effet, une grande part des contenus restent

<sup>252</sup> BREEDING, Marshall et National Information Standards Organization *Op. cit.*, p. 15-20.

<sup>253</sup> *Ibid.*, p. 18.

sous la mainmise des éditeurs. Les index de découverte commerciaux incluent du contenu propriétaire, souvent à la condition qu'il ne sera pas exposé publiquement mais réservé aux seuls abonnés. Ce contenu propriétaire peut être le texte intégral, les résumés ou l'indexation des bases de références bibliographique, voire les métadonnées elles-mêmes, même si souvent les métadonnées seules peuvent être traitées plus librement. Dans ce sens, il serait possible d'imaginer un index de découverte en *open access* indexant seulement des références bibliographiques, mais avec un périmètre plus restreint que celui des produits commerciaux. Ainsi, on pourrait postuler, d'après Marshall Breeding, que les mots sujets, les résumés, le texte intégral ne soient pas fournis dans cet index en *open access*, limité par les restrictions qu'imposent les éditeurs et les fournisseurs de données propriétaires. Dans cette configuration, la négociation des données serait complexifiée car elle impliquerait l'ensemble de bibliothèques qui gèreraient l'outil. À l'instar des outils commerciaux qui négocient avec les fournisseurs de données le versement des données dans l'index centralisé, l'index de découverte en *open access* aura lui aussi besoin d'un réseau étendu de participants pour conduire les négociations et coordonner les différents apports des éditeurs. Enfin une question subsiste : si une telle ressource en *open access* existait, est-ce que les index commerciaux y auraient eux-mêmes accès ?

Le troisième ensemble de problèmes concerne les contraintes techniques d'un tel index en *open access*. Pour Marshall Breeding, la création d'un index centralisé en *open access* implique la création d'une plateforme technique pour gérer l'index, l'intégration des données des fournisseurs de données (éditeurs ou agrégateurs) et la maintenance et mise à jour des index. Les composantes nécessaires en seraient :

- une capacité de stockage pour de multiples répliques de l'index ;
- une technologie hautement développée pour l'indexation, la récupération des données et l'affichage des résultats ;
- un logiciel pour traiter les données versées par les fournisseurs de données pour l'indexation (normalisation, détection d'erreurs, correction, stockage) ;
- le développement d'une API pour permettre l'interrogation depuis une interface de découverte ;
- un logiciel de gestion des identités pour permettre des accès authentifiés.

La création d'une plateforme technique comme celle-ci impliquerait le développement d'un logiciel pour charger et indexer les données, et afficher les résultats : Marshall Breeding signale que si des composantes *open source* sont disponibles pour l'infrastructure, il resterait beaucoup de développements nécessaires pour en faire un système fonctionnel. Techniquement, la récupération des métadonnées dans l'index en *open access* pourrait se réaliser à travers un processus de transfert de contenus semblable à ceux qui ont lieu entre les éditeurs et les outils commerciaux à index centralisé ou via un moissonnage via le protocole OAI-PMH ou ResourceSync. Il serait nécessaire pour cela que les fournisseurs de données exposent ces dernières dans un environnement technique interopérable avec ce type de moissonnages, or ce n'est pas toujours le cas. D'après Marshall Breeding, les méthodes de transfert et d'alimentation d'un index de découverte en *open access* resteraient donc à consolider.

### C. Des données de qualité pour les outils de découverte

De manière complémentaire à ces initiatives, plusieurs instances œuvrent pour permettre d'améliorer la qualité des données disponibles dans les outils commerciaux, notamment Couperin, l'ABES et ISTEEX.

## Couperin

Regroupant plus de 240 établissements d'enseignement supérieur et de recherche, le consortium Couperin a pour mission de négocier et d'organiser l'achat de ressources documentaires numériques pour les établissements participants. Si la négociation des métadonnées ne figure pas explicitement parmi ses missions<sup>254</sup>, cela n'empêche que, dans le cadre de son activité, Couperin tente d'obtenir des éditeurs des métadonnées structurées et complètes, et de faire en sorte que celles-ci puissent être disponibles pour des outils de signalement libres ou commerciaux. Ainsi, parmi les éléments énumérés dans la lettre de cadrage destinée à guider les négociateurs, se trouve clairement mentionnée la fourniture de métadonnées<sup>255</sup>. Dans le même sens, on peut signaler la liste des « 10 commandements à l'éditeur d'e-books : pour une offre idéale »<sup>256</sup> qui a été élaborée par le consortium et qui contient les recommandations suivantes :

- la fourniture des notices aux formats MARC ;
- l'utilisation des protocoles d'échanges OAI-PMH, SRU/SRW, ainsi que des API ou *web services* ;
- la mise à disposition des données de signalement pour les catalogues, ERMS, outils de découverte et notamment « l'export systématique des métadonnées des collections vers les index des Discovery tools »<sup>257</sup> ;
- le respect de la recommandation KBART.

Depuis 2013, Couperin met en œuvre également une stratégie de labellisation selon trois niveaux (« sans label », « argent » et « or »), afin de reconnaître la qualité des offres qui se conforment le plus aux critères énoncés par le consortium. Toutefois, les critères concernant la fourniture de métadonnées ne sont, pour l'heure, pas bloquants pour l'aboutissement d'une négociation ni même pour l'attribution d'un label de type « argent »<sup>258</sup>. En d'autres termes, une offre pourra être conclue et diffusée, éventuellement labellisée, sans que la fourniture de métadonnées soit entièrement garantie de manière satisfaisante pour la ressource concernée. À cet égard, le retour d'expérience du SCD de l'université de Strasbourg<sup>259</sup> sur le signalement des livres numériques Dalloz a récemment rappelé que ces recommandations en matière de fourniture de métadonnées sont encore

<sup>254</sup> COUPERIN. *Présentation - Missions* [en ligne]. [s. d.]. [Consulté le 29 novembre 2016]. Disponible à l'adresse : <http://www.couperin.org/presentation>.

<sup>255</sup> « L'éditeur accepte de fournir des métadonnées structurées, ouvertes et documentées (ex. en XML) aux fournisseurs de bases de connaissances, outils de découverte et de recherche fédérée ». COUPERIN. *Lettre de cadrage* [en ligne]. [s. d.]. [Consulté le 7 décembre 2016]. Disponible à l'adresse : <http://www.couperin.org/boite-a-outils/item/935-lettre-de-cadrage>.

<sup>256</sup> *Les 10 commandements - Relations éditeurs* [en ligne]. [s. d.]. [Consulté le 29 novembre 2016]. Disponible à l'adresse : <http://www.couperin.org/relations-editeurs>.

<sup>257</sup> *Ibid.*

<sup>258</sup> La fourniture gratuite de données MARC et la fourniture des métadonnées pour l'alimentation des bases de connaissances sont des critères conditionnels et non obligatoire pour le label « argent ». COUPERIN. *Boîte à outils - Labellisation* [en ligne]. [s. d.]. [Consulté le 7 décembre 2016]. Disponible à l'adresse : <http://www.couperin.org/boite-a-outils/item/1070-labellisation>.

<sup>259</sup> Récit d'une immersion. Traiter les ebooks Dalloz avec les données Sudoc, les données de l'éditeur et les outils du hub. Dans : *Punktokomo* ; [en ligne]. 22 juin 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/06/22/recit-dune-immersion-traiter-les-ebooks-dalloz-avec-les-donnees-sudoc-les-donnees-de-lediteur-et-les-outils-du-hub/>. Voir aussi STORNE, Catherine. *L'analyse qualitative des métadonnées et les traitements, sans collaboration de l'éditeur* [en ligne]. Montpellier, mai 2016. Disponible à l'adresse : <http://www.abes.fr/Publications-Evenements/Journees-ABES/Journees-ABES-10-11-mai-2016>.

loin d'être appliquées par tous les fournisseurs. Dans ce cas, il n'a pas été possible pour la bibliothèque d'obtenir de l'éditeur un fichier contenant les métadonnées indispensables pour un signalement *a minima* : titre, auteur, ISBN et URL d'accès. En effet, la liste de titres fournie par l'éditeur sous forme de tableau ne contenait pas l'URL ni l'auteur, tandis que l'entrepôt OAI de l'éditeur contenait certes des notices plus riches mais était incomplet en nombre<sup>260</sup>. Ainsi, en ce qui concerne la fourniture de métadonnées par les éditeurs, d'importants progrès sont encore attendus, auxquels contribue aussi l'ABES à travers son projet de base de connaissance nationale.

### **BACON**

La Base de Connaissance Nationale (BACON) mise en place par l'ABES a pour objectif de fournir des descriptifs de bouquets de périodiques ou de livres numériques, avec une granularité de niveau titre, sous licence libre<sup>261</sup>. Elle doit permettre de simplifier le travail des chargés de ressources numériques en établissement, en proposant des listes détaillées et à jour décrivant le contenu de ressources souvent mal représentées ou sous-représentées dans les bases de connaissance commerciales et les outils de découverte (en particulier les bouquets négociés par Couperin et les ressources francophones). À terme, BACON permettra d'améliorer la visibilité des données des éditeurs francophones dans les outils commerciaux, en les portant au même niveau de détail et de complétude que celles des éditeurs anglo-saxons.

Le projet de BACON consiste à inciter les éditeurs à produire des données de description de leurs bouquets qui soient fiables et structurées, afin d'améliorer les métadonnées à leur source et de construire avec eux une chaîne de confiance. Les éditeurs étant les seuls à même de procéder aux mises à jour de leurs données, l'ABES n'a pas pour finalité de produire les données à leur place mais de les accompagner dans le travail de consolidation et de rectification de ces données. En amont, l'ABES accompagne les éditeurs dans la production de leurs fichiers de données, en menant dès cette étape un contrôle sémantique sur ces derniers. Les éditeurs envoient ensuite ces fichiers de données à l'équipe de BACON qui effectue dans un premier temps un contrôle syntaxique, comportant des critères bloquants<sup>262</sup>. Si les critères bloquants sont validés, le fichier est intégré sur le site de BACON<sup>263</sup> et, à l'issue d'un contrôle de conformité plus poussé comportant une vérification sémantique, le bouquet pourra le cas échéant être labellisé « BACON », le label étant le gage de qualité de ces métadonnées.

Pour effectuer ce travail de vérification et de consolidation des données, BACON s'appuie sur la recommandation KBART (Knowledge Bases and Related Tools)<sup>264</sup>. Le format KBART est un format tabulé, structuré en 25 champs

---

<sup>260</sup> *Ibid.*

<sup>261</sup> Pour cette partie, nous nous appuyons notamment sur l'entretien téléphonique réalisé avec Cyril Leroy de l'équipe BACON de l'ABES le 04 juillet 2016.

<sup>262</sup> Voir à ce sujet : BACON et la labellisation des données : à quelle aune mesure-t-on la qualité d'un fichier KBART ? Dans : *Punktokomo* ; [en ligne]. 28 mai 2015. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2015/05/28/bacon-et-la-labellisation-des-donnees-a-quelle-aune-mesure-t-on-la-qualite-dun-fichier-kbart/>.

<sup>263</sup> BACON [en ligne]. [s. d.]. [Consulté le 20 novembre 2016]. Disponible à l'adresse : <https://bacon.abes.fr/>.

<sup>264</sup> *Knowledge Base and Related Tools - National Information Standards Organization* [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://www.niso.org/workrooms/kbart>.



comportant un intitulé normalisé, qui sert à décrire le contenu de bouquets de revues ou de livres numériques. Si le format KBART est un format relativement souple et facile à produire pour un éditeur<sup>265</sup>, il permet toutefois moins de finesse qu'une description bibliographique complète. C'est pourquoi l'ajout de deux champs supplémentaires dans les fichiers BACON est actuellement à l'étude par l'ABES : il s'agirait d'un champ indiquant la langue du document et d'un champ contenant l'identifiant correspondant à la ressource dans le SUDOC (le PPN).

Une fois les fichiers KBART, vérifiés, validés et/ou labellisés, ces derniers sont mis à disposition sur la plateforme de BACON sous forme de fichier tabulé et au format Json. Cette mise à disposition s'effectue sous licence libre CC0 (transfert dans le domaine public) : les données sont librement utilisables sans devoir en préciser l'origine. Ce régime d'ouverture des données permet notamment de faciliter les échanges avec d'autres bases de connaissance nationales ouvertes, comme KB+<sup>266</sup>, dont les données sont sous le même type de licence. Il peut être cependant l'objet de malentendus ou d'incompréhensions de la part des fournisseurs de données ou des fournisseurs d'outils, habitués à réfléchir en termes de propriété des données et de contractualisation. C'est pourquoi l'ABES travaille avec les éditeurs et les fournisseurs d'outils pour mieux faire comprendre le rôle et l'intérêt de BACON, afin de favoriser à terme l'ouverture des données le plus largement possible pour le plus grand nombre de ressources.

Si les bases de connaissance commerciales incluent progressivement les données de BACON, l'ABES ne dispose pas de visibilité sur les traitements effectués sur ces données lors de leur intégration à ces outils. Les données de BACON peuvent en effet avoir été retravaillées par le fournisseur d'outil au moment de leur import ou ultérieurement. L'ABES a ainsi constaté des écarts entre les fichiers mis à disposition dans BACON et ceux qui étaient signalés comme issus de BACON, voire labellisés comme tels, dans des bases de connaissance commerciales : même si les fichiers y figuraient comme provenant de BACON, ils présentaient des différences sur des champs clés, tels l'ISSN ou la période couverte. Par ailleurs, l'ABES a également noté qu'il pouvait y avoir un délai non connu entre la mise à disposition des fichiers dans BACON et leur intégration effective dans une base de connaissance commerciale.

