



Commission
Signalement et Systèmes
d'Information

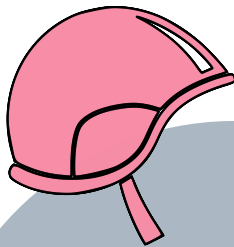
Manuel d'instruction

à l'usage

du bibliothécaire

débarquant en

INFO-DOC



Lexique et
vocabulaire
de base

Version mise à jour, mars 2024

Remerciements

Ce document est le fruit d'un travail collectif des membres de la commission Signalement et systèmes d'information de l'ADBU.

Nos remerciements spéciaux vont cependant à François-Xavier Boffy, Daniel Bourrion, Olivier Delangle, William Demet, Frédérique Joannic-Seta, Catherine Roussy, Marjolaine Simon et Frédérique Schlosser pour leur contribution majeure à ce projet.

Nous remercions également chaleureusement nos collègues David Chopard-Lallier, Agnès Leboucher et Guenno Marti, membres du Comité des services informatiques de l'enseignement supérieur et de la recherche (CSIESR), pour leur relecture attentive et avisée de nos définitions. Nous espérons que cette collaboration sera la première d'une longue série d'échanges au service du rapprochement de nos métiers.

Le projet

Vous connaissiez le *Manuel d'instruction à l'usage du soldat britannique débarquant en France en 1944* mais avez-vous déjà entendu parler du *Manuel d'instruction à l'usage du bibliothécaire débarquant en Info-doc en 2023*?

Et bien c'est chose faite!

Ce manuel de poche conçu sous forme de lexique est avant tout un guide de savoir-vivre destiné aux recrues amenées à prendre un poste relevant du domaine fonctionnel de l'informatique-documentaire. Il s'agit pour ces bibliothécaires intégrant les bataillons d'un S.I.D. d'affûter leur clavier et de réviser leurs octets. Il nous faut cependant faire preuve de pédagogie avant de les embarquer.

Sans se substituer à un dictionnaire ou à une formation spécialisée, ce lexique se veut avant tout un livret de traduction-conversation basique pour faciliter leur atterrissage. Pour commencer, quelques recommandations :

D'abord circonscrire le périmètre

L'informatique-documentaire consiste en l'application de techniques et d'outils informatiques dans le domaine spécifique de la gestion et de la recherche documentaire. Dans les faits, le champ d'action des services "info-doc" en bibliothèques universitaires n'a cessé de s'élargir au cours des dernières décennies sous l'effet, d'une part de l'informatisation et de la dématérialisation des supports, des outils et des processus de travail, d'autre part de l'élargissement du rôle et des missions des BU au sein des écosystèmes d'enseignement et de recherche.

Soutien premier des bataillons de catalogueurs pour le déploiement des catalogues en ligne, l'info-doc continue d'élargir ses terrains d'opération: RFID, appui à la recherche et à la formation, automatisation des processus métiers, développement de logiciels et bases de données etc. Si la gestion des SIGB et catalogues reste aujourd'hui une mission importante des S.I.D., les compétences attendues autant que les interlocuteurs de ces services n'ont cessé de se diversifier.

Pacifier les relations SID/DSI

Malgré notre histoire commune, notre proximité culturelle et un goût partagé des bases de données, nous paraissions toujours les uns pour les autres *so bizarre*. Des cousins un peu éloignés, aussi intelligibles qu'une notice UNIMARC en cours de LRMisation.

Il convient donc d'éviter les écueils. Afin de naviguer sereinement au sein de ces environnements informatisés multiples et ramifiés, la recrue devra apprendre à maîtriser plusieurs langues: celle de son service pour commencer, puis celle de sa D.S.I. dans sa diversité, celle des fournisseurs et prestataires ou encore celle des catalogueurs et spécialistes du signalement en poste de longue date. Astucieusement placé sur sa table de chevet ou dans ses commodités, ce lexique deviendra son meilleur allié.

S'appuyer sur ses alliés.

Débarquer en info-doc c'est aussi infiltrer de nouveaux réseaux ou en redécouvrir sous de nouveaux angles. Pénétrer les coulisses techniques de l'ABES et ses joyeuses API, se mettre au diapason de Couperin et de son EZ Family ou embarquer dans une odyssée made in CNRS avec HAL. Lignes de commande, protocoles, balises, codes... Bienvenue dans Mr Robot.

Enfin affirmer son Autorité

Une bonne recrue doit pouvoir parler avec Autorités sur les personnes, ressources, structures ou encore concepts. Il convient pour cela de maîtriser les règles de l'interopérabilité qu'il s'agisse de protocoles de communication, d'identifiants pérennes, d'ontologie ou vocabulaires contrôlés. Une belle machinerie sous-marine au service du signalement. Vous avez aimé le Nautilus, vous aimerez encore plus les thésaurus.

Pour conclure, nous espérons que ce lexique permettra à nos camarades fraîchement parachutés en terre inconnue ou méconnue de prendre rapidement langue, au sens propre comme au figuré, avec leurs nouveaux collègues et partenaires.

Table des matières

.....	1
REMERCIEMENTS	3
LE PROJET	5
PARLER AVEC SES CATALOGUEURS	9
ACTIVITES	10
NORMES ET STANDARDS DE DESCRIPTIONS BIBLIOGRAPHIQUES	11
PARLER AVEC SES FOURNISSEURS ET PRESTATAIRES	16
ERMS	17
SIG/SGBB.....	17
OUTILS DE DECOUVERTE	18
RESOLVEUR DE LIENS	20
PARLER AVEC SA DSI OU SA DNUM	22
GENERALITES	23
HEBERGEMENT	24
PROTOCOLES D'ECHANGES.....	25
ACCES.....	26
IDENTIFICATIONS	27
BASES DE DONNEES.....	28
PARAMETRAGES	30
TYPES DE MOTEURS DE RECHERCHE	30
ANNUAIRES	30
PROGRAMMATION	31
CMS.....	32
STRUCTURATION DE DONNEES.....	32
TRANSFORMATION/NETTOYAGE DE DONNEES	33
CONNECTEURS	33
RFID.....	34
PARLER AVEC SES PARTENAIRES	36
ABES.....	36
CNRS.....	40
COUPERIN	40
OCLC	41
PARLER AVEC AUTORITES	42
MAPPING	43
PID, REGISTRES ET AUTORITES.....	43
DOCUMENTS ET RESSOURCES.....	45
PERSONNES	47
STRUCTURES	48
CONCEPTS	49
WEB SEMANTIQUE, WEB DE DONNEES	50

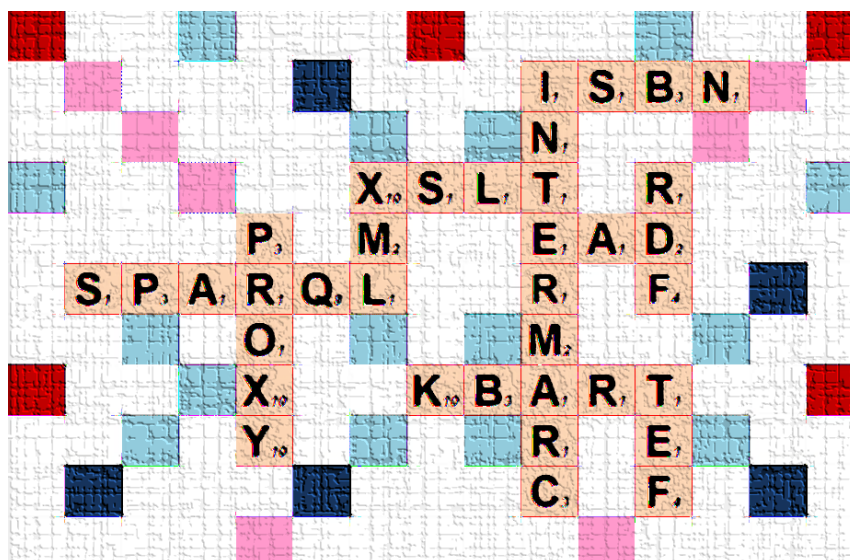
Parler avec ses catalogueurs

A l'origine, le catalogage vise à décrire un document de telle sorte qu'un lecteur puisse le retrouver dans le catalogue grâce à divers points d'entrée tels que le titre, l'auteur ou le sujet.

Ce processus d'apparence simple implique de la part du catalogueur une série d'actions allant de la description bibliographique à l'exemplarisation en passant par l'indexation, la dérivation ou encore le dédoublement en vue de garantir un signalement optimal des ressources.

Sous l'effet conjoint de la dématérialisation des contenus et des supports et de l'élargissement du champ d'actions des services documentaires, les catalogueurs sont aujourd'hui amenés à connaître, dans certains cas à maîtriser, un arsenal de normes et de standards de description et d'échange de données de plus en plus large et varié, répondant à des jargons techniques ou métiers spécifiques.

Si ces derniers peuvent paraître abscons à des oreilles non-initiées, ils permettent néanmoins de placer quelques lettres au scrabble à défaut de briller en société (yes, *KBART is a valid scrabble word!*).



ACTIVITES

— Signalement

Description de documents constituant un fonds pour en assurer la diffusion, le rendre accessible à la recherche et en favoriser la consultation. On utilise parfois signalement de préférence à catalogage dans le cas de ressources numériques pour suggérer des modalités de description par lots et/ou selon des normes différentes de l'ISBD.

— Catalogage

Description bibliographique d'un document, par la saisie des données descriptives utiles à son identification selon une structuration définie selon les prescriptions de normes nationales ou internationales, par exemple la norme ISBD (description bibliographique internationale normalisée) pour l'élaboration des notices bibliographiques.

— Circuit du document

Ensemble des opérations relatives aux documents réalisées dans une structure donnée (bibliothèque, etc.) : acquisition, réception, traitement documentaire, équipement, mise en rayon, désherbage.

— Exemplarisation

L'exemplarisation est la création d'un exemplaire en se localisant sous la notice bibliographique correspondant au document considéré. Cette opération peut être réalisée dans le catalogue local ou dans un catalogue collectif. Dans le SUDOC, l'exemplarisation peut se faire soit document par document (par exemple dans l'application Colodus) soit par lot (dans l'application ITEM dans le cas par exemple de signalements rétrospectifs).

— Dédoublonnage

Méthode qui permet d'identifier et de supprimer les doublons non souhaités présents dans une liste, un fichier, une base de données. Dans les catalogues, les doublons sont constitués de plusieurs enregistrements décrivant une même entité (document, notice bibliographique, notice autorité, etc.). Le dédoublonnage se traduit par exemple par la fusion des enregistrements identifiés, la suppression de l'un d'entre eux, etc.

— Taux de recouvrement

Le calcul du taux de recouvrement entre 2 ensembles ou 2 bases de données donne le pourcentage d'éléments du premier ensemble présent dans le second. Le taux de recouvrement SUDOC indique le pourcentage de notices d'un échantillon qui existent déjà dans le SUDOC. Il est utile de le calculer dans le cadre de projets d'exemplarisation automatique afin d'en mesurer la pertinence. Plus le taux est élevé et moins la charge de travail après exemplarisation automatique sera importante.

— Dérivation de notices

Opération consistant lors du catalogage à récupérer la notice relative au document à traiter dans un autre catalogue et à la télécharger, à l'importer dans son environnement de catalogage (et à la modifier si besoin). Par exemple, depuis WinibW, il est possible de dériver les notices d'OCLC Worldcat, du catalogue de la BnF, etc.

