

Les datas sans aléas : connaître ses métadonnées pour FRBRiser son catalogue



Cet article est le compte-rendu de la session parallèle n°1 « Les datas sans aléas : connaître ses métadonnées pour FRBRiser son catalogue » de la 2e journée professionnelle du groupe Systèmes & Données, [Métadonnées en bibliothèques : attention, travaux !](#), qui s'est tenue le mardi 14 novembre 2017 à la BnF.

Le support, les notes explicatives et les réponses aux questions posées lors de la session, ont été réalisés par les membres du groupe Systèmes & Données : Sylvain FRANCESCHI, Silvère HANGUEHARD, Annick LE FOLLIC, Sylvie LEMAIRE, Karine MENEGHETTI, François PICHENOT (coordinateur du groupe), Claudine ROVIS.

Si malgré toute l'attention apportée à l'écriture de ce billet, certaines informations étaient erronées, vous pouvez nous le signaler [en nous contactant via ce formulaire](#).

En amont d'un chantier de FRBRisation, il est nécessaire de bien connaître ses données pour les faire évoluer efficacement notamment maîtriser et harmoniser les conditions de production, identifier des lots de notices afin de préparer des corrections efficaces et automatisées. Cette session s'attachait à présenter les enjeux d'une bonne connaissance des métadonnées et a permis d'échanger sur des exemples de méthodologie.

En complément de cette présentation et lorsque cela s'avérait nécessaire pour la compréhension, les membres du groupe Systèmes & Données ont tenté de synthétiser les compléments

d'information à certaines diapositives ou les éléments de réponse aux questions posées, dans les notes ci-dessous :

Diapo 4 :

Nos catalogues sont composés de notices décrivant des ressources : livres, CDs, DVDs... Ce sont des méta-documents dont les contenus peuvent être exploités car les informations y sont structurées, bien rangées dans des zones et des champs MARC : on parle de données ou métadonnées par extension (sachant qu'une métadonnée est une donnée servant à décrire une autre donnée).

Tout l'objet de notre atelier est de montrer l'importance de mettre de l'ordre dans nos catalogues (grilles à revoir, doublons...), à l'intérieur des notices (compléter, enrichir) et d'agir sur des lots de notices grâce au repérage de données-clés.

Connaître ses métadonnées c'est comprendre le format dans lequel elles décrivent nos fonds. C'est donc être capable de vérifier si les informations cruciales à propos de nos documents (données codées par exemples) sont bien renseignées aux emplacements réservés par ce format de description.

Exemple particulièrement simple : déterminer si l'ISBN des monographies est bien renseigné en 010\$a du format UNIMARC est crucial dans la perspective d'un alignement. En tant que numéro d'identification du document décrit, la valeur contenue en 010\$a permettra d'identifier une notice équivalente dans une base de référence.

Diapo 5 :

Agir sur son catalogue, c'est organiser des chantiers de corrections sur des lots de notices isolées à partir de critères précis, après étude poussée des pratiques de catalogage : pour les textes imprimés par exemple, puis pour les BDs... On vise en premier la récupération en masse de notices proposées par BNF ou SUDOC

Alignement : il s'agit de déterminer pour une notice locale un équivalent dans une autre base. La notice identifiée dans cette base servira dans le futur à écraser ou enrichir la notice locale.

Au terme d'un processus d'alignement on obtient un tableau à deux colonnes : colonne 1 = identifiant de la notice dans la base de la bibliothèque à aligner ; colonne 2 = identifiant pérenne dans la base qui va fournir une notice plus riche (l'ARK BnF par exemple, ou le PPN du Sudoc). Pour aligner dans cet exemple il faut interroger le catalogue de la BnF à partir d'info contenues dans les notices de la base à aligner. On commence par des clefs de dédoublonnage simples (type ISBN ou EAN) avant de traiter les cas plus complexes (pas d'ISBN ou ISBN mal saisi) avec des requêtes moins fiables, plus complexes (Titre + Auteur par exemple. D'où l'intérêt de savoir où se trouvent ces informations dans la base locale.