En théorie, si la composition d'un bouquet ne figure pas en tant que telle dans une base de connaissance commerciale, il devrait être possible pour les bibliothèques d'y importer les fichiers de BACON. Néanmoins, la très grande majorité des bibliothèques interrogées n'ont pas tenté elles-mêmes cette opération. Nous signalerons ici deux bibliothèques ayant fait cette démarche : BIB01 et BIB19. La BIB01 a pu charger sans inconvénient majeur des fichiers de BACON dans sa base de connaissance commerciale, mais désormais cette opération n'est plus nécessaire car la base de connaissance les intègre directement. Le manque d'indication de langue est toutefois regretté par la bibliothèque, car en l'absence de ce champ, la langue indiquée par défaut est l'anglais, alors qu'il s'agit de ressources francophones, ce qui implique des rectifications après import. La BIB19 a pour sa part tenté de charger dans sa base de connaissance commerciale un fichier disponible sur BACON mais non labellisé : ce fichier, bien que conforme au KBART d'un point de vue structurel, présentait des problèmes sur les champs

<sup>265</sup> BACON et la labellisation des données : à quelle aune mesure-t-on la qualité d'un fichier KBART ? *Op. cit.*

<sup>266</sup> KB+ est la base de connaissance maintenue par le consortium britannique JISC. *Knowledge Base+* [en ligne]. [s. d.]. [Consulté le 6 septembre 2016]. Disponible à l'adresse : <https://www.kbplus.ac.uk/kbplus>.

auteur et titre qui se trouvaient intervertis. La BIB19 n'a pas renouvelé cette expérience, préférant désormais utiliser directement les bouquets BACON signalés dans sa base de connaissance commerciale.

Une autre tentative d'explorer l'intégration de fichiers de BACON dans une base de connaissance commerciale a été réalisée en 2016 par Cyril Leroy de l'équipe BACON, en immersion dans les bibliothèques d'Aix-Marseille Université (AMU). Ce travail a notamment fait l'objet d'un retour aux Journées de l'ABES 2016<sup>267</sup>. Il s'agissait de mettre en place le chargement de fichiers KBART de BACON dans la base de connaissance de l'AMU (360 Core), afin de rendre ces données interrogeables ensuite dans l'index centralisé de l'outil de découverte (Summon). Ce travail a porté sur quatre bouquets – *Cairn*, *Persée*, *OpenEdition*, *Techniques de l'ingénieur* – et il s'est basé sur les données tabulées fournies par BACON, dans l'espoir de pouvoir à terme poser les bases pour une alimentation automatique de l'outil de découverte de l'AMU depuis BACON.

Pour importer des fichiers KBART de BACON dans la base de connaissance de l'AMU, il a été nécessaire de les convertir au format de cette dernière, ce qui a été réalisé grâce à un script sous OpenRefine. Néanmoins, cette conversion a entraîné une perte de la qualité de certaines données car, alors que le format fidèle à la recommandation KBART contient 25 champs, le format de la base de connaissance de l'AMU n'en contenait que 18. Pour transvaser les 25 champs des fichiers KBART de BACON dans ces 18 champs, il a donc été nécessaire de faire des choix d'attribution : pour certaines données il a été possible de maintenir l'information, mais pas pour toutes. Ainsi, sur des champs importants, tels que les identifiants ISBN/ISSN, une seule colonne était disponible dans 360 Core, tandis que le KBART distingue bien les ISBN/ISSN papier et électronique. De même, dans les fichiers de BACON, les métadonnées KBART du volume et du fascicule permettent de compléter les années indiquées en date de début et date de fin d'un état de collection, afin d'avoir une information plus fine quant à la période couverte. Or dans le format de la base de connaissance 360 Core, ces données ne pouvaient pas être intégrées, ce qui se traduisait par une information moins fine au niveau de l'état de collection.

Une fois le fichier importé dans la base de connaissance, il fallait compter un délai d'environ 24h pour voir la base de connaissance mise à jour. À ce délai, s'ajoutait le délai d'indexation dans l'outil de découverte, effectuée deux fois par semaine. Les résultats d'un import ne pouvaient donc pas être constatés dans l'immédiat et il restait difficile dans ce contexte de savoir dans quelle mesure les données importées étaient effectivement exploitables dans l'outil de découverte. Là aussi, le verrouillage technique des outils ne permettait pas de disposer d'une visibilité sur l'ensemble du processus, dont seules les données d'entrée et de sortie pouvaient être connues, mais avec peu de possibilités d'effectuer des vérifications plus poussées ou des rectifications.

Etant donné la complexité de cette procédure, l'automatisation de ces imports aurait signifié un investissement considérable en temps et en moyens. Cette procédure n'a donc pas été automatisée et il apparaît difficile qu'une bibliothèque mette en place elle-même une telle procédure par ses propres moyens. Il s'agit donc plutôt pour l'ABES de faire en sorte que les fournisseurs d'outils

<sup>267</sup> ABES. *Outils de découverte : signaler grâce à BACON* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 20 juin 2016]. Disponible à l'adresse : [http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES\\_2016\\_session\\_parallele\\_1\\_06\\_BACON\\_OUTILS\\_DE\\_DECOUVERTE\\_VD](http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES_2016_session_parallele_1_06_BACON_OUTILS_DE_DECOUVERTE_VD).

commerciaux intègrent nativement les fichiers de BACON en amont et puissent les proposer en tant que tels, pour éviter que les bibliothèques aient à effectuer ces chargements par elles-mêmes. Pour cela, il est fondamental que les fournisseurs de bases de connaissances et d'outils de découverte gèrent nativement le format KBART. À terme, pour que le cercle vertueux fonctionne pleinement, l'ABES souhaiterait avoir accès à des plateformes de test des outils de découverte pour pouvoir effectuer des vérifications, suivre l'évolution des systèmes et disposer d'une visibilité sur la manière dont les outils intègrent les données de BACON, afin de mieux comprendre comment celles-ci sont exposées pour les administrateurs d'outils.

On signalera, pour finir, que depuis l'automne 2016, BACON fait désormais l'objet d'un dispositif CERCLES spécifique qui doit débiter en janvier 2017. L'objectif de ce dispositif CERCLES est similaire à celui du SUDOC, avec la particularité qu'il s'agit de métadonnées KBART et que la participation du fournisseur de données en amont reste incontournable. Dans ce dispositif CERCLES-BACON, l'établissement participant est chargé d'un chantier pour lequel il devient référent. Il sera amené à vérifier des fichiers produits par les éditeurs pour établir des rapports d'erreur mais ne sera pas amené à corriger lui-même les notices. C'est le fournisseur de données, grâce au rapport d'erreur, qui apportera les rectifications nécessaires au fichier. De cette façon, la qualité des données est toujours améliorée à la source, en s'appuyant sur l'expertise des catalogueurs des établissements pour les opérations de vérification. Comme dans le chantier BACON jusqu'à présent, le label BACON sera maintenu et confié aux référents des établissements qui seront les plus à même d'évaluer les fichiers pour les éléments qui nécessitent une vérification humaine. Pour sa part, l'ABES effectuera un suivi régulier des chantiers, tout en se consacrant davantage à améliorer la cohérence globale de l'ensemble de la base de données de BACON, par un travail d'analyse et d'alignement des métadonnées. Ce dispositif CERCLES-BACON devrait permettre de continuer d'étendre le périmètre des ressources concernées par BACON, tout en impliquant davantage les établissements du réseau dans ce travail d'amélioration de la qualité des métadonnées des ressources numériques.

### *ISTEX*

La plateforme ISTEX<sup>268</sup> gère les données des licences nationales, constituées de collections rétrospectives de revues et de livres numériques de niveau universitaire dans toutes les disciplines<sup>269</sup>. Elle travaille avec une granularité de niveau article ou chapitre, pour lesquels elle contient des métadonnées (au format MODS) et du texte intégral (au format TEI). Elle n'a pas vocation à constituer un nouveau portail de consultation de ressources mais constitue un réservoir de données qui peuvent être diffusées par différents outils : soit par l'API soit par des outils de découverte commerciaux.

La plateforme ISTEX ne se contente pas de stocker les données fournies par les éditeurs, mais elle réalise un important travail de curation de ces données. Elle effectue ainsi de nombreux traitements pour homogénéiser les contenus, en particulier en ce qui concerne les dates de publication, les mentions de langues,

<sup>268</sup> Plateforme ISTEX | - ISTEX [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://www.istex.fr/plateforme-istex/>.

<sup>269</sup> Pour cette partie, nous nous appuyons sur l'entretien téléphonique réalisé avec Laurent Schmitt, Stéphane Gully et Claude Niederlender de l'équipe ISTEX de l'INIST, le 13 juin 2016.

l'identification des titres et d'ISSN de périodiques, des éditeurs et des types de document. Ainsi, par exemple pour les ISSN, ISTEEX corrige des erreurs de saisie et établit les liens entre les titres complets et les titres abrégés. Ce travail d'alignement des contenus devrait déboucher à terme sur un référentiel afin de permettre de disposer d'un réservoir le plus homogène possible.

En complément, ISTEEX effectue un travail considérable d'enrichissement des données, qui porte notamment sur les entités nommées, le repérage des références citées et les disciplines. Ainsi, le texte intégral et les métadonnées sont traités avec un algorithme afin d'extraire les entités nommées (les noms de personne ou les noms géographiques) et d'enrichir les métadonnées pour améliorer leur exploitation par différents outils (API ou outils commerciaux). Concernant les autorités auteurs, ISTEEX s'appuie sur l'ISNI et sur IdRef, et elle travaille aussi avec le hub de métadonnées de l'ABES pour réinsérer des identifiants auteurs dans les données d'ISTEEX. Ce travail d'enrichissement normalisé devrait permettre au final de disposer d'une liste d'identifiants d'auteurs, pour obtenir plus de précision dans la recherche. Quant aux disciplines, les équipes d'ISTEEX travaillent avec des outils de *clustering* et de *data mining* pour générer des mots clés à partir des mots utilisés dans le texte intégral d'un article donné. Elles récupèrent également la catégorie des disciplines des revues indiquées dans la base Crossref ou dans Web of Science, pour alimenter les disciplines scientifiques des données ISTEEX de manière homogène.

Les ressources concernées par ISTEEX sont pour la plupart déjà indexées dans des outils commerciaux, mais à partir d'autres sources : elles ne bénéficient donc pas de ces améliorations. Ainsi, dans de nombreux cas, même si la ressource est disponible sur ISTEEX, c'est une notice moins riche et comportant un lien vers la plateforme de l'éditeur qui figure dans l'outil de découverte. Or étant donné ce travail conséquent de curation et d'enrichissement de données, il serait dommage que les données d'ISTEEX restent peu visibles dans les outils de découverte commerciaux. Il est donc important de mieux référencer ces données dans les outils commerciaux et de faire en sorte que les URL d'accès pointent bien vers le réservoir d'ISTEEX à chaque fois que cela est possible.

Pour ce faire, le comité exécutif d'ISTEEX a entamé un dialogue avec les fournisseurs d'outils de découverte afin de travailler conjointement à améliorer le signalement et l'accès aux ressources ISTEEX dans les index commerciaux. Lors d'une réunion de travail avec les fournisseurs d'outils en février 2016, le comité exécutif leur a présenté la plateforme ISTEEX, et les a invités à présenter leurs solutions pour intégrer les ressources ISTEEX, dans le but d'assurer la meilleure diffusion de celles-ci dans les outils de découverte.

Afin que les fournisseurs d'outils de découverte puissent tester l'intégration des données d'ISTEEX, la plateforme leur a fourni des échantillons de données. Il s'agit d'environ 15000 enregistrements (articles ou chapitres) pris au hasard dans le réservoir, dont environ 10000 sans traitement préalable et environ 5000 avec enrichissement des métadonnées. Cette extraction a été réalisée en TEI et a été mise à disposition des différents fournisseurs d'outils accompagnée de notices explicatives techniques. Ce jeu de test devrait permettre de déterminer si les données d'ISTEEX peuvent être intégrées par les fournisseurs d'outils, en mode brut ou en mode enrichi. Les fournisseurs d'outils seront amenés à tester ces échantillons et à l'issue de cette expérimentation, ISTEEX pourra envisager les mesures nécessaires pour faire en sorte que ses données soient mieux référencées dans les outils de découverte.

### *Le hub de métadonnées*

Le hub de métadonnées de l'ABES<sup>270</sup> travaille en collaboration avec ISTEEX pour retraiter et enrichir les données produites par les éditeurs. Lancé en 2013, le hub traite de manière semi-automatique des métadonnées non MARC afin de les normaliser, de les enrichir et de permettre leur diffusion à travers le SUDOC et les outils de découverte. Ces données peuvent concerner différents formats (XML, tabulé, csv etc.) et niveaux de granularité (livres et chapitres, périodiques et articles).

Le hub procède à des analyses et à des enrichissements des données, que ce soit par des procédures automatisées ou par un travail manuel. Les vérifications se font à chaque fois que cela est possible par des contrôles automatisés, tandis que le travail humain se concentre sur l'enrichissement intellectuel. Il s'agit par exemple d'ajouter la date d'édition électronique, des indices Dewey ou encore d'aligner les thésaurus éditeur avec les indices Dewey. Le hub a collaboré également au projet Qualinca<sup>271</sup> afin d'enrichir les liens grâce à l'utilisation des référentiels et des identifiants IdRef, VIAF, ISNI, ORCID pour associer des identifiants auteurs aux métadonnées. En ce qui concerne les données d'ISTEX, le hub travaille aussi, au niveau des périodiques, sur des développements touchant les ISSN afin de permettre de reconstruire l'historique complet d'une revue qui change de titre.

Une fois ces notices normalisées et enrichies, le hub permet la diffusion de ces données pour différents usages, à travers le SUDOC, les outils de découverte, le web de données, ou l'API de recherche ISTEEX.

## **D. La curation de données au niveau d'une bibliothèque**

Comme l'avait signalé Karen Calhoun, le rôle de la bibliothèque tend à se déplacer du contrôle bibliographique à la gestion des métadonnées, déplacement qui est encore accéléré par les outils de découverte<sup>272</sup>. Elle signale ainsi que face aux notices du catalogue (qui répondent à des règles bien maîtrisées de saisie et de traitement des données), c'est une multitude de sources, de qualité et de traitement disparates que doivent intégrer les bibliothèques. Les notices ne sont plus réalisées à la main mais générées par lots via des extractions, des conversions et des imports en grandes quantités. On est désormais dans des systèmes d'échange et de réutilisation des métadonnées entre différents entrepôts, où la production des métadonnées est répartie entre plusieurs acteurs. Les bibliothèques doivent donc repenser leur activité de contrôle bibliographique afin de se repositionner dans cette chaîne.