— Indexation (au sens informatique)

Processus informatique consistant à référencer des données (documents, mots-clés, etc.) pour faciliter leur accès ou leur traitement en stockant et en organisant leur contenu sous forme d'index. Dans les bases de données, un index est une structure de données utilisée pour retrouver rapidement les données. L'utilisation d'un index simplifie et accélère les opérations de recherche, de tri, etc.

— Indexation (au sens documentaire)

Description du contenu intellectuel d'un document à partir de l'analyse de son contenu au moyen de descripteurs, termes, mots-clé, indices et afin d'en favoriser la recherche. Par exemple, le langage encyclopédique d'indexation RAMEAU est largement utilisé en France et dans d'autres pays francophones.

NORMES ET STANDARDS DE DESCRIPTIONS BIBLIOGRAPHIQUES

— Autorité

Forme normalisée et contrôlée des points d'accès relatifs à une ressource dans un catalogue. Dans les catalogues de bibliothèques, les autorités sont principalement relatives aux "auteurs", aux "sujets" et aux "titres". Il existe des notices descriptives des autorités qui indiquent notamment la forme retenue, d'éventuelles formes rejetées, des éléments descriptifs et des liens vers les notices bibliographiques et vers d'autres notices d'autorité.

— Modèle

Représentation abstraite et schématique de données. Dans le cadre de la transition bibliographique, c'est le modèle IFLA Library Reference Model (IFLA LRM) qui devient le modèle conceptuel de référence pour les bibliothèques ; il offre une représentation schématique de l'activité des lecteurs et des usagers de catalogue de bibliothèques en structurant les métadonnées bibliographiques en s'appuyant sur trois notions : entités, attributs et relations. Le modèle OEMI (ou WEMI en anglais) désigne ainsi les quatre entités (Œuvre-Expression-Manifestation-Item) du modèle [FRBR](#) permettant de rendre compte de toute ressource bibliographique.

— Code de catalogage

Ensemble de règles de catalogage à valeur normative : le code RDA-FR, inscrit dans la logique du modèle IFLA-LRM, a vocation à remplacer les différentes normes Afnor en vigueur. Il est la transposition française du code RDA (Resource Description and Access) code de catalogage anglo-saxon à vocation internationale paru en 2010 et révisé en 2019.

— Normes

Ensemble de règles, de critères s'appliquant à un contexte spécifique. Pour le catalogage, les normes Afnor en vigueur déclinent l'ISBD au niveau national, par exemple la norme Z 44-050 d'avril 2005 intitulée : "Documentation – Catalogage des monographies – Textes imprimés – Rédaction de la description bibliographique".

— Format

Agencement structuré de données. Pour les données bibliographiques, structures de notices bibliographiques définies en vue d'échanges et de traitements informatiques : les formats MARC (MAchine-Readable Cataloging) sont aujourd'hui utilisés, UNIMARC dans le SUDOC, INTERMARC à la BnF, MARC 21 dans le monde anglo-saxon.

— UNIMARC

Format MARC créé par l'IFLA en 1977, il est utilisé en France comme format officiel d'échanges de données et sur lequel repose le format de production des données bibliographiques et d'autorité du SUDOC.

— Intermarc

Format MARC de production des données bibliographiques et d'autorité créé par la Bibliothèque nationale de France pour la majeure partie de ses collections.

— Marc 21

Format MARC de production des données bibliographiques et d'autorité largement utilisé dans le monde anglo-saxon.

— EAD

EAD (EAD-Encoding Archive Description/Description Archivistique Encodée) Format de structuration de l'information adapté au traitement des fonds d'archives et à l'informatisation des instruments de recherche archivistiques, constituant une DTD/XML Ce format respecte les principes définis par la norme générale et internationale de description archivistique, ISAD(G), dans sa seconde édition adoptée en 1999 par le Conseil international des archives.

[Pour en savoir plus](#)

— TEF (Thèses électroniques françaises)

Modèle de données élaboré dans le cadre de la recommandation AFNOR TEF – Thèses Électroniques Françaises. Le schéma TEF comporte des métadonnées descriptives (bibliographiques) et des métadonnées de gestion (administration, droits, conservation). Ce modèle est notamment utilisé dans le cadre des applications STEP et STAR, mises en œuvre par l'Abes.

— Kbart

Kbart (kаbа_Б ou kаbа_{BT}) : « Knowledge bases and related tools » : recommandation internationale qui préconise l'utilisation d'un format standard de données pour la gestion des accès et le signalement de la documentation électronique. Les fichiers Kbart sont des fichiers tabulés composés de métadonnées de titres et de données de couverture des fournisseurs de

contenus. Ils sont utilisés dans les bases de connaissances des outils de découverte. Bacon est un réservoir de fichiers Kbart.

— Dublin Core

Format de description de documents numériques ou physiques, simple et standardisé, ayant pour objectif de définir un socle commun d'éléments descriptifs et pour vocation de permettre une interopérabilité minimale entre des systèmes conçus indépendamment les uns des autres. Il prévoit à l'origine 15 éléments, tous facultatifs et tous répétables (titre, sujet, langue, créateur, etc.). Ces éléments constituent aujourd'hui un sous-ensemble des DCMI Termes plus complets. Le Dublin Core reste le format obligatoire dans le cadre du protocole OAI-PMH.

Voir : <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

— Transition bibliographique

Le programme *Transition bibliographique* a été lancé en 2015 par l'Abes et la BnF afin d'améliorer l'exposition des données bibliographiques et d'autorités dans le web de données en leur appliquant le modèle FRBR puis IFLA-LRM et en adoptant les nouvelles règles de catalogage du code RDA-FR.

— FRBR (Functional Requirements for Bibliographic Records)

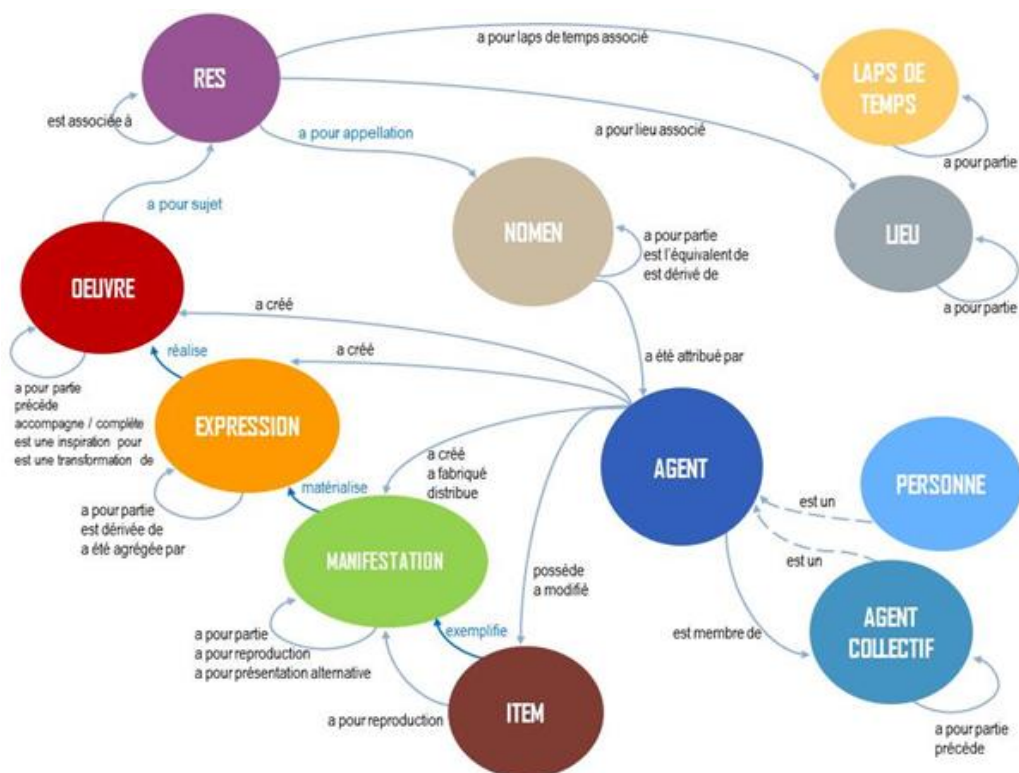
Modélisation conceptuelle des informations contenues dans les notices bibliographiques, développée dans les années 1990 puis progressivement remplacée par le modèle IFLA-LRM au mi-temps des années 2010 (Voir IFLA-LRM).

— RDA-FR

Le code RDA-FR est la transposition française du code RDA (Resource Description and Access), code de catalogage anglo-saxon à vocation internationale paru en 2010. Il s'appuie sur le modèle IFLA-LRM et définit une nouvelle approche du catalogage adaptée à l'environnement web et fondé sur un réseau de relations entre les entités de toute ressource documentaire : œuvre, expression, manifestation, item, etc. Voir aussi Code de catalogage.

Plus d'information <https://www.transition-bibliographique.fr/enjeux/definition-rda/>

— IFLA-LRM : (Library Reference Model)



Plus d'informations : <https://www.transition-bibliographique.fr/enjeux/definition-ifla-lrm/>

— Bibliostratus

Bibliostratus (Stratégie d'alignement d'URIs pour la Transition bibliographique) : logiciel d'alignement développé par le groupe Systèmes & Données du programme Transition bibliographique. Ce logiciel propose un ensemble de fonctionnalités permettant de s'aligner avec les données bibliographiques de la BnF, du Sudoc et d'IdRef. À partir de l'analyse de lots de notices bibliographiques ou d'autorité, Bibliostratus permet de trouver des identifiants ARK ou PPN correspondants et de récupérer les données des agences bibliographiques afin de les importer dans un système local (notices complètes ou identifiant seul).

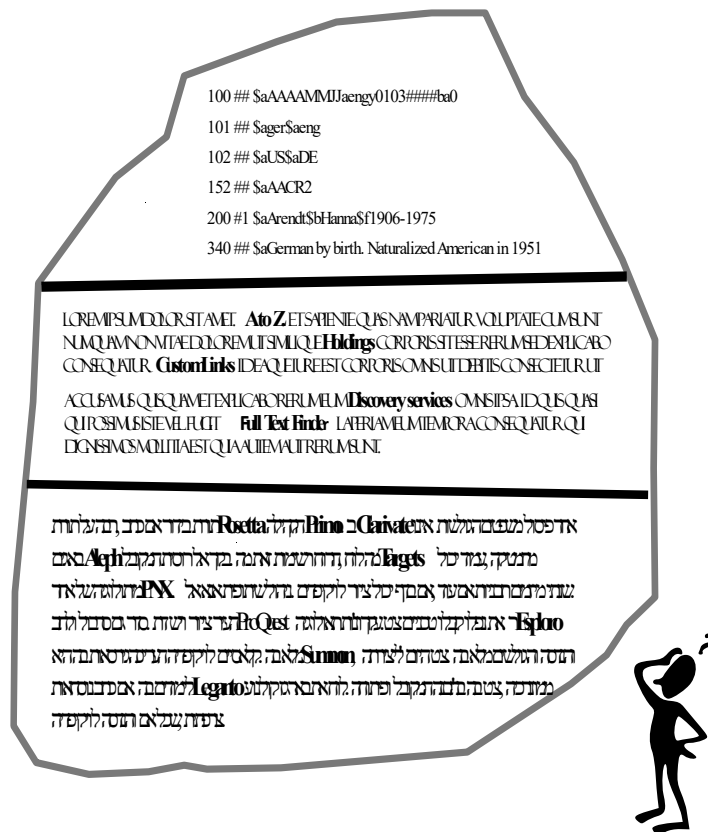
Pour en savoir plus : <https://www.transition-bibliographique.fr/systemes-et-donnees/bibliostratus/>

Parler avec ses fournisseurs et prestataires

De nombreux collègues peuvent se vanter de parler l'anglais, l'hébreu, le latin ou le grec ancien, mais seuls quelques *happy few* au sein des S.I.D. sont devenus les Champollion de l'Ebsco ou l'Ex-libris. Heureusement pour vous, nous les avons identifiés !

