
CATALOGAGE PARTAGÉ ET PRODUCTION DES MÉTADONNÉES COMMUNES AU SEIN DU RÉSEAU SUDOC ET SUDOC-PS :

UN APERÇU DES ORIENTATIONS POSSIBLES

UNE ÉTUDE VISANT À SOUTENIR LE DÉVELOPPEMENT D'UN PROGRAMME
FONCTIONNEL, EN VUE D'UNE RÉNOVATION DES SYSTÈMES EXISTANTS

Maurits van der Graaf
2014-03-31
Pleiade Management and Consultancy BV
Keizersgracht 62
1015 CS Amsterdam
Pays Bas
T: +31 20 488 9397
m.vdgraaf@pleiade.nl
www.pleiade.nl

SOMMAIRE

Sommaire	2
Résumé managérial	5
Management summary	9
1. Introduction : une stratégie Cloud intégrée pour l'ABES	13
1.1 La stratégie cohérente de l'ABES	13
1.2 L'objectif et la position de cette étude	14
2 Le design de l'étude par Pléiade	15
2.1 Les questions centrales	15
2.2 Les trois pistes d'étude	15
3 Le réseau Sudoc	16
3.1 Deux piliers : le catalogage partagé et le catalogue collectif faisant appel aux services à l'échelle nationale	16
3.2 La valeur ajoutée de SUDOC sous-utilisée	16
3.3 Les Chiffres clés et analyse des tendances de Sudoc	17
Les chiffres clés du réseau Sudoc et du SUPEB	17
Données chiffrées relatives à la production.....	18
Analyses des tendances.....	20
4 Les tendances relatives au catalogage et à la production de métadonnées	22
4.1 La majeure partie de la production de métadonnées bibliographiques a été déplacée vers les éditeurs	22
Le rôle des éditeurs	22
Le standard ONIX.....	22
L'exemple de l'éditeur Springer	23
Les métadonnées bibliographiques sont de plus en plus souvent livrées en licence d'accès libre	23
4.2 Les livres électroniques semblent remplacer les livres imprimés dans le champ de la littérature scientifique.....	24
Les grands éditeurs publient tous leurs livres au format électronique	24
Les petites éditeurs tardent à suivre.....	24
La situation actuelle relative aux livres électroniques	24
À quel moment la transition sera-t-elle complète ?	24

4.3 L'évolution des méthodes d'acquisition.....	26
L'acquisition par bouquets et par méthodes semi-automatiques	26
La conséquence sur le processus de catalogage	26
4.4 L'Évolution des fonctions des catalogues	27
La fonction de découverte des catalogues est en passe de diminuer	27
Les autres fonctions des catalogues locaux et le catalogue collectif demeurent	27
Les fonctions des données des catalogues en dehors du catalogue se multiplient.....	27
Le rôle des catalogues dans un monde de livres numériques devient insignifiant.....	28
4.5 Les grands changements dans le monde du catalogage	29
FRBR.....	29
Les règles RDA	30
BIBFRAME et le Web de données.....	32
Les traitements automatiques dans le processus de catalogage.....	32
4.6 Les changements des systèmes de catalogage partagé.....	33
Les menaces planant sur systèmes de catalogage partagé.....	33
Une nouvelle fonction pour les systèmes de catalogage partagé : la manipulation de flux de métadonnées.....	33
4.7 Conclusions	34
L'évolution possible des catalogues suite à l'influence des livres numériques	34
Un pronostic de réseau SUDOC dans 2025	35
5. Analyse et synthèse des études de cas.....	37
5.1 Les systèmes qui migreront vers le Cloud	37
5.2 Les systèmes qui migreront vers un système auto-développé.....	41
5.3 Les autres études de cas	43
5.4 Conclusions basées sur les études de cas	46
6. Les scénarios possibles pour le Sudoc	47
6.1 Les éléments moteurs du changement de mode de catalogage	47
6.2 Les conditions préalables	47
Les résultats des entretiens avec les partenaires français.....	47
Les conditions résultant des tendances relatives au catalogage et à la production de métadonnées.....	48
6.3 Les scénarios considérés pour le Sudoc	49

Une intégration complète à un système de Cloud	49
Un scénario d'alignement sur un ou plusieurs systèmes de Cloud	50
Un système auto-développé basé sur Open Linked Data	51
6.4 Les considérations issues des scénarios et des programmes de fonctionnalités pour le dialogue compétitif	52
7. Recommandations	54
Annexe A: Participants à l'étude	56
Annexe B : La formation de réseau Sudoc en cas d'une transition vers un nouveau système. 59	
La formation actuelle de réseau Sudoc	59
La formation en cas d'une transition vers un nouveau système	60
Figure 1 La stratégie cohérente de l'ABES et le positionnement de cette étude	13
Figure 2 Le catalogage partagé : les chargements et les créations nouvelles	18
Figure 3 Les tendances relatives aux documents électroniques et physiques	20
Figure 4 Les évolutions des documents relatives aux langues	20
Figure 5 Évolution possible des catalogues	34
Figure 6 Un schéma du catalogue collectif des bibliothèques publiques aux Pays-Bas	44
Figure 7 Un schéma du projet LIBOS	45
Figure 8 Un schéma d'une intégration complète au système de Cloud	49
Figure 9 Un schéma du scénario d'alignement sur un ou plusieurs systèmes de Cloud	50
Figure 10 Un schéma du scénario d'un système auto-développé basé sur <i>Open Linked Data</i>	51
Figure 11 Le calendrier des dialogues compétitifs et les projets préparatoires/ Intermédiaire recommandés	54
Table 1 Les chiffres clés du réseau Sudoc et du SUPEB	17
Table 2 La distribution de la création des nouvelles notices bibliographiques parmi les participants au réseau Sudoc	19
Table 3 les études de cas qui migreront vers le Cloud	40
Table 4 les études de cas qui migreront vers un système auto-développé	42

RÉSUMÉ MANAGÉRIAL

LE RÉSEAU SUDOC

Le système Sudoc /Sudoc -PS est un système pour le catalogage partagé et le catalogue collectif qui sert à plusieurs services bibliothécaires à l'échelle nationale. Le réseau Sudoc est utilisé par les bibliothèques universitaires en France pour toutes les collections et par 2000 autres bibliothèques pour les collections de périodiques (via 32 centres régionaux). Le réseau Sudoc et ses systèmes et services sont gérés par l'ABES - l'Agence bibliographique de l'enseignement supérieur.

LE SYSTÈME DE GESTION BIBLIOTHÉCAIRE PRENANT LA DIRECTION DU CLOUD

Plusieurs participants du réseau Sudoc doivent remplacer leur système de gestion (SGB) et envisagent une migration vers un système de gestion bibliothécaire mutualisé (SGBm) dans le Cloud. L'ABES gère cette initiative, qui va aboutir à un dialogue compétitif commençant en septembre 2014.

L'OBJECTIF DE CETTE ÉTUDE

Si les participants au réseau Sudoc migrent vers un système de Cloud, le système Sudoc, doit-il également migrer? Cette question devient urgente au moment où le système existant arrivera à la fin de son cycle de vie. Cette étude se concentre sur les possibilités d'une migration du système Sudoc/Sudoc-PS vers un système mutualisé dans le Cloud, et porte un éclairage sur les possibilités d'implémenter une migration du système Sudoc/Sudoc-PS au moyen d'un inventaire des expériences et des développements d'autres systèmes similaires. Cette étude constitue la suite logique de l'étude d'impact du système mutualisé de gestion de bibliothèques, qui a été publiée en 2013.

LES TROIS PISTES SUIVIES PAR L'ÉTUDE

L'étude a suivi trois pistes : (1) les études de cas portant sur les systèmes de catalogage partagé et les systèmes de catalogues collectifs; (2) les entretiens avec les experts portant sur les sujets spécialisés; et (3) les entretiens avec les partenaires de l'infrastructure nationale des bibliothèques françaises. L'étude est suivie par un groupe de travail constitué de représentants de l'ABES et de participants au réseau Sudoc. Enfin les résultats de l'étude ont occasionné plusieurs entretiens avec les experts au sein de l'ABES.

UNE ANALYSE DU RÉSEAU SUDOC

Une analyse du réseau Sudoc a mené aux observations suivantes :

- La production des nouvelles notices bibliographiques par an sont au nombre d'environ 700.000 : environ 37% sont importées au moyen d'un chargement provenant des éditeurs, environ 34% sont créées par les participants au réseau, et environ 29% sont récupérées et adaptées « titre par titre » par les catalogueurs d'autres systèmes (comme WorldCat).
- La valeur ajoutée du réseau Sudoc est généralement sous-utilisée : plus de 80 % des participants récupèrent uniquement les modifications des notices de catalogage qu'ils ont opérées eux-mêmes.
- Au cours des dernières années, les ajouts de documents électroniques ont exponentiellement augmenté, au détriment des ajouts de documents physiques et constituent actuellement environ 30 % du nombre total d'ajouts.

LES ÉLÉMENTS MOTEURS DU CHANGEMENT DE MODE DE CATALOGAGE

Les entretiens avec les spécialistes du monde des éditeurs et bibliothèques ont identifié les éléments moteurs du changement de mode de catalogage :

- La grande majorité des éditeurs produisent les métadonnées et livrent ces métadonnées aux bibliothèques. Par conséquent, les systèmes de catalogage partagé s'orientent vers la possibilité de manipulation des flux de métadonnées provenant des éditeurs.
- Les méthodes d'acquisition des bibliothèques changent profondément. La méthode traditionnelle - l'acquisition par sélection manuelle des titres - décline et sera partiellement remplacée par l'acquisition par lots ou bouquets de livres électroniques, ou par une méthode d'acquisition semi-automatique : *patron-driven acquisition* ou *demand-driven acquisition*. Le phénomène d'acquisitions par bouquets et les livraisons simultanées de métadonnées par les éditeurs rendra, en grande partie, le catalogage manuel superflu.
- Les livres numériques vont remplacer la plupart des livres scientifiques imprimés à moyen terme. La transition dans le domaine des sciences exactes avancera vite et sera peut-être quasi-complète. La transition dans le domaine des sciences humaines et sociales avancera plus lentement et ne sera probablement pas complète.
- L'importance des catalogues des bibliothèques comme outil de découverte a déjà diminué et diminuera encore en raison de l'existence d'outils de découverte qui peuvent utiliser le texte intégral des publications numériques.
- Les règles de catalogage et le format bibliographique seront adaptés en réponse aux exigences d'Internet (FRBR et les règles RDA) et de l'évolution vers le Web de données (le développement du nouveau format bibliographique BIBFRAME). En outre, les méthodes de catalogage utilisent de plus en plus souvent les traitements automatiques dans le processus de catalogage.

Ces tendances ont mené à deux conclusions :

- On prévoit deux vitesses de transition vers les publications numériques :
 - Dans le domaine des sciences exactes, cette transition avancera rapidement et sera probablement quasi complète.
 - Dans le domaine des sciences humaines et sociales, cette transition se fera plus lentement et ne sera pas complète, on envisage un certain rôle dévolu aux livres imprimés.
- Un pronostic portant sur le réseau Sudoc en 2025 - basé sur les éléments moteurs du changement évoqué ci-dessus - montre que 70% à 80% des métadonnées proviendront des éditeurs et seront importées dans le système, parfois avec les données « holdings » et avec les codes des notices d'autorités internationales (à titre de comparaison : 37% en 2013).

LES ÉTUDES DE CAS

Dix études de cas ont été menées dans le cadre de cette étude : cinq études de cas concernent une migration vers le Cloud et deux études de cas concernent une migration vers un système auto-développé, alors que les trois autres études illustrent d'autres aspects. En se basant sur ces études de cas, les fonctionnalités des ressources partagées peuvent être réparties en 4 catégories distinctes : (1) les fonctionnalités concernant le système mutualisé de gestion de bibliothèques; (2) les fonctionnalités concernant le catalogage partagé; (3) les fonctionnalités concernant les services communs basés sur des métadonnées du catalogage collectif; (4) les fonctionnalités concernant la gestion partagée de ressources électroniques.

Les cinq études de cas relatives à la migration vers un système de Cloud montrent que BIBSYS en Norvège est le seul projet qui envisage une migration de l'ensemble des 4 catégories de fonctionnalités vers le même système de Cloud. Les bibliothèques norvégiennes utilisent d'ailleurs déjà un système de gestion mutualisé depuis 15 ans. Les bibliothèques universitaires des Pays-Bas sont les seules qui prennent la direction d'une intégration complète de leur système de catalogage partagé et de leur catalogue collectif sur la plate-forme WorldShare. Orbis Cascade continue d'utiliser WorldCat pour son système de catalogage partagé, malgré sa migration vers le système d'Alma. Les métadonnées seront transférées vers une base bibliographique au sein du système d'Alma, qui sert aux services du consortium Orbis Cascade, comme par exemple le PEB. Les autres projets (CSUC, CIB) envisagent un environnement de catalogage partagé qui se trouve dans le système de Cloud, connecté aux flux de métadonnées internationaux, mais avec une identité propre.

Les études de cas relatives à l'auto-développement de systèmes de catalogage partagé portent sur le système suédois LIBRIS XL (qui utilise un format bibliographique en RDF auto-développé et est basé sur un logiciel auto-développé) et le système finlandais planifié, qui utilisera le logiciel issu de l'OPEN SOURCE, mais qui maintiendra le format Marc 21 comme standard internationalement accepté.

Deux aspects pertinents sont illustrés par les autres études de cas :

- L'étude de cas de Libraries Australia montre que les bibliothèques qui migrent leurs SGB vers un système de Cloud choisissent dans la majorité des cas d'utiliser également le système de catalogage partagé de ce système de Cloud. Pour le catalogue collectif d'Australie, cela signifie qu'on doit importer les métadonnées de ces bibliothèques. Généralement, ce phénomène illustre le fait que les systèmes de Cloud commerciaux et globaux deviennent des concurrents des systèmes nationaux ou régionaux de catalogage partagé et des catalogues collectifs.
- Le nouveau catalogue collectif des bibliothèques des Pays-Bas illustre les possibilités d'Open Linked Data pour une plate-forme de découverte.

LES SCÉNARIOS RETENUS POUR LE SUDOC

Les études de cas ont mené à des scénarios qui sont jugés applicables au système Sudoc/Sudoc-PS :

Un scénario d'alignement sur un ou plusieurs systèmes de Cloud existants : Un scénario possible pour le Sudoc serait un alignement sur un ou plusieurs systèmes de Cloud. Ce scénario est inspiré par le projet CIB Allemagne. Ce scénario comporte les données suivantes :

- La création d'un espace de données français (*data space*). L'idée est qu'il existera un *data space* dans chaque système de Cloud utilisé en France et que ces multiples *data spaces* seront synchronisés. L'objectif du *data space* est de servir au catalogage partagé et aux flux de métadonnées, surtout aux flux internationaux
- La création d'un *data pool* français séparé des systèmes de Cloud, c'est-à-dire une base de données pour les services à l'échelle nationale : le catalogue collectif, le système de PEB, les services de gestion des collections (*collection management*) et les autres services à l'échelle nationale qui pourraient se développer dans le futur.

Un système auto-développé basé sur *Open Linked Data* : Un autre scénario possible pour le Sudoc envisage une migration vers le Web de données (*Open Linked Data* ; RDF). Ce scénario est inspiré par le développement de LIBRIS XL en Suède. Ce scénario implique l'auto-développement du logiciel pour cette plate-forme (par un prestataire) et l'attente du développement du format BIBFRAME de la *Library of Congress*. De plus, ce scénario implique : (1) le développement d'un nouveau circuit de travail pour le catalogage (*'from cataloguing to cataloging'*) utilisant les avantages de la structure

RDF; (2) l'auto-développement des services comme le PEB utilisant les données au format RDF; et (3) le développement de méthodes permettant le partage des ressources bibliographiques à l'échelle mondiale.

LES CONSIDÉRATIONS ET LES RECOMMANDATIONS

L'ABES a décidé de dissocier le dialogue compétitif pour le SGBm et le dialogue compétitif pour le système Sudoc/Sudoc-PS. Cela signifie que le dialogue compétitif au sujet du SGBm commencera en septembre 2014, et que le dialogue portant sur le système Sudoc/Sudoc-PS commencera en automne 2015. Ce calendrier laisse de la marge pour une étude ultérieure sur les deux scénarios retenus, avec pour objectif de faire un choix. Ce choix permettra d'affiner le programme fonctionnel pour le dialogue compétitif.

En fonction du calendrier, on peut s'attendre à ce que l'implémentation du nouveau système Sudoc/Sudoc-PS aboutisse au début de l'année 2019. En considérant les éléments moteurs du changement de mode de catalogage, l'immobilité durant la période 2014-2018 ne sera pas une option envisageable pour le réseau Sudoc. Il y a trois projets préparatoires au nouveau système qui sont recommandés pour cette période intermédiaire :

- L'élargissement par l'ABES de l'importation, de la gestion et de l'enrichissement des métadonnées des ressources électroniques provenant des éditeurs dans le système existant;
- La rénovation de la base XML prenant la direction d'un data pool national qui servira aux services à l'échelle nationale comme par exemple le PEB;
- L'avancement du développement du système IDREF prenant la direction d'un système indépendant, et basé sur *Open Linked Data*.

MANAGEMENT SUMMARY

THE SUDOC NETWORK

The Sudoc/Sudoc-PS system is a system for shared cataloging and for a union catalogue that is used for a number of national library services. The Sudoc network is used by the university libraries in France for their entire collections and by another 2000 libraries for their journal collections (via 32 regional centers). The Sudoc network and their systems and services are managed by ABES – the French bibliographic agency for higher education.

LIBRARY MANAGEMENT SYSTEMS TO THE CLOUD

Several participants of the Sudoc network must replace their management system (LMS) and consider migrating to a shared library management system in the cloud. ABES supports this initiative, which will lead to a competitive dialogue beginning September 2014.

THE OBJECTIVE OF THIS STUDY

If participants of the Sudoc network migrate to a cloud system, should the Sudoc system also migrate to the cloud? This question is more urgent as the existing system is nearing the end of its life cycle. This study focuses on the possibilities of a migration of the Sudoc/Sudoc-PS system to a shared system in the cloud by making an inventory of experiences and developments of other similar systems. This study builds on the 'impact study' by ABES, which has been published in 2013.