Plusieurs bibliothèques travaillent ou envisagent de travailler sur la qualité des données de leur outil de découverte (BIB01, BIB26). Pour la BIB01, le besoin d'exhaustivité et la fraîcheur de mise à jour sont prioritaires par rapport à celui de la qualité. Mais la bibliothèque tente aussi d'agir sur la qualité des données,

<sup>270</sup> ABES : *Hub de métadonnées – Projets en cours* [en ligne]. [s. d.]. [Consulté le 10 février 2016]. Disponible à l'adresse : <http://www.abes.fr/Projets-en-cours/Hub-de-metadonnees>.

<sup>271</sup> ABES : *Qualinca – Projets en cours* [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://www.abes.fr/Projets-en-cours/Qualinca>.

<sup>272</sup> CALHOUN, Karen. Supporting digital scholarship: bibliographic control, library co-operatives and open access repositories. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 143-177.

notamment pour repérer des manques qui seraient bloquants et les corriger quand le niveau de qualité est trop insuffisant.

La BIB01 a ainsi défini des besoins minimaux en matière de métadonnées qui concernent les champs suivants : le titre, l’auteur, la langue, la date du contenu et l’URL. Ces critères sont considérés comme fondamentaux car ils sont très utilisés dans les recherches à travers les facettes. Si ces informations sont manquantes ou erronées, l’utilisation des facettes dans la recherche risquerait d’exclure des notices des résultats. Ces critères identifiés comme fondamentaux doivent donc être vérifiés et actualisés. En effet la mise à jour rapide des données dans l’index centralisé et dans la base de connaissance peut être une contrainte lourde en matière d’amélioration de la qualité des données : les mises à jour continues effaceront peut-être les efforts réalisés. La perspective possible d’un désabonnement à telle ou telle ressource peut aussi à terme limiter la recherche d’amélioration qualitative.

Dans ce cas, les champs sujets ne font pas partie des critères de qualité minimale requise (en raison de leur manque potentiel et de leur très grande disparité), cependant la BIB01 a entrepris un travail d’homogénéisation des sujets de ses ressources locales. Dans sa base de connaissance, les sujets des ressources numériques ont été convertis en mots RAMEAU. Cette indexation RAMEAU est ensuite mappée avec la classification CDU utilisée pour les notices du catalogue pour qu’elle puisse être ajoutée à ces dernières. Ainsi, toutes les notices provenant des ressources locales (SIGB et base de connaissance) auront une indexation sujet homogène commune avec les mêmes catégories. Dans sa base de connaissance, la BIB01 enrichit également les données à travers des traitements pour ajouter des auteurs grâce à l’ISBN de l’édition imprimée, ce qui contribue à l’homogénéisation des données.

On peut penser qu’à terme la curation des données et la question de la définition des seuils de qualité est appelée à se poser de manière de plus en plus large dans le cadre de l’extension du projet CERCLES à BACON et, au-delà, du SGBm<sup>273</sup>. D’autres bibliothèques envisagent aussi de travailler sur cette thématique (BIB26).

---

<sup>273</sup> Cf. *supra* I, « Outils propriétaires, outils libres, outils mutualisés ».

## CONCLUSION

---

Dans ce travail, nous avons tenté de préciser les définitions et les contours des outils de découverte, afin de clarifier le positionnement de ces derniers vis-à-vis d'autres outils et ressources proposés par les bibliothèques universitaires : s'ils s'apparentent à des ressources numériques par leur modèle économique et les données qu'ils exposent, ils relèvent de l'informatique documentaire par les interactions complexes qu'ils ont avec l'ensemble des systèmes locaux de la bibliothèque.

Il ressort des entretiens que nous avons pu mener que l'articulation de ces outils avec l'environnement informatique des bibliothèques est un point crucial de leur mise en œuvre et une condition préalable à leur bon fonctionnement. C'est pourquoi Marshall Breeding conseillait aux bibliothèques de ne pas réfléchir simplement en termes de fonctionnalités offertes : « les checklists de fonctionnalités peuvent être trompeuses quant à la performance effective d'un outil donné dans son fonctionnement quotidien au sein d'une bibliothèque »<sup>274</sup>. Cette articulation avec les systèmes locaux est souvent complexe et peut entraîner la mise en place de nouveaux *workflows* ou la modification des *workflows* existants. Même dans le cas d'outils appartenant à la suite d'un même fournisseur, la « brique » de découverte n'est pas un simple ajout à une infrastructure préexistante. La mise en place de ces outils implique un travail non négligeable de configuration, d'ajustement et de rééquilibrage au sein de l'ensemble de l'environnement technique de la bibliothèque. Les contraintes, voire les difficultés, que les bibliothèques ont rencontrées pour leur mise en œuvre contrastent fortement avec l'idéal unificateur de ces outils qui se veulent de prime abord « clé-en-main ». L'interopérabilité et le niveau de paramétrabilité de ces outils demeurent donc des questions de première importance pour les bibliothèques qui envisagent de les déployer.

À cette complexité de mise en œuvre s'ajoute celle du modèle économique qui sous-tend les services de découverte à index général centralisé. Plus particulièrement, le caractère confidentiel de la négociation des flux de données avec les fournisseurs de celles-ci (éditeurs, agrégateurs) et la faible transparence sur les caractéristiques de l'indexation limitent la connaissance et la compréhension qu'une bibliothèque peut avoir des outils qu'elle met en place. À cet égard, les recommandations de l'ODI restent toujours d'actualité et la marge d'amélioration dans ce domaine est encore très grande. De même, d'importantes améliorations sont encore possibles en ce qui concerne la qualité des métadonnées de l'index centralisé, leur niveau de normalisation et la visibilité qu'ont les bibliothèques sur les traitements effectués sur ces données.

À long terme, la confidentialité sur les données négociées, ajoutée au verrouillage technique des outils, peut risquer de conduire à une rente de situation pour les fournisseurs, réduisant d'autant la marge de manœuvre des bibliothèques. Cette situation est d'autant plus paradoxale que, dans de nombreux cas, l'outil de

---

<sup>274</sup> « Checklists of functionality can be misleading relative to the actual performance of any given system in its daily operation in a library. » BREEDING, Marshall. Introduction and Concepts. *Op. cit.*, p. 5.

découverte devient véritablement le *front door*<sup>275</sup> ou porte d'entrée principale de la bibliothèque.

En ce qui concerne l'utilisation des outils et l'impact de ces derniers sur la consultation de la documentation numérique, les données dont nous avons disposé sont encore peu nombreuses et il serait imprudent d'en tirer des généralisations. Les outils de mesure de l'utilisation de ces services ainsi que de leur impact sur la consultation de la documentation numérique sont appelés à être consolidés et enrichis à l'avenir, afin de mieux cerner cet aspect du problème. En matière d'utilisabilité et d'*information literacy*, ces outils présentent des fonctionnalités et des potentialités intéressantes. Ils suscitent toutefois des réticences au niveau de la décontextualisation de l'information, à l'image des outils du web dont ils s'inspirent.

Dans une perspective plus large, les outils de découverte font apparaître la curation des données comme un enjeu crucial pour la documentation au sein de l'ESR. Comme l'ont rappelé Jason Thomale, William Hicks et Mark Phillips, il n'y a pas d'outil de découverte qui puisse suppléer à la fragmentation des données issues de différentes sources et dans ce sens, ce que peut faire la bibliothèque c'est d'améliorer la qualité des données (*improve our data*)<sup>276</sup>. Le travail de curation de données, piloté au niveau national par des institutions telles que l'ABES et l'INIST, est amené à se développer et à s'étendre, comme le montre la récente mise en place d'un dispositif CERCLES pour le projet BACON.

Ce travail de curation ne peut être pleinement efficace et profitable à toute la communauté que dans une perspective d'ouverture des données. Il semble en effet difficile pour une bibliothèque de consacrer beaucoup de moyens à améliorer la qualité des données des ressources numériques lorsque ces données restent l'objet d'accords commerciaux. C'est pourquoi l'ouverture des données apparaît être le corrélat nécessaire du travail de curation. À travers des projets nationaux comme BACON et ISTEEX, l'ABES et le comité exécutif d'ISTEEX œuvrent déjà en ce sens, en mettant leurs données à disposition sous licence ouverte, afin que tout établissement puisse librement les réutiliser. À terme, ce modèle d'ouverture des données, semble le plus viable économiquement et le plus transparent pour les bibliothèques, tout en s'inscrivant dans la perspective des biens communs de la connaissance.

---

<sup>275</sup> BOLLINGER, Stephen et SILTON, Kate. *Op. cit.*, p. 27.

<sup>276</sup> THOMALE, Jason, HICKS, William et PHILLIPS, Mark. *Op. cit.*, p. 134



## SOURCES

Nous détaillons ici la liste des entretiens effectués auprès des bibliothèques, des experts dans le domaine des métadonnées et des fournisseurs.

### 1. BIBLIOTHEQUES

Type d'entretien	Interlocuteur(s)	Etablissement	Date	Outil concerné
Téléphonique	Perrine Helly	Université de Bretagne Occidentale	18 avril 2016	Primo - Primo Central Index
Présentiel	Carole Vinot, Michaël Villanova et Vincent Baas	Bibliothèque Diderot de Lyon	19 avril 2016	Primo - Primo Central Index
Présentiel	Dominique Rouger	Université Jean Monnet Saint-Etienne	22 avril 2016	Vufind
Visioconférence	Loïc Ducasse	SICD Toulouse	27 avril 2016	Summon
Téléphonique	Aurélie Bec	Université de Picardie Jules Verne	29 avril 2016	Enterprise - Ebsco Discovery Service
Téléphonique	William Demet	Université de Versailles Saint-Quentin	29 avril 2016	Primo - Primo Central Index
Téléphonique	Raphaële Bussemey	Bibliothèque Clermont-Université	2 mai 2016	Blacklight
Téléphonique	Johnatan Joly	Sorbonne Universités	3 mai 2016	Ebsco Discovery Service
Téléphonique	Laurent Lhuillier	Aix-Marseille Université	3 mai 2016	Summon
Présentiel	Gianna Giorgi	Bibliothèque interuniversitaire de la Sorbonne	11 mai 2016	Encore Duet
Téléphonique	Julien Sicot	Université Rennes 2	13 mai 2016	Ebsco Discovery Service
Téléphonique	Anne-Sophie Pascal	Université de Poitiers	17 mai 2016	Summon

<b>Type d'entretien</b>	<b>Interlocuteur(s)</b>	<b>Etablissement</b>	<b>Date</b>	<b>Outil concerné</b>
Téléphonique	Anne-Claire Bernaudin	Université Rennes 1	27 mai 2016	Ebsco Discovery Service
Présentiel	Brigitte Crayssac	Université Toulouse 1 Capitole	31 mai 2016	Summon
Téléphonique	Yves Tomic	Université Paris Dauphine	3 juin 2016	Primo - Primo Central Index
Présentiel	Sonia Bouis	Université Jean Moulin Lyon 3	10 juin 2016	Summon
Téléphonique	Gaël Revelin	Université de Savoie Mont Blanc	16 juin 2016	Ebsco Discovery Service
Présentiel	Yohann Davrieux	Université Lumière Lyon 2	24 juin 2016	Summon
Présentiel	Pascale Nalon	Mines ParisTech	27 juin 2016	Primo - Primo Central Index
Téléphonique	Thomas Jouneau	Université de Lorraine	7 juillet 2016	Ebsco Discovery Service
Téléphonique	François Renaville	Université de Liège	8 juillet 2016	Primo - Primo Central Index
Téléphonique	Ewa Nieszkowska	Université Paris Est Créteil Val de Marne	12 juillet 2016	Primo - Primo Central Index
Téléphonique	Laurent Nabias	Université Paris 3	12 juillet 2016	Primo - Primo Central Index
Visioconférence	Laurent Aucher	Paris Sciences et Lettres	18 juillet 2016	Primo - Primo Central Index
Présentiel	Laurent Quinson	ENSSIB	19 juillet 2016	Ebsco Discovery Service
Téléphonique	Dorothee Benhamou-Suesser et Emilie Cosson	Université Paris Descartes	26 août 2016	Ebsco Discovery Service

## 2. EXPERTS DANS LE DOMAINE DES METADONNEES

Type d'entretien	Interlocuteur(s)	Etablissement	Date	Projet concerné
Téléphonique	Laurent Schmitt, Stéphane Gully et Claude Niederlender	INIST	13 juin 2016	ISTEX
Téléphonique	Cyril Leroy	ABES	4 juillet 2016	BACON

## 3. FOURNISSEURS

Type d'entretien	Interlocuteur(s)	Société	Date	Produit concerné
Téléphonique	Mathieu Ponthenier	Ebsco	16 septembre 2016	Ebsco Discovery Service
Téléphonique	Pierre Persinette-Gautrez	Ex Libris - Proquest	26 septembre 2016	Summon



# BIBLIOGRAPHIE

---

## 1. INFORMATIQUE DOCUMENTAIRE : CONTEXTE ET EVOLUTIONS

BERMÈS, Emmanuelle. Enabling your catalogue for the semantic web. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 117-142.

BERTHIER, Sandrine. *Le SIGB: pilier ou élément désormais mineur de l'informatique documentaire ?* [en ligne]. Villeurbanne, Rhône, France : [s. n.], 2012. [Consulté le 23 février 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/60267-le-sigb-pilier-ou-element-desormais-mineur-de-l-informatique-documentaire.pdf>

BREEDING, Marshall. *Next-gen library catalogs*. London : Facet Pub, 2010. The tech set, 1. ISBN 978-1-85604-721-0.

CALLEWAERT, Rosemie. FRBRizing your catalogue. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 93-115.