Et pour cause, nos outils de gestion et de signalement désormais nombreux au sein de systèmes d'information, tantôt de plus en plus intégrés tantôt de plus en plus morcelés, appellent une collaboration renforcée avec divers fournisseurs et prestataires de logiciels ou de services. Dans cette nébuleuse en perpétuelle évolution, le bon fonctionnement de nos services est tributaire de la qualité de nos échanges avec ces fournisseurs et prestataires.

Comprendre et se faire comprendre, d'un environnement à l'autre, d'un prestataire à l'autre, c'est bien souvent à cela que tient le secret d'une opération !



ERMS

— ERMS

Les ERMS (Electronic Resource Management System) sont des outils de gestion des ressources électroniques. Ils sont apparus au début des années 2000 pour pallier les difficultés de gestion des ressources électroniques dans les SIGB classiques conçus pour les collections papier. Ils permettent en particulier de gérer les licences et les accès aux ressources électroniques. Des communications sont généralement prévues avec les SIGB pour permettre une intégration des ressources électroniques dans le catalogue.

SIG/SGBB

— SIGB/SGB

Un SGB (système intégré de gestion de bibliothèque) dit également système de gestion de bibliothèque est un logiciel qui permet de gérer de manière informatique et intégrée le fonctionnement d'une bibliothèque. Le même outil permet d'acquérir, de cataloguer, de diffuser (catalogue), de gérer toutes les opérations de prêt, réservations incluses. Il permet également de gérer les documents de l'entrée à la sortie des fonds et de produire des indicateurs statistiques. Ces différentes activités sont généralement organisées en modules intégrés au sein d'un même logiciel. Malgré son nom, certaines activités restent hors périmètre, par exemple la gestion du planning ou la bibliothèque numérique. Le SIGB est alors complété par des outils (briques) avec lesquels il peut éventuellement communiquer.

— Zone/Champs/Étiquette

Ces termes sont utilisés pour décrire la structuration des données dans un système informatique, ils définissent la manière dont doivent être saisies les informations soit sous forme de code soit sous forme de texte libre.

— SGBM

Le SGBM (système de Gestion de Bibliothèque mutualisé) est le nom d'un projet piloté par l'ABES qui avait pour but de rechercher en commun (bibliothèques du réseau Sudoc et Sudoc lui-même) un SIGB qui faciliterait les échanges entre catalogues locaux et Sudoc, et intégrerait également la gestion de tous les types de ressources, en particulier les collections électroniques, et la mise en œuvre des préconisations et possibilités induites par la Transition Bibliographique. La mutualisation devait permettre d'obtenir ce SIGB de nouvelle génération à un coût financier maîtrisé. Lancé en 2012, ce projet a évolué, avec un recentrage sur les seules bibliothèques volontaires du réseau, soit 49 bibliothèques à l'issue d'un accord cadre en 2016, qui retenait 4 solutions logicielles : Data management de Décalog, Alma d'Ex-Libris (ProQuest), Linagora de Biblibre et OCLC, elles sont progressivement passées en production de 2018 à 2022. A ce jour, la solution Ex-Libris a été choisie par la grande majorité des établissements.

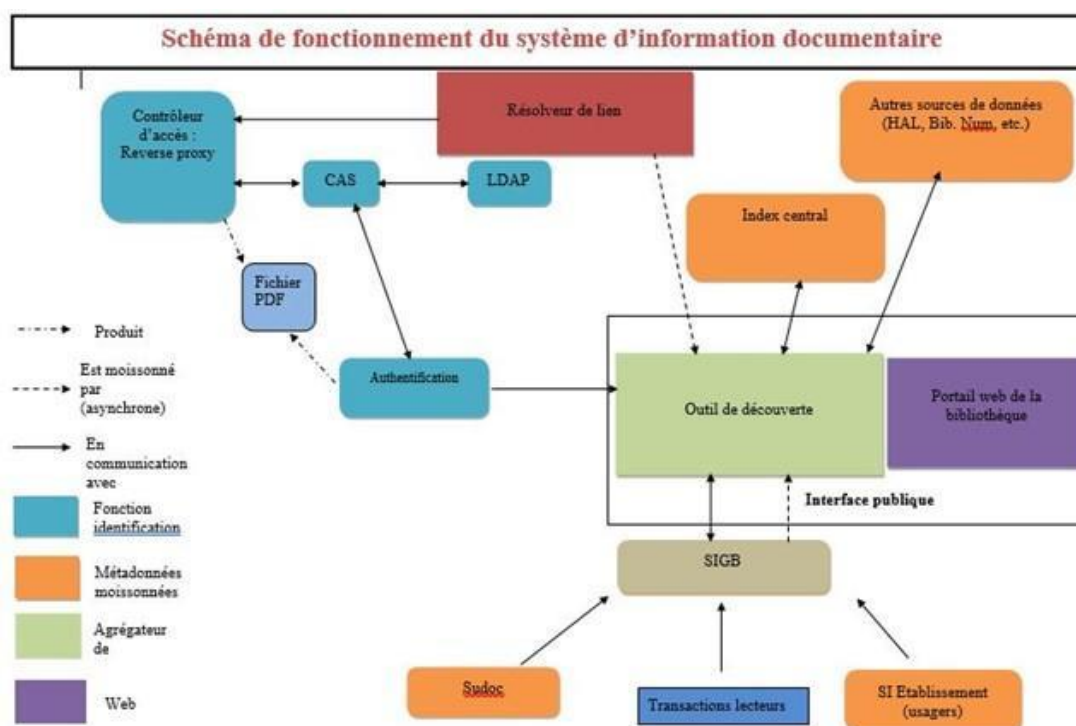
OUTILS DE DECOUVERTE

— Outil de découverte / *discovery tool*

Un outil de découverte est un système de recherche de bibliothèques basés sur la technologie des moteurs de recherche. Ce terme générique regroupe en réalité plusieurs catégories :

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

— Schéma global



— Base de connaissances

Base de données rassemblant de façon structurée des informations sur des ressources en lignes (bases de données, périodiques électroniques). Y sont en particulier mentionnées les informations permettant d'accéder à la ressource titres, périmètres des bouquets, états de

collection, URL d'accès. Exemple de base de connaissances : BACON, Base de connaissance nationale développée par l'ABES.

— **Index central**

Index présent dans un outil de découverte et référençant les ressources contenues dans la base de connaissances.

— **Index local**

Index présent dans le S(I)GB et référençant les ressources locales possédées par l'établissement.

— **Activation des ressources**

Les outils d'administration des ressources en ligne sont la plupart du temps reliés à un index central très large, tendant à l'exhaustivité. Pourtant, qu'on soit à Harvard ou à la Scuola Normale Superiore di Pisa, il n'y a jamais accès à toute la documentation en ligne. Activer les ressources, dans un outil de ce type, consiste à signaler celles qui sont effectivement accessibles pour les usagers de l'institution. Cela constitue l'un des paramétrages essentiels de l'index local.

— **Granularité**

Les bases de connaissance, bases de données documentaires, peuvent donner des informations et des capacités d'action sur des unités documentaires isolées, mais aussi sur des regroupements (par collection, série, publication périodique, bouquet de revues ou de monographies). On parle de granularités différentes pour désigner ces différents ensembles et sous-ensembles. La gestion de la granularité est un aspect essentiel de la fonction de découverte et a un effet direct sur le bruit ou le silence des résultats de recherche.

— **Vocabulaire spécifique Ex-Libris**

- PNX : Primo normalized XML – fichier au format XML structuré et normalisé par ExLibris afin de contenir les métadonnées des notices de ressources documentaires en vue d'une indexation et d'un affichage dans leur outil de découverte Primo
- Targets : Regroupement de revues ou d'ebooks au sein d'un groupe artificiel ou correspondant à une négociation afin de gérer plus facilement le signalement et l'accès au texte intégral. Équivalent des holdings chez EBSCO.

— **Vocabulaire spécifique Ebsco**

- *Holdings* : Ressources ou regroupements de ressources ("Packages") avec des granularités variables que les administrateurs des solutions de découverte EBSCO déclarent comme accessibles pour leurs usagers. Les métadonnées de ces ressources s'ajoutent alors à la base spécifique de l'institution cliente. Équivalent de targets chez Ex-libris.
- *CustomLinks* : Lorsqu'une ressource accessible (holding) ne fait pas partie de la base de connaissance d'EBSCO ou nécessite un lien ou un protocole spécifique, les administrateurs de l'outil peuvent préciser manuellement les paramètres. Ces liens sont aussi un moyen de proposer des alternatives à l'accès direct au texte intégral (orientation vers le service du PEB par exemple).

RESOLVEUR DE LIENS

— Résolveur de liens

Outil informatique se basant sur le protocole Open URL pour faire le lien entre la référence bibliographique d'une ressource et son texte intégral.

— Open URL

Protocole standardisé d'URL ayant pour but un accès facilité aux ressources en ligne qu'un usager a le droit de consulter, que celles-ci soient libres ou disponibles via un abonnement pris par sa bibliothèque. Le protocole contient les métadonnées et des éléments de contexte comme le serveur d'origine de la requête.

— CrossRef

Organisation à but non lucratif enregistrant les DOI (Digital Object Identifier) afin de permettre de mieux identifier les objets numériques grâce à un identifiant unique et normalisé. Émergence dans les années 2000.

— A to Z

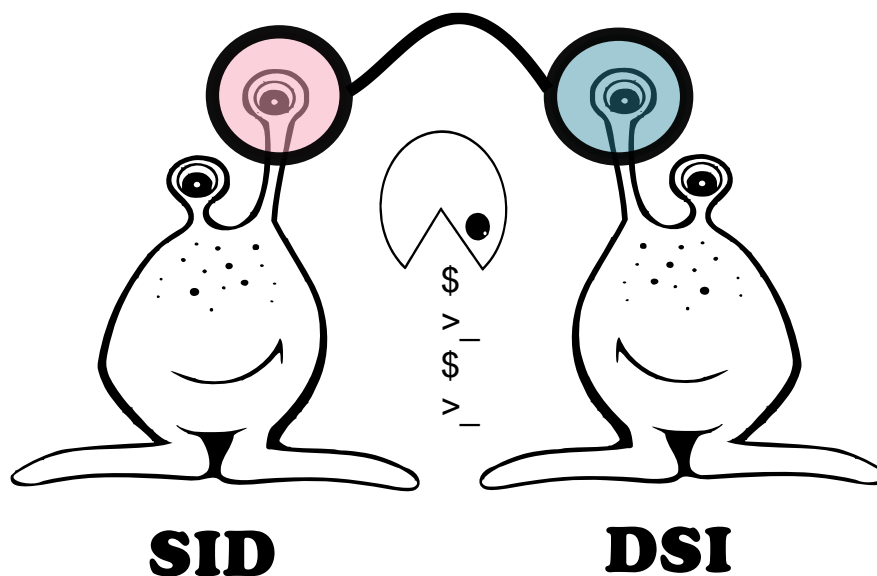
On appelle « A to Z » la liste alphabétique des ressources indexées dans une base de connaissance.

Parler avec sa DSI ou sa DNUM

Nous percevons parfois à tort les directions des services informatiques ou numériques comme des entités homogènes, restreintes au domaine de la programmation ou à la fourniture des machines alors qu'elles se composent de multiples entités. Ces dernières reflètent la diversité des services, technologies et compétences que revêtent aujourd'hui les termes « informatique » et « numérique ».