THE THREE TRACKS OF THE STUDY

The study has followed three tracks: (1) case studies on systems shared cataloging / union catalogues (2) interviews with experts on specialized topics, and (3) interviews with partners of the French national library infrastructure. The study was supported by a working group consisting of representatives of ABES and participants of the Sudoc network. In addition, the results of this study have led to a number of interviews with experts within ABES.

AN ANALYSIS OF NETWORK SUDOC

An analysis of Sudoc network led to the following observations:

- Approximately 700,000 new bibliographic records are produced annually: about 37% of these new records are imported to the system and produced by publishers, about 34% is created by the Sudoc participants and about 29% is retrieved and adapted 'title by title' from other systems (such as WorldCat) by the cataloguers from the SUDOC network.
- The added value of Sudoc network is underused as over 80 % of participants recover only those adaptations of the bibliographic records that they adapted themselves.
- The influx of electronic documents grows exponentially in recent years (replacing physical documents) and is currently about 30% of the total number of added documents.

THE DRIVERS OF CHANGE OF CATALOGING

Numerous interviews with specialists from the world of publishers and libraries have identified the following drivers of change mode of cataloging:

- The vast majority of publishers produce metadata and deliver these to the libraries. Consequently, the shared cataloging systems are changing into shared systems for the management of metadata flows from the publishers.

- The acquisition methods libraries are changing profoundly. The traditional method - the acquisition by manual selection of titles – is declining and will be partially replaced by the acquisition of packages of electronic books and/or by semi-automatic acquisition methods called patron-driven acquisition or demand-driven acquisition. A combination of the acquisition of packages and the simultaneous deliveries of metadata by the publishers will make manual cataloguing largely superfluous.
- Electronic books will replace most printed scientific books in the medium term. The transition in the field of exact sciences appears to advance rather quickly and will probably leave little room for printed books. However, the transition in the field of the humanities and social sciences will advance more slowly and probably will not completely replace the printed books.
- The importance of library catalogues as a discovery tool has already declined and will decline further due to the existence of discovery tools that can make use of the full text from digital publications in that search engines.
- The cataloguing rules and the bibliographic format are adapted in response to the requirements of Internet (FRBR and RDA rules) and to the requirements of the Web of Data (the development of new bibliographic format BIBFRAME). In addition, increasingly, automatic processing is used in the cataloguing process.

These trends have led to two conclusions:

- A two-speed transition into digital publications:
 - In the exact sciences, this transition progresses rapidly and is likely to be virtually complete.
 - In the area of humanities and social sciences, this transition will be slower and will probably see a continuing role for printed books.
- A prognosis for the Sudoc network in 2025 - based on the drivers of change mentioned above - shows that 70 % to 80 % of the metadata will come from publishers (compared to 37% in 2013) and will be imported into the system, sometimes with "holdings" data and with the codes of international authority files provided by the publisher).

THE CASE STUDIES

Ten case studies have been conducted for this study: five case studies involve a migration to the cloud and two case studies relate to a migration to a self-developed system, while the other three studies illustrate other aspects. Based on these case studies, the functionalities with regard to shared library resources can be divided into four distinct categories: (1) functionalities with regard to the (shared) library management systems (2) functionalities with regard to shared cataloging (3) functionalities with regard to the union catalogue and its services (4) functionalities with regard to the shared management of electronic resources.

From the five case studies relating to a migration to cloud systems, the case study of BIBSYS shows that Norway is the only case, which envisages a migration of all four categories of functionalities to the same system in the cloud. However, Norwegian libraries already used a shared library management system for 15 years. The university libraries in the Netherlands formed only case study that will lead to a full integration of their shared cataloging system and their collective catalogue in the WorldShare platform.

The other case studies do not migrate all four categories of functionality to the same system. Orbis Cascade will continue using WorldCat as their shared cataloging system, while they migrate the other categories of functionality to the Alma system. The metadata records produced by their cataloguing

efforts will be transferred to a bibliographic database within the Alma system that serves their services such as ILL. The other projects (CCUC and CIB) envisage a shared cataloging environment that is in the cloud system, connected to the flow of international metadata, but with their own identity.

The two case studies that foresee the self-development of shared cataloging systems consist of the Swedish LIBRIS XL system (which uses a self-developed bibliographic format using the RDF structure and a self-developed software system) and the planned Finnish system (that will make use of open source software for its system and will keep the Marc 21 format as this is the accepted international bibliographic standard).

Two relevant aspects are illustrated by the other case studies:

- The case study on Libraries Australia shows that libraries that have migrated their library management system to a shared system in the cloud in most cases also choose to use the system for their shared cataloging. This means for the union catalogue of Australia that they must import the metadata of these libraries to keep the union catalogue comprehensive. In addition, it illustrates the fact that the Cloud systems of commercial, global vendors become competitors of national or regional systems for shared cataloging and union catalogues.
- The new union catalogue of public libraries in the Netherlands illustrates the many possibilities of Open Linked Data for discovery.

THE SCENARIOS RETAINED FOR SUDOC

The case studies have led to two scenarios that are considered feasible for the Sudoc/Sudoc-PS system:

A scenario of an alignment with one or more existing systems Cloud: A possible scenario for the Sudoc would align with one or more systems Cloud. This scenario is inspired by the CIB Germany project. This scenario includes the following data:

- The creation of a French data space in the cloud: the idea is that there will be a French data space in each cloud system used by French libraries and that these multiple data spaces will be synchronized. The data space is used for the purposes of shared cataloging and shared management of (international) metadata flows.
- The creation of a French data pool outside cloud: this data pool will consist of the union catalogue, the various services that are based on the union catalogue, such as interlibrary loan, collection management and other services at the national level that might be developed in the future.

A self-developed system based on Open Linked Data: this other possible scenario for Sudoc considers a migration to the Web of Data (Open Linked Data, RDF). This scenario is inspired by the development of the LIBRIS XL in Sweden. This scenario involves the self-development of software for this platform (by a provider) and waiting for the finalisation of the development of the BIBFRAME format by the Library of Congress. This scenario involves (1) the development of a new workflow for cataloging ('from cataloging to catalinking') using the possibilities of the RDF structure, (2) the self-development of services such as interlibrary loan using the RDF data structure and (3) the development of methods for sharing bibliographic resources worldwide.

CONSIDERATIONS AND RECOMMENDATIONS

ABES decided to separate the competitive dialogue for a shared library management system and the competitive dialogue for the Sudoc/Sudoc-PS system. This means that the competitive dialogue for the shared library management system will start in September 2014 and the competitive dialogue for the Sudoc/Sudoc-PS system in the second half of 2015. This time schedule leaves room for a further study on the two above-mentioned scenarios with the purpose of making a choice between the two. Such a choice will facilitate the finalisation of the functional specifications for the competitive dialogue.

Based on the time schedule it can be expected that the implementation of a new Sudoc/Sudoc-PS system will be completed at the start of 2019. Within an eye on the drivers of change of cataloging, a standstill during the period of 2014- 2018 cannot be an option for the SUDOC network. Therefore, the following projects are recommended for this interim period:

- An extension of the import, management and enrichment by ABES of metadata of electronic resources from publishers in the present SUDOC system
- The development of the present XML database - managed by ABES - into a national datapool that will serve as a base for the national library services in France.
- The development of the authority system of the SUDOC network (IDREF) into an independent system that is based on Open Linked Data.

1. INTRODUCTION : UNE STRATÉGIE CLOUD INTÉGRÉE POUR L'ABES

1.1 LA STRATÉGIE COHÉRENTE DE L'ABES

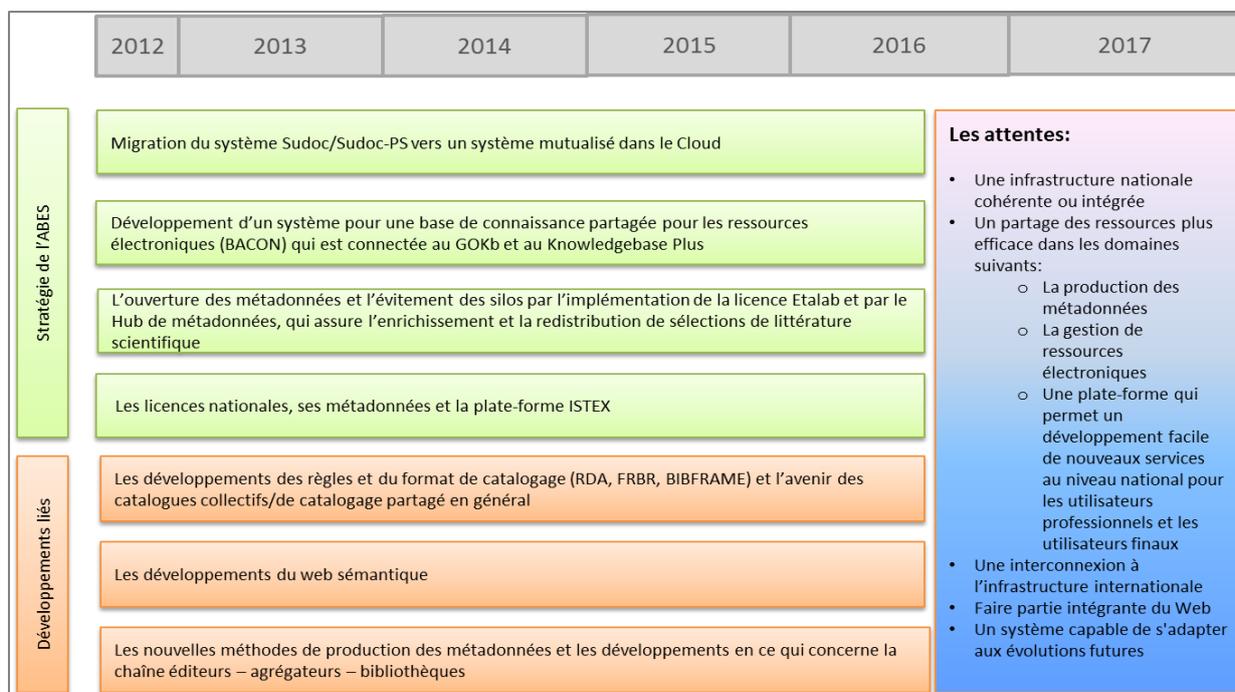


FIGURE 1 LA STRATÉGIE COHÉRENTE DE L'ABES ET LE POSITIONNEMENT DE CETTE ÉTUDE

La figure 1 présente un schéma de la stratégie cohérente de l'ABES et le positionnement de cette étude :

- L'ABES envisage une migration du système Sudoc /Sudoc-PS vers un système mutualisé, qui sera probablement dans le Cloud.
- La stratégie de l'ABES est orientée vers l'ouverture des métadonnées et l'évitement des silos par l'implémentation de la licence Etalab et par le Hub de métadonnées, qui assure l'enrichissement et la redistribution de sélections de la littérature scientifique.
- L'ABES travaille au développement d'un système pour une base de connaissances partagée pour les ressources électroniques (BACON). Il est envisagé que BACON soit connecté à GOKb et à Knowledgebase Plus.
- La stratégie de l'ABES inclut les licences nationales, les métadonnées de ces mêmes licences nationales, et la plate-forme ISTEK.

Le futur du système Sudoc/Sudoc-PS sera étroitement lié :

- aux développements des règles de catalogage (RDA) et du format de catalogage (FRBR, BIBFRAME) et à l'avenir des catalogues collectifs et du catalogage partagé en général
- aux développements du web sémantique

- aux nouvelles méthodes de production des métadonnées et aux développements relatifs à la chaîne éditeurs – agrégateurs – bibliothèques.

Les attentes d'une migration éventuelle vers un système mutualisé dans le Cloud sont :

- Une infrastructure nationale cohérente et intégrée
- Un partage des ressources plus efficace dans les domaines suivants :
 - La production des métadonnées
 - La gestion de ressources électroniques
 - Une plate-forme qui permet le développement facile de nouveaux services au niveau national pour les utilisateurs professionnels et les utilisateurs finaux
- Une interconnexion avec l'infrastructure internationale
- Faire partie intégrante du Web
- Un système capable de s'adapter aux évolutions futures

1.2 L'OBJECTIF ET LA POSITION DE CETTE ÉTUDE

Cette étude se concentre sur les possibilités d'une migration du système Sudoc/Sudoc-PS vers un système mutualisé dans le Cloud. Elle apporte un éclairage quant aux possibilités d'implémentation d'une migration du système Sudoc /Sudoc-PS. Ceci au moyen d'un inventaire des expériences réalisées et des développements d'autres systèmes similaires.

Cette étude constitue la suite logique de [l'étude d'impact](#) du système de gestion de bibliothèque mutualisé. L'étude d'impact a abouti à la formation de deux groupes de travail, qui ont travaillé d'octobre 2013 à avril 2014. Un groupe de travail avait pour but le développement et l'affinage d'un programme de fonctionnalités destiné aux besoins locaux, l'autre groupe de travail avait pour but le développement et l'affinage du programme de fonctionnalités pour les besoins relatifs au catalogage partagé et au catalogue collectif de réseau Sudoc. La présente étude a été conduite afin de soutenir ce dernier groupe de travail.

2 LE DESIGN DE L'ÉTUDE PAR PLÉIADE

2.1 LES QUESTIONS CENTRALES

Les questions centrales de l'étude réalisée par Pléiade sont les suivantes :

- Un inventaire des tendances et des développements dans les domaines du catalogage partagé, des catalogues collectifs et des services à l'échelle nationale basés sur les catalogues collectifs, comme par exemple le prêt entre bibliothèques et la gestion de collections.
- Un inventaire des tendances et des développements dans les domaines de la production, des formats, des licences et de la distribution en ce qui concerne les catalogues et les métadonnées liées aux publications scientifiques
- L'identification des possibilités permettant de modifier le système du Sudoc dans le cadre de l'infrastructure nationale des bibliothèques françaises.

2.2 LES TROIS PISTES D'ÉTUDE

L'étude a suivi trois pistes :

- Les études de cas sur les systèmes de catalogage partagé/de catalogues collectifs (chapitre 5)
- Une série d'entretiens avec les experts, portant sur les sujets spécialisés (comme les méthodes nouvelles de production des métadonnées, le Web sémantique, RDA ; chapitre 4)
- Une série d'entretiens avec les partenaires de l'infrastructure nationale des bibliothèques françaises (section 6.2).

L'étude est accompagnée par le groupe de travail et les résultats obtenus sont discutés au sein de ce groupe de travail, dans le cadre de cinq ateliers. Les résultats des études de cas ont occasionné plusieurs entretiens avec les experts au sein de l'ABES.

L'étude est menée par Maurits van der Graaf sous la supervision de Jean Bernon de l'ABES. L'annexe A comporte une liste de tous les interlocuteurs participant à l'étude.

3 LE RÉSEAU SUDOC

3.1 DEUX PILIERS : LE CATALOGAGE PARTAGÉ ET LE CATALOGUE COLLECTIF FAISANT APPEL AUX SERVICES À L'ÉCHELLE NATIONALE

La valeur ajoutée du réseau Sudoc repose sur deux piliers :

- Le catalogage partagé, qui fait partie intégrante des circuits de travail des bibliothèques déployées. Le catalogage partagé permet aux bibliothèques de gagner grandement en efficacité dans le processus de catalogage, ainsi qu'en qualité des notices bibliographiques.
- Le catalogue collectif, inclut les métadonnées des catalogues des bibliothèques déployées et celles des catalogues des bibliothèques non-déployées (qui peuvent ajouter leurs données exemplaires grâce à l'outil Colodus, uniquement pour des périodiques ; Colodus sert également aux bibliothèques déployées pour mettre à jour les données exemplaires des monographies autant que des périodiques. Le catalogue collectif, qui est lui-même un outil de découverte pour les utilisateurs finaux, constitue aussi la base des services à l'échelle nationale : le système de prêt entre bibliothèques SUPEB, l'outil de gestion des collections PÉRISCOPE et les services destinés aux bibliothèques visant à améliorer leurs catalogues locaux sont les plus importants d'entre eux.

3.2 LA VALEUR AJOUTÉE DE SUDOC SOUS-UTILISÉE

Le système Sudoc est sous-utilisé par les participants eux-mêmes et ce de différentes manières, en ce qui concerne la mise à jour des systèmes locaux :

- La fréquence :
 - 144 systèmes locaux de gestion des bibliothèques sont mis à jour chaque nuit ;
 - 20 systèmes locaux de gestion des bibliothèques sont mis à jour chaque semaine.
 - autrement dit : environ 10 % de participants sont en retard en moyenne de quatre jours, pour ce qui est de leur mise à jour.
- Les données :
 - 132 bibliothèques déployées récupèrent seulement les modifications des notices de catalogage qu'elles ont elles-mêmes opérées ;
 - 32 bibliothèques déployées récupèrent toutes les modifications de notices pertinentes pour leur catalogue.
 - autrement dit : plus de 80 % de participants au système Sudoc n'utilisent pas complètement la valeur ajoutée offerte par le catalogage partagé en ignorant la valeur ajoutée à leurs notices lorsque ces ajouts sont effectués par d'autres participants au réseau. Les raisons évoquées sont (1) limiter le volume des fichiers de mise à jour/ limiter la lourdeur des changements (2) stabiliser la notice dans le catalogue local (3) l'impossibilité de contrôler toutes les notices entrant dans le catalogue local.
- Les notices d'autorités du Sudoc :
 - 37 bibliothèques ne chargent pas les données d'autorités dans leurs systèmes de gestion locaux. 127 bibliothèques chargent les données d'autorités dans leurs systèmes de gestion locaux. Néanmoins, des doutes fondés demeurent sur la capacité de plusieurs systèmes locaux à vraiment tirer parti des notices d'autorités pour les utilisateurs finaux.