CHAMBERS, Sally (dir.). *Catalogue 2.0: the future of the library catalogue*. London : Facet Publ, 2013. ISBN 978-1-85604-716-6

COUDRIN, Delphine. Les dispositifs de signalement en université. Dans : *Intégrer des ressources numériques dans les collections*. Villeurbanne, France : Presses de l'enssib, 2014, p. 112-125. Boîte à outils, 29.

COUDRIN, Delphine. *Faire évoluer un SID : enjeux / retour d'expérience - Intervention DCB 25* [en ligne]. ENSSIB, 29 juin 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : [http://moodle.enssib.fr/pluginfile.php/20764/mod\\_resource/content/0/Intervention%20Delphine%20Coudrin.pdf](http://moodle.enssib.fr/pluginfile.php/20764/mod_resource/content/0/Intervention%20Delphine%20Coudrin.pdf)

COUDRIN, Delphine et HATT, Guillaume. Intégrer, signaler les ressources électroniques : 20 questions clés. Dans : *Intégrer des ressources numériques dans les collections*. Villeurbanne, France : Presses de l'enssib, 2014, p. 90-102. Boîte à outils, 29.

DEMANGE, Sophie. L'informatisation documentaire : combien ça coûte ? *Arabesques*. 2013, Vol. 70, p. 18-20

JESTAZ, Laure. Réinformatiser à l'heure du SGBM : Journée ADBU/ABES – 20 février 2014. *Bulletin des bibliothèques de France* [en ligne]. Avril 2014, n° 1. [Consulté le 19 janvier 2016]. Disponible à l'adresse : [http://bbf.enssib.fr/tour-d-horizon/reinformatiser-a-l-heure-du-sgbm\\_64250](http://bbf.enssib.fr/tour-d-horizon/reinformatiser-a-l-heure-du-sgbm_64250)

MACHEFERT, Sylvain. *L'OpenURL dans les institutions françaises, une chance pour la valorisation des ressources électroniques ?* [en ligne]. [S. l.] : Institut national des techniques de la documentation du CNAM, 6 novembre 2007. [Consulté le 2 décembre 2016]. Disponible à l'adresse : [https://memsic.ccsd.cnrs.fr/mem\\_00000613/document](https://memsic.ccsd.cnrs.fr/mem_00000613/document)

MULLER, Catherine. *Étude et repérages sur les interfaces documentaires innovantes* [en ligne]. Villeurbanne, France : EnssibLab, 2014. [Consulté le 21 février 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/64911-etude-et-reperages-sur-les-interfaces-documentaires-innovantes.pdf>

REVELIN, Gaël. *Signaler les ressources électroniques* [en ligne]. 1 janvier 2012. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://bbf.enssib.fr/consulter/bbf-2012-03-0087-003>

TOSCA CONSULTANTS. *Tosca Consultants : Les logiciels métier destinés aux bibliothèques* [en ligne]. [s. d.]. [Consulté le 2 mars 2016]. Disponible à l'adresse : <http://www.toscaconsultants.fr/logiciel.htm>

VAISMAN, Jean-Baptiste. *Qu'est-ce qu'un portail de BU réussi ?* S.l. : s.n., 2014. [Consulté le 29 mars 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/64153-qu-est-ce-qu-un-portail-de-bu-reussi.pdf>

WENZ, Romain. *L'avenir des catalogues* [en ligne]. S.l. : s.n., 2009. [Consulté le 30 avril 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/21205-l-avenir-des-catalogues.pdf>

## 2. NORMALISATION, CURATION ET ECHANGES DE METADONNEES

ABES. BACON et la labellisation des données : à quelle aune mesure-t-on la qualité d'un fichier KBART ? Dans : *Punktokomo* ; [en ligne]. 28 mai 2015. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2015/05/28/bacon-et-la-labellisation-des-donnees-a-quelle-aune-mesure-t-on-la-qualite-dun-fichier-kbart/>

ABES. CERCLES : 4 nouveaux corpus. Dans : *Punktokomo* ; [en ligne]. 17 juillet 2015. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2015/07/17/cercles-4-nouveaux-corpus/>

ABES. CERCLES : Retour d'expérience de La Sorbonne & l'École Française d'Athènes. Dans : *Punktokomo* ; [en ligne]. 15 janvier 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/01/15/cercles-retour-dexperience-de-la-sorbonne-lecole-francaise-dathenes/>

ABES. CERCLES : retour d'expérience du SCD de l'Université de Picardie Jules Verne sur le corpus CAIRN. Dans : *Punktokomo* ; [en ligne]. 7 septembre 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/09/07/cercles-retour-dexperience-du-scd-de-luniversite-de-picardie-jules-verne-sur-le-corpus-cairn/>

ABES. Mettre nos données en réseau – un démonstrateur. [4f] Matrice des fascicules pour conservation partagée. Dans : *Punktokomo* ; [en ligne]. 16 mai 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/05/16/mettre-nos-donnees-en-reseau-un-demonstrateur-4f-matrice-des-fascicules-pour-conservation-partagee/>

ABES. Mettre nos données en réseau – un démonstrateur. [4g] Le Bouquet des ebooks dalloz. Dans : *Punktokomo* ; [en ligne]. 16 mai 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/05/16/mettre-nos-donnees-en-reseau-un-demonstrateur-4g-le-bouquet-des-ebooks-dalloz/>

ABES. OpenRefine au service de BACON : quelle évaluation pour les fichiers KBART ? [1] – Introduction. Dans : *Punktokomo* ; [en ligne]. 12 septembre 2016. [Consulté le 13 septembre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/09/12/openrefine-au-service-de-bacon-quelle-evaluation-pour-les-fichiers-kbart-1-introduction/>

ABES. OpenRefine au service de BACON : quelle évaluation pour les fichiers KBART ? [2] – Un outil : OpenRefine. Dans : *Punktokomo* ; [en ligne]. 12 septembre 2016. [Consulté le 13 septembre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/09/12/openrefine-au-service-de-bacon-quelle-evaluation-pour-les-fichiers-kbart-2-un-outil-openrefine/>

ABES. OpenRefine au service de BACON : quelle évaluation pour les fichiers KBART ? [3] – Cas pratique. Dans : *Punktokomo* ; [en ligne]. 23 septembre 2016. [Consulté le 4 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/09/23/openrefine-au-service-de-bacon-quelle-evaluation-pour-les-fichiers-kbart-3-cas-pratique/>

ABES. OpenRefine au service de BACON : quelle évaluation pour les fichiers KBART ? [4] – Dispositif CERCLES dans le cadre de BACON. Dans : *Punktokomo* ; [en ligne]. 4 octobre 2016. [Consulté le 4 octobre 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/10/04/openrefine-au-service-de-bacon-quelle-evaluation-pour-les-fichiers-kbart-4-dispositif-cercles-dans-le-cadre-de-bacon/>

ABES. *Outils de découverte : signaler grâce à BACON* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 20 juin 2016]. Disponible à l'adresse : [http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES\\_2016\\_session\\_parallele\\_1\\_06\\_BACON\\_OUTILS\\_DE\\_DECOUVERTE\\_VD](http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES_2016_session_parallele_1_06_BACON_OUTILS_DE_DECOUVERTE_VD)

ABES. Récit d'une immersion. Traiter les ebooks Dalloz avec les données Sudoc, les données de l'éditeur et les outils du hub. Dans : *Punktokomo* ; [en ligne]. 22 juin 2016. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://punktokomo.abes.fr/2016/06/22/recit-dune-immersion-traiter-les-ebooks-dalloz-avec-les-donnees-sudoc-les-donnees-de-lediteur-et-les-outils-du-hub/>

ABES. Retour sur l'enquête « Programme de travail BACON septembre 2015-décembre 2015 ». Dans : *Fil ABES* [en ligne]. 19 octobre 2015. [Consulté le 22 juin 2016]. Disponible à l'adresse : <https://fil.abes.fr/2015/10/19/retour-sur-lenquete-programme-de-travail-bacon-septembre-2015-decembre-2015/>

ABES. *Vers un cercle vertueux : Comment les établissements Sudoc, les éditeurs et l'ABES collaborent pour proposer des données de signalement de qualité* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 20 juin 2016]. Disponible à l'adresse : [http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES\\_2016\\_session\\_parallele\\_1\\_01\\_intro\\_VD](http://www.abes.fr/Media/Fichiers/Footer/Journees-ABES/jabes16/JABES_2016_session_parallele_1_01_intro_VD)

CALHOUN, Karen. Supporting digital scholarship: bibliographic control, library cooperatives and open access repositories. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 143-177. ISBN 978-1-85604-716-6

CLARAC, Cécile. Les bases de connaissance à la croisée des chemins. *Arabesques*. 2013, n° 70, p. 8-9

E-DATA QUALITY WORKING GROUP OCLC. *Success strategies for electronic content and access: a cross-industry white paper* [en ligne]. Dublin, OH : OCLC, 2014. [Consulté le 20 mars 2016]. ISBN 978-1-55653-483-6. Disponible à l'adresse : <http://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>

ELLERO, Nadine. Integration or Disintegration: Where Is Discovery Headed? *Journal of Library Metadata* [en ligne]. Octobre 2013, Vol. 13, n° 4, p. 311-329. DOI 10.1080/19386389.2013.831277

GROUPE DE TRAVAIL NISO/UKSG KBART. *KBART: Bases de connaissances et outils associés : Rapport préparé par le groupe de travail NISO/UKSG KBART* [en ligne]. Trad. par Thomas PORQUET et Karine DELVERT. [S. l.] : Couperin, 2011. [Consulté le 1 avril 2016]. Disponible à l'adresse : [http://www.couperin.org/images/stories/KBART/kbart\\_couperin\\_fr.pdf](http://www.couperin.org/images/stories/KBART/kbart_couperin_fr.pdf)

HAN, Myung-Ja K. et WEATHERS, William Fletcher. Know Thy Metadata: Metadata Challenges in Discovery Services. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 267-276.

HARPEL-BURKE, Pamela. Catalog Maintenance in the Discovery System: Opinions of Early Adopters. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 126-143.

ICOLC. *Revised Statement on the Global Economic Crisis and Its Impact on Consortial Licenses / ICOLC Website* [en ligne]. 14 juin 2010. [Consulté le 30 septembre 2016]. Disponible à l'adresse : <http://icolc.net/statement/revised-statement-global-economic-crisis-and-its-impact-consortial-licenses>

*Knowledge Base and Related Tools - National Information Standards Organization* [en ligne]. [s. d.]. [Consulté le 14 novembre 2016]. Disponible à l'adresse : <http://www.niso.org/workrooms/kbart>

LEROY, Cyril. Les rouages de BACON. *Arabesques*. 2015, n° 79, p. 26



NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO) et OPEN DISCOVERY INITIATIVE WORKING GROUP. *Open discovery initiative: promoting transparency in discovery: a recommended practice of the National Information Standards Organization*. Baltimore, MD : National Information Standards Organization (U.S.), 2014. ISBN 978-1-937522-42-1. Disponible à l'adresse : [http://www.niso.org/apps/group\\_public/download.php/14820/rp-19-2014\\_ODI.pdf](http://www.niso.org/apps/group_public/download.php/14820/rp-19-2014_ODI.pdf)

NATIONAL FEDERATION OF ABSTRACTING AND INFORMATION SERVICES (NFAIS). *Recommended Practices: Discovery Services* [en ligne]. Philadelphia, PA : National Federation of Abstracting and Information Services, 2013. [Consulté le 20 mars 2016]. Disponible à l'adresse : [https://nfais.memberclicks.net/assets/docs/BestPractices/recommended\\_practices\\_final\\_aug\\_2013.pdf](https://nfais.memberclicks.net/assets/docs/BestPractices/recommended_practices_final_aug_2013.pdf)

PASTOR MARTÍNEZ, Nuria. *Un cercle vertueux: E-books Licence Nationale Springer* [en ligne]. Montpellier, 10 mai 2016. [Consulté le 10 octobre 2016]. Disponible à l'adresse : <http://www.abes.fr/Publications-Evenements/Journees-ABES/Journees-ABES-2016-10-11-mai-2016>

RENAVILLE, François. Open Access and Discovery Tools : How do Primo Libraries Manage Green Open Access Collections ? Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 233-256. ISBN 978-0-8389-1414-4

WALKER, Jenny. The NISO Open Discovery Initiative: promoting transparency in discovery. *Insights* [en ligne]. Mars 2015, Vol. 28, n° 1. [Consulté le 28 octobre 2016]. DOI 10.1629/uksg.186

ZHU, Julie et KELLEY, Jalyn. Collaborating to Reduce Content Gaps in Discovery: What Publishers, Discovery Service Providers, and Libraries Can Do to Close the Gaps. *Science & Technology Libraries* [en ligne]. Octobre 2015, Vol. 34, n° 4, p. 315-328. DOI 10.1080/0194262X.2015.1102677

### 3. OUTILS DE DECOUVERTE : ETUDES GENERALES

AHARONY, Noa et PREBOR, Gila. Librarians' and Information Professionals' Perspectives Towards Discovery Tools — An Exploratory Study. *The Journal of Academic Librarianship* [en ligne]. Juillet 2015, Vol. 41, n° 4, p. 429-440. DOI 10.1016/j.acalib.2015.05.003

ÁVILA-GARCÍA, Lorena, ORTIZ-REPISO, Virginia et RODRÍGUEZ-MATEOS, David. Herramientas de descubrimiento: ¿una ventanilla única? *Revista Española de Documentación Científica*. Mars 2015, Vol. 38, n° 1, p. 1-17

BOLLINGER, Stephen et SILTON, Kate. When You Rent Your Front Door: A Mid-Size Library's Experiences Relying on Discovery Vendors. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 27-34.

BREEDING, Marshall. Chapter 2: Major Discovery Product Profiles. *Library Technology Reports*. Février 2014, Vol. 50, n° 1, p. 33-52

BREEDING, Marshall. Chapter 3. Relationship with Discovery. *Library Technology Reports*. Mai 2015, Vol. 51, n° 4, p. 22-25

BREEDING, Marshall. Discovery product functionality. *Library Technology Reports*. 2015, Vol. 50, n° 1, p. 5-32

BREEDING, Marshall. Introduction and Concepts. *Library Technology Reports*. Juin 2015, Vol. 51, n° 4, p. 5-19

BREEDING, Marshall. Chapter 2: Major Discovery Product Profiles. *Library Technology Reports*. Février 2014, Vol. 50, n° 1, p. 33-52

BREEDING, Marshall. Next-generation discovery: an overview of the European Scene. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 37-64.