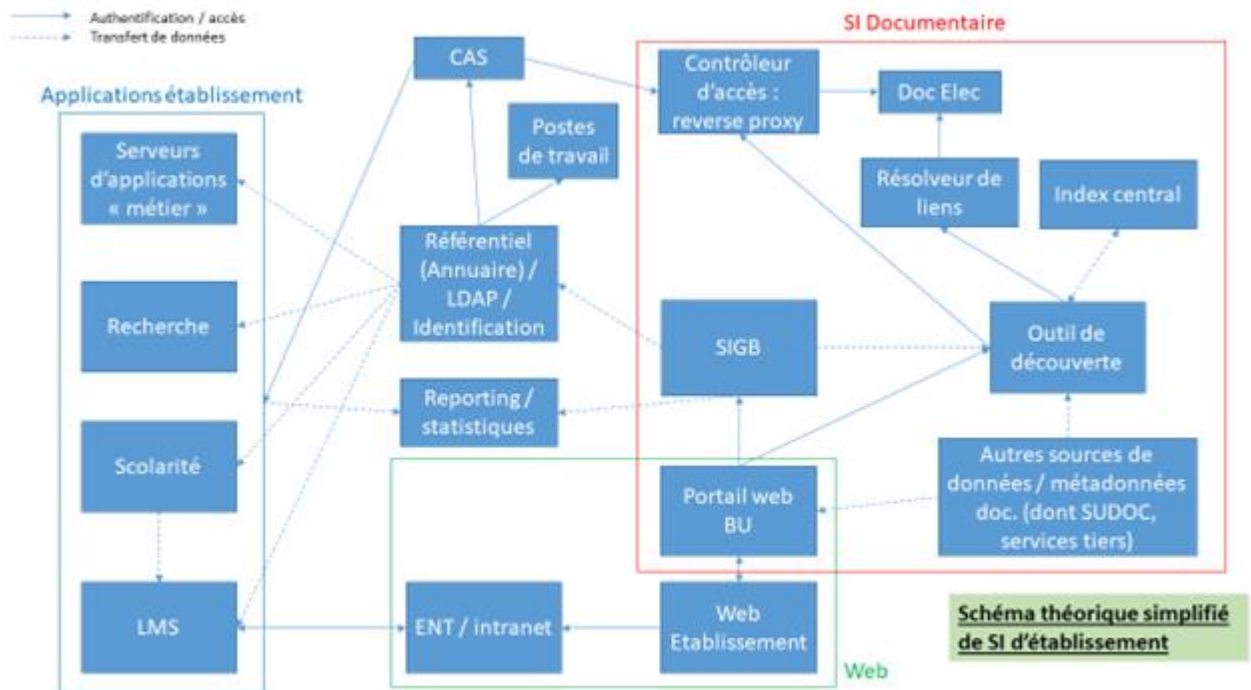
Il est donc essentiel à toute coopération vertueuse avec ces services d'en identifier les différentes composantes et leur champ d'action. Pourtant, tel *Pac-man* dans son labyrinthe, il n'est pas toujours aisé de naviguer dans ce dédale : infrastructures, support, parc, applicatifs, sécurité...

Nous espérons cependant que cette rubrique permettra d'éclairer le squelette de ces directions aux visages multiples mais ô combien nécessaires pour donner corps à nos services. Entrez donc dans le monde SolR et urbain des systèmes informatiques !



— **Système d'Information**

Cartographie théorique du Système d'Information



— **Urbanisation du SI**

Évolution des applications informatiques d'une organisation (un SCD ou une université par exemple), souvent anciennes et fonctionnant historiquement en vase clos (en silos) vers un système global où ces applications communiquent entre elles et échangent les informations de façon automatique pour former alors un système global et intégré.

— **Base de production/ Base de préproduction**

(Synonyme de : "base réelle/base de test")

La base de pré-production est une base alimentée par un échantillon des données (qui peut toutefois aller jusqu'à la totalité des données de la base réelle !) et qui possède l'essentiel des fonctionnalités (voire toutes les fonctionnalités) de cette dernière. La base test permet de vérifier les fonctionnements des divers paramètres de la base réelle avant de mettre les nouvelles fonctionnalités qui en découlent à la disposition des utilisateurs.

La base de recette : similaire à la base de test et parfois identique à elle, la base de recette sert de validation avant mise en production d'un produit logiciel. Au début de la recette, le fournisseur vérifie que l'application peut fonctionner avec les paramètres définitifs et dans le SI de destination ; ensuite le "client" vérifie que les exigences du cahier des charges sont bien toutes remplies et que le logiciel fonctionne comme attendu.

— Implémentation vs installation

L'installation est une mise en route simple d'un logiciel. L'implémentation (ici) est une installation et un paramétrage du logiciel qui prend en compte les besoins spécifiques de l'organisation qui utilise ce logiciel et, souvent, un accompagnement à sa mise en production.

HEBERGEMENT

— Serveur web

Un serveur web est un ordinateur qui stocke, traite et fournit des fichiers de sites internet aux navigateurs web. Il se compose de matériel et de logiciels utilisant le protocole HTTP (Hypertext Transfer Protocol) afin de répondre aux requêtes des utilisateurs web effectuées via le [World Wide Web](#). Parmi les serveurs fréquemment utilisés, on peut citer le serveur Apache, un logiciel gratuit et open source développé par Apache Software Foundation. Il permet de répondre aux demandes envoyées par les navigateurs des visiteurs d'un site web en renvoyant des réponses correctes et sécurisées grâce au protocole HTTPS. Apache fonctionne à la fois sur les serveurs Unix et Windows et est réputé pour sa stabilité et sa sécurité. Il existe d'autres serveurs comme Nginx par exemple.

— Serveur SFTP

Un serveur Secure File Transfer Protocol (SFTP) permet de transférer ou télécharger des fichiers à distance avec un client FTP, sur le principe d'un répertoire distant. FileZilla, Cyberduck sont des logiciels permettant d'accéder à un serveur SFTP.

— Serveur OAI

Serveur OAI : serveur web permettant d'exposer des métadonnées selon le protocole d'échange OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting) afin de permettre leur collecte par d'autres institutions. L'entrepôt ainsi disponible peut être moissonné par des clients OAI afin que les données collectées soient utilisées dans d'autres sites web afin de valoriser des collections numériques ; il peut également collecter les données d'autres institutions.

— SaaS

Software As A Service ou logiciel en tant que service. Application disponible en ligne grâce à un hébergement dans le Cloud. L'accès à l'application est en général fourni contre la souscription d'un abonnement. Exemples : Alma, Microsoft Office 365, etc.

— CaaS

Containers As A Service ou Conteneurs en tant que service. Service de Cloud permettant aux développeurs de logiciels de télécharger, d'organiser, d'exécuter, de gérer, de mettre à l'échelle et d'arrêter des containers en utilisant l'interface web ou l'API d'un fournisseur.

— IaaS

Infrastructure As A Service ou Infrastructure en tant que service. Service de Cloud dans lequel l'infrastructure informatique est hébergée dans un cloud (matériel serveur, stockage, virtualisation, réseaux).

— FaaS

Function As a Service ou fonction en tant que service. Service de Cloud permettant de développer des applications sans avoir à gérer les serveurs et l'infrastructure technique. Le serveur ne fonctionne que si une fonction est lancée.

— PaaS

Platform As A Service ou plateforme en tant que service. Type d'offre de services dans le Cloud mettant à disposition des utilisateurs une plateforme permettant d'installer et gérer les applications de son choix sans avoir à installer et gérer l'infrastructure nécessaire pour + Middlewares. Exemples : OVH, Google App Engine, Microsoft Azure,...

— On premise

Ou "sur site", infrastructure mise en place sur un serveur ou des serveurs appartenant à l'établissement et géré par sa DSI. À opposer au modèle "Cloud Computing" où les applicatifs sont installés et gérés sur des serveurs distants, administrés par l'hébergeur.

PROTOCOLES D'ECHANGES

— Z39.50

Protocole, en mode synchrone, permettant la recherche et le transfert d'informations. Il est défini par la norme ISO 23950.

— ISO 2709

Norme de référence pour la définition du MARC. Elle spécifie "les exigences d'un format d'échange général qui contiendra des notices donnant les descriptions des documents de tous types susceptibles d'être inclus dans une bibliographie, ainsi que d'autres sortes de notices. Elle ne précise ni la taille ni le contenu des enregistrements et n'attribue aucune valeur aux étiquettes, aux indicateurs ou aux identificateurs, ces spécifications correspondant aux fonctions d'un format d'application."

— OAI-PMH, SRU, SRW

Il s'agit de protocoles informatiques d'échange de métadonnées.

OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting) est un protocole créé pour la collecte de métadonnées de l'initiative pour les Archives ouvertes. Il repose sur une communication de client à serveur. Le client envoie des requêtes au serveur en http, le serveur répond par un flux de données en XML. Le protocole SRU (Search/Retrieval via URL) est basé sur l'usage d'URLs. Il utilise techniquement le protocole de communication client/serveur HTTP

(HyperText Transfer Protocol) en mode GET ou en mode POST. Son protocole jumeau SRW (Search/Retrieve Web service est quant à lui basé, non pas sur l'usage d'URLs mais sur l'offre de web services faisant transiter les requêtes http encapsulées dans un format xml. Il utilise le protocole de communication SOAP (Simple Object Access Protocol).

— API et Webservice

Interface logicielle qui permet :

- d'accéder aux fonctions d'une application et de développer de nouvelles applications utilisant ces fonctions
- et/ou l'échange de données entre applications

— Interopérabilité

Capacité d'une solution technique à fonctionner avec d'autres solutions techniques ; cela peut aussi s'appliquer aux données : capacité d'extraire des données d'une solution A afin de les importer dans une solution B en les gardant intactes

ACCES

— Port (logiciel)

Point d'accès donnant accès au système d'exploitation, chaque programme informatique pouvant écouter ou émettre sur un port donné, caractérisé par un numéro. Le port est un point virtuel où les connexions réseau commencent et se terminent.

— DNS

Domain Name System ou système de nom de domaine. Le processus de résolution DNS implique la conversion/traduction d'un nom d'hôte en adresse IP qui permettra de contacter les ressources web.

— Passerelle

Dispositif informatique permettant de relier deux réseaux informatiques utilisant des protocoles différents.

— DMZ

Demilitarized zone ou zone démilitarisée. Dans un réseau local, sous-réseau regroupant les machines contactables depuis Internet mais n'accédant pas au réseau local, isolé à l'aide d'un pare-feu. La DMZ protège le réseau local qui n'est ainsi pas exposé.

— Firewall/Pare-feu

Logiciel ou matériel assurant la sécurité d'un réseau en contrôlant les applications et flux de données. Le pare-feu définit le type de communications autorisées sur le réseau et les accès aux serveurs.

— IP

IP : *Internet Protocol address*. Numéro d'identification attribué à un appareil connecté au réseau Internet. L'attribution de l'IP peut être permanente ou provisoire. Une IP peut être unique si elle est publique car son exposition sera alors mondiale. Une IP peut être attribuée plusieurs fois si elle est privée sur un réseau local car alors non exposée sur Internet.

— LAN

Local Area Network. Réseau local dans lequel les appareils et leurs périphériques peuvent communiquer entre eux en général via le Wifi ou via une connexion filaire directe. Réseau lié à une zone géographique précise et unique mais pouvant être de taille variable.

— VLAN

Virtual Local Area Network. Réseau local indépendant réunissant virtuellement un ensemble de plusieurs appareils sans nécessité que ceux-ci soient proches physiquement. On peut ainsi opérer des regroupements logiques pour leur attribuer la même configuration.