3.3 LES CHIFFRES CLÉS ET ANALYSE DES TENDANCES DE SUDOC

LES CHIFFRES CLÉS DU RÉSEAU SUDOC ET DU SUPEB

Le réseau Sudoc		
Bibliothèques déployées	164 systèmes locaux déployés (établissements) avec 1450 bibliothèques (Sudoc)	
Bibliothèques non-déployées	32 centres régionaux cataloguent les périodiques pour 2000 bibliothèques non-déployées en utilisant WinIBW (Sudoc PS); les bibliothèques non déployées peuvent ajouter les données exemplaires elles-mêmes grâce à Colodus (depuis mi-2013)	
Catalogueurs utilisant le système		
Connexions WinIBW (en 2012)	890000	
Nombre de licences WinIBW	3687	
SUPEB (chiffres de 2012)		
Bibliothèques déployées	338	
Bibliothèques chaînées (seulement fournisseur)	176	
Demandes émises	150000	
Demandes satisfaites, total (2012)	134000	
Demandes satisfaites, monographies (2012)	75000	56%
Dont satisfaites par prêt du document	70000	93%
Demandes satisfaites, ressources continues (2012)	59000	44%
Dont satisfaites par reproduction	51000	86%

TABLE 1 LES CHIFFRES CLÉS DU RÉSEAU SUDOC ET DU SUPEB

Les chiffres du réseau Sudoc et le système PEB (SUPEB) sont présentés dans la table 1. On peut en déduire les conclusions suivantes :

- Il y a environ 3000 à 4000 connexions à WinIBW par jour de travail.
- Le système SUPEB peut répondre à 89% de demandes émises.
- Plus de la moitié (56%) des demandes satisfaites concernent les monographies. La grande majorité des demandes de monographies est satisfaite par le prêt du document.
- 44% des demandes satisfaites concernent les ressources continues. La grande majorité des demandes de ressources continues est satisfaite par reproduction.

DONNÉES CHIFFRÉES RELATIVES À LA PRODUCTION

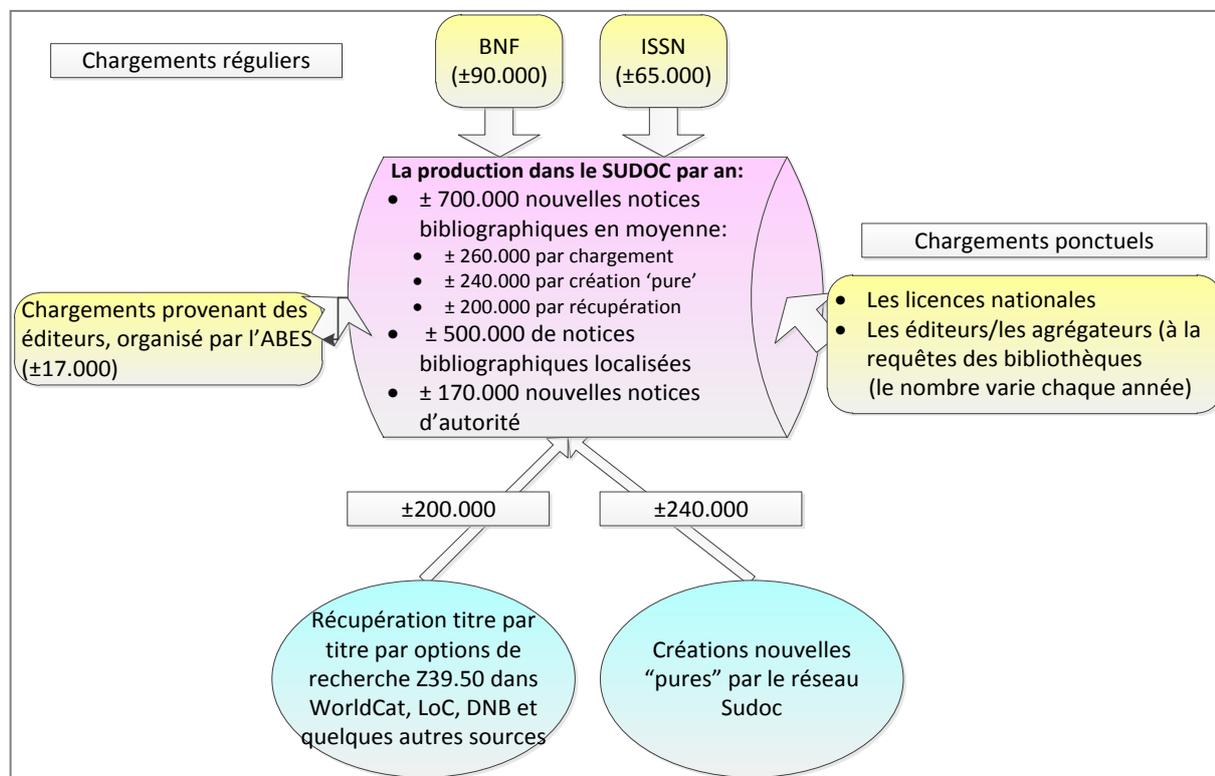


FIGURE 2 LE CATALOGAGE PARTAGÉ : LES CHARGEMENTS ET LES CRÉATIONS NOUVELLES

Dans la figure 2 la production du système de catalogage partagé du Sudoc est esquissée à grands traits. La production annuelle est constituée d'environ 440,000 nouvelles notices bibliographiques, d'environ 500,000 notices bibliographiques localisées et d'environ 170,000 nouvelles notices d'autorités. Cette production annuelle est réalisée en employant trois méthodes :

- Les chargements de nouvelles notices :
 - Les chargements réguliers : provenant de la BnF (notices bibliographiques et d'autorités) ; provenant d' ISSN (notices bibliographiques), provenant de quelques éditeurs (notices bibliographiques).
 - Les chargements ponctuels : les métadonnées relatives aux licences nationales et relatives à certains fonds ou agrégations de métadonnées, provenant des éditeurs ou des agrégateurs, à la requête des bibliothèques.
- La récupération titre par titre par option de recherche Z39.50 : les catalogueurs peuvent formuler une recherche pour récupérer une notice de catalogue d'autres sources. Les autres sources, qui sont connectées par Z39.50, sont : WorldCat, la *Library of Congress*, la *Deutsche Nationalbibliothek* et quelques autres catalogues.
- La création, dite « pure », de nouvelles notices : c'est-à-dire que les catalogueurs du réseau Sudoc créent une nouvelle notice bibliographique ou une nouvelle notice d'autorité.

Qui procède aux récupérations titre par titre des notices bibliographiques ou aux créations dites « pures » de nouvelles notices ? La distribution parmi les participants de réseau Sudoc de cette production des notices bibliographiques est présentée dans table 2. La majeure partie de cette

production est réalisée par un nombre limité de participants : 22 établissements réalisent plus de 60 % de la production annuelle, 55 établissements réalisent plus de 30 % des 40% restants. La majorité des participants (60%, soit 119 établissements) réalise moins de 10 % de la production annuelle. Parmi eux, 40 établissements (dont 26 centres régionaux) ne créent quasiment pas de notices bibliographiques.

Nombre de notices bibliographiques créées total	Pourcentage de la production annuelle totale	Nombre d'établissements	Nombre de notices bibliographiques créées par établissement
245534	60,8%	22	5,000 - 40,000
125298	31,0%	55	1,000 - 5,000
32359	8,0%	79	100 - 1,000
629	0,2%	40	0 - 100
403820		196 ¹	

TABLE 2 LA DISTRIBUTION DE LA CRÉATION DES NOUVELLES NOTICES BIBLIOGRAPHIQUES PARMIS LES PARTICIPANTS AU RÉSEAU SUDOC

¹ Ce total de 196 inclut l'ABES, 30 centres régionaux et 164 bibliothèques déployées (2 centres régionaux n'ont pas travaillé en 2013; un système détaillé a disparu courant 2013).

ANALYSES DES TENDANCES

Les chiffres relatifs à l'évolution du Sudoc pour la période 2010-2013 sont présentés dans la table 3. Deux évolutions pertinentes sont également présentées dans les figures 3 et 4. On peut en tirer les conclusions suivantes :

- Le nombre de nouveaux documents électroniques croît exponentiellement, au détriment des documents physiques.
- Le nombre de nouveaux documents en langue anglaise croît au détriment des documents en langue française et en langues autres que le français et l'anglais.
- Le nombre de recherches couronnées de succès par le biais des interfaces publiques de consultation (PSI) croît depuis 2011.

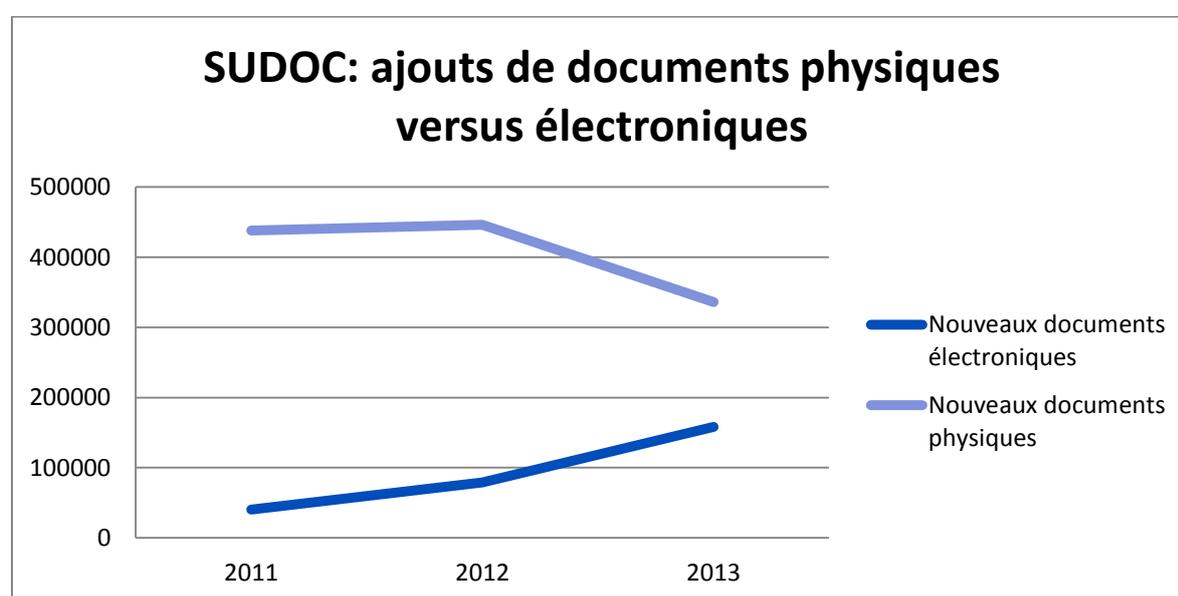


FIGURE 3 LES TENDANCES RELATIVES AUX DOCUMENTS ÉLECTRONIQUES ET PHYSIQUES

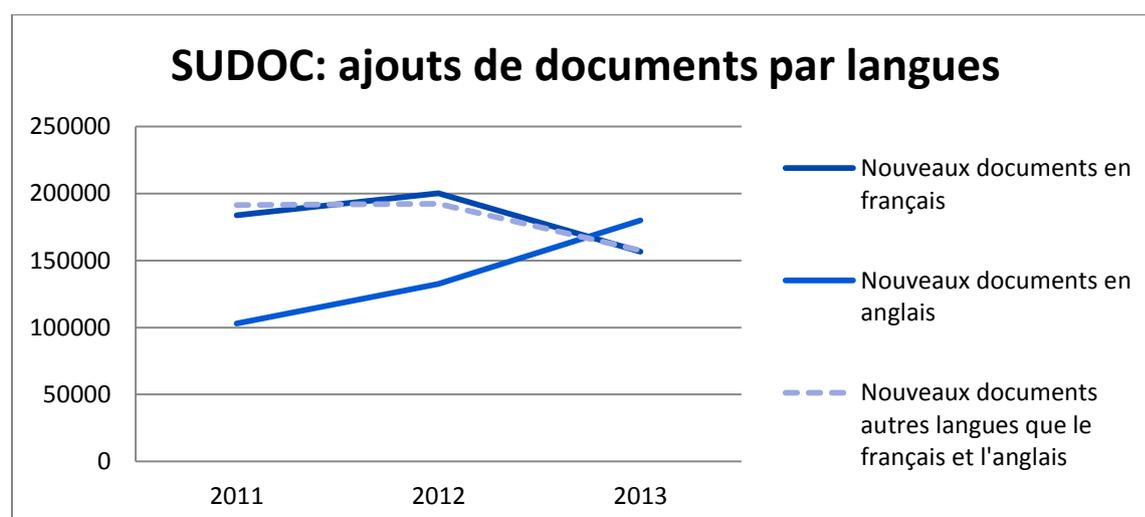


FIGURE 4 LES ÉVOLUTIONS DES DOCUMENTS RELATIVES AUX LANGUES

TABLE 3 CHIFFRES DE LA PRODUCTION SUDOC	2010	2011	2012	2013
Recherches réussies (dans PSI)	17.683.514 ²	13.602.163	13.963.140	14.664.567
Créations de notices bibliographiques (chiffres arrondis)	418400	473.500	467.000	403.800
Créations de notices d'autorité (chiffres arrondis)	164.000	171.700	186.000	175.700
Notices bibliographiques localisées	9.875.421	10.353.501	10.878.572	11.372.620
Nouvelles notices bibliographiques localisées		478.080	525.071	494.048
Documents électroniques	263.390	303.818	382.925	540.947
Nouveaux documents électroniques		40.428	79.107	158.022
Documents physiques	9.612.031	10.049.683	10.495.647	10.831.673
Nouveaux documents physiques		437.652	445.964	336.026
Documents en français	4.579.464	4.763.266	4.963.435	5.120.034
Nouveaux documents en français		183.802	200.169	156.599
Documents en anglais	2.128.140	2.231.193	2.363.809	2.543.643
Nouveaux documents en anglais		103.053	132.616	179.834
Documents autres langues (que les précédentes)	3.167.817	3.359.042	3.551.328	3.708.943
Nouveaux documents autres langues		191.225	192.286	157.615
Documents publiés en France	4.088.376	4.261.441	4.448.288	4.609.674
Nouveaux documents publiés en France		173.065	186.847	161.386
Documents publiés en Allemagne	1.037.561	1.119.344	1.205.312	1.264.979
Nouveaux documents publiés en Allemagne		81.783	85.968	59.667
Documents publiés en UK	784.017	816.025	877.778	1.015.350
Nouveaux documents publiés en UK		32.008	61.753	137.572
Documents publiés aux USA	797.174	829.972	867.582	891.320
Nouveaux documents publiés aux USA		32.798	37.610	23.738
Documents publiés dans d'autres pays (que les précédents)	3.168.293	3.326.719	3.479.612	3.591.297
Nouveaux documents publiés dans d'autres pays		158.426	152.893	111.685

² Les chiffres relatifs à l'année 2010 incluent les robots, les autres chiffres les excluent.

4 LES TENDANCES RELATIVES AU CATALOGAGE ET À LA PRODUCTION DE MÉTADONNÉES

4.1 LA MAJEURE PARTIE DE LA PRODUCTION DE MÉTADONNÉES BIBLIOGRAPHIQUES A ÉTÉ DÉPLACÉE VERS LES ÉDITEURS

LE RÔLE DES ÉDITEURS

La majeure partie de la production de métadonnées bibliographiques a été déplacée vers les éditeurs et a de plus en plus tendance à se déplacer vers le début de la chaîne : les premières métadonnées sont souvent créées lors du processus de soumission de la publication à l'éditeur. On produit les métadonnées au moyen d'une combinaison d'ordinateurs et d'efforts humains.

Un interlocuteur officiant chez un éditeur scientifique distingue les métadonnées que l'on peut qualifier de « structurales » (comme par exemple le titre, l'auteur, les citations, les références, le sommaire, les organisations qui ont financé la recherche etc.) et les métadonnées « sémantiques », qui renseignent sur la signification d'un élément du texte présent sur la notice bibliographique. Les métadonnées « structurales » sont souvent détaillées et souvent au format XML dans les systèmes internes³.

LE STANDARD ONIX

ONIX for BOOKS est un standard pour l'échange des métadonnées dans la chaîne logistique des livres, destiné aux acteurs commerciaux. Ce standard existe depuis 14 ans et il a commencé avec la version 1.0. Actuellement, les versions 2.1 et 3.0 sont les plus utilisées. ONIX est développé et maintenu par EDitEUR, une organisation comportant plus de cent membres, incluant quelques bibliothèques majeures comme la *British Library*. La majorité des membres sont des éditeurs, des distributeurs et des détaillants.

Les métadonnées d'ONIX contiennent des éléments de métadonnées pertinents pour les bibliothèques comme le titre, l'auteur et le synopsis mais également les métadonnées utiles aux fonctions commerciales (comme par exemple le stockage, le prix etc.). Comparé à ce qui se pratique dans le monde des bibliothèques, cet échange de métadonnées est dynamique, ayant parfois une fréquence quotidienne ou hebdomadaire.

ONIX est seulement utilisé par quelques grandes organisations bibliothécaires : OCLC, la *Library of Congress*, la *British Library*. Ils convertissent des éléments d'ONIX vers le format Marc 21. ONIX est d'ailleurs basé sur l'INDECS, un schéma similaire au FRBR, mais comportant une différence importante (voir 4.5).

³ De même, une grande partie des éditeurs scientifiques de petite dimension n'ont pas encore atteint ce stade : selon une enquête de l'Association of American University Presses' parmi les éditeurs universitaires (*University presses*) 39 % n'utilisent pas XML dans leur circuits de traitement (Digital Book Publishing in the AAUP community ; survey report : spring 2013).

L'EXEMPLE DE L'ÉDITEUR SPRINGER

L'éditeur Springer produit trois flux de données pour les bibliothèques et leurs systèmes :

- Les métadonnées au format Marc 21 : Les bibliothèques peuvent télécharger ces notices directement depuis le site Internet de Springer (sans frais supplémentaires ; licence CC0 ; avec un DOI et les codes des systèmes d'autorités de DDC, LCC et LCSH) ou recevoir via OCLC ces notices qui sont alors plus riches (par exemple les codes des systèmes d'autorités sont plus détaillés). Dans les deux cas les notices contiennent la table des matières, qui mentionne les titres de chapitres.
- Les métadonnées au format KBART pour les résolveurs de liens : ces métadonnées sont livrées aux prestataires qui maintiennent les résolveurs de liens. Ces métadonnées sont généralement plus pauvres que les notices au format Marc 21. Par exemple, elles ne comportent pas la table des matières ou les codes des systèmes d'autorités.
- Les données au format XML pour les outils de découverte : les prestataires qui maintiennent les outils de découverte peuvent obtenir les données au format XML (ainsi que le texte intégral).

Récemment, Springer a développé deux nouveaux services liés aux périodiques numériques. Les bibliothèques peuvent également télécharger les notices Marc 21 des titres de périodiques et des articles de périodiques.

L'interlocuteur rencontré chez Springer n'envisage aucun rôle pour les bibliothèques dans la production des métadonnées, sinon un rôle d'enrichissement de ces métadonnées. Springer serait-il intéressé par ces métadonnées enrichies par les bibliothèques ? L'interlocuteur a expliqué que leurs systèmes de production ne sont pas encore capables de les intégrer.