BREEDING, Marshall et NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). *The future of library resource discovery: a white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee*. Baltimore : NISO, 2015. [Consulté le 21 février 2016]. ISBN 978-1-937522-41-4. Disponible à l'adresse : [http://www.niso.org/apps/group\\_public/download.php/14487/future\\_library\\_resource\\_discovery.pdf](http://www.niso.org/apps/group_public/download.php/14487/future_library_resource_discovery.pdf)

COUPERIN. *Discovery-tools - Analyse de l'enquête 2013* [en ligne]. [S. l.] : Couperin, 2013. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.couperin.org/groupe-de-travail-et-projets-deap/acces-aux-ressources-cat/discovery/item/1083-analyse-de-l-enquete-discovery-tools-2013>

COUPERIN. *Discovery-tools - Journée d'étude Discovery tools libres du 24 janvier 2012* [en ligne]. [s. d.]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.couperin.org/groupe-de-travail-et-projets-deap/acces-aux-ressources-cat/discovery/item/954-journee-detude-discovery-tools-libres-du-24-janvier-2012>

DEMPSEY, Lorcan. Thirteen ways of looking at libraries, discovery and the catalogue: scale, workflow, attention. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 179-202.

DEODATO, Joseph. Evaluating Web-Scale Discovery: A Step-by-Step Guide. *Information Technology & Libraries*. Juin 2015, Vol. 34, n° 2, p. 19-75. DOI 10.6017/ital.v34i2.5745

GROUPE DE TRAVAIL DU RÉSEAU DES BIBLIOTHÈQUES DE L'UNIVERSITÉ DE MONCTON. *Les outils de découverte : Rapport du groupe de travail du Réseau des bibliothèques de l'Université de Moncton. Version abrégée destinée au public* [en ligne]. [S. l.] : [s. n.], mars 2011. [Consulté le 21 février 2016]. Disponible à l'adresse : [http://www.umoncton.ca/umcm-bibliotheque-champlain/files/umcm-bibliotheque-champlain/wf/wf/pdf/Outils\\_decouverte\\_Rapport\\_final\\_abrege\\_mars2011.pdf](http://www.umoncton.ca/umcm-bibliotheque-champlain/files/umcm-bibliotheque-champlain/wf/wf/pdf/Outils_decouverte_Rapport_final_abrege_mars2011.pdf)

HOEPPNER, Athena. The Ins and Outs of Evaluating Web-Scale Discovery Services. *Computers in Libraries*. Avril 2012, Vol. 32, n° 3, p. 6-40

HUDSON, Cole et HUKILL, Graham. One-to-Many: Building a Single-Search Interface for Disparate Resources. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 141-153.

KABASHI, Arta, PETERSON, Christine, PRATER, Tim, et al. *Discovery services : a white paper for the Texas State Library and Archives' Commission* [en ligne]. [S. l.] : [s. n.], 2014. [Consulté le 20 mars 2016]. Disponible à l'adresse : [https://www.tsl.texas.gov/sites/default/files/public/tslac/lot/TSLAC\\_WP\\_discovery\\_\\_final\\_TSLAC\\_20140912.pdf](https://www.tsl.texas.gov/sites/default/files/public/tslac/lot/TSLAC_WP_discovery__final_TSLAC_20140912.pdf)

KERICO, Juliet, ANTHONY, Paul, BULOCK, Chris, et al. Fostering Discovery through Web Interface Design: Perpetual Beta as the New Norm. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 288-302.

KEYSER, Piet de. Discovery Services, the Next Tool Libraries Must Have, or... Should Discard? Dans : *E-Discovery Tools and Applications in Modern Libraries*. Hershey, PA : Information Science Reference, 2016, p. 56-65.

KINSTLER, Till. Making search work for the library user. Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 17-36.

KORNBLAU, Amy I., STRUDWICK, Jane et MILLER, William. How Web-Scale Discovery Changes the Conversation: The Questions Librarians Should Ask Themselves. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 24-42.

LE NÉZET, Romain. Les outils de découverte : apports et prospective. *Arabesques*. 2013, n° 70, p. 10-11

LUTHER, Judy et KELLY, Maureen C. The Next Generation of Discovery. *Library Journal*. Mars 2011, Vol. 136, n° 5, p. 66-71

NAMEI, Elizabeth et YOUNG, Christal A. *Measuring Our Relevancy: Comparing Results in a Web-Scale Discovery Tool, Google & Google Scholar* [en ligne]. Portland, OR, mars 2015. [Consulté le 3 octobre 2016]. Disponible à l'adresse : [http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2015/Namei\\_Young.pdf](http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2015/Namei_Young.pdf)

POPP, Mary Pagliero et DALLIS, Diane. *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012. ISBN 978-1-4666-1821-3

SINGLEY, Emily. To Bento or Not to Bento – Displaying search results. Dans : *usable libraries* [en ligne]. 4 janvier 2016. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://emilysingley.net/to-bento-or-not-to-bento-displaying-search-results/>

SMET, Egbert de et DHAMDHERE, Sangeeta N. (dir.). *E-discovery tools and applications in modern libraries*. Hershey, PA : Information Science Reference, 2016. ISBN 978-1-5225-0474-0

TAY, Aaron. *Musings about librarianship: 8 things we know about web scale discovery systems in 2013* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://musingsaboutlibrarianship.blogspot.fr/2013/06/8-things-we-know-about-web-scale.html#.VqUqWFKD6yx>

THOMALE, Jason. *You Gotta Keep 'em Separated: The Case for « Bento Box » Discovery Interfaces* [en ligne]. Portland, OR, 5 décembre 2014. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://code4lib.org/conference/2015/thomale>. Présentation filmée disponible sur <https://www.youtube.com/watch?v=i853dLWavZA>

THOMALE, Jason, HICKS, William et PHILLIPS, Mark. The Bento Box Design Pattern. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 117-139. ISBN 978-0-8389-1414-4

THOMPSON, JoLinda. *Implementing web-scale discovery services: a practical guide for librarians*. Lanham, Md : Rowman & Littlefield, 2014. Practical guides for librarians series, no. 9. ISBN 978-0-8108-9126-5

THOMSETT-SCOTT, Beth et REESE, Patricia E. Academic Libraries and Discovery Tools: A Survey of the Literature. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 3-23.

VARNUM, Kenneth J. *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016. ISBN 978-0-8389-1414-4

VAUGHAN, Jason. Chapter 1: Web Scale Discovery What and Why? *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 5-11

WEBSTER, Peter. The Web-Scale Discovery Environment and Changing Library Services and Processes. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 646-661.

YONGMING WANG et DAWES, Trevor A. The Next Generation Integrated Library System: A Promise Fulfilled. *Information Technology & Libraries*. Septembre 2012, Vol. 31, n° 3, p. 76-84

#### 4. UTILISATION, INFORMATION LITERACY ET ETUDES D'UTILISABILITE

ABBS, Fran et COLLIER-WILSON, Rachel. *How is our Discovery System Perceived by Users?* [en ligne]. Budapest, 2015. Disponible à l'adresse : [http://igelu.org/wp-content/uploads/2015/08/11.22How-is-our-Discovery-System-Perceived-by-Users\\_RCW\\_FA.pps](http://igelu.org/wp-content/uploads/2015/08/11.22How-is-our-Discovery-System-Perceived-by-Users_RCW_FA.pps)

ACRL. Framework for Information Literacy for Higher Education. Dans : *Association of College & Research Libraries (ACRL)* [en ligne]. 9 février 2015. [Consulté le 21 octobre 2016]. Disponible à l'adresse : <http://www.ala.org/acrl/standards/ilframework>

AHARONY, Noa et PREBOR, Gila. Librarians' and Information Professionals' Perspectives Towards Discovery Tools — An Exploratory Study. *The Journal of Academic Librarianship* [en ligne]. Juillet 2015, Vol. 41, n° 4, p. 429-440. DOI 10.1016/j.acalib.2015.05.003

ALLEN, Natasha Danae. Utilizing discovery tools for classrooms: how do librarian attitudes on discovery impact tools they teach? *Library Hi Tech News* [en ligne]. Janvier 2015, Vol. 32, n° 1, p. 8-12. DOI 10.1108/LHTN-09-2014-0078

ASHER, Andrew, DUKE, Lynda et WILSON, Suzanne. Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources. *College & Research Libraries*. Septembre 2013, Vol. 74, n° 5, p. 464-488

ASSOCIATION OF COLLEGE & RESEARCH LIBRARIES (ACRL). Information Literacy Competency Standards for Higher Education. Dans : *Association of College & Research Libraries (ACRL)* [en ligne]. 1 septembre 2006. [Consulté le 23 septembre 2016]. Disponible à l'adresse : <http://www.ala.org/acrl/standards/informationliteracycompetency>

BETZ, Sonya et ROBERTON, Ian. Integrating Discovery to Improve the User Experience. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 155-168.

BREITBACH, William. Web-Scale Discovery: A Library of Babel? Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 637-645.

CHAPMAN, Suzanne, DENNIS, Scott, FOLGER, Kathleen, et al. Developing a User-Centered Article Discovery Environment. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 194-224.

CHEVILLOTTE, Sylvie. Bibliothèques et Information Literacy. *Bulletin des bibliothèques de France*. Janvier 2005, n° 2, p. 42-48

CHRISTENSEN, Anne. Next-generation catalogues: what do users think? Dans : CHAMBERS, Sally (dir.), *Catalogue 2.0. The future of the library catalogue*. London : Facet Publ, 2013, p. 1-14.

CMOR, Dianne et LI, Xin. Beyond boolean, towards thinking: discovery systems and information literacy. *Library Management*. Octobre 2012, Vol. 33, n° 8/9, p. 450-457. DOI 10.1108/01435121211279812

COMEAX, David J. Usability Testing of a Web-Scale Discovery System at an Academic Library. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 69-86.

DAHL, David et MACDONALD, Patricia. Implementation and Acceptance of a Discovery Tool: Lessons Learned. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 366-387.

DECOSTER, Sara. *Relever le défi des outils « discovery » : les formations orientées vers l'usager* [en ligne]. Mai 2015. [Consulté le 12 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/182048>

DECOSTER, Sara et RENAUVILLE, François. *Défis et opportunités des outils de découverte : côté professionnel, côté usager* [en ligne]. Mars 2016. [Consulté le 14 juin 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/194784>

DJENNO, Mireille, INSUA, Glenda, GREGORY, Gwen M., et al. Discovering Usability: Comparing Two Discovery Systems at One Academic Library. *Journal of Web Librarianship*. Juillet 2014, Vol. 8, n° 3, p. 263-285. DOI 10.1080/19322909.2014.933690

DURANTE, Kim et WANG, Zheng. Creating an Actionable Assessment Framework for Discovery Services in Academic Libraries. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 95-108.

EMANUEL, Jennifer. Usability of the VuFind Next-Generation Online Catalog. *Information Technology & Libraries*. Mars 2011, Vol. 30, n° 1, p. 44-52

FAGAN, Jody Condit. Discovery Tools and Information Literacy. *Journal of Web Librarianship*. Juillet 2011, Vol. 5, n° 3, p. 171-178. DOI 10.1080/19322909.2011.598332

FALCIANI-WHITE, Nancy. Understanding Information Seeking Behavior of Faculty and Students: A Review of the Literature. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 1-21.

FAWLEY, Nancy et KRYSAK, Nikki. Information Literacy Opportunities within the Discovery Tool Environment. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 87-94.

FRANSEN, Janet, FRIEDMAN-SHEDLOV, Lara, THEIS-MATHON, Nicole, et al. Setting a direction for discovery: a phased approach. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*: Hershey, PA : IGI Global, 2012, p. 174-193.

GARRISON, Scott, PRESTAMO, Anne et RODRIGUEZ, Juan Carlos. Putting Library Discovery Where Users Are. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*. Hershey, PA : IGI Global, 2012, p. 388-406.

GUTHRIE, Ana et MCCOY, Rhonda. A Glimpse at Discovery Tools within the HBCU Library Landscape. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 177-191.

HANRATH, Scott et KOTTMAN, Miloche. Use and Usability of a Discovery Tool in an Academic Library. *Journal of Web Librarianship* [en ligne]. Janvier 2015, Vol. 9, n° 1, p. 1-21. DOI 10.1080/19322909.2014.983259

HOLMAN, Lucy, DARRAJ, Elias, GLASER, Jonathan, et al. How users approach discovery tools. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*: Hershey, PA : IGI Global, 2012, p. 252-267.

JOUNEAU, Thomas et PORQUET, Thomas. Evaluer l'usage des ressources numériques : quels enjeux, quelles solutions? Dans : *Evaluer la bibliothèque par les mesures d'impact*. Villeurbanne, France : ENSSIB, 2016, p. 140-149.

JOUNEAU, Thomas et PORQUET, Thomas. *Exploitation des traces de passage vers la documentation numérique : ezPAARSE AnalogIST* [en ligne]. Enssib, 14 janvier 2016. [Consulté le 31 octobre 2016]. Disponible à l'adresse : [http://www.enssib.fr/sites/www/files/documents/manifestations/2016/ezpaarse\\_\\_\\_analogist\\_-\\_enssib\\_-\\_14\\_01\\_2016.pdf](http://www.enssib.fr/sites/www/files/documents/manifestations/2016/ezpaarse___analogist_-_enssib_-_14_01_2016.pdf)

KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. Discovering the Discovery Tool: The Introduction and Impact on Research and Instruction at Seminole State College of Florida. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 158-176.