Diapo 6 :

De plus, ce n'est pas si facile « d'enquêter » sur les pratiques de catalogage de chaque équipe sur plusieurs décennies : celles du livre ancien, livre jeunesse, vidéo... ni sur , ni sur les modes d'entrée et la provenance des notices où plusieurs 801 par exemple ont disparu au profit de celui de la dernière agence de catalogage !

Diapo 7 :

Liées aux changements de priorités dans les services. De moins en moins de temps à consacrer à l'application rigoureuse du format.

Problème lié à la récupération de notices à partir de réservoirs gratuits mais de mauvaise qualité.

C'est souvent le cas des bibliothèques qui ont besoin de notices pour faire leurs bons de commande à partir de leurs modules « acquisition » et qui n'ont pas Electre. Les « nouveautés » ne se trouvant pas encore à la BNF, elles n'ont pas le temps de faire autrement.

– Pratiques différentes selon les services ou les supports

Le manque de formation à l'UNIMARC ou les variantes dans son interprétation (de la part des bibliothécaires comme des éditeurs SIGB) ont pu donner lieu à des strates de catalogages.

Exemple, au remplacement de tel collègue, on a cessé de saisir telle information, jugée inutile car non indexée pour le public.

Exemple de non respect de l'UNIMARC lié à l'affichage public où l'on devient irrespectueux de l'ordre des 200 \$e, \$i, et \$h sur l'OPAC

– Présence de différentes strates chronologiques

Diapo 8 :

– disparités des sources

La dérivation de notices ayant été de pratiques communes, surtout en lecture publique, depuis de nombreuses années, les sources ont pu varier au grès des marchés publics, notamment pour les documents multimédia. Les métadonnées fournies ne sont pas nécessairement structurées de la même façon selon les cas. EAN en 345\$b ou pas, etc.

– mauvais alignement des données

Ces variations dans l'usage du format par les fournisseurs successifs ont pu être corrigées ou mal corrigées par des chargeurs d'import. La volonté d'une simplification d'affichage public aura pu inciter à des choix abandonnés plus tard... (exemple : catalogage à niveaux pour la BD)

– données mal reprises lors d'un changement de SIGB

Par exemple des informations conservées dans un champ local suite à une migration précédente, et « oubliées » lors de la migration présente car préparées sans prise en compte de ces spécificités.

Diapo 9 :

– Difficulté à obtenir les évolutions des passerelles et filtres d'import de nos SIGB par les fournisseurs

Paramétrage souvent complexes et rarement accessibles en interface utilisateurs ; ils sont pourtant indispensable pour obtenir que les données livrées soit placées aux endroits choisis dans le format.

– Veille parfois pas assurée sur l'évolution des formats et des référentiels de travail

Aussi bien du côté des bibliothécaires que du côté des SIGB.

Diapo 10 :

Cet état des lieux passe par la délimitation de grands ensembles en fonction de la présence ou non des données dans les champs en question (présence ou non de l'ISBN en 010\$a) et par l'édition de listing permettant de catégoriser les ensembles avant d'avancer dans le choix d'autres clefs de calculs (chaînes de caractères pour le titre ou l'auteur, nombre de pages, etc.)

Données codées, champs UNIMARC, Numéros d'identification, ARK...

Nouveaux champs 181, 182...

Diapo 11 :

Les processus d'entrée des données dans un SIGB seront présentés par un retour d'expérience des actions menées à la Bpi dans le cadre du projet Concordance.

A l'instar des bibliothèques de lecture publique, la Bpi ne catalogue pas dans un catalogue collectif.

Diapo 12 :

Nous allons proposer des pistes de réflexion en abordant différentes étapes :

- identifier les sources d'entrée dans le catalogue
- mettre à jour le format dans le SIGB
- Manipuler les données
- paramétrer les imports

Diapo 13 :

Pour commencer, un rappel du contexte de production de nos données.