LES MÉTADONNÉES BIBLIOGRAPHIQUES SONT DE PLUS EN PLUS SOUVENT LIVRÉES EN LICENCE D'ACCÈS LIBRE

En ce qui concerne les licences des métadonnées livrées par les éditeurs, les métadonnées de base (comme le titre et l'auteur) sont considérées comme des données factuelles libres de droits et ne peuvent pas faire l'objet d'une licence. La plupart des éditeurs livrent ces métadonnées de base ainsi que les synopsis et les références bibliographiques. Cette tendance vaut pour l'ensemble des métadonnées, qui prennent de plus en plus souvent la direction de l'accès libre.

4.2 LES LIVRES ÉLECTRONIQUES SEMBLENT REMPLACER LES LIVRES IMPRIMÉS DANS LE CHAMP DE LA LITTÉRATURE SCIENTIFIQUE

LES GRANDS ÉDITEURS PUBLIENT TOUS LEURS LIVRES AU FORMAT ÉLECTRONIQUE

Les livres électroniques semblent remplacer les livres imprimés dans le champ de la littérature scientifique. Les grands éditeurs scientifiques publient actuellement tous leurs livres (également) au format électronique et plusieurs bibliothèques ont défini une stratégie d'achat donnant la priorité aux livres électroniques.

L'éditeur Springer est le plus grand éditeur de livres scientifiques. Il produit tous ses livres au format numérique et les vend aux bibliothèques uniquement en bouquets. On peut commander un livre imprimé, mais cette commande sera redirigée par Springer vers un prestataire qui va imprimer ce livre grâce à un système d'« impression à la demande ». Autrement dit, l'éditeur Springer a opéré en interne une transition complète vers le monde numérique.

LES PETITES ÉDITEURS TARDENT À SUIVRE

De l'autre côté, les éditeurs scientifiques de dimension moindre n'ont apparemment pas encore atteint ce stade : selon l'enquête d'ALSPS, 58% des éditeurs de très petite dimension, 34% des éditeurs de petite dimension et 21% des éditeurs de dimension moyenne ne publient pas de livres numériques⁴.

LA SITUATION ACTUELLE RELATIVE AUX LIVRES ÉLECTRONIQUES

La production mondiale annuelle de livres est estimée à 2,2 millions de titres. Il n'existe pas de spécification du nombre de livres scientifiques produits chaque année, mais l'on peut estimer ce nombre à plusieurs centaines de milliers. Combien parmi ces nouveaux titres sont publiés au format électronique ? Les chiffres concernant la production de livres au Royaume-Uni montrent que 34% des titres de livres publiés en 2012 le sont au format électronique⁵. Il n'existe pas de statistiques disponibles pour la France, seulement un chiffre portant sur les ventes de livres numériques : en 2011 celles-ci constituaient 2,1% du chiffre d'affaires de ventes de livres⁶.

À QUEL MOMENT LA TRANSITION SERA-T-ELLE COMPLÈTE ?

Pour ce qui est de l'évolution des livres numériques, l'ensemble des chiffres collectés ci-dessus indique que le remplacement quasi-total des livres scientifiques imprimés par des livres numériques constitue un processus qui durera de nombreuses années. Par analogie aux périodiques scientifiques

⁴ Ces chiffres datent de 2009 : ces pourcentages sont, sans doute, actuellement moindres. L'étude de l'ALSP ne fournit pas de données portant sur la part respective des livres publiés chaque année par chaque catégorie d'éditeurs. On peut déduire que les éditeurs de grande dimension (de cet ensemble de données) publient environ 75% des titres.

⁵ The UK Book Publishing Industry in Statistics 2012; the Publishers Association

⁶ Chiffres clés du secteur du livre 2011/2012 (mars 2013) ; Service du livre et de la lecture ; Observatoire de l'économie du livre.

(qui se sont transformés en documents au format numérique en une période ayant duré de 10 à 15 années), on peut estimer que la quasi-totalité des livres scientifiques seront publiés au format numérique en 2025. De même, on peut envisager une évolution à deux vitesses : dans le domaine des sciences exactes (où les périodiques scientifiques constituent le support de communication dominant) on peut s'attendre à ce que le passage aux livres numériques progresse plus rapidement (voir l'exemple de la bibliothèque de l'Université technologique d'Eindhoven ; paragraphe 4.6). Au contraire, dans le domaine des sciences humaines et sociales (où les livres constituent un support de communication important) on peut s'attendre à ce que cette évolution soit plus lente.

4.3 L'ÉVOLUTION DES MÉTHODES D'ACQUISITION

L'ACQUISITION PAR BOUQUETS ET PAR MÉTHODES SEMI-AUTOMATIQUES

Les méthodes d'acquisition des bibliothèques changent profondément. La méthode traditionnelle - l'acquisition par sélection manuelle des titres - décline et sera partiellement remplacée par un certain nombre d'autres méthodes d'acquisition.

Le cas de l'éditeur Springer peut illustrer ce changement : il publie actuellement tous ses livres scientifiques au format numérique et ne les vend aux bibliothèques que par bouquets.

L'acquisition par lots ou bouquets de livres électroniques est accompagné d'une autre méthode d'acquisition semi-automatique : *patron-driven acquisition or demand-driven acquisition*. Cette méthode suit la procédure suivante : les livres appartenant au fonds donné d'un éditeur donné sont rendus accessibles aux utilisateurs finaux de la bibliothèque acheteuse pour une période déterminée. À l'issue de cette période, les livres qui sont utilisés sont achetés par la bibliothèque acheteuse. Les livres non utilisés sont retirés. Selon une récente [présentation](#) d'un vendeur de livres aux États-Unis (Yankee Book Peddler), les éditeurs ont enregistré un recul important de ventes de livres imprimés, un recul important de ventes de livres électroniques titre par titre et une explosion des ventes de type *demand-driven acquisition*.

Il existe également une méthode d'acquisition semi-automatique pour les livres imprimés : le *book approval plan*. Selon cette méthode, un éditeur livre tous les ouvrages portant sur un sujet donné à la bibliothèque acheteuse et la bibliothèque peut retourner les livres dont elle ne veut pas.

LA CONSÉQUENCE SUR LE PROCESSUS DE CATALOGAGE

Les conséquences sur le processus de catalogage sont évidentes : les métadonnées des fonds achetés sont livrées par les éditeurs aux bibliothèques (directement ou par l'intermédiaire d'OCLC ; voir 3.4). Cette procédure rend dans une large mesure superflu le catalogage manuel. La méthode de « *patron-driven acquisition* » crée par ailleurs un nouveau problème : il devient nécessaire de supprimer les métadonnées des livres non retenus pour figurer au catalogue local (et de supprimer les données de « holding » dans le catalogue collectif).

4.4 L'ÉVOLUTION DES FONCTIONS DES CATALOGUES

LA FONCTION DE DÉCOUVERTE DES CATALOGUES EST EN PASSE DE DIMINUER

Les fonctions du catalogue dédiées à la découverte peuvent être, du point de vue des utilisateurs, réparties entre les catégories suivantes : (1) trouver (2) identifier (3) sélectionner et (4) obtenir/avoir accès. À cet égard, les catalogues des bibliothèques ont perdu du terrain en comparaison d'autres outils de découverte : les moteurs de recherche sur Internet (surtout Google et Google Scholar), les bases bibliographiques scientifiques (par exemple Web of Science) et les outils de découverte à l'échelle du Web acquis par un nombre de plus en plus grand de bibliothèques (par exemple Primo, WorldCat Local ou Summon). Néanmoins, les catalogues des bibliothèques continuent d'être utilisés de manière intensive.

LES AUTRES FONCTIONS DES CATALOGUES LOCAUX ET LE CATALOGUE COLLECTIF DEMEURENT

Les catalogues locaux remplissent plusieurs autres fonctions :

- La gestion d'inventaire/la gestion de collection
- La fonction de base pour d'autres services locaux : la circulation des documents, le prêt entre bibliothèques etc.

De point de vue des bibliothèques locales, les catalogues collectifs (nationaux ou régionaux) remplissent les fonctions suivantes :

- Une source de métadonnées pour le catalogue local.
- La visibilité au sein d'un périmètre national/régional pour mettre en place des services de PEB et de gestion des collections.

À l'échelle nationale/régionale, le catalogue collectif peut aussi fournir des index pour les systèmes de recherche ou les systèmes de présentation spécialisée et/ou en combinaison avec l'index des catalogues de musées ou d'archives, etc.

LES FONCTIONS DES DONNÉES DES CATALOGUES EN DEHORS DU CATALOGUE SE MULTIPLIENT

De même, il existe de multiples fonctions offertes par les notices bibliographiques des catalogues des bibliothèques. Lorcan Dempsey⁷ distingue ainsi :

- Les utilisations pour les besoins institutionnels ou personnels (par exemple : listes de lecture, guide de ressources, bibliographies personnelles, bibliographies institutionnelles, les pages du Web sur la recherche et/ou l'expertise établissements, pour le logiciel de gestion bibliographique etc.)
- L'extension des modes d'accès au catalogue par la création d'accès au sein d'autres environnements (flux RSS, widgets, barres d'outils, applications mobiles, etc.)

⁷ Lorcan Dempsey; 'Thirteen ways of looking at libraries, discovery and the catalogue scale, workflow, attention', Catalogue 2.0, The future of the library catalogue; Editor S. Chamber; 2013; ISBN 978-1-55570-943-3

- L'effet de levier : l'utilisation d'un environnement de découverte hors de contrôle de la bibliothèque afin d'attirer les utilisateurs vers les catalogues des bibliothèques (voir chapitre 5 l'exemple des bibliothèques publiques des Pays-Bas, qui espèrent attirer les internautes par le biais des liens entre Wikipédia et leur catalogue).

LE RÔLE DES CATALOGUES DANS UN MONDE DE LIVRES NUMÉRIQUES DEVIENT INSIGNIFIANT

Quand la quasi-totalité des livres scientifiques seront produits au format numérique et également achetés sous cette forme par la quasi-totalité des bibliothèques, on peut imaginer qu'alors les catalogues traditionnels des bibliothèques seront arrivés à la fin de leur cycle de vie. D'autres outils - comme des outils actuels de découverte à l'échelle du Web - reprendront les tâches dévolues aux catalogues traditionnels en utilisant les possibilités offertes par le texte intégral numérique (comme par exemple la recherche dans le texte intégral, le *text mining* et l'analyse sémantique).

4.5 LES GRANDS CHANGEMENTS DANS LE MONDE DU CATALOGAGE

FRBR

FRBR distingue trois niveaux hiérarchiques de notices bibliographiques : l'entité « manifestation », l'entité « expression » et l'entité « œuvre ». L'objectif de FRBR est d'améliorer la fonction de découverte des catalogues. INDECS - développé par le monde des éditeurs et la base du standard ONIX- est un schéma similaire à celui de FRBR, mais comporte une différence importante : INDECS distingue seulement deux niveaux hiérarchiques de notices bibliographiques : l'entité « manifestation » d'INDECS est similaire à l'entité « manifestation » de FRBR, l'entité « œuvre » d'INDECS est similaire à l'entité « expression » de FRBR et il n'y a pas d'entité similaire à l'entité « œuvre » de FRBR.

De même, le format bibliographique en développement BIBFRAME distingue seulement deux niveaux (l'entité « œuvre » et l'entité « instance »).

Dans la pratique, la distinction entre le niveau « expression » et le niveau « œuvre » de FRBR est apparemment difficile à concevoir. Les bibliothèques publiques des Pays-Bas ont développé un nouveau catalogue national (NBC) utilisant le système CBS d'OCLC. Dans le cadre du système CBS, un algorithme a été développé pour procéder à la création des entités 'œuvre' des notices bibliographiques existantes (qui constituent l'équivalent des entités 'manifestation'). On a choisi d'ignorer l'entité 'expression', par ce qu'on a estimé que cette entité ne serait pas très utile aux utilisateurs finaux. Maintenant, l'algorithme a créé un thésaurus d'environ 500 000 notices d'autorités d'œuvre, à comparer aux 3 millions de notices bibliographiques existantes.

SUDOC et FRBR

Il y a deux ans que l'ABES et l'OCLC Leiden ont étudié les possibilités d'une implémentation de FRBR en utilisant le format Unimarc. Les résultats de cette étude sont positifs : il est possible d'utiliser le format Unimarc pour procéder à une « FRBRisation » du Sudoc. En vue de procéder à une implémentation, il est nécessaire de reparamétrer le système CBS. La proposition financière d'OCLC était conséquente et arrivait au moment où l'ABES démarrait le projet SGM. Il était inutile d'engager des dépenses alors que l'avenir du Sudoc au sein du système CBS n'était pas sûr.

Plusieurs bibliothèques nationales s'occupent des projets de « FRBRisation » des leurs catalogues (comme la BnF et la *Deutsche Nationalbibliothek*). La *Deutsche Nationalbibliothek* a opté pour une implémentation de deux niveaux et ne créera pas d'entité « expression ». VIAF (www.viaf.org) est en train de créer des notices de grappes (*cluster records*) pour les entités « œuvre » et « expression ». Ils ont déclaré que les notices de grappes d'entité « expression » sont surtout des traductions⁸.

⁸ Authority files: breaking out of the library silo to become signposts for research information; Maurits van der Graaf et Leo Waaijers, report for Knowledge Exchange; to-be-published
PLEIADE MANAGEMENT & CONSULTANCY

LES RÈGLES RDA

Une explication portant sur les règles RDA ainsi qu'une description de la situation actuelle en France sont présentées dans l'encadré jaune ci-dessous. Les entretiens menés lors de cette étude ont apporté les observations suivantes :

- Les bibliothèques européennes sont en général en avance sur la *Library of Congress* en ce qui concerne les liens entre les différentes entités (bibliographiques et autorités). Le catalogue de la *Library of Congress* ne propose pas de liens entre les notices bibliographiques et les notices d'autorités. C'est pourquoi l'implémentation du système RDA tel qu'il est appliqué actuellement à la *Library of Congress* constituerait un pas en arrière par rapport aux règles de catalogage actuelles en France.
- Les notices bibliographiques créées selon les règles RDA seraient plus riches que les notices bibliographiques créées selon les règles traditionnelles et peuvent offrir de meilleures fonctionnalités selon un expert de RDA. Mais le même interlocuteur a souligné que ces avantages des notices bibliographiques « RDA » ne sont pas encore disponibles pour la majorité des utilisateurs finaux parce que les systèmes des bibliothèques ne sont pas encore optimisés.
- Un argument fort en faveur d'une migration vers le système RDA en France est que, si d'autres pays choisissent d'effectuer une telle migration, les bibliothèques françaises se trouveraient isolées. Actuellement, en Europe, le Royaume Uni, l'Irlande, l'Allemagne, la Lettonie, les Pays-Bas, l'Autriche, la Suède et la Finlande et hors d'Europe, la *Library of Congress*, la Bibliothèque nationale de Canada et la Bibliothèque nationale d'Australie ont décidé de migrer vers RDA.

Quelles seraient des conséquences d'une éventuelle implémentation de RDA en France selon les différents interlocuteurs ?

- L'efficacité du système RDA serait équivalente à celle du système actuel. La même efficacité du processus de catalogage constituait d'ailleurs une exigence primordiale du projet d'implémentation de RDA en Allemagne.
- La connaissance et la formation : actuellement, il n'y a aucun catalogueur en France qui maîtrise les règles de catalogage exigées par le système RDA. Au sein de l'ABES le système RDA est connu, mais pas à un niveau suffisant pour permettre la formation des catalogueurs à son utilisation. Il serait en premier lieu nécessaire de l'enseigner aux formateurs, et puis d'organiser la formation des catalogueurs. Pour ces derniers, le système RDA constitue un changement d'attitude considérable. C'est une révolution dans le domaine des techniques de catalogage. Les conséquences pour la formation des bibliothécaires français sont profondes. En Allemagne, on prévoit 7 heures de formation pour chaque catalogueur quand le système d'autorités GND a opéré la transition vers les règles RDA. La formation prévue pour l'application des règles RDA sur les notices bibliographiques (planifiée au troisième trimestre 2015) est encore inconnue.
- Le profil d'application : L'application de RDA signifie que beaucoup de règles existantes doivent être adaptées. Cela représentera un travail énorme. Il sera difficile de former un groupe de bibliothécaires qui disposeraient du temps et de la motivation nécessaires pour changer les règles du catalogage français en fonction du système RDA. Cet effort doit aboutir entre autres à la conception d'un nouveau *Guide Méthodologique*.
- L'adaptation des systèmes : de même, les systèmes de catalogage et les catalogues doivent être adaptés aux nouvelles règles de catalogage.
- Les notices existantes : la perspective de devoir transformer les notices existantes pour les adapter aux nouvelles règles RDA suscite des inquiétudes.

Quelle est l'interdépendance entre une migration éventuelle vers un système de Cloud et l'application des règles RDA ? Les deux systèmes coexistent actuellement dans le système de catalogage partagé de *Libraries Australia*, qui utilise CBS, et dans WorldCat. Mais il a été dit que la réutilisation d'une notice produite par les règles RDA au sein d'un environnement de règles classiques exigerait un effort supplémentaire modéré.

Les règles RDA expliquées

RDA est le nouveau code de catalogage élaboré par la communauté des bibliothèques anglo-américaines pour remplacer les AACR2. Ce nouveau code de catalogage définit une nouvelle approche du catalogage adaptée à l'environnement actuel des bibliothèques, dominé par le web. RDA place les modèles de l'information bibliographique FRBR et FRAD au cœur des règles de catalogage et prépare une nouvelle organisation des catalogues reflétant le modèle FRBR. Les notices bibliographiques sont amenées à être remplacées par un réseau de relations entre des « notices » correspondant aux entités du modèle FRBR, selon une structure proche de celle qui caractérise l'information sur le web de données. Toutefois, pour favoriser une transition en douceur, RDA autorise aussi, dans un premier temps du moins, le maintien de la structure traditionnelle des catalogues, tout en renforçant le rôle des notices d'autorité et en rendant systématique la création de notices d'autorité pour les œuvres [condensé du texte de la source : rda-en-france.enssib.fr].