LOGRE, Julien. *Autour de COUNTER: l'utilisation des Ressources électroniques onéreuses spécialisées dans l'enseignement supérieur et la recherche en France mesures, indicateurs, pilotage* [en ligne]. S.l. : s.n., 2013. [Consulté le 15 avril 2016]. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/documents/60371-autour-de-counter-l-utilisation-des-ressources-electroniques-onereuses-specialisees-dans-l-enseignement-superieur-et-la-recherche-en-france-mesures-indicateurs-pilotage.pdf>

LOWN, Cory, SIERRA, Tito et BOYER, Josh. How Users Search the Library from a Single Search Box. *College & Research Libraries*. Mai 2013, Vol. 74, n° 3, p. 227-241

MISCHO, William H., SCHLEMBACH, Mary C., BISHOFF, Joshua, et al. User search activities within a library gateway: implications for web-scale discovery systems. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*: Hershey, PA : IGI Global, 2012, p. 153-173.

MOULAISON, Sandy Heather et BOSSALLER, Jenny. Les études d'utilisabilité et les outils de découverte en bibliothèque. (French). *Argus (Montreal, Quebec)*. Mai 2016, Vol. 44, n° 2, p. 35-41

RACE, Tammera. Resource discovery: supporting serendipity. Dans : *Planning and Implementing Resource Discovery Tools in Academic Libraries*: Hershey, PA : IGI Global, 2012, p. 139-152.

ROUGER, Dominique. « Don't let me be Miss Understood » ou les bibliothécaires lisent-ils le COUNTER dans le chiffre ? Dans : *L'information scientifique et technique dans l'univers numérique*. Paris : ADBS, 2010, p. 113-128. Sciences et techniques de l'information.

ROUGER, Dominique. *Regard critique sur les statistiques « d'utilisation » de la documentation numérique*. ENSSIB, 13 mars 2012, p. 1-32

SCHONFELD, Roger C. et LONG, Matthew P. *Ithaka S+R Library Survey 2013* [en ligne]. Octobre 2014. [Consulté le 1 mai 2016]. DOI <http://doi.org/10.3886/ICPSR35352.v1>

SCHONFELD, Roger et ITHAKA S+R. *Does Discovery Still Happen in the Library? Roles and Strategies for a Shifting Reality* [en ligne]. [S. l.] : [s. n.], 2014. [Consulté le 20 mars 2016]. Disponible à l'adresse : [http://www.sr.ithaka.org/wp-content/mig/files/SR\\_Briefing\\_Discovery\\_20140924\\_0.pdf](http://www.sr.ithaka.org/wp-content/mig/files/SR_Briefing_Discovery_20140924_0.pdf)

SWANSON, Edward et JOUDREY, Daniel N. Book Review: The Changing Nature of the Catalog and Its Integration with Other Discovery Tools: Final Report, and Rethinking How We Provide Bibliographic Services for the University of California: Final Report. *Library Resources & Technical Services* [en ligne]. Octobre 2006, Vol. 50, n° 4, p. 295-297. DOI 10.5860/lrts.50n4.295

SWANSON, Troy A. et GREEN, Jeremy. Why We Are Not Google: Lessons from a Library Web site Usability Study. *The Journal of Academic Librarianship* [en ligne]. Mai 2011, Vol. 37, n° 3, p. 222-229. DOI 10.1016/j.acalib.2011.02.014

THOMALE, Jason. Resource Discovery System Usage Report, 02-01-2012 to 08-25-2014. Dans : *UNT Digital Library* [en ligne]. septembre 2014. [Consulté le 22 septembre 2016]. Disponible à l'adresse : <http://digital.library.unt.edu/ark:/67531/metadc499075/>

THOMSETT-SCOTT, Beth et REESE, Patricia E. Academic Libraries and Discovery Tools: A Survey of the Literature. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 3-23.

WANG, Yongming et MI, Jia. Searchability and Discoverability of Library Resources: Federated Search and Beyond. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 109-125. ISBN 978-0-415-70666-7

WAY, Doug. The Impact of Web-scale Discovery on the Use of a Library Collection. *Serials Review* [en ligne]. Décembre 2010, Vol. 36, n° 4, p. 214-220. DOI 10.1016/j.serrev.2010.07.002



WILLIAMS, Sarah C. et FOSTER, Anita K. Promise Fulfilled? An EBSCO Discovery Service Usability Study. *Journal of Web Librarianship* [en ligne]. Juillet 2011, Vol. 5, n° 3, p. 179-198. DOI 10.1080/19322909.2011.597590

## 5. SITOGRAPHIE DES EDITEURS D'OUTILS

*EBSCO Discovery Service / EBSCO* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <https://www.ebscohost.com/discovery>

*Encore Discovery Solution / Innovative* [en ligne]. [s. d.]. [Consulté le 23 février 2016]. Disponible à l'adresse : <https://www.iii.com/products/sierra/encore>

*Ex Libris the bridge to knowledge, Présentation* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.exlibrisgroup.com/fr/category/PrimoOverview>

*Home — Blacklight* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://projectblacklight.org/>

*ProQuest AquaBrowser / Discovery layer that combines popular and scholarly content discovery in an superior user experience - AquaBrowser®* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.proquest.com/products-services/AquaBrowser.html>

*ProQuest Summon / Web-scale discovery service increases the value of your library through an unprecedented research experience - The Summon® Service* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://www.proquest.com/products-services/The-Summon-Service.html>

SirsiDynix Enterprise. Dans : *SirsiDynix.com* [en ligne]. 7 juillet 2015. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.sirsidynix.com/products/enterprise>

*VuFind - Search. Discover. Share.* [en ligne]. [s. d.]. [Consulté le 24 janvier 2016]. Disponible à l'adresse : <http://vufind-org.github.io/vufind/>

*WorldCat Discovery Services* [en ligne]. [s. d.]. [Consulté le 23 février 2016]. Disponible à l'adresse : <https://www.oclc.org/worldcat-discovery.en.html>

## 6. SITOGRAPHIE DES OUTILS DE DECOUVERTE EN BU

*Aix-Marseille université - Bibliothèques universitaires* [en ligne]. [s. d.]. [Consulté le 23 février 2016]. Disponible à l'adresse : <http://univ-amu.summon.serialssolutions.com/#/>

*Archipel +* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : [http://fr.slideshare.net/slideshow/embed\\_code/18685608](http://fr.slideshare.net/slideshow/embed_code/18685608)

*Bibliothèque Clermont-Université* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : <http://bibliotheque.clermont-universite.fr/>

*Bibliothèques de l'université Rennes 2 Catalogue* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : <http://catalogue.bu.univ-rennes2.fr/>

*BU Lorraine* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <https://bu.univ-lorraine.fr/>

*Bu Reims* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.univ-reims.fr/site/bibliotheques,9200.html?>

*BU Savoie Mont Blanc* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.scd.univ-smb.fr/wordpress41/index.php/9000>

*BU UPJV | Bibliothèque Universitaire de L'Université de Picardie Jules Verne* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : [https://amiensscd.ent.sirsidynix.net.uk/client/fr\\_FR/default](https://amiensscd.ent.sirsidynix.net.uk/client/fr_FR/default)

*BU - UVSQ* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : <http://www.bib.uvsq.fr/>

*Catalogue | Sciences Po ./ la bibliothèque* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://www.sciencespo.fr/bibliotheque/fr/rechercher/inmedia>

*Catalogue des bibliothèques universitaires de Limoges* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://catalogue.unilim.fr/cgi-bin/koha/opac-main.pl>

*Encore - BIS* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://encore.biu.sorbonne.fr/iii/encore/?lang=frf>

*Encore - Université Tours* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : <http://portail.scd.univ-tours.fr/iii/encore/?jsessionid=6C8CC013FB647D41D536F4314789F381?lang=frf>

*LILLIAD - Moteur de recherche documentaire* [en ligne]. [s. d.]. [Consulté le 2 décembre 2016]. Disponible à l'adresse : [http://lilliad-primos.hosted.exlibrisgroup.com/primos\\_library/libweb/action/dlDisplay.do?vid=33BUBLIL\\_VU1&afterPDS=true&docId=](http://lilliad-primos.hosted.exlibrisgroup.com/primos_library/libweb/action/dlDisplay.do?vid=33BUBLIL_VU1&afterPDS=true&docId=)

*Odyssée - Outil de découverte de la bibliothèque Paris-Dauphine* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : [http://primos.bu.dauphine.fr/primos\\_library/libweb/action/search.do?vid=DAU](http://primos.bu.dauphine.fr/primos_library/libweb/action/search.do?vid=DAU)

*Portail Bibliothèque Diderot de Lyon* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse : [http://rechercher.bibliotheque-diderot.fr/primos\\_library/libweb/action/search.do?dscnt=0&dstmp=1453056672190&vid=BDL&fromLogin=true](http://rechercher.bibliotheque-diderot.fr/primos_library/libweb/action/search.do?dscnt=0&dstmp=1453056672190&vid=BDL&fromLogin=true)

*Portail documentaire Sorbonne Universités* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://documentation.sorbonne-universites.fr/>

*Primo by Ex Libris* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse : <http://virtuoseplus.univ->

paris3.fr/primo\_library/libweb/action/dlDisplay.do?vid=default&afterPDS=true&docId  
=

*PSL-Explore* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse :  
[http://beryl.univ-  
psl.fr/primo\\_library/libweb/action/dlDisplay.do?vid=33PSL\\_V1&afterPDS=true&docId](http://beryl.univ-psl.fr/primo_library/libweb/action/dlDisplay.do?vid=33PSL_V1&afterPDS=true&docId)  
=

*Recherche - Université de Bordeaux - Documentation* [en ligne]. [s. d.].  
[Consulté le 17 janvier 2016]. Disponible à l'adresse : [https://babordplus.univ-  
bordeaux.fr/](https://babordplus.univ-bordeaux.fr/)

*Recherche simple: Lille 2* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à  
l'adresse : [http://eds.b.ebscohost.com/eds/search/basic?sid=79d78f23-a5c0-4cc0-b3b2-  
5a087feaa17e%40sessionmgr104&vid=0&hid=112](http://eds.b.ebscohost.com/eds/search/basic?sid=79d78f23-a5c0-4cc0-b3b2-5a087feaa17e%40sessionmgr104&vid=0&hid=112)

*Recherche simple: Supernova - Rennes 1* [en ligne]. [s. d.]. [Consulté le 9 mai 2016].  
Disponible à l'adresse : [http://eds.a.ebscohost.com/eds/search/basic?sid=07a23073-  
8c5a-4264-9c84-696ccea2559e%40sessionmgr4003&vid=0&hid=4208](http://eds.a.ebscohost.com/eds/search/basic?sid=07a23073-8c5a-4264-9c84-696ccea2559e%40sessionmgr4003&vid=0&hid=4208)

*Recherche simple: Université Paris Descartes* [en ligne]. [s. d.].  
[Consulté le 9 mai 2016]. Disponible à l'adresse :  
[https://eds.b.ebscohost.com/eds/search/basic?sid=91ed270d-378f-4cfc-95ff-  
ff99d6f0eeb4%40sessionmgr104&vid=0&hid=112](https://eds.b.ebscohost.com/eds/search/basic?sid=91ed270d-378f-4cfc-95ff-ff99d6f0eeb4%40sessionmgr104&vid=0&hid=112)

*Ressources+ - portail des ressources numériques et papier Télécom Bretagne*  
[en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à l'adresse :  
<http://telecombretagne.summon.serialssolutions.com/>

*Service Commun de la Documentation - Paris Diderot* [en ligne]. [s. d.].  
[Consulté le 9 mai 2016]. Disponible à l'adresse : [http://bibliotheque.univ-paris-  
diderot.fr/](http://bibliotheque.univ-paris-diderot.fr/)

*Service Commun de la Documentation Paris Ouest - Accueil* [en ligne]. [s. d.].  
[Consulté le 17 janvier 2016]. Disponible à l'adresse : [http://scd.u-  
paris10.fr/scd/accueil/](http://scd.u-paris10.fr/scd/accueil/)

*Université de Bretagne-Sud - La Bibliothèque universitaire : Service Commun de la  
Documentation* [en ligne]. [s. d.]. [Consulté le 17 janvier 2016]. Disponible à l'adresse :  
[http://www.univ-ubs.fr/la-bibliotheque-universitaire-service-commun-de-la-  
documentation-274349.kjsp](http://www.univ-ubs.fr/la-bibliotheque-universitaire-service-commun-de-la-documentation-274349.kjsp)

*Université de Poitiers - Service Commun de Documentation - Accueil* [en ligne]. [s. d.].  
[Consulté le 9 mai 2016]. Disponible à l'adresse : <http://scd.univ-poitiers.fr/>

*Université Lyon 2 - Bibliothèque électronique* [en ligne]. [s. d.].  
[Consulté le 9 mai 2016]. Disponible à l'adresse : [http://www.univ-  
lyon2.fr/bibliotheques/collections/etu-bibliotheque-electronique-  
426114.kjsp?RH=WWW1010](http://www.univ-lyon2.fr/bibliotheques/collections/etu-bibliotheque-electronique-426114.kjsp?RH=WWW1010)

*UPEC – Bibliothèque* [en ligne]. [s. d.]. [Consulté le 9 mai 2016]. Disponible à  
l'adresse : [http://athena.u-  
pec.fr/primo\\_library/libweb/action/dlDisplay.do?vid=upec&afterPDS=true&docId=](http://athena.upec.fr/primo_library/libweb/action/dlDisplay.do?vid=upec&afterPDS=true&docId=)

## 7. ETUDES DE CAS SUR DES OUTILS SPECIFIQUES

ARNETT, Barbara et FORRESTAL, Valerie. Bridging the Gap from Wikipedia to Scholarly Sources: A Simple Discovery Tool. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 56-68.