— SIP2 (protocole)

Standard Interchange Protocol, protocole de communication utilisé en bibliothèque pour la communication avec les automates de prêt. L'automate envoie une requête au SIGB qui enverra une réponse pour indiquer si l'opération est permise ou non.

— Serveur proxy

Serveur qui joue le rôle d'intermédiaire entre deux appareils ou réseaux. L'utilisation d'un proxy permet entre autres d'avoir des informations sur les connexions via les logs, de filtrer l'accès afin de bloquer certains contenus, de protéger le réseau local qui n'est pas exposé en direct.

— Reverse proxy

Serveur placé devant les serveurs web pour intercepter les requêtes des utilisateurs, il garantit ainsi l'absence de communication entre le serveur d'origine et les clients afin de le protéger, en particulier parce que l'IP des serveurs n'est jamais exposée. Son usage est courant en bibliothèques universitaires pour garantir le contrôle et sécuriser la gestion des accès distant à la documentation électronique. Voir EZ Proxy.

— VPN

Virtual Private Network ou réseau privé virtuel. Applicatif installé sur un appareil relié à Internet qui lui permet de se connecter à un serveur VPN qui va masquer son IP pour la remplacer par celle du serveur VPN. Grâce au VPN, les données sont chiffrées et les échanges sécurisés.

IDENTIFICATIONS

— SSO (Single Sign-On)

Méthode qui permet de s'authentifier une seule fois pour accéder ensuite à plusieurs applications ou sites web sécurisés sans avoir à re-rentre son identifiant/mot de passe.

Exemple : la méthode SSO permet aux étudiants de s'identifier une fois sur l'ENT de l'université puis de naviguer de leur plateforme Moodle à leur compte-lecteur de BU et de sauter d'un article de revue en ligne à l'autre sans avoir à re-montrer patte blanche.

— CAS (Central Authentication Service)

Système de SSO (voir définition SSO) qui permet de s'authentifier une seule fois sur un site puis d'être automatiquement authentifié sur tous les sites qui dépendent du même serveur CAS. Il sert à faire du SSO local : entre les applications déployées dans le système d'information de l'établissement par exemple.

— Fédération d'identités

Il s'agit d'une association entre fournisseurs de services extérieurs à l'établissement (comme les éditeurs de revues en ligne par exemple) et fournisseurs d'identités (les établissements de l'ESR avec leurs annuaires et référentiels, voir définition LDAP et Référentiel). Le mécanisme mis en place permet à l'utilisateur de se connecter à une ressource ou un service en sélectionnant son établissement d'appartenance. Il sera ensuite redirigé vers la page d'authentification de son établissement pour s'identifier ce qui lui ouvrira la porte de la ressource car le fournisseur d'identité qu'est l'établissement aura transmis au fournisseur de service qu'est l'éditeur un laissez-passer via un token (voir définition Token) qui précise que l'utilisateur a un compte valide et des attributs lui permettant d'accéder au service voulu.

— Shibboleth

Système de SSO (voir définition SSO) pour les applications situées en dehors du système d'information de l'établissement et qui repose sur la fédération d'identités (voir définition Fédération d'identités). Shibboleth se branche sur l'annuaire LDAP de l'établissement pour en extraire des attributs utilisateur (validité du compte, autorisations d'accès, informations sur l'utilisateur...) qui permet l'authentification vers des services tiers (éditeurs de revues en ligne, Opac d'un SIGB commun à plusieurs établissements...).

— Token

Jeton virtuel unique à durée limitée qui permet à un utilisateur d'accéder à un service après avoir confirmé son identité. Il est invalidé dès que l'utilisateur se déconnecte ou quitte le service.

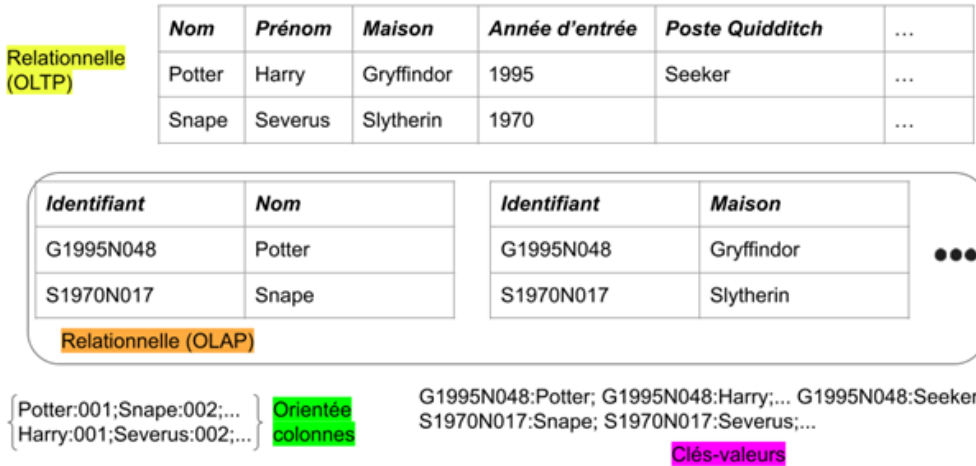
BASES DE DONNEES

Une base de données est un ensemble d'informations organisé afin de les rendre (facilement) accessibles, de les gérer et de les mettre à jour.

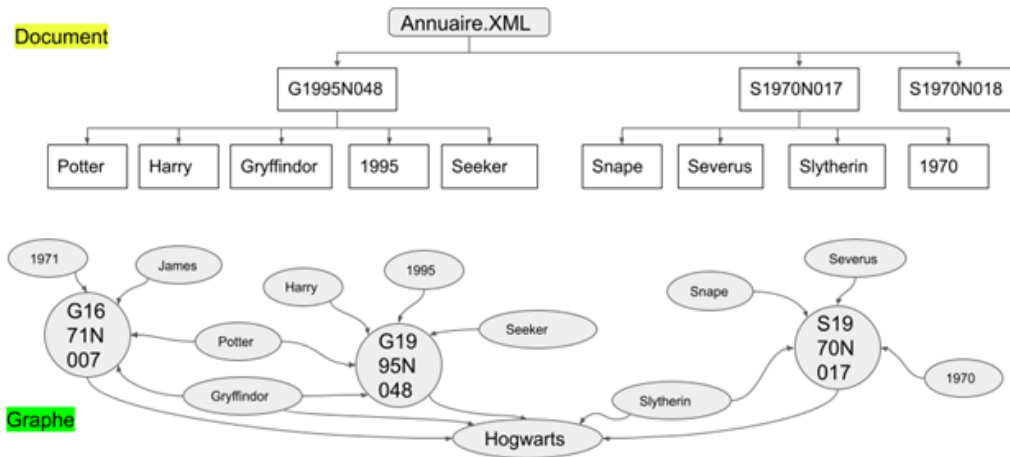
— Typologie des bases de données

On distingue la plupart du temps les bases de données relationnelles et les bases de données.. non relationnelles. En informatique documentaire les bases de données de type document et plus récemment de type graphe forment la tendance majoritaire même si les bases de données relationnelles restent très présentes dans les SIGB par exemple. Du type de base de données dépend l'efficacité des traitements appliqués aux données.

Modèles de bases de données



Modèles de bases de données - 2



— SQL (Structured Query Language)

Langage d'interrogation et de traitement des données dans une base de données relationnelle. Les différents systèmes de gestion de base de données (MySQL, PostgreSQL, Oracle...) impliquent des instructions et variables différentes en SQL. SQL peut être utilisé directement ou au travers de tiers (logiciel plus orienté utilisateur, langage spécifique du type de base de données).

— Dump

Avec différentes variations techniques, un dump de la base de données d'un logiciel est la copie brute de la base à un instant donné, permettant de revenir à cet état en cas de problème dans une opération impliquant des corruptions ou erreurs de données.

PARAMETRAGES

— Crontab

Un crontab est un utilitaire qui permet de planifier l'exécution d'un programme ou d'une commande, soit au niveau du système d'exploitation, soit au niveau d'une application.

— Syspref

Préférences Système spécifiques

TYPES DE MOTEURS DE RECHERCHE

— SolR

Moteur de recherche open source développé par Apache, spécialisé dans l'analyse de données non structurées stockées dans des silos, basé sur la bibliothèque de recherche Lucène.

— ElasticSearch

Moteur de recherche open source développé par l'éditeur Elastic, spécialisé dans l'analyse de données structurées en temps réel, basé sur la bibliothèque de recherche Lucène.

— Zebra

Moteur d'indexation et de récupération de texte structuré très performant et universel présent par exemple dans Koha. Il lit les notices dans une variété de formats d'entrée (par ex. e-mail, XML, MARC) et leur fournit l'accès par une combinaison puissante d'expressions de recherche Booléennes et des requêtes de texte libre, et évalue la pertinence des occurrences.

ANNUAIRES

— LDAP (Annuaire)

Lightweight Directory Access Protocol (LDAP) – Protocole d'interrogation et de modification d'annuaire. Par extension, annuaire compatible avec ce protocole. Organisé généralement en arborescence, avec différents domaines et sous-domaines, divisés en unités d'organisation, subdivisées en personnes ou objets (les machines peuvent faire partie de l'annuaire).

— Active Directory (AD)

Service d'identification et d'authentification des entités informatiques répertoriant dans un annuaire LDAP tous les éléments connectés (utilisateurs mais aussi ordinateurs, imprimantes, serveurs) à un réseau local. De technologie Microsoft, ce service s'appuie sur des contrôleurs de domaine qui définissent les droits d'accès et d'interaction. Par exemple, c'est l'AD qui dit si

un usager peut ouvrir une session sur un PC Windows sur le campus, et avec quels identifiant et mot de passe.

— Référentiel

Ici, base de données relationnelle intégrant toutes les informations sur tous les membres d'une communauté d'établissement, y compris le cas échéant les usagers inscrits à la BU. Ses interconnexions avec les autres applications (Apogee, logiciels scolarité, ...) permettent l'exploitation de ces données personnelles.

PROGRAMMATION

— Script

Ensemble de commandes ou d'instructions exécutées par un ordinateur et rédigées au préalable dans un langage de programmation. Les scripts sont souvent utilisés pour automatiser des tâches répétitives (sauvegardes, traitement de données).

— Langage

Le langage informatique est un ensemble de règles et de conventions qui permettent à un programmeur de communiquer avec un ordinateur et de lui donner des instructions comme créer des programmes, contrôler le fonctionnement d'un système informatique ou effectuer des tâches automatisées. Exemple : Java, Python, JavaScript, PHP, Perl...

— Algorithme

Suite d'instructions finies et non ambiguës permettant d'effectuer une tâche précise ou de résoudre un problème : recherche, tri de données, compression de fichiers, analyse de données... Un algorithme peut être écrit dans différents langages de programmation.

— Github

Plateforme de développement de logiciels qui permet aux développeurs de collaborer sur des projets, de partager du code-source et de gérer des versions de logiciels. Exemple : sur Github ou d'autres plateformes de type Git, on trouve des modèles de requêtes, de scripts ou de versions pour le SIGB libre et open source Koha qui ont été partagés par les développeurs pour être ensuite utilisés librement.

— PHP

C'est un langage de script côté serveur conçu pour être intégré dans des pages HTML (voir HTML) et qui peut être utilisé pour générer du contenu dynamique, traiter des formulaires, accéder à des bases de données. Très utilisé dans les CMS tels WordPress, Drupal ou Joomla.