La situation actuelle en France

Le Groupe stratégique a décidé de ne pas adopter dès maintenant RDA en France, car les règles de RDA ne sont pas satisfaisantes au regard de l'analyse et de la pratique catalographiques françaises. Une adoption de RDA en l'état ne permettrait pas de répondre aux besoins des bibliothèques françaises, voire impliquerait une régression par rapport à la situation actuelle des catalogues en France.

Toutefois, la France a pour objectif de rejoindre un jour RDA et de contribuer à l'évolution du code, de manière à améliorer les règles, les rendre réellement internationales et permettre ainsi leur adoption en France. Le Groupe stratégique a jugé préférable de conduire ce travail en vue d'une évolution de RDA dans le cadre d'EURIG, afin de porter une position européenne vis à vis du JSC.

[Source : www.bnf.fr]

BIBFRAME ET LE WEB DE DONNÉES

Le Web de données (*Open Linked Data* ; une structure RDF) suscite beaucoup d'intérêt du fait de ses capacités à rassembler les ensembles de données sur Internet et des nombreuses possibilités de réutilisation de ces données. Le Web de données existe déjà mais toujours à un stade embryonnaire.

Plusieurs bibliothèques ont déjà publié leurs catalogues ou leurs systèmes d'autorités dans le cadre d'une structure RDF (comme la BnF et la *Library of Congress*). Cette dernière a pris l'initiative de développer BIBFRAME - un nouveau format des données des catalogues en remplaçant les formats MARC existants par une structure RDF. BIBFRAME envisage d'ailleurs deux niveaux bibliographiques (l'entité « œuvre » et l'entité « instance »). On prévoit que le développement de BIBFRAME sera fini dans 2015.

Au cours de cette étude, plusieurs interlocuteurs issus du monde des éditeurs se sont exprimés au sujet du Web de données :

- 'Les éditeurs craignent *Open Linked Data* parce qu'ils perdront le contrôle du contexte de leurs publications'.
- 'On peut observer une décélération des développements d'*Open Linked Data* due à l'inertie des systèmes'.
- Il y a un projet pour convertir ONIX (le standard d'échange des données des éditeurs) au format *Open Linked Data*.
- Un interlocuteur représentant un éditeur scientifique décrit un éventail allant de la 'récupération de l'information' à la 'récupération de la connaissance'. Google est un exemple de récupération de l'information, l'utilisation d'un thésaurus dans le processus de recherche est un exemple de récupération de la connaissance. Cet éditeur engage des ressources pour développer des produits qui sont basés sur la récupération avancée de la connaissance - utilisant les concepts liés. L'avantage majeur de ces produits pour les utilisateurs finaux sera un résultat de recherche plus limité et plus pertinent (et une économie de temps). Il est important de remarquer que cet éditeur utilise surtout les *Linked Data (RDF)* pour la production de produits orientés vers la *récupération de la connaissance*.

LES TRAITEMENTS AUTOMATIQUES DANS LE PROCESSUS DE CATALOGAGE

Il est important de noter ici l'étude portant sur la faisabilité et le positionnement d'un *hub* de métadonnées ABES⁹. Cette étude décrit le résultat d'un prototype de *hub* de métadonnées ABES - un outil permettant de convertir, d'agréger, d'évaluer, de corriger, et d'enrichir les métadonnées bibliographiques. Une des conclusions de cette étude est la suivante : « S'il parvient à produire certaines informations qui sont traditionnellement créées par les catalogueurs, il ne rendra pas caduque l'activité de catalogage. Il faut imaginer comment le Web et les catalogueurs cohabiteront voire collaboreront ». Autrement dit : les flux de métadonnées en provenance des éditeurs vers les systèmes de catalogage feront de plus en plus souvent l'objet d'une combinaison de traitements automatique *et* manuel.

⁹ Étude sur la faisabilité et le positionnement d'un hub de métadonnées ABES, 2013

4.6 LES CHANGEMENTS DES SYSTÈMES DE CATALOGAGE PARTAGÉ

LES MENACES PLANANT SUR LES SYSTÈMES DE CATALOGAGE PARTAGÉ

Les développements des systèmes de gestion des bibliothèques mutualisés dans le Cloud et la digitalisation progressive des littératures scientifiques menacent particulièrement le catalogage partagé. Deux types de menaces existent :

- **La compétition exercée par les systèmes de Cloud** : cette menace est illustrée par la bibliothèque d'Ifremer, qui a migré vers le système de Cloud de l'OCLC et utilise maintenant WorldCat au lieu du Sudoc pour son système de catalogage partagé. Ce phénomène est aussi décrit dans l'étude de cas portant sur *Libraries Australia* (voir chapitre 5). L'effet de ce phénomène sur *Libraries Australia* est qu'on demande les données des bibliothèques qui utilisent WorldCat pour les ajouter au catalogue collectif de *Libraries Australia*. Autrement dit : les flux de métadonnées entre le système national et ces bibliothèques sont inversés.
- **L'absence de besoin d'un système de catalogage partagé** : cette menace est illustrée par la bibliothèque de l'Université Technologique d'Eindhoven. L'acquisition de livres imprimés par cette bibliothèque a été presque réduite à zéro, lorsque la littérature relative aux sciences appliquées et à la technologie est devenue principalement électronique. C'est la raison pour laquelle cette bibliothèque n'a pas de besoin d'un système de catalogage partagé. C'est la seule bibliothèque universitaire des Pays-Bas qui ne migre pas vers le système de Cloud d'OCLC.

UNE NOUVELLE FONCTION POUR LES SYSTÈMES DE CATALOGAGE PARTAGÉ : LA MANIPULATION DE FLUX DE MÉTADONNÉES

La conséquence de ce déplacement de la production des métadonnées vers les éditeurs est que les systèmes de catalogage partagé s'orientent vers la possibilité de manipulation des flux de métadonnées.

Plusieurs interlocuteurs ont souligné que le système de catalogage partagé se transformait en un système permettant la manipulation des flux de métadonnées. Il est clair que l'OCLC, avec ses services de la gamme WorldCat, occupe une position presque monopolistique à cet égard. L'exemple donné est souvent l'acquisition d'un fonds de livres électroniques par une bibliothèque. L'éditeur peut livrer à WorldCat les métadonnées des livres électroniques *ainsi que* les données de '*holding*' de l'établissement acheteur. Ensuite, un service de la plate-forme WorldShare assure que ces métadonnées sont automatiquement chargées dans le système de gestion de l'établissement acheteur. La liste des éditeurs collaborant avec l'OCLC dans le cadre de ce service ('*cataloguing partners*') est constituée d'environ 50 éditeurs ou autres fournisseurs pertinents.

Un service similaire existe d'ailleurs déjà au sein du système Sudoc : l'ABES importe actuellement des notices provenant de cinq éditeurs et étudie les possibilités de collaboration similaire avec six autres éditeurs. Les notices sont actuellement importées, mais les éditeurs ne fournissent pas les données d'abonnement : une deuxième opération permet de localiser les bibliothèques possédant les *e-books*.

4.7 CONCLUSIONS

L'ÉVOLUTION POSSIBLE DES CATALOGUES SUITE À L'INFLUENCE DES LIVRES NUMÉRIQUES

Quelle seront les conséquences de la transition vers les livres numériques pour le catalogage partagé, le catalogue collectif à l'échelle nationale et des catalogues des bibliothèques individuelles ?

Dans le schéma ci-dessous (figure 5) sont présentées les conséquences éventuelles pour les bibliothèques de sciences exactes et pour les bibliothèques de sciences humaines et sociales. Le schéma distingue par ailleurs les besoins du catalogage partagé, ceux du catalogue collectif à l'échelle nationale et ceux des catalogues locaux.

L'argumentation est la suivante :

- Dans le domaine des sciences exactes, la transition vers les publications numériques avancera rapidement et sera peut être quasi-complète. Si cela devait être le cas, les bibliothèques spécialisées en sciences exactes auront de plus en plus besoin d'un système de gestion de flux de métadonnées provenant des éditeurs. Leurs catalogues locaux peuvent être remplacés par un outil de découverte. Le besoin d'un catalogue collectif à l'échelle nationale demeure, en raison de la nécessité de disposer de services à l'échelle nationale, mais son importance diminue.
- Dans le domaine des sciences humaines et sociales, la transition vers les publications numériques se fera plus lentement. De même, on peut s'attendre à ce que la transition ne soit pas complète et à ce qu'un certain rôle dévolu aux livres imprimés subsiste. En ce qui concerne le besoin d'un système de catalogage partagé, le besoin d'un système de catalogage partagé demeurera et la nécessité d'un système de gestion de flux de métadonnées provenant des éditeurs augmentera. Le besoin d'un catalogue collectif à l'échelle nationale demeurera.

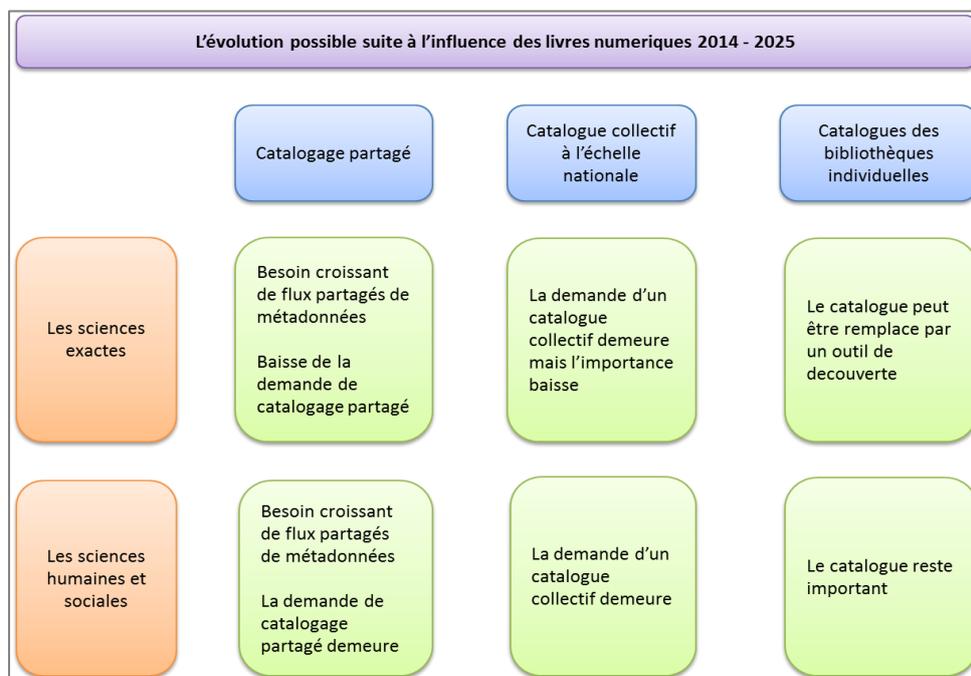


FIGURE 5 ÉVOLUTION POSSIBLE DES CATALOGUES

UN PRONOSTIC DE RÉSEAU SUDOC DANS 2025

Le nombre des tendances présentées dans ce chapitre 4 rend difficile d'obtenir une vision claire des conséquences pour le réseau Sudoc. Pour apporter un éclairage sur ces conséquences, voici ci-dessous un pronostic pour l'année 2025, basé sur les estimations de l'auteur de ce rapport :

La production de métadonnées :

- 70% - 80% des métadonnées proviendront des éditeurs et seront importées dans le système, parfois avec les données « holdings » et avec les codes des notices d'autorités internationales. Ces métadonnées sont relatives à l'acquisition par bouquets ou par les méthodes de *demand-driven acquisition* :
 - la majeure partie de ce flux de métadonnées partagé proviendra des éditeurs internationaux et la gestion de ce flux sera assurée par un prestataire international.
 - une petite partie de ce flux de métadonnées partagé proviendra des éditeurs français et la gestion de ce flux pourra être assurée par l'ABES.
 - Les traitements de modification en masse seront assurés par l'ABES, surtout ceux relatifs aux ajouts de codes de système d'autorités français.
- 20% - 30% des métadonnées sont relatives à l'acquisition « titre par titre » et surtout aux livres imprimés. Le catalogage sera fait manuellement.

Ce pronostic quantifié

La production annuelle de nouvelles notices bibliographiques au sein du réseau Sudoc a été d'environ 700.000 notices en 2013.

La croissance annuelle est estimée à 3 %.

Cette croissance annuelle signifie qu'en 2025 la production annuelle au sein du réseau Sudoc sera environ 1.000.000 notices :

- le nombre de métadonnées créées au moyen du catalogage partagé en 2025 sera de 200.000 à 300.000 notices (440.000 notices en 2013)
- le nombre de métadonnées importées provenant des éditeurs en 2025 sera 700.000 à 800,000 notices (260.000 notices en 2013).

Le réseau Sudoc :

- La plupart des SGB des bibliothèques participantes auront migré vers quelques systèmes de gestion de bibliothèques mutualisés et internationaux dans le Cloud.
- Une minorité de SGB des bibliothèques participantes auront un SGB non-mutualisé.
- Quelques bibliothèques universitaires françaises auront opté pour l'utilisation d'un système de catalogage partagé d'un prestataire international et auront quitté le réseau Sudoc pour le catalogage. La majorité d'entre elles continuera de participer au catalogue collectif de l'ABES afin de pouvoir continuer à assurer les services à l'échelle nationale. Pour cela, leurs métadonnées seront importées dans le catalogue collectif par l'ABES.

Le PEB :

- En raison de développement du libre-accès pour les articles de périodiques et de développement de la transition vers les livres numériques, on peut s'attendre à ce que l'utilisation de service de prêt entre bibliothèques diminue et se concentre sur les livres imprimés. Du point de vue des utilisateurs finaux, on peut dire que la fréquence d'utilisation diminuera mais que la valeur ajoutée croîtra.

Nouveaux services et nouvelles utilisations des métadonnées :

- On peut s'attendre à ce que les systèmes du réseau Sudoc facilitent d'utilisation des métadonnées du Sudoc au format RDF (par exemple par le biais du format BIBFRAME).
- Les nouveaux services seront développés par les acteurs du monde de bibliothèques, à l'échelle nationale et à l'échelle locale, en utilisant des métadonnées du réseau Sudoc.
- Les métadonnées de réseau Sudoc seront utilisées par les autres acteurs et surtout par les utilisateurs finaux suite à l'influence du libre-accès et des possibilités offertes par le Web de données.

5. ANALYSE ET SYNTHÈSE DES ÉTUDES DE CAS

5.1 LES SYSTÈMES QUI MIGRERONT VERS LE CLOUD

Cinq études de cas concernent une migration vers le Cloud. Les principaux résultats sont présentés en table 3. Les études de cas sont les suivantes :

- Orbis Cascade Alliance
- CIB – le projet de l’Allemagne
- Les bibliothèques universitaires des Pays-Bas
- CSUC et CCUC – le système de catalogage partagé de la Catalogne
- BIBSYS en Norvège.

Une description de cinq études de cas qui migreront vers un système de Cloud :

Orbis Cascade Alliance :

- *Orbis Cascade Alliance* migrera les systèmes de gestion des 37 bibliothèques universitaires vers le système de Cloud d’Ex-Libris (Alma). Il continuera à utiliser WorldCat pour le système de catalogage partagé et pour les flux partagés de métadonnées comme les métadonnées de fonds des livres électroniques (*e-book packages*). La raison invoquée pour le maintien de l’emploi de WorldCat est qu’il offre des avantages majeurs aussi bien qualitatifs que quantitatifs dans le domaine des métadonnées. Ensuite, les notices de catalogage et de métadonnées seront chargées dans la ‘zone cooperative’ du système d’Alma (*community zone*). Ce sera une base commune de données bibliographiques que l’on utilisera pour le PEB à l’intérieur du consortium. *Orbis Cascade Alliance* accorde de l’importance à la possibilité de partager des ressources en utilisant le développement des circuits de traitement collaboratifs. Un groupe de travail s’occupe du développement des circuits de traitement collaboratifs.

Le projet CIB en Allemagne :

- On veut établir un *German data space*. Pour alimenter ce *data space*, deux options existent : (1) on peut utiliser WorldCat comme seule source de métadonnées ou (2) on peut utiliser WorldCat et Alma comme deux sources de métadonnées.
- L’idée est de créer un *German data space* dans chaque système de Cloud utilisé en Allemagne et de synchroniser entre eux les différents *German data spaces*. La mise en œuvre est en cours de développement – OCLC et Ex-Libris ont approuvé cette idée par le biais d’une lettre d’intention. L’unique objectif du *German data space* est de servir au catalogage partagé et aux flux partagés des métadonnées. Les principales raisons de cette migration sont (1) l’élimination de la fragmentation du paysage des bibliothèques d’Allemagne (il existe actuellement 6 systèmes collectifs) et (2) la connexion aux systèmes internationaux. En ce qui concerne les services nationaux, on veut créer un *German data pool*, dont les fonctions sont similaires à celles de la base de données bibliographiques communes d’Orbis Cascade Alliance.
- Le projet CIB envisage que les données des *German data spaces* dans les systèmes de Cloud soient liées au système d’autorités allemand (GND). Une éventuelle intégration de GND à d’autres systèmes d’autorités internationales est considérée comme un développement éventuel indépendant de la migration vers un système de Cloud.

Les bibliothèques universitaires des Pays-Bas :

- Elles migreront vers la plate-forme WorldShare et utiliseront WorldCat. En même temps, elles passeront aux règles RDA. Le système de catalogage partagé GGC n'existera plus, le catalogue national (NCC) deviendra une vue de WorldCat (*window*). La raison avancée pour justifier cette migration est l'attente d'un gain de productivité considérable en matière de catalogage.
- Les bibliothèques universitaires des Pays-Bas utiliseront les notices d'autorités de la *Library of Congress* au lieu des notices d'autorités nationales (une classification et un thesaurus de termes contrôlés). Quant à l'avenir de ces catégories d'autorités, la décision n'est pas encore prise sur les noms d'auteurs et sur deux autres catégories d'autorités nationales, qui ont une portée limitée.
- Le catalogue collectif deviendra une vue de WorldCat. Le système de prêt entre bibliothèques – connecté à WorldCat - sera un système national externe à la plate-forme de WorldShare. Ce système devra être développé. Le système actuel de PEB aux Pays-Bas facilite des demandes de documents par les utilisateurs finaux qui ont un compte (*account*) pour le paiement. Le système est caractérisé par un tarif uniforme. Il comporte aussi une fonctionnalité de « chambre de compensation » (*clearinghouse*) - c'est-à-dire que la convention financière entre fournisseurs et demandeurs permet un règlement cumulatif annuel.