Celles-ci sont intégrées à des notices descriptives structurées dans le respect des normes et des formats Marc, Unimarc le plus souvent.

Diapo 14 :

Chaque donnée d'une notice fait donc partie d'un ensemble structuré : elle ne se suffit pas à elle-même, ne porte pas une information à elle seule.

Elle est qualifiée par la zone qu'elle occupe dans le format.

Elle peut être liée à des référentiels (codes de fonctions, codes de langue).

Elle peut être liée à d'autres notices lorsqu'elle se trouve dans une zone de liens, ou dans une zone soumise à autorité.

Diapo 15 :

Pour améliorer l'entrée de données de qualité dans le catalogue, il est important de se mettre toujours dans la perspective d'un enrichissement ultérieur.

Se donner dès le départ les meilleures conditions possibles pour préparer les données à une opération d'enrichissement ou d'alignement sur d'autres sources.

L'identifiant est bien sûr une donnée pivot mais il ne suffit pas toujours : par exemple pour identifier des doublons, les informations contenues dans le bloc de description standard (zones 2XX : titre, éditeur, date de publication...) sont nécessaires pour lever les ambiguïtés. Par ailleurs, certaines notices sont dépourvues d'identifiants : on ne peut alors s'appuyer que sur les données descriptives.

Diapo 16 :

La fonction de copie de notice est proposée dans nos SIGB en fonction standard mais il faut la proscrire absolument du circuit classique de production. Le risque est de ne pas identifier toutes les zones à corriger dans la notice dupliquée et de laisser dans la notice des informations qui ne concernent pas la ressource décrite.

Des erreurs peuvent ainsi se faire sur les identifiants, mais aussi les dates, les titres... (dans le cas d'une nouvelle édition, la copie d'une notice d'édition précédente risque de ne pas correctement mettre à jour un changement d'accès auteur, une modification du titre, la zone fixe de la date, les identifiants...).

La duplication est une action rapide en elle-même mais les erreurs liées à des copies de notices peuvent être lourdes à corriger. Exemple vu à la Bpi : des séries de bandes dessinées avaient été cataloguées par duplication d'une notice prise en modèle, sans en modifier l'ISBN. L'ISBN étant

le critère permettant l'affichage des vignettes de couverture sur le catalogue public, on imagine assez bien le résultat... La correction de cette erreur liée à la duplication a été fastidieuse et longue. Si on avait laissé cette erreur dans le catalogue, cela aurait de plus compromis toute opération d'enrichissement ultérieur.

Pour les mêmes raisons, il faut proscrire également le recyclage de notices : cela consiste à réutiliser toujours la même notice, en modifiant certaines données pour signaler l'évolution de cette ressource (changement de titre de périodique ou nouvelle édition de manuel, catalogage de séries...)

A la Bpi, nous avons remédié à des cas de recyclage de notices dans le cadre de l'import des notices de périodiques du Sudoc (chantier post-Concordance) : auparavant une seule et même notice était utilisée pour décrire un titre de périodiques : celle-ci comprenait les informations relatives à l'état le plus actuel du titre. Dans le Sudoc, différentes notices sont utilisées pour signaler les différents états de ce titre (selon les changements de titre et d'ISSN). Nous avons donc fait évoluer nos pratiques de description des périodiques en cohérence avec celles du Sudoc. Un mot d'ordre : une publication = une notice

Diapo 17 :

Le fonds de la Bpi est constitué de collections très diverses, du CD jusqu'à la monographie papier mais aussi bandes dessinées, didacticiels, cartes,...

Le catalogage ex nihilo est encore présent pour une part des collections, même si du fait des récupérations de notices (que nous verrons au point suivant), il représente une part plus réduite de la production des données.

Il s'agit d'apporter un soin particulier à cette description dans le respect des normes et du format afin de faciliter un enrichissement futur de la notice.