— JAR

Format de fichier utilisé pour emballer des fichiers Java afin d'exécuter l'application Java (comme un fichier .zip mais réservé à du Java, quoi).

CMS

— CMS

Content Management System, logiciel en ligne permettant de créer facilement des sites web, de les structurer et de les alimenter sans expertise en développement. Exemples : WordPress, Drupal, Joomla

— Framework

Littéralement "cadre de travail", sorte de boîte à outils et modules pour développer des applications ou des sites web. Le Framework va permettre de disposer d'un cadre structuré et de composants standardisés afin de faciliter la tâche de création.

— Javascript (JS)

Langage de programmation à objets utilisé principalement pour les pages, applications et sites web.

— HTML

HyperText Markup Language ou langage de balises pour l'hypertexte utilisé pour créer, alimenter et structurer des pages web. Le langage HTML utilise des balises pour annoter les contenus et permettre leur affichage dans les navigateurs web.

— CSS

Cascading Style Sheets ou « feuilles de style en cascade », langage de mise en forme des pages web en HTML ou XML utilisant des feuilles de style et du code.

STRUCTURATION DE DONNEES

— Balises

Éléments de repérage insérés dans un texte, un document ou une série de caractères, permettant aux machines de repérer la structure, de reconnaître des entités, et/ou de déclencher des actions ou traitements prédéfinis. Dans beaucoup de cas, elles sont composées de 2 éléments différents, une balise ouvrante et une balise fermante. Les plus courantes et les plus communes utilisent le caractère "<" pour introduire le nom de la balise, le caractère ">" pour clore un élément de balisage, et le caractère "/" avant le nom d'une balise pour dire que c'est une balise fermante. Exemple : <auteur>Victor Hugo</auteur> place "Victor Hugo" à l'intérieur d'une balise "auteur".

— JSON

JavaScript Object Notation. Modèle de représentation des données rattachées à des objets, s'appuyant sur une formalisation très légère. Il est tout particulièrement adapté au stockage et aux transferts de données. Prononcez "Jason" comme le monstre de Vendredi 13 si vous avez

peur de ce format taiseux et que vous commencez à retrouver à chaque détour du web. Prononcez "Jason" comme le médecin dans la série "Quoi de neuf docteur?" si vous vous intéressez à sa capacité d'adaptation à tous les contextes et à son ouverture aux autres (= aux informaticiens, qui l'aiment beaucoup).

— XML

Famille de langages à balises initiée par Tim Bray, alliant rigueur et lisibilité pour les humains. Les différents formats XML sont généralement définis par des règles contenues dans des DTD ou des XML Schemas : quelles balises sont autorisées, quel est leur contenu, peuvent-elles être répétées... Pour le stockage et les transferts de métadonnées, cette structuration permet une meilleure fiabilité, c'est pourquoi XML est choisi fréquemment comme format pivot. Mais attention il y a XML (les formats reconnus, Dublin Core, TEI, EAD, METS...) et XML (obscur format maîtrisé par le seul prestataire d'un logiciel avec des "exports XML").

TRANSFORMATION/NETTOYAGE DE DONNEES

— XSLT

Langage de transformation qui permet de transformer un document XML dans un autre format XML correspondant à un schéma différent.

— OpenRefine

Outil permettant d'explorer, corriger, transformer, restructurer et enrichir de manière interactive des données.

CONNECTEURS

— Généralités

On désigne souvent par "connecteur" des micro services qui assurent les transmissions d'information entre les applications du SI, et/ou qui déclenchent des interactions de ces applications de façon planifiée ou par réaction à une intervention humaine. Dans le cas des bibliothèques, les connecteurs autour du SIGB sont monnaie courante.

— Talend

Nom d'une série logicielle spécialisée dans la création de connecteurs entre les applications. Les "jobs Talend" sont des séries d'instructions qui collectent, trient, transforment les données issues d'une application ou d'une base pour les injecter dans une autre. Par exemple, pour récupérer les données d'un étudiant à l'université pour en faire un compte lecteur dans le SIGB, en ajoutant le numéro de sa carte issu d'une autre base.

— Plug-in

Mot à mot "(qui se) branche dedans". Module complémentaire et non nécessaire d'une application lui apportant des fonctionnalités qui n'étaient pas présentes au départ. Le plug-in

peut être proposé par le développeur de l'application principale, ou par un tiers. Exemples : ajout d'une recherche dans WorldCat à partir de l'outil de découverte, ou d'une fonction d'impression 3D du texte des e-mails (on le voit un plug-in n'est pas forcément utile, dans ce cas éviter de l'installer).

RFID

— RFID

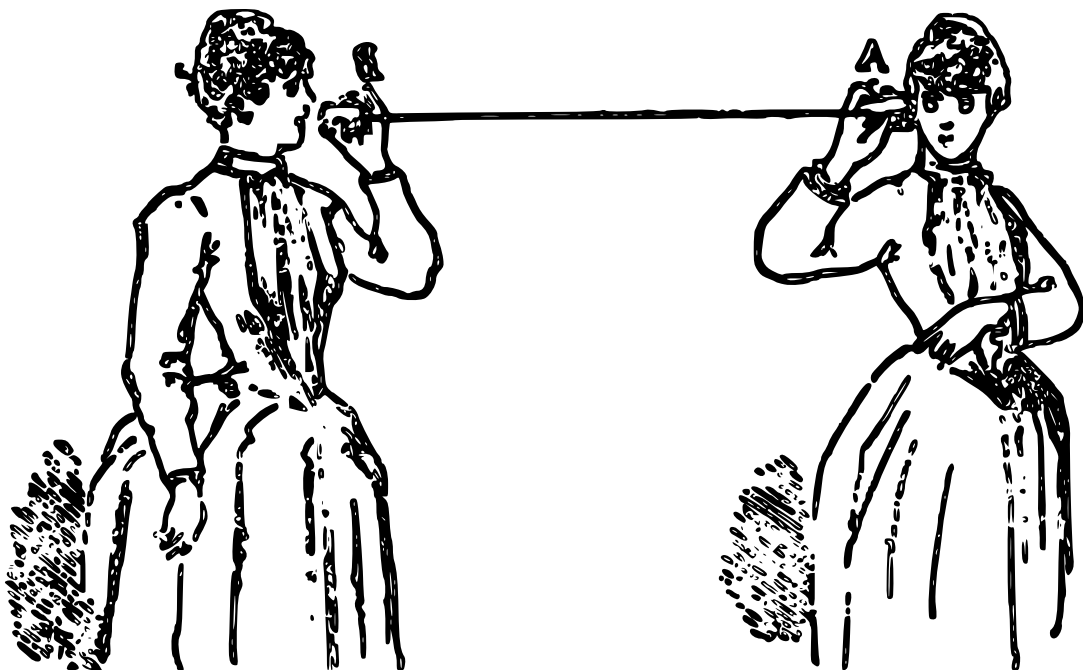
(“Radio Frequency Identification” ou en français “Identification par Fréquence Radio”) Technologie sans contact qui permet d'identifier et de suivre des objets à distance en utilisant des étiquettes électroniques équipées d'une puce et d'une antenne qui permettent de transmettre des informations à distance à un lecteur RFID. En bibliothèque, l'installation RFID se traduit ainsi : étiquettes pour identification sur les documents, platines RFID pour l'encodage et la lecture des étiquettes, portiques antivols et automates/robots de prêt-retour.

— Mifare

Marque déposée pour les cartes à puce sans contact basés sur la technologie RFID. Les produits Mifare comprennent les badges (ou cartes) et les lecteurs de badges associés ; ils sont utilisés notamment pour l'identification (par exemple sur un automate de prêt) et le contrôle d'accès. Produit non réservé aux « Lady ».

Parler avec ses partenaires

Communiquer avec ses alliés n'est pas toujours chose aisée. Rejoindre un SID c'est infiltrer de nouveaux réseaux et intégrer leurs codes et modes de communication. Les allemands avaient Enigma nous avons Paprika, ils ont créé BMW, nous roulons avec WiniBW. Cette section est donc consacrée aux sigles et outils de nos réseaux et partenaires. Sans prétendre à l'exhaustivité, elle se veut un premier pied à l'étrier.



ABES

Sudoc / Sudoc PS

Système Universitaire de Documentation – constitue le dispositif national de catalogage partagé des ressources documentaires à la disposition de l'ensemble des bibliothèques universitaires et de recherche. Le Sudoc PS en constitue un sous-ensemble consacré aux publications en série. Il repose sur un socle technique, le système Central Bibliographic System (CBS), développé par la société hollandaise Pica et désormais maintenu par OCLC. Le Sudoc se décline en plusieurs outils et services présentés ci-dessous.

[Pour en savoir plus sur le Sudoc](#)

— **Catalogue Sudoc**

Interface publique donnant accès à la base de données bibliographique nationale pour l'Enseignement Supérieur et la Recherche, décrivant tous types de ressources documentaires : documents imprimés, thèses de doctorat soutenues depuis 1981, cartes géographiques, partitions, ressources audiovisuelles...

— **Répertoire des centres de ressources (RCR)**

Base de données publique permettant la consultation des données d'autorité relatives aux bibliothèques et centres de documentation membres des réseaux Sudoc et Sudoc PS

— **WiniBW**

Interface commune de catalogage dans le Sudoc. Accessible aux catalogueurs des établissements membres du réseau, cette application développée par OCLC en mode client/serveur comporte des modules communs et des modules spécifiques (ex. Prêt entre Bibliothèques (PEB)). Elle repose sur le format Unimarc.

— **Cidemis**

Application professionnelle dans le cadre d'un partenariat avec le Centre ISSN France et le CIEPS, Cidemis est utilisée pour déposer les demandes d'attribution de numéros ISSN et/ou de correction de notices ISSN auprès des 90 centres du réseau ISSN présents dans le monde, en permettant la transmission dématérialisée des justificatifs accompagnant toute demande de numérotation ou de correction d'ISSN.

— **Colodus**

Application professionnelle utilisée pour l'administration des données d'exemplaire (création, modification, suppression) liées aux notices bibliographiques du Sudoc, sans passer par l'interface WiniBW.

— **Periscope**

Application professionnelle utilisée principalement dans le cadre de la mise en œuvre de Plans de Conservation partagée des Périodiques (PCPP). En accès libre sur le web, elle permet notamment l'identification des établissements possédant un exemplaire du périodique recherché ; la visualisation et comparaison des états de collection de périodiques disponibles

dans les bibliothèques membres des réseaux Sudoc / Sudoc PS ; l'identification des PCPP dans lequel s'inscrit un périodique et du statut des différentes collections.

— **SELF Sudoc**

Application professionnelle pour la conception, en autonomie, de produits bibliographiques comme des listes de notices d'unica, des listes d'identifiants rattachés à un établissement, ...

— **Item (Interface de Traitement des exemplaires en masse)**

Application professionnelle utilisée pour les traitements en masse des exemplaires signalés dans le Sudoc, comme obtenir des taux de recouvrement, réaliser des exemplarisations en masse ou modifier des informations par lots d'exemplaires.