CSUC:

- CSUC est le consortium des bibliothèques universitaires de Catalogne et de la Bibliothèque nationale de Catalogne. Les dix participants ont en commun un système de catalogage partagé et un catalogue collectif (CCUC), un serveur partagé avec 10 implémentations du SGB Millennium et d'autres projets partagés.
- Le but du projet de migration vers un système de Cloud est de remplacer les 10 implémentations de SGB par un seul SGB mutualisé, puis ensuite d'harmoniser les circuits de travail dans les 10 bibliothèques concernées. De même, la gestion de ressources électroniques partagées est un objectif important.
- Le système de catalogage partagé et le catalogue collectif sont le fruit d'efforts considérables déployés depuis 15 ans. C'est pourquoi CSUC veut maintenir cette base de données dans un environnement séparé au sein du système de Cloud choisi. Autrement dit, une intégration complète n'est pas envisagée.

BIBSYS :

- BIBSYS est le consortium de toutes bibliothèques académiques et de recherche et la Bibliothèque nationale de Norvège. Depuis de plusieurs années, toutes les bibliothèques utilisent le même SGB mutualisé. Le système consiste en un module dédié au catalogage partagé/catalogue collectif. Il n'y a pas de catalogues locaux.
- La migration vers le système de Cloud d'ALMA a commencé début 2014 et va durer deux ans. Elle consiste surtout en un remplacement du système mutualisé existant.
- On a choisi le système d'ALMA pour le circuit de travail unifié pour les publications imprimées et numériques.
- Le catalogue collectif norvégien sera une base de données séparée au sein du système d'Alma, mais connectée aux flux de métadonnées international. Ce flux de métadonnées sera géré par l'Ex-Libris.
- Les métadonnées des publications numériques seront livrées par les éditeurs et on n'a pas envisagé d'enrichissement par les bibliothèques.

- On estime que le rôle de catalogage par les bibliothèques norvégiennes sera limité aux collections nationales ou locales. Pour cela, il existe un système d'autorités norvégien.

Ressources partagées	Orbis Cascade Alliance	CIB projet d'Allemagne	Bibliothèques universitaires des Pays-Bas	CSUC	BIBSYS en Norvège
Établissements concernés	37 bibliothèques universitaires	Les bibliothèques universitaires d'Allemagne	12 des 13 bibliothèques universitaires des Pays-Bas	10 bibliothèques universitaires	105 bibliothèques académiques et de recherche, et la Bibliothèque nationale
Statut actuel	Migration à mi-chemin	Le projet a débuté en septembre 2013	Le projet démarre incessamment	Stade initial	Un SGM mutualisé national avec le catalogage partagé et le catalogue collectif existe déjà (il n'y a pas de catalogues locaux)
Système de gestion des bibliothèques	Migration vers ALMA	Au choix	Au choix	Migration vers un SGB mutualisé	Migration vers ALMA
Système de catalogage partagé/flux partagés de métadonnées	Utilisation de WordCat	Création d'un <i>German data space</i> dans les systèmes de Cloud multiples Les <i>German data space</i> sont synchronisés D'autres systèmes de Cloud peuvent rejoindre les <i>German data space</i> Migration ou connexion de notices d'autorités allemandes (GND)	Migration du système de catalogage partagé et catalogage collectif vers la plate-forme WorldShare Utilisation des notices d'autorités de la <i>Library of Congress</i> Fin de l'utilisation des notices d'autorités néerlandaises	Ils veulent maintenir leur propre environnement de catalogage partagé	L'environnement du catalogage partagé/le catalogue collectif sera une base de données séparée au sein du système d'ALMA, mais connectée aux flux de métadonnées internationaux gérée par ExLibris Le système d'autorités norvégien demeure mais sera surtout utilisé pour le catalogage de collections spéciales
Services basés sur des métadonnées de catalogage collectif	Une base de données bibliographique commune dans Alma basée sur les métadonnées de WordCat	Création de <i>German data pool</i> à l'extérieur des systèmes de Cloud	Le catalogage collectif néerlandais devient une fenêtre du WorldCat Développement d'un système de PEB distinct de WorldShare	Ils veulent maintenir leur propre catalogue collectif	Le catalogue collectif sera une base de données séparée au sein du système d'ALMA
Gestion partagée de ressources électroniques	Alma	Externe au projet	Externe au projet		On a choisi ALMA en raison du circuit de travail unifié pour les publications imprimées et numériques. L'outil découverte (ORIA) existe déjà et est basée sur Primo d'ExLibris

TABLE 3 LES ÉTUDES DE CAS QUI MIGRERONT VERS LE CLOUD

5.2 LES SYSTÈMES QUI MIGRERONT VERS UN SYSTÈME AUTO-DÉVELOPPÉ

Deux études de cas concernant une migration vers un système auto-développé. Ces études de cas sont :

- le système LIBRIS-XL de la Bibliothèque nationale de la Suède
- Le nouveau système de bibliothèques de la Bibliothèque nationale de Finlande.

Voici maintenant une description de ces études de cas :

LIBRIS en Suède :

- La Bibliothèque Nationale de Suède a développé sa propre plate-forme et son propre format bibliographique en RDF (qui peut être défini comme un prédécesseur de BIBFRAME). Le nouveau système sera implémenté en 2014. En même temps, on passera aux règles RDA. Cette migration est inspirée par la perspective d'emploi d'*Open Linked Data*. Selon la Bibliothèque nationale de Suède, *Open Linked Data* constitue la meilleure manière de partager d'une façon efficace les ressources entre les bibliothèques et les autres acteurs.
- Le principal changement dans la manière de cataloguer, que ce soit avec le format BIBFRAME ou avec un autre format basé sur RDF (par exemple le format en JSON-LD développé par la Bibliothèque nationale de la Suède pour LIBRIS XL), est que le catalogueur choisit des options dans une liste au lieu de saisir les chaînes de caractères.
- Il est toutefois possible d'avoir une interface de travail utilisant les chaînes de caractères et de « taper au kilomètre » pour les catalogueurs expérimentés (comme ils le font actuellement pour cataloguer en Marc21 ou en Unimarc) alors que le système utilise un format en RDF pour le stockage. Le nouveau système LIBRIS XL aura aussi une interface de travail qui utilisera le format Marc 21, converti « en coulisses » au format RDF/JSON-LD.
- Au sein de LIBRIS XL on utilisera le système d'autorités suédois sans exclure l'utilisation d'autres systèmes d'autorités internationaux.
- La Bibliothèque nationale de Suède a également développé son propre système de PEB.
- La Bibliothèque Nationale de Suède va arrêter son service de résolveur de liens national en 2014. La raison avancée est que la gestion du service représente un travail conséquent et que ce service semble dépasser les missions de la Bibliothèque nationale suédoise. La bibliothèque a pris contact avec l'équipe de KB+ afin de remplacer ce service.

Le système envisagé en Finlande :

- Melinda est le système de catalogage partagé et le catalogue collectif des bibliothèques universitaires et des bibliothèques des écoles polytechniques en Finlande. Le système est maintenu par la Bibliothèque nationale de Finlande. Melinda est actuellement basée sur la plate-forme Aleph d'Ex-Libris. L'alignement des bibliothèques publiques sur le système Melinda est envisagé.
- Melinda utilise actuellement le format Marc 21 car c'est le standard international. On peut envisager une transition vers le format BIBFRAME à moyen terme si BIBFRAME devient le standard international accepté.
- Le système planifié sera un système de gestion mutualisé et basé sur une plate-forme auto-développée avec des logiciels Open Source. La nouvelle plate-forme incorpore le système de catalogage partagé et le catalogue collectif Melinda. Ce nouveau système sera maintenu par la Bibliothèque nationale de la Finlande, qui jouera un rôle de prestataire. L'idée de nouveau système de gestion mutualisée provient des bibliothèques universitaires, qui doivent ré-

informatiser leurs systèmes actuels périmés. C'est pourquoi il n'y a pas de libre choix pour les bibliothèques universitaires et les bibliothèques des écoles polytechniques. Seules les bibliothèques publiques peuvent librement choisir d'adopter le nouveau système ou un autre système.

- Il y a actuellement un système de notices d'autorités finlandais, qui demeurera au sein du nouveau système.
- Open Source et Open Access sont considérés comme les conditions prépondérantes en Finlande. C'est pour cette raison qu'on utilise la licence CC0 et pas de flux de métadonnées de l'OCLC (dont les conditions juridiques sont considérées comme non acceptables).
- L'outil de découverte nationale (FINNA) utilise déjà les métadonnées de Melinda

Ressources partagées	LIBRIS XL en Suède	Finlande
Établissements concernés	180 bibliothèques (en majorité universitaires ou de recherche)	Les bibliothèques des universités et des écoles polytechniques ; on envisage l'alignement des bibliothèques publiques à l'étape suivante
Statut actuel	Implémentation en mars 2014	Melinda est l'environnement de catalogage partagé et le catalogue collectif sur une plate-forme d'Aleph
Système de gestion des bibliothèques	Externe au projet	Développement d'une plate-forme pour SGBm et Melinda basée sur le logiciel Open Source ; Les bibliothèques publiques peuvent maintenir une autre SGB
Système de catalogage partagé/flux partagés de métadonnées	Migration vers un système de catalogage partagé utilisant un format en RDF développé par la Bibliothèque nationale	Licences CC 0 Pas d'usage de flux de métadonnées d'OCLC mais de Bookwhere Maintien de format Marc21 Un système de notices d'autorités Finlandais
Services basés sur des métadonnées de catalogage collectif	Un système de PEB distinct, développé par la Bibliothèque nationale	L'outil de découverte nationale (FINNA) utilise déjà les métadonnées de Melinda
Gestion partagée de ressources électroniques	Externe au projet Approche de KB+	1. Inclus dans le projet

TABLE 4 LES ÉTUDES DE CAS QUI MIGRERONT VERS UN SYSTÈME AUTO-DÉVELOPPÉ

5.3 LES AUTRES ÉTUDES DE CAS

Trois autres études de cas ont été effectuées :

- le réseau *Libraries Australia*
- le nouveau catalogue collectif des bibliothèques publiques des Pays-Bas
- le projet LIBOS en Allemagne, qui a été rejeté.

Voici une description de ces études de cas :

Le réseau *Libraries Australia* :

- Le catalogue collectif *Australian National Bibliographic Database (ANBD)* utilise le système CBS. Environ 700 bibliothèques contribuent à ANBD, environ 100 bibliothèques utilisent également la possibilité de catalogage partagé. Les autres bibliothèques téléchargent leurs métadonnées « holding » sur le catalogue collectif, mais ce flux de métadonnées n'est ni complet ni fréquent et mène à des lacunes dans le catalogue collectif.
- Il y a déjà plusieurs bibliothèques (environ 30) en Australie qui ont migré vers un système de Cloud, principalement la plate-forme WorldShare. Quelques unes des ces bibliothèques ont choisi d'utiliser WorldCat comme système de catalogage partagé, ce qui a pour conséquence d'inverser les flux de métadonnées entre la bibliothèque et le catalogue collectif.
- Pour le service de PEB on utilise le logiciel VDX, mais OCLC a annoncé que ce logiciel serait abandonné en 2016 ou en 2017.
- Le réseau *Libraries Australia* est géré par la Bibliothèque nationale d'Australie. Pour l'instant celle-ci adopte une attitude attentiste vis-à-vis du développement des systèmes de Cloud et elle n'a aucun projet concret de migration. La position stratégique de *Libraries Australia* est considérée comme forte en raison du succès de l'outil national de découverte, nommé *Trove*.

Le nouveau catalogue collectif des bibliothèques publiques des Pays-Bas :

Le système du catalogue collectif des bibliothèques publiques aux Pays-Bas (figure 6) est intéressant, pour les raisons suivantes :

- Elles ont « FRBRisé » leur catalogue utilisant le système CBS à l'aide d'un algorithme, qui a créé l'entité « œuvre » en groupant les « manifestations » (on a jugé inutile de créer une entité « expression »).
- Les métadonnées sont converties vers un format bibliographique en RDF basé sur BIBFRAME et transférées sur une plate-forme de découverte. On crée les liens entre le format bibliographique en RDF et les sources d'Internet populaires comme Wikipedia. L'idée est d'attirer les internautes qui utilisent Wikipedia vers la plate-forme des bibliothèques publiques.

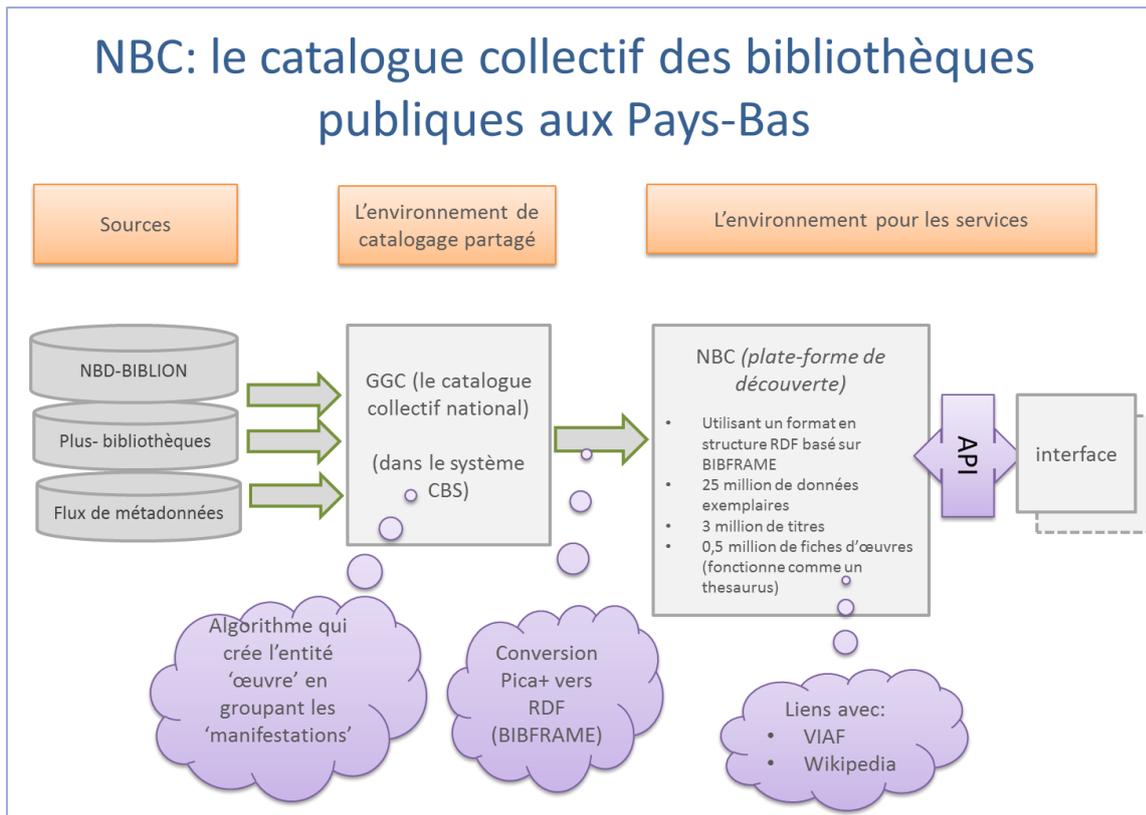


FIGURE 6 UN SCHÉMA DU CATALOGUE COLLECTIF DES BIBLIOTHÈQUES PUBLIQUES AUX PAYS-BAS

Le projet LIBOS en Allemagne, qui a été rejeté :

Ce projet a été rejeté par la *Deutsche Forschungs Gemeinschaft*, qui lui a préféré le projet CIB. Le projet (voir figure 7) est intéressant parce qu'il a proposé :

- Un Datalab (laboratoire de données) pour les expériences d'Open Linked Data, de groupement et de création des entités œuvre, d'enrichissement automatique des métadonnées etc.
- Un Datahub, un environnement bibliographique national qui sert aux services à l'échelle nationale comme la recherche, les index etc.
- Une plate-forme de catalogage partagé national qui est connecté aux sources mondiales comme WordCat.