La Bpi fait des recommandations « qualité » sur les zones essentielles à renseigner pour une bonne exposition des données sur le catalogue public (présentation, pertinence des résultats liée à une bonne indexation) et aussi pour favoriser l'enrichissement ultérieur. Les recommandations portent sur des points de vigilance particuliers.

L'idée n'est pas d'établir un guide Unimarc maison, mais d'apprendre aux chargés de collection à se servir de la documentation complète qui existe en ligne sur les formats (guides méthodologiques du Sudoc et documentation sur le site de la BnF).

Nous privilégions la consultation des sources d'information les plus complètes sur les règles de description afin de favoriser la compréhension et l'utilisation du format dans toute sa richesse.

Ces recommandations se font lors des formations et la documentation est accessible dans la « boîte à outils catalogue » sur l'intranet.

Diapo 18 :

En terme de qualité des données, le choix de la récupération de notices d'une source fiable apparaît comme la meilleure option.

Il s'agit d'encourager la récupération de notices, en particulier celles des agences bibliographiques en mettant à la disposition des catalogueurs des outils ergonomiques pour la récupération de notices.

Choix de la Bpi, deux modes de récupération adaptés aux besoins, et surtout ne générant pas de doublons (protocole z39.50 mais avec processus d'import avec filtres dans la base du sigb). Un mode destiné au chargé de collection (sélection de notices pour alimenter un panier que l'on charge ensuite dans la base), l'autre plus automatisé pour gérer des lots de notices à enrichir (possibilité de constitution de lots sur des critères de données bibliographiques ou d'exemplaire). Cet outil automatisé a ses limites car il ne permet de travailler que sur des notices contenant des identifiants (en l'occurrence ISBN). Les lots doivent être constitués avec attention pour éviter des mises à jour inadaptées.

Diapo 19 :

Si l'on souhaite promouvoir la récupération de notices issues des agences bibliographiques et profiter de la richesse de leurs descriptions, il s'agit d'embarquer dans le SIGB le format Unimarc de manière complète.

Pour mettre à jour le format dans le SIGB, on intègre tout d'abord les caractéristiques du format standard ce qui permet de constituer le socle commun du format.

On s'appuiera ensuite sur les formats de diffusion des agences bibliographiques.

Par exemple, le format Unimarc dans lequel la BnF livre ses données (après conversion du format natif Intermarc) :

http://www.bnf.fr/fr/professionnels/recuperation_donnees_bnf_boite_outils/a.produits_format_unimarc.html

Il peut y avoir des différences entre les formats standards et les formats de diffusion. Pour faciliter la réception des données des agences au sein de notre SIGB, il est conseillé d'intégrer également ces particularités dans le format du SIGB (ou du moins de se poser la question de ce que l'on souhaite intégrer, comment et où).

Exemple à la Bpi : dans les notices diffusées en Unimarc par la BnF, certains indicateurs sont remplacés par des barres verticales. Ces cas ont été entrés dans le paramétrage des zones Unimarc du SIGB pour que la présence de barres verticales pour certains indicateurs ne soit pas considérée comme une erreur dans les notices importées. Notre SIGB détecte les erreurs sur les indicateurs et les signale : il s'agit de lui faire accepter la barre verticale comme une valeur correcte afin de ne pas générer un message d'erreur qui pourrait gêner les catalogueurs en travail courant (et donc les encourager à ne pas tenir compte des messages d'erreur qui apparaissent lorsqu'ils enregistrent leur notice).

La mise à jour du format est un travail conjoint entre l'administrateur et le fournisseur du SIGB :

-déterminer ce que chacun peut faire, ce sur quoi on a la main en tant qu'administrateur.

-demander l'aide du fournisseur pour bien comprendre comment est embarqué le format dans le SIGB et le faire intervenir pour des aspects que l'on ne peut pas prendre en charge soi-même.