— **QualiMarq**

Application professionnelle utilisée pour analyser la qualité des notices bibliographiques produites avec WinIBW. Elle permet notamment d'analyser des notices par lot ou à l'unité et d'obtenir un signalement des erreurs ou des enrichissements nécessaires).
<https://qualimarc.sudoc.fr/>

— **IdRef (Identifiants et Référentiels pour l'Enseignement supérieur et la Recherche)**

Développée par l'Abes et interconnectée aussi bien avec les applications gérées par l'Abes (Sudoc, Calames, STEP/STAR) qu'avec des gisements documentaires autres (Persée, CAIRN, catalogues du réseau Frantiq, des bibliothèques francophones en Suisse, de l'Université de Liège en Belgique...) la plateforme IdRef constitue tout à la fois :

- une base vivante et administrée de notices d'autorité (personnes physiques, collectivités, congrès, noms communs RAMEAU ou FMesh, forme ou genre RAMEAU, nom géographique, titre, auteur-titre ...) dotée d'un identifiant IdRef unique et pérenne ;
- un registre de personnes et de structures avec un fort tropisme pour les auteurs scientifiques exerçant en France, du fait du signalement à caractère obligatoire des thèses de doctorat (auteurs, directeurs de recherche, membres de jury, établissements, écoles doctorales, unités..);
- une interface publique de recherche et de consultation des données d'autorité ;
- un module autonome de catalogage de notices d'autorité

Sont liées spécifiquement à IdRef deux applications :

- data.idref.fr : Triple store des données d'autorités et de leurs citations bibliographiques (Sudoc, HAL...) exposées en RDF.
- Paprika : Application professionnelle de curation de données, dédiée au contrôle qualité des liens entre notices bibliographiques et notices d'autorité de type personne, s'appuyant sur un service d'intelligence artificielle co-développé par l'Abes et l'équipe de recherche GraphIK (LIRMM).

— Calames (Catalogue en ligne des archives et manuscrits de l'enseignement supérieur)

Interface publique donnant accès à la base de données bibliographique nationale pour l'Enseignement Supérieur et la Recherche décrivant les archives et manuscrits conservés dans les établissements de l'Enseignement supérieur et de la Recherche. Depuis 2017, Calames Pro est une application exécutable sur le bureau Windows et à installer en suivant une procédure spécifique

Produits liés au signalement des thèses

[Pour en savoir plus sur les outils de l'Abes pour le signalement des thèses de doctorat](#)

— STEP (Signalement des thèses en préparation)

Application professionnelle nationale permettant la saisie des métadonnées descriptives des thèses en préparation et des événements les concernant (ex : date de soutenance, changement d'établissement de soutenance, abandon). En cohérence avec le système d'information de l'Abes pour le signalement des thèses, l'application STEP est interconnectée avec STAR.

— STAR (Signalement des thèses Archivage)

Application professionnelle nationale mise à la disposition des établissements habilités à délivrer le diplôme de doctorat et leur permettant d'assurer le signalement et l'archivage des thèses de doctorat sous forme électronique soutenues en France. STAR constitue l'intermédiaire entre les établissements et la plateforme d'archivage du CINES (PAC). Les métadonnées produites alimentent différentes applications nationales de signalement (catalogue Sudoc, archive ouverte HAL, site theses.fr). Un entrepôt OAI-PMH est disponible pour moissonner l'ensemble des métadonnées des thèses traitées dans STAR

— theses.fr

Moteur de recherche alimenté par les applications nationales STEP, STAR et Sudoc. Il donne accès à l'ensemble des métadonnées des thèses de Doctorat Françaises soutenues depuis 1985, avec lien vers le texte intégral, sauf en cas de confidentialité exigée par l'établissement."

Produits liés aux ressources électroniques

— SciencesPlus

[En savoir plus](#)

Triple store de métadonnées de publications scientifiques numériques (revues, ebooks, articles, chapitres) pour lesquelles l'Abes a obtenu le droit de réutilisation de la part des éditeurs. Entrepôt proposé à ce jour en version beta.

— BACON (Base de connaissance nationale)

[En savoir plus](#)

Entrepôt de métadonnées enrichies relatives à 850 bouquets de ressources électroniques, mis à disposition des établissements sous la forme de fichiers KBART. Idéal le matin, avec une bonne tasse café et des œufs brouillés.

[Pour plus d'information sur les outils de l'ABES](#)

CNRS

— HAL/TEL/DUMAS

Ensemble de plateformes mises à disposition par le CCSD (Centre pour la Communication Scientifique Directe) pour signaler et héberger de manière pérenne les publications scientifiques, les thèses et les mémoires.

— Persée

Portail offrant "un accès libre et gratuit à des collections complètes de publications scientifiques (revues, livres, actes de colloques, publications en série, sources primaires, etc.) associé à une gamme d'outils de recherche et d'exploitation"

COUPERIN

— ezPaarse

Logiciel libre et gratuit permettant de générer des statistiques d'usages des plateformes des éditeurs de revues électroniques. Il se base pour cela sur les logs du reverse-proxy. Cela lui permet de connaître les profils des usagers (en fonction du paramétrage local du proxy et des informations d'annuaire récupérées) mais aussi leurs pratiques de recherche (grâce à l'analyse collaborative des URLs des plateformes, centralisées sur <http://analyses.ezpaarse.org/> et transformées en parseurs, voir infra)

— Parseur

Un parseur pour ezPAARSE est un script (ou petit programme) permettant d'analyser les URLs d'une plateforme donnée afin d'en déduire les actions réalisées sur les ressources disponibles (articles, ouvrages, etc.) en fonction des éléments que ces URLs contiennent (en particulier : un DOI, une année de publication, un issn, un identifiant propriétaire, un format, etc.). Un parseur distinct doit être développé pour chaque plateforme car les URLs ne se construisent pas de la même manière.

— ezMesure

Plateforme nationale basée sur les outils Elastic/Kibana et opérée par l'Inist-CNRS en partenariat avec Couperin où les établissements membres du consortium Couperin peuvent :

- reverser les données de leur ezPAARSE déployé dans leur établissement pour les exploiter sous la forme de tableaux de bords dynamiques
- déclarer leurs identifiants SUSHI pour faire moissonner les rapports d'usages COUNTER 5 des éditeurs et fournisseurs conformes (voir ezCOUNTER infra)

— **ezCOUNTER**

Dispositif dans ezMESURE pour moissonner tous les rapports d'usages COUNTER 5 via SUSHI directement auprès des éditeurs et fournisseurs, de façon automatisée

OCLC

— **EZ Proxy**

Logiciel d'accès et d'authentification utilisé pour garantir l'accès aux ressources électroniques souscrites par les bibliothèques peu importe le lieu et le moment de la connexion. Après l'authentification fournie par l'établissement (en général le CAS), EzProxy connecte l'utilisateur à la ressource en présentant une IP autorisée. EzProxy génère des statistiques intéressantes sur les accès aux Ressources éducatives libres (REL).

— **WorldCat**

Base de données bibliographiques de l'OCLC (Online Computer Library Center, organisation à but non lucratif basée aux États-Unis) qui contient les données de milliers de catalogues de bibliothèques (dont le SUDOC) partout dans le monde. Accessible en ligne à l'adresse <https://www.worldcat.org/fr>.

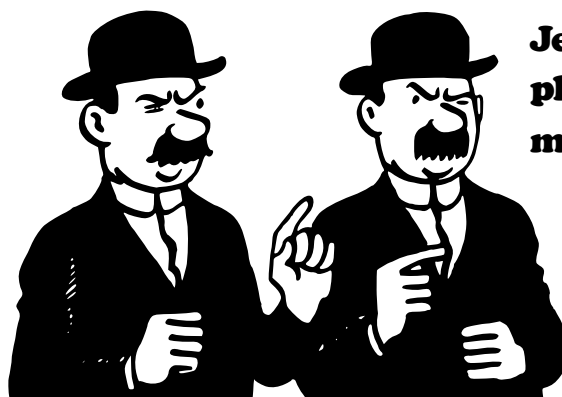
Parler avec Autorités

En informatique, l'interopérabilité se caractérise par la capacité de plusieurs systèmes à communiquer entre eux grâce à un protocole commun.

Il existe plusieurs niveaux d'interopérabilité répondant à des normes spécifiques. Au-delà des protocoles techniques d'échange de données, les systèmes d'information documentaire se sont de longue date intéressés aux enjeux sémantiques de l'interopérabilité à travers la mise au point de thésaurus et de vocabulaires contrôlés. Ces derniers reposent sur des logiques d'Autorités et d'identification pérennes qui trouvent aujourd'hui une place majeure dans les développements du web de données moteur de la transition bibliographique.

À la sémantique s'ajoute également un enjeu juridique, permettant, à travers l'adoption progressive des licences ouvertes, de consolider la dimension légale de ce cadre d'interopérabilité que représente aujourd'hui le web de données.

**PIDs, ORCID,
DOI, IdRef et ARK !**



**Je dirais même
plus, Bibliostratus
mon cher Dupont !**

MAPPING

— Mapping

Le mapping des données consiste à mettre en correspondance les champs de plusieurs bases de données. C'est la première étape pour faciliter la migration puis l'intégration des données d'un SIGB à un autre.

— Correspondance/Conversion

Les tables de correspondance/de conversion sont les outils principaux du mapping. Elles consistent en une liste des valeurs contenues dans certains champs-sources des notices du SIGB d'origine, à laquelle on fait correspondre une liste des champs-cibles du SIGB d'accueil, dans lesquels les données d'origine devront être reprises ou converties. Dans le cadre d'une migration d'un SIGB à l'autre, ce sont particulièrement les notices d'exemplaires qui feront l'objet de ces tables de conversion.

PID, REGISTRES ET AUTORITES

— Autorité (données d')

Données permettant d'identifier, de manière univoque et pérenne, les entités qui contrôlent la description des ressources et l'accès à celles-ci. Les données d'autorité comprennent : les points d'accès autorisés et rejetés, les identifiants pérennes relatifs à l'entité décrite et des données d'information et de source.

[Pour en savoir plus](#)

— FNE (Fichier national d'entités)

Projet lancé en 2017, co-piloté par l'ABES et la BnF, ayant pour objet la mise en place d'une plate-forme de production mutualisée d'entités, données d'autorité structurées selon le modèle IFLA LRM, entre la BnF et les réseaux d'utilisateurs des produits de l'ABES. En juin 2023, le Comité Stratégique Bibliographique (CSB) a réaffirmé l'importance stratégique de l'ambition du FNE mais a pris la décision d'en reporter la réalisation,

en tenant compte des projets importants déjà en cours de réalisation à la BnF et de ceux à venir à l'Abes d'une part, et de l'impossibilité d'utiliser Wikibase comme socle technique d'autre part. [Pour en savoir plus](#)

— Référentiel

Un référentiel est une liste d'éléments formant un système de référence participant de la bonne interopérabilité des systèmes d'information. Appliqué aux données, il s'agit d'une base ou d'un ensemble de bases permettant la collecte, la gestion et le stockage, à des fins de réutilisation ou d'analyse, d'un ensemble de données servant de référence à un système d'information. Les données y sont décrites et définies de telle sorte à être interrogeables et utilisables par des applications chaque fois que nécessaire. Un référentiel peut également désigner un ensemble d'informations techniques relatives à l'administration d'une base de données. Il vise alors à référencer des normes et standards relatifs aux formats, propriétés, conditions de mise à jour etc. des données et applications d'un système d'information.

— PID (Persistant Identifier)

Un identifiant est un numéro ou une étiquette alphanumérique, opaque ou explicite, lisible par des machines et par des humains, permettant de désigner et de retrouver de manière univoque et pérenne un objet, un document, une personne, un lieu, un organisme, ou toute entité, dans le monde physique ou numérique. Pour faciliter l'accès à une entité numérique sur internet, l'identifiant est souvent associé à une URI (Uniform Resource Identifier), du type `http(s)://url/identifiant`. Un identifiant pérenne (Persistent identifier ou PID) est disponible et gérable à long terme ; il ne changera pas si l'objet est renommé ou déplacé (changement de site, d'entrepôts de données...).