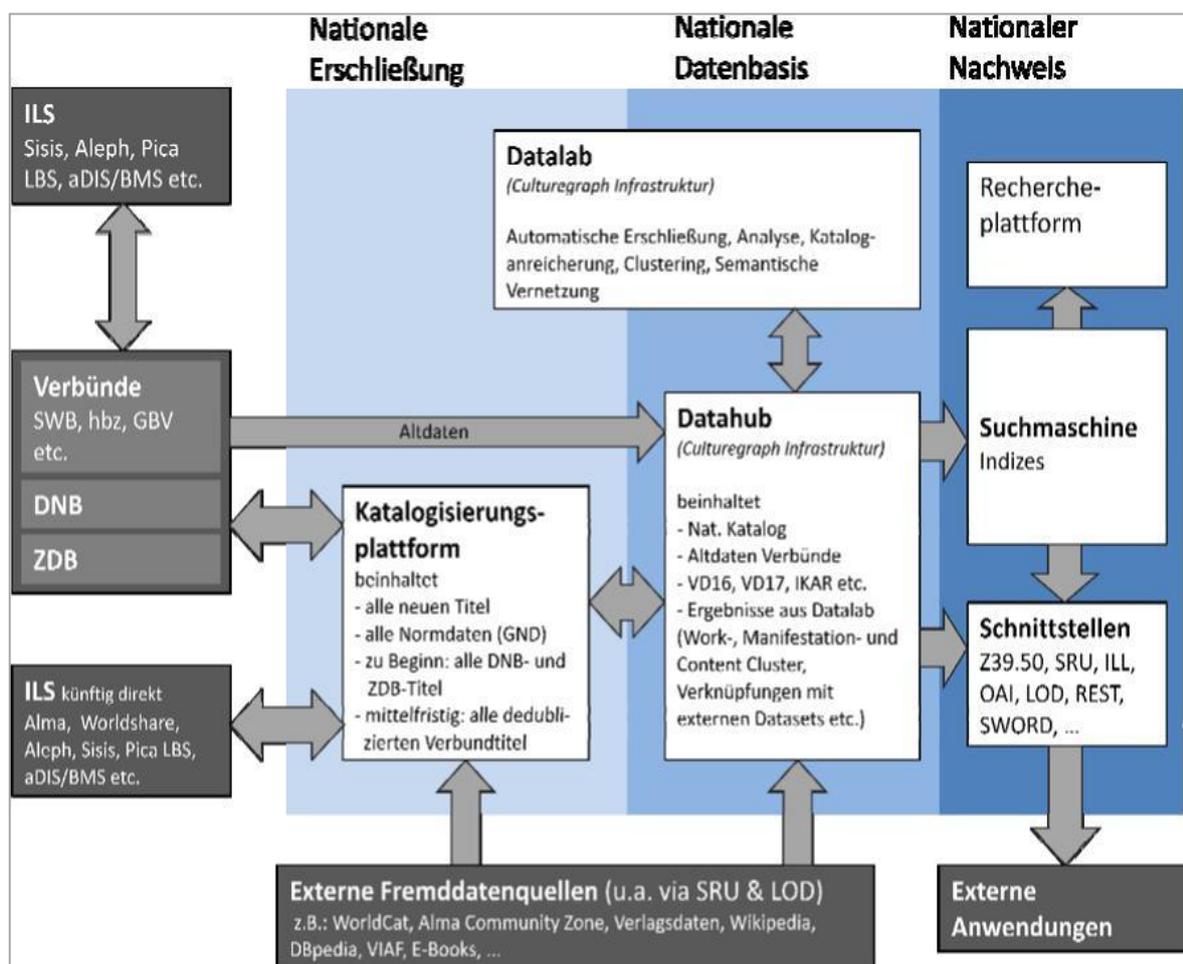


FIGURE 7 UN SCHÉMA DU PROJET LIBOS

5.4 CONCLUSIONS BASÉES SUR LES ÉTUDES DE CAS

Les études de cas menées dans le cadre de cette étude et décrites ci-dessus mènent aux conclusions suivantes :

Les fonctionnalités des ressources partagées peuvent être réparties en 4 catégories distinctes :

- Système de gestion de bibliothèques mutualisé
- Système de catalogage partagé
- Services communs basés sur des métadonnées du catalogage collectif
- Gestion partagée de ressources électroniques

Les études de cas relatives à la migration vers un système de Cloud montrent que :

- BIBSYS en Norvège est le seul projet qui envisage une migration de l'ensemble des 4 catégories de fonctionnalités vers le même système de Cloud. Les bibliothèques norvégiennes utilisent déjà un système de gestion mutualisé depuis 15 ans.
- Les bibliothèques des Pays-Bas sont les seules qui s'orientent vers une intégration complète de leur système de catalogage partagé et de leur catalogue collectif à la plate-forme WorldShare.
- Orbis Cascade continue d'utiliser WorldCat comme système de catalogage partagé, malgré leur migration vers le système d'Alma. Les métadonnées seront transférées vers une base bibliographique dans le système d'Alma qui sert à leurs services comme le PEB.
- Les autres projets envisagent un environnement de catalogage partagé qui se trouve dans le système de Cloud, connecté aux flux de métadonnées internationaux, mais avec une identité propre.
- Le projet CIB envisage un alignement de plusieurs systèmes de Cloud sur un espace de données national dans le Cloud et un "catalogue collectif" en dehors du Cloud.

Les études de cas relatives à l'auto-développement de systèmes du catalogage partagé montrent que :

- Le système suédois LIBRIS XL utilise un format bibliographique en RDF auto-développé et il est basé sur un logiciel auto-développé.
- Le système finlandais planifié sera développé à partir de logiciels Open Source, mais maintiendra le format Marc 21 comme standard international accepté.
- Les deux systèmes soulignent l'importance des données en libre accès (*CCO licence*). Cette position semble compliquer la connexion aux flux de métadonnées internationaux des deux systèmes.

Les autres études de cas montrent que :

- Les bibliothèques qui migrent leurs SGB vers un système de Cloud choisissent dans la majorité des cas d'utiliser aussi le système de catalogage partagé de ce système de Cloud (*Libraries Australia*).
- Le nouveau catalogue collectif des bibliothèques publiques des Pays-Bas montre les possibilités d'*Open Linked Data* pour une plate-forme de découverte.
- Le projet LIBOS souligne l'importance de disposer d'un environnement « laboratoire » pour tester et développer les innovations en matière de traitements automatiques et de Web sémantique, etc.

6. LES SCÉNARIOS POSSIBLES POUR LE SUDOC

6.1 LES ÉLÉMENTS MOTEURS DU CHANGEMENT DE MODE DE CATALOGAGE

En se basant sur cette étude, on peut identifier six éléments moteurs du changement de mode de catalogage :

- La grande majorité des éditeurs produisent les métadonnées et livrent ces métadonnées aux bibliothèques. Par conséquent, les systèmes de catalogage partagé s'orientent vers la possibilité de manipulation des flux de métadonnées provenant des éditeurs.
- Les méthodes d'acquisition des bibliothèques évoluent : l'acquisition « titre par titre » diminue et les acquisitions par bouquets ou par méthodes semi-automatiques (de type *demand-driven acquisition* ou *book approval plan*) augmentent. Le phénomène d'acquisitions par bouquets et les livraisons simultanées de métadonnées par les éditeurs rendront le catalogage manuel pour une large part superflu.
- Les livres numériques vont remplacer la plupart des livres imprimés scientifiques à moyen terme. La transition dans le domaine des sciences exactes avancera vite et sera peut être quasi-complète. La transition dans le domaine des sciences humaines et sociales avancera plus lentement et ne sera probablement pas complète.
- L'importance des catalogues de bibliothèque comme outils de découverte a déjà diminué et diminuera encore du fait de l'existence des nouveaux outils de découverte qui peuvent utiliser le texte complet issu des publications numériques.
- Les règles de catalogage, le format bibliographique seront adaptés en réponse aux exigences d'Internet (FRBR et les règles RDA) et de l'évolution vers le Web de données (le développement du nouveau format bibliographique BIBFRAME). En outre, les méthodes de catalogage utilisent de plus en plus les traitements automatiques dans le processus de catalogage.
- Les systèmes de Cloud commerciaux et globaux deviennent des concurrents des systèmes nationaux ou régionaux de catalogage partagé et des catalogues collectifs.

6.2 LES CONDITIONS PRÉALABLES

LES RÉSULTATS DES ENTRETIENS AVEC LES PARTENAIRES FRANÇAIS

La bibliothèque universitaire des langues et civilisations (BULAC) catalogue les titres des publications acquises en langue originale. Le système CBS actuel prend en charge la translittération automatique de ces titres depuis la langue originale vers l'alphabet latin. De plus, le format Unimarc permet la double écriture du titre (mais le format Marc 21 peut également soutenir cette double écriture). Or FRBR est très pertinent pour la BULAC en ce qui concerne les traductions de livres en langue orientales. C'est pourquoi la BULAC formule les exigences suivantes à l'égard du nouveau système de Sudoc :

- La translittération automatique de titres
- Un format bibliographique qui permet la double écriture
- La « FRBRisation » du nouveau système Sudoc.

La Bibliothèque nationale de France a exprimé un vif intérêt pour la création de l'interopérabilité de l'information bibliographique produite par les bibliothèques françaises. Depuis 2014 la BnF utilise la licence Etalab pour les métadonnées qu'elle produit, comme le réseau Sudoc.

Donc l'interopérabilité juridique des métadonnées des bibliothèques françaises est assurée. Si le réseau Sudoc venait à migrer vers un nouveau système, il lui faudrait conserver l'interopérabilité des formats bibliographiques (la BnF utilise actuellement le format INTERMARC, qui est compatible avec le format Unimarc). L'application éventuelle de règles RDA pour le catalogage doit être synchronisée. Les autres éléments d'interopérabilité des métadonnées des bibliothèques françaises sont l'interopérabilité des notices d'autorités et les identifiants. C'est pourquoi la BnF préconise une maîtrise du nouveau système par l'ABES en matière d'interopérabilité des métadonnées bibliographiques.

Les autres interlocuteurs, le bureau ISSN de France et l'INIST, n'ont pas formulé d'exigences quant au nouveau système du réseau Sudoc. Le bureau ISSN a souligné que le statut légal actuel du numéro ISSN peut empêcher l'accès libre à cette information dans les notices du réseau Sudoc aux autres utilisateurs d'un système de Cloud, qui ne participent pas au réseau Sudoc. L'interlocuteur de l'INIST a souligné les problèmes juridiques actuels de leur service RefDoc en ce qui concerne la fourniture de documents.

LES CONDITIONS RÉSULTANT DES TENDANCES RELATIVES AU CATALOGAGE ET À LA PRODUCTION DE MÉTADONNÉES

Les tendances relatives au catalogage et à la production de métadonnées, telles qu'elles sont décrites au chapitre 4, amènent à poser les conditions suivantes pour le nouveau système du Sudoc :

- Le nouveau système Sudoc doit être connecté aux flux de métadonnées internationaux et français.
- La gestion des flux de métadonnées issus des éditeurs internationaux doit être exécutée par un prestataire extérieur, car cette tâche dépasse la capacité de l'ABES.
- L'adhésion aux standards internationaux facilite la connexion aux flux de métadonnées internationaux.
- La dépendance croissante aux flux de métadonnées provenant des éditeurs implique aussi un choix de standards internationaux utilisés par la majorité des éditeurs.
- La gestion de flux de métadonnées des éditeurs français peut être exécutée par l'ABES en coopération avec la BnF et Couperin.
- Le système doit permettre les traitements de modification en masse et les traitements d'enrichissements en masse par l'ABES.
- Le système doit permettre de créer de nouveaux services à l'échelle nationale et de nouveaux services à l'échelle locale.
- Le système doit prévoir un environnement pour le catalogage collectif et un mécanisme d'importation des métadonnées provenant des bibliothèques françaises qui ont choisi d'utiliser un autre système de catalogage partagé. Une solution évidente peut être que le catalogue collectif et les services basés sur le catalogue collectif soient maîtrisés par l'ABES.

6.3 LES SCÉNARIOS CONSIDÉRÉS POUR LE SUDOC

UNE INTÉGRATION COMPLÈTE À UN SYSTÈME DE CLOUD

Un scénario possible pour le Sudoc peut être une intégration complète au même système de Cloud que celui choisi par les sites pilotes. Ce scénario est inspiré par l'exemple des bibliothèques universitaires aux Pays-Bas. Ce scénario peut impliquer les données suivantes :

- Le catalogage collectif constituerait un sous-ensemble d'une base de données mondiale.
- On utiliserait l'anglais comme langue de catalogage.
- On utiliserait un système d'autorités international
- Les participants au réseau Sudoc devraient suivre les développements de ce système de Cloud en ce qui concerne les règles de catalogage et le format bibliographique etc.

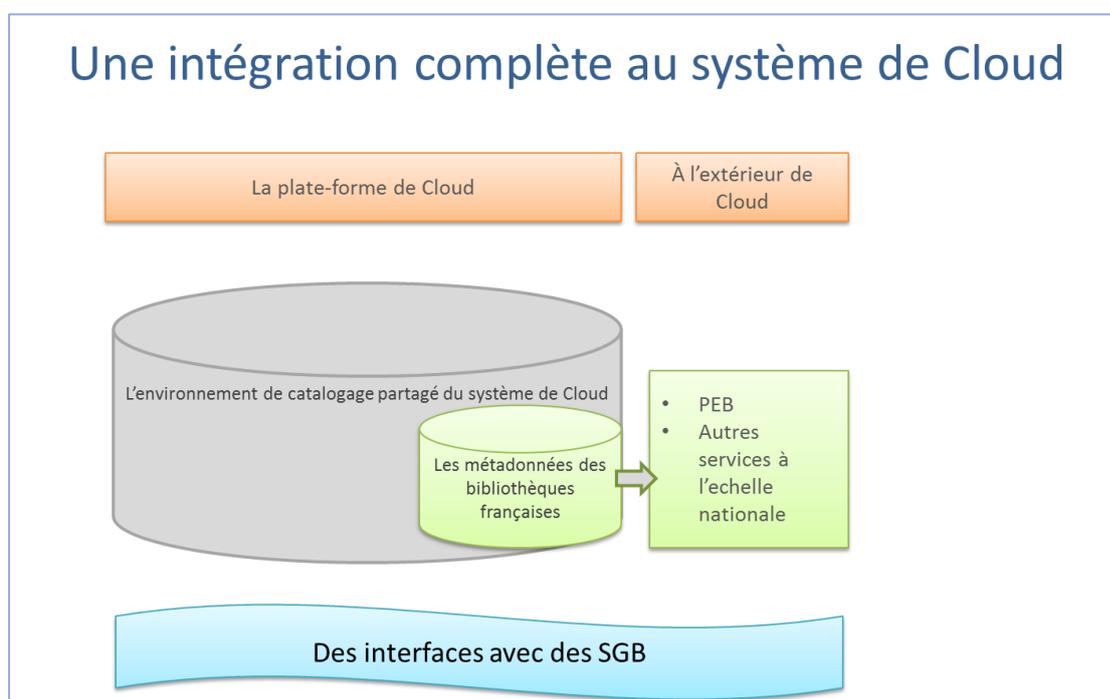


FIGURE 8 UN SCHÉMA D'UNE INTÉGRATION COMPLÈTE AU SYSTÈME DE CLOUD

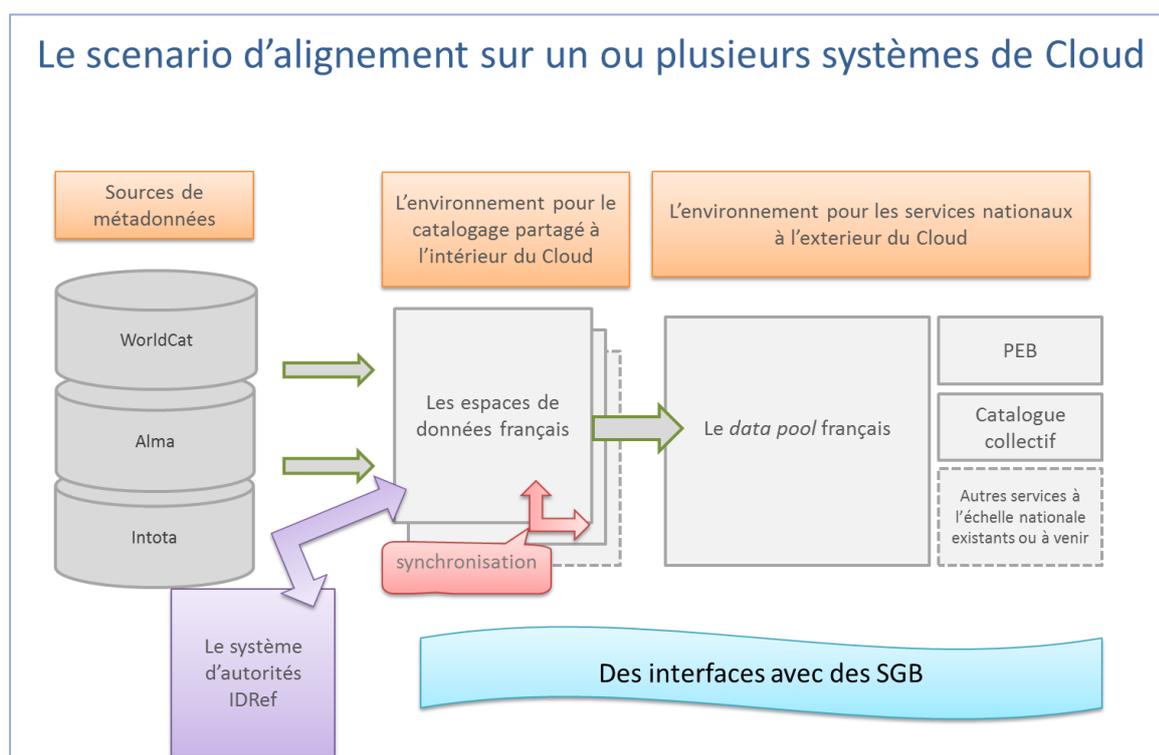


FIGURE 9 UN SCHEMA DU SCENARIO D'ALIGNEMENT SUR UN OU PLUSIEURS SYSTEMES DE CLOUD

Un autre scénario possible pour le Sudoc serait un alignement sur un ou plusieurs systèmes de Cloud. Ce scénario est inspiré par le projet CIB Allemagne. Ce scénario comporte les données suivantes :

- La création d'un espace de données françaises (*data space*)
 - L'idée est qu'il existera un *data space* dans chaque système de Cloud utilisé en France et que ces multiples *data spaces* seront synchronisés.
 - Pour alimenter ce *data space*, il existe deux options : (1) on peut utiliser WorldCat comme seule source de métadonnées, ou (2) on peut utiliser plusieurs systèmes de Cloud comme WorldCat et Alma comme sources de métadonnées.
 - L'objectif du *data space* est de servir au catalogage partagé et aux flux de métadonnées, surtout aux flux internationaux
- La création d'un *data pool* français séparé des systèmes de Cloud avec les fonctionnalités suivantes :
 - Une base de données pour les services à l'échelle nationale :
 - le catalogue collectif
 - le système de PEB
 - les services de gestion des collections (*collection management*)
 - les autres services à l'échelle nationale qui pourraient se développer dans le futur
 - Une garantie à l'égard des fournisseurs des systèmes de Cloud afin de préserver l'indépendance et la sécurité des métadonnées françaises

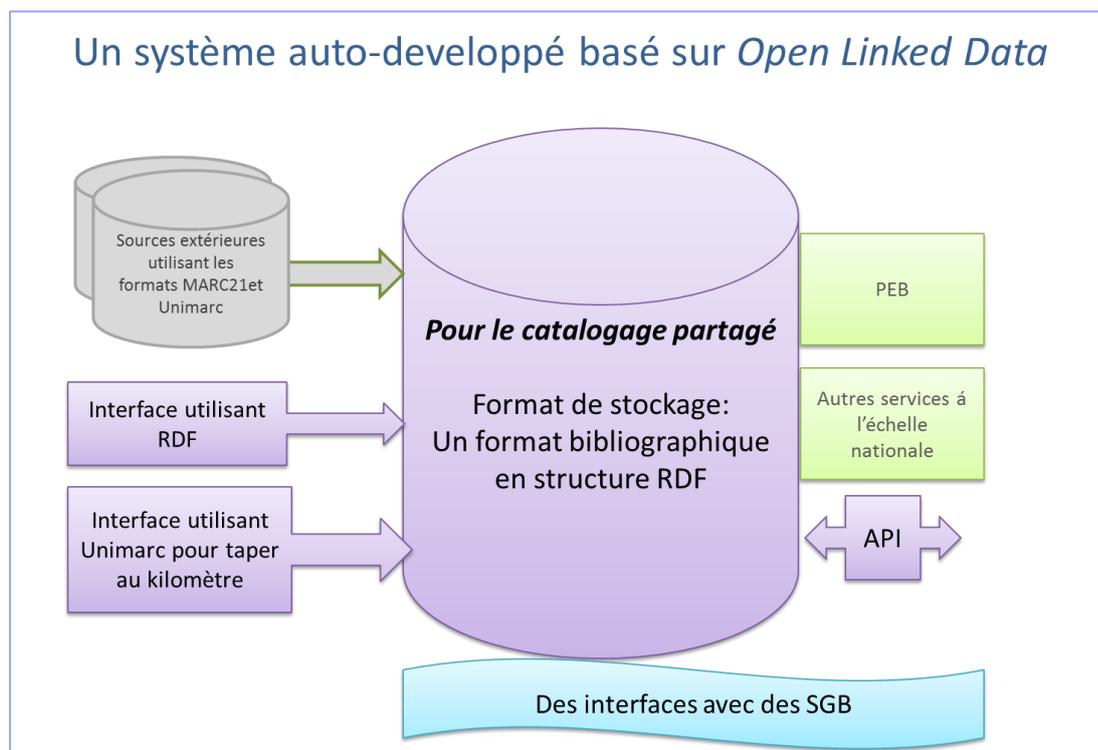


FIGURE 10 UN SCHEMA DU SCENARIO D'UN SYSTÈME AUTO'-DEVELOPPÉ BASÉ SUR *OPEN LINKED DATA*

Un troisième scénario possible pour le Sudoc envisage une migration vers le Web de données (*Open Linked Data* ; RDF). Ce scénario est inspiré par le développement de LIBRIS XL en Suède. Ce scénario implique :

- L'auto-développement du logiciel pour cette plate-forme
- L'auto-développement du format bibliographique en structure RDF ou l'attente du développement du format BIBFRAME de la Library of Congress
- Le développement d'un nouveau circuit de travail pour le catalogue ('from cataloguing to catalinking') utilisant les avantages de la structure RDF.
- L'auto-développement des services comme le PEB utilisant les données en structure RDF.
- Le développement de méthodes permettant le partage des ressources bibliographiques à l'échelle mondiale.