BACAULT, Mathieu. *Blacklight : déploiement avec Ruby On Rails* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/3456>

BACAULT, Mathieu. *Blacklight, deuxième partie : les données d'exemplaires* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/2114>

BACAULT, Mathieu. *Blacklight, troisième partie : les services au lecteur* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/2785>

BACAULT, Mathieu. *Outil de découverte libre : un schéma pour y voir plus clair* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4641>

BACAULT, Mathieu. *Présentation de Blacklight* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/1855>

BOYER, Fabrice. Le SGBM et la BCU : « Techpoint et Checkpoint ». Dans : *L'Alambic numérique* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://alambic.hypotheses.org/5635>

BUSSEMEY, Raphaële. Intégration des revues électroniques dans le catalogue. Dans : *L'Alambic numérique* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4656>

BUSSEMEY, Raphaële. *Mise en place de Blacklight : comment ?* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4172>

BUSSEMEY, Raphaële. *Mise en place de Blacklight : pourquoi ?* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://alambic.hypotheses.org/4163>

BUSSEMEY, Raphaële. *Un outil de découverte libre : poster pour les JABES* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4436>

CALVERT, Kristin. Maximizing Academic Library Collections: Measuring Changes in Use Patterns Owing to EBSCO Discovery Service. *College & Research Libraries* [en ligne]. Janvier 2015, Vol. 76, n° 1, p. 81-99. DOI 10.5860/crl.76.1.81

CICCONE, Karen et VICKERY, John. Summon, EBSCO Discovery Service, and Google Scholar: A Comparison of Search Performance Using User Queries. *Evidence Based Library & Information Practice*. Janvier 2015, Vol. 10, n° 1, p. 34-49

COTTA-SCHØNBERG, Michael. *Primo* [en ligne]. 7 novembre 2010. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://hal-hprints.archives-ouvertes.fr/hprints-00531590/document>

COUDRIN, Delphine. BABORD+ : Une interface de recherche documentaire innovante pour les universités bordelaises. *Arabesques*. 2012, n° 66, p. 20

CURMIN, Frédéric et DOUILLY, Laurence. *Une autre interface publique en lien avec Koha Summon à Telecom Bretagne* [en ligne]. Vitré, 16 juin 2015. [Consulté le 10 juillet 2016]. Disponible à l'adresse : [http://koha-fr.org/sites/default/files/interface\\_summon.pdf](http://koha-fr.org/sites/default/files/interface_summon.pdf)

DANIELS, Jeffrey, ROBINSON, Laura et WISHNETSKY, Susan. Results of Web-Scale Discovery: Data, Discussions, and Decisions. *Articles* [en ligne]. Janvier 2013. Disponible à l'adresse : [http://scholarworks.gvsu.edu/library\\_sp/48](http://scholarworks.gvsu.edu/library_sp/48)

DIGBY, Todd et ELFSTRAND, Stephen. Discovering Open Source Discovery: Using VuFind to Create MnPALS Plus. *Computers in Libraries*. Mars 2011, Vol. 31, n° 2, p. 6-10

FAHEY, Sue, GORDON, Shannon et ROSE, Crystal. Seeing Double at Memorial University: Two WorldCat Local Usability Studies. *Partnership: The Canadian Journal of Library and Information Practice and Research; Vol 6, No 2 (2011)* [en ligne]. 2011. Disponible à l'adresse : <https://journal.lib.uoguelph.ca/index.php/perj/article/view/1552>

FU, Li et THOMES, Cynthia. Implementing discipline-specific searches in EBSCO Discovery Service. *New Library World* [en ligne]. Mars 2014, Vol. 115, n° 3/4, p. 102-115. DOI 10.1108/NLW-01-2014-0003

GAINES, David J. et CONDIT FAGAN, Jody. OAIster on EBSCO Discovery Service, FirstSearch, and OAIster.worldcat.org. *Charleston Advisor* [en ligne]. Avril 2015, Vol. 16, n° 4, p. 5-12. DOI 10.5260/chara.16.4.5

GEORGE, Lisa. *OpenEdition enters EBSCO Discovery* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://oep.hypotheses.org/702>

HELLY, Perrine. Catalogue+ : arrêt du service. Dans : *Ubodoc* [en ligne]. 9 juillet 2015. [Consulté le 25 mars 2016]. Disponible à l'adresse : <http://ubodoc.univ-brest.fr/catalogue-arret-du-service/>

HELLY, Perrine. Catalogue ou Catalogue+? les résultats de l'enquête. Dans : *Ubodoc* [en ligne]. 30 juin 2015. [Consulté le 25 mars 2016]. Disponible à l'adresse : <http://ubodoc.univ-brest.fr/catalogue-ou-catalogue-les-resultats-de-lenquete/>

KAUFMANN, Karen, LARSEN, Jeanne et DESALVO, Patricia. Discovering the Discovery Tool: The Introduction and Impact on Research and Instruction at Seminole State College of Florida. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 158-176.

KORNBLAU, Amy I., STRUDWICK, Jane et MILLER, William. How Web-Scale Discovery Changes the Conversation: The Questions Librarians Should Ask Themselves. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 24-42.

KORTEKAAS, Simone et KRAMER, Bianca. Thinking the unthinkable – doing away with the library catalogue: Based on a paper presented by Simone Kortekaas at the 37th UKSG Annual Conference, Harrogate, April 2014. *Insights: the UKSG journal* [en ligne]. Novembre 2014, Vol. 27, n° 3, p. 244-248. DOI 10.1629/2048-7754.174

LAWTON, Aoife. Use of ESBCO Discovery Tool at One University Reveals Increased Use of Electronic Collections but Decreased Use in Circulation of Print Collections. *Evidence Based Library & Information Practice* [en ligne]. Décembre 2015, Vol. 10, n° 4, p. 244-246. DOI 10.5860/crl.76.1.81

LEEBAW, Danya, CONLAN, Brian, GONNERMAN, Kasia, et al. Improving Library Resource Discovery: Exploring the Possibilities of VuFind and Web-Scale Discovery. *Journal of Web Librarianship* [en ligne]. Avril 2013, Vol. 7, n° 2, p. 154-189. DOI 10.1080/19322909.2013.785825

LEGENDRE, Olivier. *L'OPAC avant les RAMEAU* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <https://alambic.hypotheses.org/4496>

MAHONEY, Jessica et LEACH-MURRAY, Susan. Implementation of a Discovery Layer: The Franklin College Experience. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 207-223.

MEADOW, Kelly et MEADOW, James. Search Query Quality and Web-Scale Discovery: A Qualitative and Quantitative Analysis. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 43-55.

MILBERG, Craig I. A Tale of Two Systems: A Case Study on the Implementation of Two Discovery Systems at Davidson College. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 144-157.

MORTIMORE, Jeffrey M., DUNN, Leah McGinnis, MCNAUGHTON, Drew A., et al. Smoke Gets in Your Eyes: Trials and Triumphs of Implementing the Ex Libris Primo Discovery Service in a Small Regional Consortium. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 224-246. ISBN 978-0-415-70666-7

NABIAS, Laurent, JESTIN, André et RENAVILLE, François. *Primo dans le Cloud ?* [en ligne]. Juin 2015. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/183397>

ORDUÑA-MALEA, Enrique, AYLLÓN, Juan M., MARTÍN-MARTÍN, Alberto, et al. Methods for estimating the size of Google Scholar. *Scientometrics* [en ligne]. Septembre 2015, Vol. 104, n° 3, p. 931-949. DOI 10.1007/s11192-015-1614-6

OURS, Michael G. The Evaluation of Discovery Services at Lynchburg College: 2009–2010. Dans : *Discovery tools: the next generation of library research*. London : Routledge/Taylor & Francis, 2014, p. 267-277.

PEEL, Lisa. KOHA/EDS application at César Ritz Colleges. Dans : *Symposium Koha 2016* [en ligne]. Université Jean Moulin Lyon 3, 28 juin 2016. [Consulté le 11 novembre 2016]. Disponible à l'adresse : <https://youtu.be/-T5D48lilqM>

POPOWICH, Sam. Agile open-source discovery : Blacklight with Ebsco Discovery Service. Dans : *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content*. Chicago, IL : ALA Editions, 2016, p. 73-82.

RENAVILLE, François. *D'un Primo en local à un Primo dans le cloud: deux déploiements en deux ans d'une solution discovery à l'Université de Liège* [en ligne]. Novembre 2015. [Consulté le 14 juin 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/handle/2268/188168>

RENAVILLE, François et RICHELLE, Laurence. Déploiement de la solution discovery Primo à l'Université de Liège. *Cahiers de la Documentation = Bladen voor Documentatie*. Décembre 2013, Vol. 67, n° 4, p. 41-53

RENAVILLE, François et RICHELLE, Laurence. *Primo @ ULg: Implémentation de Primo 4.1 à l'ULg* [en ligne]. Juin 2013. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/jspui/handle/2268/150658>

RICHELLE, Laurence, RENAVILLE, François et LHOEST, Rémy. *Primo @ ULg: formation à destination du personnel des Bibliothèques de l'Université de Liège* [en ligne]. Octobre 2013. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://orbi.ulg.ac.be/jspui/handle/2268/154830>

ROUGER, Dominique. Intégration des services web de recherche dans des bases bibliographiques au sein d'un catalogue Koha. Dans : *Symposium Koha 2016* [en ligne]. Université Jean Moulin Lyon 3, 28 juin 2016. Disponible à l'adresse : <https://youtu.be/-T5D48lilqM>

SÉVIGNY, Martin. *Index centraux à l'UdeM: Primo et Primo Central* [en ligne]. Québec, 27 avril 2012. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://fr.slideshare.net/nst2011/index-centraux-ludem-primo-et-primo-central>

SUNG, Aline Dang Van. *OpenEdition fait son entrée dans EBSCO Discovery* [en ligne]. [s. d.]. [Consulté le 1 mai 2016]. Disponible à l'adresse : <http://leo.hypotheses.org/6956>

VAN NOORDEN, Richard. Google Scholar pioneer on search engine's future. *Nature* [en ligne]. Novembre 2014. [Consulté le 11 novembre 2016]. DOI 10.1038/nature.2014.16269

VAUGHAN, Jason. Chapter 2: OCLC WorldCat Local. *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 12-21

VAUGHAN, Jason. Chapter 3: Serials Solutions Summon. *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 22-29

VAUGHAN, Jason. Chapter 4: Ebsco Discovery Services. *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 30-38

VAUGHAN, Jason. Chapter 5: Ex Libris Primo Central. *Library Technology Reports*. Janvier 2011, Vol. 47, n° 1, p. 39-47

WALKER, Shaundra et SIMS, Iyanna L. Implementing a Discovery Tool at Two HBCUs. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 192-206.

WILLIAMS, Sarah C. et FOSTER, Anita K. Promise Fulfilled? An EBSCO Discovery Service Usability Study. *Journal of Web Librarianship* [en ligne]. Juillet 2011, Vol. 5, n° 3, p. 179-198. DOI 10.1080/19322909.2011.597590

WINTER, Joost C. F. de, ZADPOOR, Amir A. et DODOU, Dimitra. The expansion of Google Scholar versus Web of Science: a longitudinal study. *Scientometrics* [en ligne]. Février 2014, Vol. 98, n° 2, p. 1547-1565. DOI 10.1007/s11192-013-1089-2

WROSCH, Jackie, ROGERS-COLLINS, Karen, BARNES, Michael, et al. Search Me: Eastern Michigan University's Journey through the Highs and Lows of Implementing the Summon Discovery Tool. Dans : *Discovery tools: the next generation of library research*. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord : Routledge/Taylor & Francis, 2014, p. 247-266.



## ANNEXES

---

### *Table des annexes*

<b>ANNEXE 1 : TABLE DE CORRESPONDANCE DES CODES DE BIBLIOTHEQUES .....</b>	<b>162</b>
<b>ANNEXE 2 : GRILLE DE RECUEIL D'INFORMATIONS PREALABLE A L'ENTRETIEN AVEC LES BIBLIOTHEQUES .....</b>	<b>163</b>
<b>ANNEXE 3 : GRILLE D'ENTRETIEN AVEC LES BIBLIOTHEQUES SUR LEURS OUTILS DE DECOUVERTE.....</b>	<b>165</b>
<b>ANNEXE 4 : QUESTIONNAIRE UTILISE LORS DES ENTRETIENS AVEC LES FOURNISSEURS.....</b>	<b>168</b>

# **ANNEXE 1 : TABLE DE CORRESPONDANCE DES CODES DES BIBLIOTHEQUES**

Cette annexe est confidentielle.

## ANNEXE 2 : GRILLE DE RECUEIL D'INFORMATIONS PREALABLE A L'ENTRETIEN AVEC LES BIBLIOTHEQUES

Etablissement	
Date d'entretien	
Outils préexistants à l'outil de découverte	
SIGB	
Liste A-Z	
Résolveur de liens	
Erms	
Moteur de recherche fédérée	
Les raisons d'installer l'outil de découverte	
Offrir un point d'accès unique aux collections	
Améliorer la visibilité et l'utilisation des ressources électroniques	
Couverture des données proposée par l'outil de découverte	
Vitesse de recherche et d'affichage des résultats	
Ergonomie de l'outil	
Coût de l'outil	
Autre	
Intégration de l'outil de découverte et de l'opac	
Outil de découverte en interface unique	
Outil de découverte dans onglet spécifique	
Outil de découverte et OPAC en parallèle	
Installation de l'outil de découverte	
En cloud	
Sur serveurs BU	
Les moyens mobilisés	
Coûts d'installation / de paramétrage	
Coûts de renouvellement / d'abonnement	
Coûts de maintenance	
Nombre d'ETP mobilisés pour l'outil de découverte	
Planning	
Durée de la phase de mise en œuvre et paramétrage	
Année de mise en production	
Tests et enquêtes	
Test menés avant l'installation par la bibliothèque	
Test utilisateurs après installation / enquêtes	
Relations avec le fournisseur	
Ecoute des besoins de la bibliothèque	
Développements/ modifications spécifiques à la demande	
Réactivité	
Qualité des réponses et des solutions fournies	

Bilan	
Bilan de l'utilisation de l'outil, le cas échéant par rapport au catalogue public classique	
Impact de l'outil sur la consultation de la documentation numérique	
Le cas échéant, les principales difficultés rencontrées	
Observations	