Diapo 20 :

Le principe de départ est d'intégrer dans le SIGB toutes les caractéristiques du format dans leur intégralité.

Que ce soit pour les listes de codes ISO ou pour les zones et sous-zones : ne pas faire de tri mais tout saisir (/reprendre).

Au niveau de l'outil de production des données, il s'agit de disposer du format le plus complet possible, pour avoir une mine d'information riche à exploiter ensuite sur les interfaces de recherche (l'adaptation de l'affichage se fera donc au niveau de l'interface et non de l'outil de production des données).

Il est important d'avoir l'assistance du fournisseur pour la mise à jour du format car cela relève d'opérations minutieuses et répétitives.

Exemple de la Bpi :

- les zones variables sont paramétrables par l'administrateur, mais pas les zones fixes.
- aide du fournisseur pour la saisie des codes (langues, fonctions) : lorsqu'un code est associé à plusieurs sous-zones, nous n'avons fait le travail de paramétrage que sur une sous zone. Le fournisseur a ensuite reporté par programme les codes dans les autres sous-zones concernées.
- développement réalisé par le fournisseur pour intégrer le caractère obligatoire/facultatif des zones et sous-zones.

Il s'agit aussi de mesurer les impacts possibles sur le fonctionnement du SIGB : l'introduction de nouvelles zones peut amener à réviser les index de recherche et les modalités de contrôle des zones soumises à autorités. A la Bpi, suite à la mise à jour du format, nous avons révisé de manière significative les index et les contrôles d'autorité.

Diapo 21 :

Le format Unimarc évolue pour s'adapter à la modélisation FRBR. Une veille sur l'évolution de ces formats est donc nécessaire (aussi bien sur le format standard que sur les formats de diffusion) pour mettre à jour le format intégré au SIGB.

Une piste serait de faire entrer cette mise à jour du format intégré au SIGB dans le contrat de tierce maintenance applicative.

Faire aussi appel au club utilisateur du logiciel pour mutualiser cette préoccupation autour du format et de sa nécessaire mise à jour.

Exemple des zones nouvellement créées : comment intégrer dans le SIGB les zones 181/182. Le paramétrage de ces zones pose question aux fournisseurs de SIGB de par leurs caractéristiques : zones et sous-zones répétables, zones à positions, avec des listes de valeurs fermées. Les modalités de saisie de ces zones sont donc assez inédites et peuvent supposer le développement d'outils d'aide à la saisie.

Diapo 22 :

Après avoir parlé des évolutions du format, nous revenons aux opérations faites sur les données dans le cadre du projet Concordance.

Une fois la structure du format mise en place dans le SIGB, il s'agit d'examiner comment les données sont renseignées.

Après analyse des pratiques, des orientations peuvent être données pour améliorer le traitement courant.

Des interventions sont également à prévoir sur les données déjà saisies afin d'harmoniser le traitement sur le rétrospectif.

Pour ne pas se laisser décourager par l'ampleur de la tâche, travailler sur des lots précis en mettant en oeuvre les nouvelles consignes en production courante même si on ne corrige pas le rétrospectif dans l'immédiat. Cela permet d'amorcer des nouvelles pratiques et de s'engager dans une dynamique d'évolution.

Diapo 23 :

De manière concomitante à la mise au carré du format dans l'outil de production, les pratiques de description et de saisie des données sont à réviser.

Méthodologie proposée :

Se focaliser d'abord sur la production courante : inventorier les pratiques, mettre en œuvre de nouvelles recommandations pour la production courante et après seulement, s'intéresser au rétrospectif.

Acter les nouvelles règles et prendre le risque d'avoir des strates (en attendant de pouvoir les traiter par des chantiers de correction automatisés ou non).

Mettre en place des contrôles de saisie et des aides dans le module de catalogage : à la Bpi cela s'est traduit par de nouvelles alertes au moment du catalogage ("attention, zone obligatoire devant être renseignée") et la documentation contextuelle des messages d'erreur (explicitier l'erreur en donnant des pistes de résolution).