— Thésaurus

Liste organisée de termes contrôlés et normalisés représentant les concepts d'un domaine de la connaissance. Les thésauri constituent une catégorie spécifique de langage documentaire. La norme ISO 25964 définit les thésauri et les conditions d'interopérabilité avec d'autres vocabulaires.

DOCUMENTS ET RESSOURCES

— ARK (Archival Resource Key)

Format d'identifiant créé en 2001 par la [California Digital Library](#) (CDL) qui a vocation à identifier des ressources physiques, numériques ou même immatérielles (concepts). Son but est de fournir des identifiants adaptés aux besoins des producteurs et diffuseurs de données sur le web, mais également capables de durer sur le long terme.

[Pour en savoir plus](#)

— DOI (Digital object identifier)

Identifiant international régi par la norme ISO 26324, à la fois mécanisme de nommage des ressources et protocole de résolution des identifiants en adresses plus concrètes. Il identifie de manière univoque toute ressource stable, et en particulier de ressources numériques comme des articles scientifiques ou des rapports. L'attribution d'un DOI est le fait d'agences d'enregistrement : CrossRef pour les DOI des publications et DataCite pour les DOI des données scientifiques et des produits de la recherche (l'INIST est l'agence française des DOI DataCite).— EAN (European Article Numbering)

Système global destiné à l'identification univoque d'objet, qui attribue un code-barre (le code EAN) à tout produit ayant vocation à entrer dans une chaîne logistique commerciale.

— ISAN (International Standard Audiovisual Number)

Identifiant international normalisé régi par les normes ISO 15706:2002 et ISO 15706-2. Il permet d'identifier de manière unique les œuvres audiovisuelles. Il est géré en France par l'Agence française ISAN.

[Pour en savoir plus](#)

— ISBN (International Standard Book Number)

Identifiant international normalisé régi par la norme [ISO](#) 2108. Il permet d'identifier de manière univoque une monographie quel qu'en soit le support de publication : imprimé ou multimédia. En France, l'ISBN est attribué par l'Agence Francophone pour la Numérotation Internationale du Livre (AFNIL).

[Pour en savoir plus](#)

— ISCI (International Standard Identifier for Collections)

Identifiant international normalisé, régi par la norme ISO 27730. Il permet d'identifier de manière unique les collections et fonds documentaires. Il est principalement destiné aux établissements de patrimoine culturel, tels que les bibliothèques, les musées et les services d'archives, aussi bien que d'autres organismes, tels que les maisons d'édition travaillant dans la chaîne de fourniture de biens culturels.

— ISMN (International Standard Music Number)

Identifiant international normalisé régi par la norme ISO 10957. Il permet d'identifier de manière unique les publications musicales imprimées. La Société des Éditeurs et Auteurs de Musique (SEAM) est agence française pour l'ISMN

[Pour en savoir plus](#)

— ISRC (International Standard Recording Code)

Identifiant international normalisé régi par la norme ISO 3901. Il permet d'identifier de manière unique les enregistrements sonores et vidéo. La Société civile des producteurs phonographiques (SCPP) est l'agence française pour l'ISRC.

[Pour en savoir plus](#)

— ISSN (International Standard Serial Number)

Identifiant international normalisé régi par la norme ISO 3297 : 2017. Il permet d'identifier de manière univoque et non ambiguë toute publication en série. L'e-ISSN s'applique plus spécifiquement aux publications numériques. Il est attribué pour les publications françaises par le Centre ISSN-France.

[Pour en savoir plus](#)

— ISWC (International Standard Musical Work Code)

Identifiant international normalisé régi par la norme ISO 15707. Il permet d'identifier de manière unique les œuvres musicales. La Société des auteurs, compositeurs et éditeurs de musique (SACEM) est agence française pour l'ISWC.

[Pour en savoir plus](#)

— PPN (Pica production number)

Identifiant interne des notices gérées dans les bases utilisant la solution PICA d'OCLC. En France, il est particulièrement utilisé dans les notices Sudoc. On le retrouve notamment en affichage public dans l'URL des notices Sudoc.

PERSONNES

— IdRef

Identifiant attribué à toute notice d'autorité gérée via la plateforme IdRef développée par l'ABES. Stable, pérenne et ouvert, l'identifiant IdREF est interopérable avec les principaux identifiants internationaux (ISNI, VIAF, ORCID, ...).

-> Pour en savoir plus : <https://abes.fr/reseaux-idref-orcid/outils-et-services-autorites/plateforme-idref/>

— ISNI (International Standard Name Identifier)

Identifiant international normalisé, défini par la norme ISO 27729:2012, servant à identifier de manière univoque les identités publiques des personnes ou des organismes impliqués dans la création, la production ou la gestion et la distribution de contenus intellectuels et artistiques.

-> Pour en savoir plus : <https://www.bnf.fr/fr/isni-international-standard-name-identifier>

— ORCID (Open Researcher and Contributor ID)

Identifiant international, servant à identifier de manière équivoque les chercheurs et auteurs de contributions académiques et scientifiques. Mise en place en 2012, il est compatible avec la norme ISO 27729 et est géré par l'organisme à but non lucratif Orcid.

-> Pour en savoir plus : <https://orcid.org/>

— VIAF (Virtual international Authority File)

Fichier d'autorité international virtuel fédérant en un fichier unique des notices d'autorité (personnes, collectivités, noms géographiques, œuvres, expressions) issues des principaux fichiers nationaux. Il est mis en œuvre par OCLC. L'identifiant VIAF, généré par le système, constitue de fait un identifiant international stable et pérenne d'un cluster de notices d'autorités décrivant un même objet. VIAF est utilisé dans IdRef comme source de dérivation pour faire gagner du temps au catalogueur.

-> pour en savoir plus : <https://viaf.org/>

STRUCTURES

— ILN (Internal Library Number)

Identifiant propre à l'Abes et délivré à chaque établissement (ou regroupement administratif d'établissement) ayant signé une convention et doté d'un même SGB, au moment de son déploiement dans le réseau Sudoc. Il sert principalement dans le cadre des échanges de données entre le Sudoc et les réseaux locaux. Il constitue de fait un identifiant national stable et pérenne.-> Pour en savoir plus : https://documentation.abes.fr/sudoc/manuels/administration/gestion_bibliotheques/index.html#IlnRcr

— ISIL (International Standard Identifier for Libraries and Related Organizations)

Identifiant international normalisé, reposant sur la norme ISO 15511:2019, identifiant les bibliothèques et les organismes apparentés. L'ABES est agence d'enregistrement ISIL pour la France : la structure de l'ISIL en France reprend le numéro RCR précédé de « FR- »

-> Pour en savoir plus :

https://documentation.abes.fr/sudoc/manuels/administration/gestion_bibliotheques/index.html#IlnRcr

— RCR (Répertoire des centres de ressources)

Ce répertoire, administré par l'ABES, recense les composantes documentaires des établissements de France et de l'étranger déployés dans l'un ou l'autre des réseaux coordonnés par l'ABES. L'identifiant RCR, délivré par l'ABES, constitue de fait un identifiant national stable et pérenne.

-> Pour en savoir plus :

https://documentation.abes.fr/sudoc/manuels/administration/gestion_bibliotheques/index.html#lInRcr

— RNSR (Répertoire national des structures de recherche)

Ce répertoire, administré par le ministère chargé de la recherche, référence les structures de recherche publiques et privées au niveau national. L'identifiant RNSR, généré par le système, constitue de fait un identifiant national stable et pérenne.

-> Pour en savoir plus : <https://appliweb.dgri.education.fr/mnr/>

— ROAR (Registry of open access repositories)

Ce registre, administré par l'Université de Southampton, référence les plateformes institutionnelles d'archives ouvertes. L'identifiant ROAR, généré par le système, constitue de fait un identifiant international stable et pérenne. Il grogne mais ne mord pas !

— IdRef : dRef : Identifiant attribué à toute notice d'autorité gérée via la plateforme IdRef développée par l'ABES. Stable, pérenne et ouvert, l'identifiant IdRef est interopérable avec les principaux identifiants internationaux (ROR) et nationaux (RNSR...)-> Pour en savoir plus : <http://roar.eprints.org/>

CONCEPTS

— RAMEAU (Répertoire d'autorité-matière encyclopédique et alphabétique unifié)

Langage d'indexation matière utilisé, en France, par la Bibliothèque nationale de France, les bibliothèques universitaires, de nombreuses bibliothèques de lecture publique ou de recherche ainsi que plusieurs organismes privés. Il est élaboré depuis 1980, de façon autonome, en relation avec le « Répertoire de vedettes-matière » de l'Université Laval à Québec, et avec la liste de vedettes-matière de la Bibliothèque du Congrès (Library of Congress Subject Headings). Le langage d'indexation précoordonné RAMEAU se compose d'un vocabulaire de termes reliés entre eux et d'une syntaxe indiquant les règles de construction des vedettes-matière à l'indexation. L'ensemble des notices d'autorité – dont le noyau est formé de noms communs et de noms géographiques – constitue les autorités RAMEAU. Elles sont complétées par un Guide d'indexation RAMEAU qui en assure le bon usage.

-> Pour en savoir plus : <https://rameau.bnf.fr/>

— IdRef

Identifiant attribué à toute notice d'autorité gérée via la plateforme IdRef développée par l'ABES. Stable, pérenne et ouvert, l'identifiant IdRef est interopérable avec les principaux identifiants nationaux (ARK BnF).

WEB SEMANTIQUE, WEB DE DONNEES

— Web sémantique, web de données

Évolution du web qui permet une plus grande interopérabilité et une meilleure compréhension par les machines des ressources présentes sur le web et de leurs relations. Il est basé sur une structuration et une mise en relation de données ouvertes (open data) et permet entre autres une meilleure visibilité des catalogues de bibliothèque sur le web. Voir Emmanuelle Bermès, Le web sémantique en bibliothèque

— RDF

Langage du web sémantique et un modèle de graphe. Il structure des données sous forme de triplets (Voir Triplets) afin de décrire des ressources web ainsi que leurs métadonnées. Le but est que ces données soient lisibles et exploitables par les machines.

— Triplets

Structurés sous la forme sujet-prédicat-objet, les triplets permettent de définir les ressources et leurs propriétés suivant le modèle RDF. Chacun des 3 éléments est une URI (Voir URI) extraite d'un référentiel, comme par exemple celui de Wikidata.

Sujet	Prédicat	Objet
Mona Lisa	a pour créateur	Léonard de Vinci
https://www.wikidata.org/entity/Q12418	https://www.wikidata.org/property/direct/P170	https://www.wikidata.org/entity/Q762

Triple store

Base de données qui stocke et permet de récupérer des données RDF décrites sous forme de triplets. Exemple : data.idref.fr, data.istex.fr, data.persee.fr, data.hal.science

SPARQL

Le langage SPARQL permet de réaliser des requêtes sur des réservoirs de données décrites en RDF, comme les triplestores. Les réservoirs mettent souvent à disposition des endpoints afin que les utilisateurs puissent les interroger, comme ici : <http://data.bnf.fr/sparql/>

— URI

Identifiant qui permet de localiser de manière unique et pérenne un document ou une ressource sur Internet. Attention, l'URL pour sa part est un type d'URI. Voir aussi : DOI, ARK.

— Ontologie

Une ontologie est un modèle de données contenant les concepts d'un domaine et leurs relations, dans un langage adapté aux machines. L'ontologie RDA-FR permet de définir les règles qui vont structurer l'information contenue dans les notices et qui seront ensuite comprises et utilisées par les programmes informatiques. D'autres ontologies existent : SKOS (thésaurus), FOAF (personnes)..

Pour toutes remarques, questions ou suggestions

commission-ssi@groupes.renater.fr