## ANNEXE 3 : GRILLE D'ENTRETIEN AVEC LES BIBLIOTHEQUES SUR LEURS OUTILS DE DECOUVERTE

<b>Caractéristiques générales de l'outil</b>	
Feuille de style <i>responsive</i> / application mobile / interface mobile	
Traduction en français des interfaces	
Disponibilité d'une documentation à jour	
Intégration dans portail de la bibliothèque	
<b>Métadonnées et indexation</b>	
Intégration des données bibliographiques du catalogue local	
Choix / adaptation des <i>mappings</i>	
Fréquence de la synchronisation avec le SIGB pour les données d'exemplaires	
Ressources indexées : périodiques	
Ressources indexées : ebooks	
Ressources indexées : licences nationales	
Ressources indexées : open access	
Ressources indexées : bibliothèques numériques locales	
Ressources indexées : ressources francophones	
Ressources indexées en texte intégral	
Ressources non indexées	
Fréquence du moissonnage des fournisseurs de données	
Dédoublonnage bibliographique	
Dédoublonnage des autorités	
Qualité des métadonnées	
Intéropérabilité KBART - Bacon	
Observations sur les métadonnées et l'indexation	
<b>Recherche</b>	
Autocomplétion	
Suggestions orthographe / synonymes	
Configuration du moteur d'indexation (mots vides, diacritiques etc.)	
Choix des facettes à afficher	
Facettes à représentation graphique	
Facette base de données / source	
Facette type de document (format)	
Facette sujet	
Facette date publication	
Facette éditeur	
Facette langue	
Facette bibliothèque	
Facette auteur	
Facettes pour exclure	

Observations sur la recherche	
<b>Résultats</b>	
Choix / adaptation des critères de l'algorithme de pertinence	
Mise en avant des résultats du catalogue local	
Mise en avant de résultats francophones	
Affichage distinct en rayon / en ligne	
Résultats frbérés	
Gestion des caractères non latins	
Gestion de la granularité : lien revue > article	
Affichage droits d'accès aux ressources num	
Lien vers le texte intégral dès liste résultats	
Affichage couvertures de livres	
Contenus enrichis (Amazon, Wikipedia etc.)	
Intégration d'outils de bibliométrie	
Observations sur les résultats	
<b>Gestion des accès</b>	
Intégration avec la base de connaissance du résolveur de liens	
Fréquence de la synchronisation avec la base de connaissance du résolveur de liens	
Gestion et signalement des différents types d'accès	
Gestion des accès via proxy	
Gestion de l'authentification Shibboleth / CAS LDAP	
Accès invité	
Gestion de différents systèmes d'authentification coexistants si outil mutualisé	
Observations sur la gestion des accès	
<b>Compte lecteur</b>	
Cryptage des données personnelles	
Compte lecteur : gestion prêts, renouvellements, réservations, demandes magasin	
Dossier de gestion des références biblio	
Tags / commentaires utilisateurs	
Téléchargement / impression / envoi mail des références	
Alertes mail / rss	
Gestion de l'historique de recherche /recherches sauvegardées /liens permanents	
Recommandations	
Fonctionnalités sociales	
Observations sur le compte lecteur	
<b>Réception / utilisation de l'outil</b>	
Etudiants L : Réception / utilisation / retour des formateurs	
Etudiants M/D : Réception / utilisation / retour des formateurs	

Enseignants-chercheurs : Réception / utilisation	
Professionnels : Réception / utilisation / retour des formations professionnelles	
Observations sur la réception et l'utilisation	
Usages professionnels	
Customisation	
Statistiques d'utilisation de l'outil	
Statistiques de consultation de la documentation numérique	
Exploitation des logs	
Statistiques par outil extérieur (Google analytics, piwik)	
Gestion de l'anonymisation données lecteurs	
Observations sur les usages professionnels	
Bilan global	
Impression générale de l'outil	

## ANNEXE 4 : QUESTIONNAIRE UTILISE LORS DES ENTRETIENS AVEC LES FOURNISSEURS

Questions contractuelles concernant les accords commerciaux avec les fournisseurs de données (éditeurs ou agrégateurs) :

1. Y a-t-il des fournisseurs de données avec lesquels vous avez des accords qui limitent l'accès à leurs données dans l'index central uniquement aux établissements abonnés à leur ressource ?

2. Y a-t-il des fournisseurs de données avec lesquels vous avez des accords qui impliquent de pondérer davantage leurs données par rapport à d'autres dans les résultats issus de l'index ?

3. Y a-t-il des fournisseurs de données avec lesquels vous avez des accords exclusifs qui leur interdisent de mettre leurs données à disposition d'outils concurrents ?

4. Les éléments suivants font-ils l'objet de conditions contractuelles dans les accords avec les fournisseurs de données ?

- La qualité minimale des données versées
- Le format et le protocole de versement
- La fréquence des mises à jour
- Le traitement des données dans l'index central après versement (normalisation, dédoublement...)

5. Lorsque pour une ressource il n'existe pas d'accord contractuel avec le fournisseur de données, par quels moyens estimez-vous le taux de couverture de cette ressource dans l'index central ?

Questions techniques concernant le traitement des données versées dans l'index central :

6. Sous quels formats les fournisseurs de données versent-ils leurs données ?

7. Le fichier versé par le fournisseur doit-il contenir des champs obligatoires ? Si oui lesquels ?

8. Effectuez-vous un contrôle lors de la réception des données des fournisseurs pour vérifier la conformité des données versées ?

9. De quelle manière les données des différents fournisseurs sont-elles normalisées dans le schéma de données de l'index central ?

10. Effectuez-vous un dédoublement des notices bibliographiques dans l'index central ? Si oui de quelle manière ?

11. Effectuez-vous un contrôle sur les autorités (sujets et auteurs) dans l'index central ? Si oui de quelle manière ?

Questions techniques concernant l'articulation avec l'environnement informatique des bibliothèques universitaires :

12. En tant que prestataire de service, comment envisagez-vous l'interfaçage de vos outils avec l'environnement informatique des bibliothèques universitaires françaises (réseaux, systèmes d'authentification, logiciels de bibliothèque etc.) ?



13. Y a-t-il des éléments dans cet environnement qui vous posent régulièrement problème pour l'installation ou la maintenance de votre outil de découverte ? Si oui lesquels ?

14. D'après vous, quelles mesures les bibliothèques universitaires pourraient-elles prendre pour améliorer les conditions d'interopérabilité entre leur environnement technique et vos applications ?



## TABLE DES MATIERES

---

<b>SIGLES ET ABBREVIATIONS .....</b>	<b>11</b>
<b>INTRODUCTION.....</b>	<b>13</b>
<i>La nébuleuse discovery .....</i>	<i>13</i>
<i>Une réponse au double défi du web et de la documentation numérique</i>	<i>15</i>
<i>Des attentes initiales très fortes ... et de nombreux questionnements....</i>	<i>16</i>
<i>Méthodologie.....</i>	<i>19</i>
<b>I. CONTEXTE ET ENJEUX.....</b>	<b>21</b>
<b>1. A la recherche d'un outil centralisateur .....</b>	<b>21</b>
A. <i>La multiplication d'outils de signalement et de gestion de ressources numériques .....</i>	<i>21</i>
B. <i>Les moteurs de recherche fédérée .....</i>	<i>22</i>
<b>2. Les grandes catégories d'outils de découverte.....</b>	<b>23</b>
A. <i>Les interfaces de découverte .....</i>	<i>23</i>
B. <i>Les services de découverte à index général centralisé .....</i>	<i>24</i>
C. <i>Des outils mixtes : une interface et un index d'origines différentes .....</i>	<i>25</i>
<b>3. Quel modèle économique pour les services de découverte à index centralisé? .....</b>	<b>26</b>
A. <i>Un marché fortement concurrentiel, évolutif et anglo-saxon.....</i>	<i>26</i>
Concentrations et convergences .....	26
Des produits évolutifs.....	27
Des produits très liés au monde anglo-saxon.....	28
B. <i>Les flux des données .....</i>	<i>29</i>
La contractualisation .....	29
Le cas des bases d'A&I .....	31
L'open access.....	33
C. <i>Le modèle SaaS .....</i>	<i>33</i>
Des services en cloud .....	33
La tarification .....	35
D. <i>La maintenance .....</i>	<i>36</i>
L'écoute et la réactivité .....	37
Les développements et les solutions.....	38
<b>4. Les spécificités du paysage français.....</b>	<b>40</b>

A.	<i>Outils propriétaires, outils libres, outils mutualisés</i> .....	40
B.	<i>En amont de l'installation</i> .....	41
	Les raisons d'installer un outil de découverte .....	41
	Les tests avant la mise en œuvre .....	42
C.	<i>Les besoins en termes de ressources humaines</i> .....	44
	Une forte mobilisation des équipes .....	44
	Des clubs utilisateurs inégalement développés .....	45
<b>II.</b>	<b>LES SERVICES DE DECOUVERTE A L'EPREUVE DE LEUR IMPLEMENTATION</b> .....	<b>47</b>
<b>1.</b>	<b>Appréhender l'index général centralisé</b> .....	<b>47</b>
A.	<i>L'index général centralisé et ses inconnues</i> .....	47
B.	<i>La granularité de l'indexation</i> .....	49
C.	<i>Des variations selon les éditeurs, les langues et les disciplines</i> ..	50
D.	<i>Adapter l'index général centralisé à la bibliothèque</i> .....	52
<b>2.</b>	<b>L'intégration avec les systèmes locaux</b> .....	<b>53</b>
A.	<i>L'intégration avec le SIGB</i> .....	53
	L'intégration des notices bibliographiques .....	53
	La récupération des données d'exemplaires en temps réel .....	56
	Intégration avec des services .....	58
B.	<i>L'intégration avec la(les) base(s) de connaissance</i> .....	59
	L'intégration d'une liste alphabétique de ressources .....	59
	Signalement et gestion des accès .....	61
C.	<i>L'intégration avec les systèmes d'authentification</i> .....	63
D.	<i>L'articulation avec d'autres bases</i> .....	65
	Intégration via OAI-PMH.....	65
	Complémentarité avec un moteur de recherche fédérée, des widgets, des systèmes de recommandation, des ERMS .....	65
<b>3.</b>	<b>Une recherche unifiée ?</b> .....	<b>66</b>
A.	<i>Une recherche unifiée, des métadonnées distinctes</i> .....	66
	Des métadonnées très hétérogènes .....	66
	Le dédoublement .....	69
	La gestion des autorités .....	71
B.	<i>Les facettes</i> .....	71
	Bibliothèque.....	72
	Type de document .....	73
	Sujet .....	74
	Auteur.....	75
	Editeur, date de publication .....	76

Langue .....	76
C. <i>L'algorithme de pertinence</i> .....	77
Données locales .....	78
Données de l'index .....	79
D. <i>Des interfaces multiples</i> .....	81
Interface unique ou interface principale ? .....	81
Une seule recherche, plusieurs affichages possibles .....	81
<b>III. LES OUTILS DE DECOUVERTE ET LEUR UTILISATION .....</b>	<b>85</b>
<b>1. Du côté du public .....</b>	<b>85</b>
A. <i>Quel public cible ?</i> .....	85
B. <i>Outils de découverte et information literacy</i> .....	86
Des fonctionnalités prometteuses pour l'information literacy .....	87
Un gain incontesté : la découverte d'articles .....	89
Des points de vigilance .....	90
Un besoin d'accompagnement renforcé .....	93
C. <i>Une évaluation de l'utilisation encore balbutiante</i> .....	94
Mesures de l'utilisation de l'outil de découverte .....	94
Mesures de l'impact sur la consultation de la documentation numérique .....	96
Enquêtes et tests sur les outils de découverte .....	101
D'autres retours qualitatifs .....	102
<b>2. Du côté des utilisateurs professionnels .....</b>	<b>103</b>
A. <i>Un bilan mitigé</i> .....	103
B. <i>Accompagner l'évolution des pratiques professionnelles</i> .....	105
<b>IV. PERSPECTIVES .....</b>	<b>107</b>
<b>1. Des alternatives aux interfaces commerciales : les outils libres ....</b> .....	<b>107</b>
A. <i>La mise en place</i> .....	107
Blacklight à la Bibliothèque Clermont-Université .....	107
Babord+ .....	109
B. <i>Les moyens mobilisés</i> .....	111
<b>2. Des alternatives aux index centralisés .....</b>	<b>112</b>
A. <i>Paramétrer un résolveur de liens avec Google Scholar</i> .....	112
B. <i>Intégrer des API d'autres bases à un catalogue public</i> .....	113
C. <i>Exploiter les données de Crossref ?</i> .....	115
D. <i>Des corpus et des réseaux : le dispositif CERCLES</i> .....	115
<b>3. Des index de découverte à la curation des données .....</b>	<b>117</b>
A. <i>Des bonnes pratiques à l'Open Discovery Initiative</i> .....	117

Les recommandations de bonnes pratiques .....	117
L'Open Discovery Initiative .....	119
B. <i>Quid d'un index de découverte en open access ?</i> .....	125
C. <i>Des données de qualité pour les outils de découverte</i> .....	126
Couperin .....	127
BACON .....	128
ISTEX .....	131
Le hub de métadonnées .....	133
D. <i>La curation de données au niveau d'une bibliothèque</i> .....	133
<b>CONCLUSION</b> .....	<b>135</b>
<b>SOURCES</b> .....	<b>137</b>
1. <b>Bibliothèques</b> .....	<b>137</b>
2. <b>Experts dans le domaine des métadonnées</b> .....	<b>139</b>
3. <b>Fournisseurs</b> .....	<b>139</b>
<b>BIBLIOGRAPHIE</b> .....	<b>141</b>
1. <b>Informatique documentaire : contexte et évolutions</b> .....	<b>141</b>
2. <b>Normalisation, curation et échanges de métadonnées</b> .....	<b>142</b>
3. <b>Outils de découverte : études générales</b> .....	<b>145</b>
4. <b>Utilisation, <i>Information literacy</i> et études d'utilisabilité</b> .....	<b>149</b>
5. <b>Sitographie des éditeurs d'outils</b> .....	<b>153</b>
6. <b>Sitographie des outils de découverte en BU</b> .....	<b>153</b>
7. <b>Etudes de cas sur des outils spécifiques</b> .....	<b>156</b>
<b>ANNEXES</b> .....	<b>161</b>
<b>TABLE DES MATIERES</b> .....	<b>171</b>