Reprendre les bordereaux de saisie :

Les SIGB intègrent différents dispositifs de saisie des notices en Unimarc. A la Bpi, cela se traduit par des grilles de saisie (appelées "bordereaux") qui comportent un certain nombre de zones affichées par défaut lorsque l'on crée une notice ex nihilo et des zones que l'on peut ajouter selon les besoins. Les bordereaux comportent les zones Unimarc que l'administrateur leur affecte. Ils ne contiennent pas toutes les zones mais celles les plus utilisées pour un type de document donné (à la Bpi, les bordereaux correspondent à des types de document : monographies, documents sonores, films...).

Dans le cadre de la révision du format dans le SIBG, il faut aussi mettre à jour ces bordereaux de saisie.

Les bordereaux sont comme une surcouche intermédiaire entre le format présent dans le SIGB et les catalogueurs. Il est donc nécessaire de les sensibiliser à ce fonctionnement afin qu'ils ne se bornent pas aux zones présentes par défaut dans le bordereau mais appellent des zones supplémentaires si besoin. De même, si une zone n'est pas présente du tout dans le bordereau, les catalogueurs pourront demander à l'administrateur de l'ajouter. Il est préférable de mettre à disposition des catalogueurs les zones adéquates selon les besoins qu'ils expriment pour ne pas risquer des détournements d'usage de zones non prévues à cet effet.

Soulignons que le format Unimarc est très riche et prévoit des zones très précises selon les cas à traiter.

Selon les besoins des catalogueurs, il s'agit de trouver la zone la plus adaptée. Une vigilance particulière est à avoir pour bien distinguer les informations descriptives (qui relèvent notamment du bloc 2XX) des données de gestion propres à chaque établissement (nom de fournisseur, états de collections, modalités d'accès au document). Les données de gestion ne doivent pas figurer dans les zones standard mais dans des zones locales (bloc 9XX) ou dans des données d'exemplaire.

Exemple Bpi, certains services mettaient le fournisseur en zone d'édition (210) alors que le format propose une zone de note 345 plus adaptée à cette mention (note sur les modalités d'acquisition).

Diapo 24 :

– Pour le traitement du rétrospectif, nous nous concentrons sur la connaissance des données (et non de la composition des collections, le format étant le même quel que soit le support). L'analyse porte sur le contenu des zones et l'identification de strates de traitement.

– Pour localiser les données placées dans des zones inappropriées (analyse quantitative et qualitative), la Bpi utilise l'outil pgAdmin. On interroge ainsi la base catalographique qui est au format PostgreSQL.

– En parallèle, est menée une analyse qualitative basée sur les connaissances des chargés de collection. Dans une démarche de recueil des pratiques, il s'agit de donner du sens aux constats faits dans les extractions informatiques de données.

– Avec l'aide du fournisseur, la Bpi a aussi faire du ménage préalable dans des notices coquilles vides pour réduire les masses à traiter.

– L'inventaire des pratiques permet d'identifier des typologies de notices à traiter. Pour faciliter les interventions de correction, la Bpi a affecté des marqueurs aux notices, dans une zone locale

indexée, pour établir des lots ayant les mêmes propriétés (noms de famille, titre uniformes musicaux...).

– L’analyse des pratiques et des données peut se faire avec des outils externes au SIGB comme pgAdmin. La limite étant que ceux-ci ne permettent que l’identification des cas à traiter mais il n’y a pas de moyens directs d’intervention sur les données à corriger, celles-ci étant “captives” du SIGB (dans le cas des SIGB propriétaires comme le nôtre). C’est pourquoi la Bpi a développé au sein du SIGB des outils de détection des erreurs (doublons, identification de zones en erreur) et de correction directs.