6.4 LES CONSIDÉRATIONS ISSUES DES SCÉNARIOS ET DES PROGRAMMES DE FONCTIONNALITÉS POUR LE DIALOGUE COMPETIF

Le réseau Sudoc utilise actuellement deux systèmes interconnectés : le système CBS d'OCLC Leiden et la base miroir en XML maintenue par l'ABES. Le scénario consistant à utiliser les systèmes existants pour affronter les éléments moteurs du changement de catalogage existe-il dans la réalité ? OCLC a souligné qu'ils perfectionneront la plate-forme WorldShare et ses services. Les innovations de la plate-forme WorldShare peuvent être connectées au système CBS dans plusieurs cas, mais il y a des limites. C'est pourquoi la suppression progressive du système CBS est envisagée à moyen terme. Ce scénario n'est donc plus examiné et les considérations se concentrent sur les trois scénarios décrits ci-dessous :

Le scénario d'une intégration complète à un système de Cloud ne semble pas bien cadrer avec l'infrastructure nationale. L'utilisation des notices d'autorités issues d'un système international ne semble pas acceptable dans le contexte français. Le scénario comporte d'autres contraintes : les traitements de modification en masse par l'ABES seraient impossibles ; les services nationaux seraient basés sur un système international sans avoir guère leur mot à dire au sujet des données internationales ; enfin le scénario implique une dépendance à l'égard d'un prestataire qui peut aboutir à une situation de marché captif. Si l'on récapitule brièvement, le scénario d'intégration complète à un système de Cloud ne donne pas assez de maîtrise au réseau Sudoc sur plusieurs éléments importants.

Le scénario d'alignement sur un ou plusieurs systèmes de Cloud semble aller de concert avec le scénario 2 de l'Étude d'Impact. Un autre avantage de ce scénario est que les bibliothèques françaises conservent leur liberté de choix pour la migration vers un système de Cloud. Ce scénario leur évite de se retrouver en situation de marché captif. L'environnement du *data pool français* leur donnerait toute la liberté de créer des services à l'échelle nationale. Néanmoins il demeure quelques soucis :

- La synchronisation en temps réel de deux ou trois espaces de données est mise en doute par certains membres du groupe de travail, en raison de mauvaises expériences dans ce domaine.
- Deux ou trois systèmes qui auraient deux ou trois interfaces de travail différentes, des formats différents, des règles de catalogage différentes, rendraient difficile la coordination du réseau Sudoc.
- Les systèmes de Cloud devraient aussi synchroniser leurs développements en matière de format bibliographique, etc.

Le scénario d'un système auto-développé basée sur *Open Linked Data* a éveillé beaucoup d'intérêt au sein du groupe de travail, car une telle migration vers le Web de données semble pouvoir relever les défis du futur (*future proof*). Néanmoins plusieurs interlocuteurs ont estimé que ce scénario était risqué. De même l'entretien avec un expert du Web de données laisse place aux doutes sur ce scénario pour le réseau Sudoc. Les soucis sont les suivants :

- Il manque un standard international accepté pour le format bibliographique en structures RDF
- Il manque les outils logiciels suffisants et performants pour permettre la migration d'un système de catalogage partagé et de catalogue collectif et pour supporter les services basés sur des métadonnées en RDF.

Une autre considération concerne les changements survenant dans le monde de catalogage : FRBR, les règles RDA et le développement de BIBFRAME. Dans cette étude il apparaît clairement qu' une incertitude persiste à l'égard de l'implémentation de FRBR : la théorie de FRBR prescrit trois niveaux, le schéma des éditeurs INDECS prescrit deux niveaux et BIBFRAME en développement prévoit également deux niveaux. Le calendrier du dialogue compétitif prévoit un choix du nouveau système pour le catalogage partagé et le catalogue collectif en 2015-2016. À ce moment-là - un choix du scénario effectué et le développement de BIBFRAME achevé - le réseau Sudoc pourra décider de la manière dont il implémentera FRBR et les règles RDA. Cela veut dire qu'une transition vers les règles RDA peut être implémentée simultanément au nouveau système. Du point de vue de la formation cela semble faisable (voir Annexe B).

7. RECOMMANDATIONS

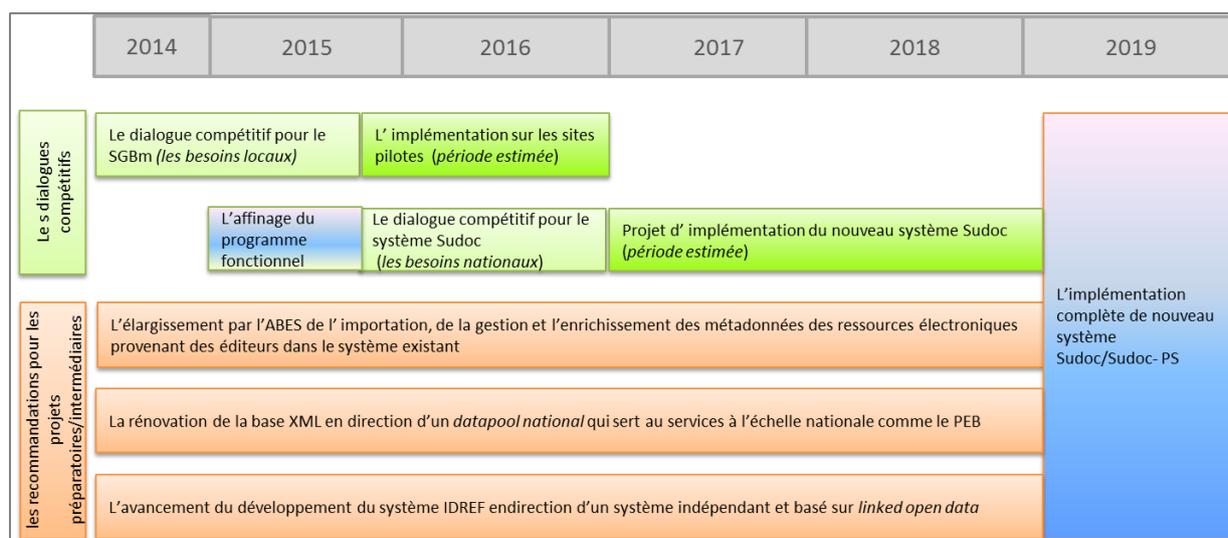


FIGURE 11 LE CALENDRIER DES DIALOGUES COMPÉTITIFS ET LES PROJETS PRÉPARATOIRES/ INTERMÉDIAIRES RECOMMANDÉS

Dans figure 11, le calendrier envisagé des dialogues compétitifs est présenté :

- Le dialogue compétitif pour le SGBm commencera en octobre 2014 et débouchera sur une décision à l'automne 2015. Ensuite, un projet d'implémentation sera mené. Ce dernier pourra occuper l'année 2016.
- Après la fin du dialogue compétitif concernant les besoins locaux, le dialogue compétitif concernant les besoins nationaux débutera. On peut estimer que ce dialogue compétitif occupera la majeure partie de l'année 2016. Ensuite un projet d'implémentation sera prévu sur une durée de deux ans.
- Ce calendrier implique qu'au début de l'année 2019 l'implémentation du nouveau système Sudoc/Sudoc-PS sera achevée.
- En ce qui concerne les éléments moteurs du changement de mode de catalogage (voir paragraphe 6.1) l'immobilité durant la période 2014-2018 n'est pas une option envisageable pour l'ABES. Quatre projets préparatoires sont recommandés durant cette période intermédiaire :
 - **L'affinage du programme fonctionnel** : cette étude a retenu deux scénarios faisables pour le nouveau système Sudoc : le scénario d'alignement sur un ou plusieurs systèmes de Cloud (s'inspirant du projet CIB en Allemagne) et un système basé sur *Open Linked Data* (s'inspirant du développement du système LIBRIS XL par la Bibliothèque nationale de Suède). Il est recommandé d'entreprendre un voyage d'études visant à découvrir ces deux projets au printemps 2015. À ce moment-là le système suédois sera implémenté depuis quelques mois et le projet allemand sera plus avancé. Il est envisagé que l'ABES fasse le choix entre ces deux scénarios, en se basant sur ces voyages d'études. Ce choix sera important pour affiner le programme fonctionnel qui servira au dialogue compétitif, mais également à l'autre projet recommandé, portant sur la rénovation de la base XML (voir ci-dessous).

- **L'élargissement par l'ABES de l'importation, de la gestion et de l'enrichissement des métadonnées des ressources électroniques provenant des éditeurs dans le système existant** : actuellement l'ABES gère quelques flux de métadonnées provenant des éditeurs. Il est proposé que l'ABES élargisse cette importation et cette gestion de flux de métadonnées et qu'elle élargisse en même temps son expérience des algorithmes d'enrichissement de ces métadonnées. Cette augmentation des efforts fournis est estimée cruciale pour le réseau Sudoc en raison de la croissance prévue des publications électroniques.
- **La rénovation de la base XML en direction d'un *data pool national* qui sera la base des services à l'échelle nationale, comme par exemple le PEB** : dans ce projet recommandé, il est envisagé d'utiliser la base XML comme un point de départ pour créer le *data pool* national et les services qui lui seront associés. Elle pourrait être alimentée en temps réel par un ou plusieurs nouveaux services de catalogage, comme elle l'est actuellement par CBS. L'interface de PÉRISCOPE utilise déjà la base XML. Un moteur SOLR lui est déjà associé. Toutefois les modèles de données qui alimenteront le data pool seront différents et plus hétérogènes et donc une profonde reconfiguration de la base XML sera sans doute nécessaire, sur la base des technologies actuellement maîtrisées par l'ABES. En conclusion le data pool et ses services à l'échelle nationale seront basés sur la base XML renouvelée et seront gérés par l'ABES. L'orientation de la rénovation de la base XML dépendra en partie du scénario retenu pour le nouveau système Sudoc (voir l'affinage du programme fonctionnel, mentionné ci-dessus).
- **L'avancement du développement du système IDREF en direction d'un système indépendant et basé sur *Open Linked Data*** : le système d'autorités IdRef jouera un rôle très important dans l'enrichissement des métadonnées provenant des éditeurs. Ce rôle important d'IdRef peut être élargi à d'autres domaines si le système devient indépendant du système de catalogage et basé sur *Open Linked Data*.

ANNEXE A : PARTICIPANTS Á L'ÉTUDE

Études de cas		
Finlande	La Bibliothèque nationale de Finlande	Kristiina Hormia, Ninan Hyvonen, Petteri Kivimaki
Suède	LIBRIS XL	Martin Malmsten
Norvège	BIBSYS	Arve Olaussen
Les bibliothèques universitaires des Pays-Bas	La bibliothèque de l'université Amsterdam	Marion van Brunshot; Martine van de Burg
	La bibliothèque de l'université Tilburg	Marc van den Berg
	La bibliothèque de l'université Erasmus	Maarten Steenhuis
Les bibliothèques publiques des Pays-Bas	Fondation Bibliotheek.nl	Johan Stapel
Orbis Cascade	Orbis Cascade Alliance	Al Cornish
Libraries Australia	La Bibliothèque nationale d'Australie	Monika Szunejko
CBUC	CSUC de Catalonia	Lluís Anglada; Ramon Ros
Projet CIB	Hebis	Uwe Risch
Les éditeurs et prestataires		
Springer	Miriam Kessler	
Elsevier	Maarten Cleeren	
STM Association	Eefke Smit	
EDiTEUR	Graham Bell; Tim Davenport	
OCLC Leiden	Eric van Lubeek	
Les autres experts		
GBV	Jakob Voss	
SURFMarket	Pim Slot	
Deutsche Nationalbibliothek	Renate Polak-Bennemann; Reinhard Altenhoener	

L'université de Pittsburgh	Karen Calhoun	
OCLC	Lorcan Dempsey	
Joint Steering Committee RDA	Gordon Dunsire	
Les partenaires du réseau SUDOC		
BULAC	Marie-Lise Tsagouria; Christophe Péralès; Benjamin Guichard	
BnF	Gildas Illien	
INIST	Raymond Bérard	
ISSN France	Philippe Cantié	

Les entretiens internes	Sujet
Olivier Rousseaux	Le catalogage partagé
Philippe Le Pape	Le catalogage partagé
Yann Nicolas	Le catalogage partagé
Laurent Piquemal	La formation de réseau Sudoc
Corinne Ceroni	SUPEB
Marianne Giloux	Les chiffres clés
Marie Lombard	Les chiffres clés
Deux ateliers avec l'équipe de direction	
Jérôme Kalfon	
Martine Brunet	
Camille Dumont	
Stéphane Rey	
Carole Melzac	
Yann Nicolas	

Participants de groupe de travail sur le catalogage partagé

Jean Bernon
Olivier Rousseaux
Philippe Le Pape
Yann Nicolas
Marianne Giloux
Mireille Teissède
Gregory Miura
Emilie Liard
Alexis Jean-Laurent
Julian Sempéré
Isabelle Mauger Perez
Jean-Marie Feurtet
Valerie Néouze
Frédérique Blondelle

ANNEXE B : LA FORMATION DU RÉSEAU SUDOC EN CAS DE TRANSITION VERS UN NOUVEAU SYSTÈME

LA FORMATION ACTUELLE DU RÉSEAU SUDOC

Les cours organisés par l'ABES qui sont pertinents pour une migration éventuelle sont listés dans le tableau au-dessous.

Cours ¹⁰	Durée	Nombre par année
Formation pour nouveaux catalogueurs entrants dans le réseau	3,5 jours	1 fois par année ; à l'ABES ; 12 participants par cours. Présentiel / à distance
Formation pour les nouveaux catalogueurs déjà dans le réseau(1)	3,5 jours	Il y a un relais de 14 formateurs qui prennent soin des cours présentiels dans 14 lieux différents ; 25-30 cours par année ; nombre des participants total : 200-250 catalogueurs par année
Formation continue pour les catalogueurs de réseau Sudoc (2)	1 heure	2 fois par mois ; session à distance ; interactif (visioconférence)
Formation SUPEB	2,5 jours	3 - 4 fois année ; à l'ABES ; 12 participants par cours

Les remarques suivantes sur la formation actuelle sont pertinentes :

- Il y a une transition lente de la formation en présentiel vers la formation en ligne via le *learning management system Moodle*.
- Deux fois par mois il y a une formation en ligne par visioconférence qui est très interactive.
- La formation à l'ABES pour les catalogueurs nouveaux est considérée importante pour la création de l'esprit du réseau Sudoc. Si on migre vers la formation en ligne, une autre méthode pour la création de l'esprit du réseau Sudoc doit être développée.
- Les participants au cours dispensé à l'ABES pour les catalogueurs nouveaux préféreraient une méthode de formation plus modulaire (par niveaux). Actuellement, le cours contient une formation exhaustive et souvent les participants n'utilisent pas tous les possibilités. La formation par module est envisagée, mais elle est compliquée par les problèmes logistiques.
- La formation pour l'utilisation de l'interface Colodus est entièrement en ligne.
- Les catalogueurs du réseau Sudoc attendent les règles RDA et la « FRBRization ».

¹⁰ Il y a des autres cours pour le rôle de coordinateur et le rôle de responsable de centre régional.

LA FORMATION EN CAS D'UNE TRANSITION VERS UN NOUVEAU SYSTÈME

Si le système du Sudoc migre vers un nouveau système avec une nouvelle interface de travail et si simultanément le réseau Sudoc commence à utiliser les règles RDA (voir section 6.4), il y a besoin :

- (1) d'un profil RDA (probablement développé à l'échelle nationale) avec des consignes pour le réseau Sudoc (une tâche pour l'ABES)
- (2) d'un guide méthodologique adapté (une tâche pour l'ABES)
- (3) de matériel pour la formation à la nouvelle interface de travail (une tâche pour l'ABES avec le soutien du prestataire).

La formation pour accompagner cette transition est possible si l'ABES utilise une combinaison de moyens éducatifs :

- La formation par module en ligne pour les catalogueurs qui produisent surtout les données d'exemplaires (comme la formation actuelle d'interface Colodus ; voir table 2)
- Une combinaison de formation en présentiel et de formation en ligne pour les catalogueurs qui créent de nouvelles notices bibliographiques (la majorité de ces catalogueurs appartient à un nombre limité d'établissements (voir table 2).