– Recommandations pour les chantiers rétrospectifs : les SIGB proposent des fonctions de correction en lot. Certaines corrections peuvent être faites soi-même si les volumes ne sont pas trop importants. Ces fonctions sont néanmoins limitées à certains types de corrections. Pour d’autres interventions plus complexes sur les données (notamment des déplacements de zones ou des changement d’étiquette pour des autorités), la Bpi a fait appel à une prestation du fournisseur. Enfin, certaines modifications ne sont pas automatisables et nécessitent une analyse au cas par cas. A la Bpi, nous envisageons d’organiser ce type de correction sous forme de chantiers participatifs : une première séance se déroulera fin novembre pour la correction de zones de lien contenant des titres uniformes musicaux. L’idée est de réunir plusieurs collègues volontaires sur des créneaux précis pour corriger ensemble les notices. L’enjeu est de trouver du temps et des moyens humains pour commencer ces corrections sans trop impacter la charge de travail déjà lourde des chargés de collection et des services support.

Diapo 25 :

Dans cet exemple, on cherche à savoir comment est renseignée la zone 305 : le contenu de cette zone apparaît dans la colonne B.

La restitution se fait sous forme de tableur, ce qui permet de faire des tris et d’affiner l’analyse des résultats.

Diapo 26 :

Pour améliorer la qualité des données en traitement courant mais aussi sur le rétrospectif, nous avons vu qu’il était intéressant de recourir à l’import de notices issues des agences bibliographiques notamment.

Selon les sources, le format de diffusion des données ainsi que leur qualité peuvent varier : il est donc important de porter une attention particulière au paramétrage des imports pour arriver à un résultat satisfaisant.

Il est parfois nécessaire de faire un paramétrage particulier pour une source donnée, cela implique d’avoir différentes modalités d’import selon les sources.

Ce ne sont pas les sources les plus riches qui posent le plus de problème, mais plutôt celles où les données sont moins fiables.

Exemple à la Bpi, pour l’import des notices autorités de la BnF, le paramétrage des imports est très léger. On importe toutes les zones, il n’y a pas de filtrage ou d’exclusion de zones.

Le paramétrage des imports passe par la mise en place de clés de mise à jour, de filtres et d’options permettant d’intervenir sur les données au moment de l’import.

Diapo 27 :

Quand les imports ont pour fonction d’enrichir une notices déjà présente dans le SIGB, la mise en à jour se fait sur la base d’une clé, un identifiant numérique. Quand cette clé est reconnue comme

identique dans le SIGB et dans la source externe, la mise à jour se fait sur ce critère.

Pour les notices bibliographiques, la Bpi utilise l'ISBN comme clé de mise à jour et pour les autorités, il s'agit du numéro FRBNF en zone 035.

Les identifiants pérennes ARK sont importés dans les notices mais n'ont pas d'usage fonctionnel pour le moment à la Bpi. Pour préparer de futures exploitations, il est important de bien stocker ces identifiants dans le SIGB (zone 033).

Dans les paramètres de l'import, il s'agit de bien spécifier les zones de destination de ces identifiants. Dans la notice source, l'identifiant des agences bibliographiques est en zone 001 et doit être importé en zone 035.

Les identifiants ARK sont en 003 et doivent être importés en zone 033.

La mise en place de clé de mise à jour peut supposer une intervention du fournisseur pour qu'une zone soit bien prise en compte comme pivot de l'import.

Diapo 28 :

Les SIGB proposent de définir des règles d'import.

A la Bpi, ce paramétrage se fait dans des tables de conversion. On y détermine zone à zone comment l'import se fait : si une zone est rejetée complètement, ou importée et selon quelles modalités (on précise quelles sont les sous-zones importées, ce qui permet d'exclure aussi certaines sous-zones non citées).

(Exemple : ne pas importer la zone 676 car la Dewey n'est pas notre système d'indexation).

Des filtres plus fins peuvent aussi exister : selon la présence ou l'absence de certaines données dans des zones ou sous-zones précises, on peut demander au système d'exclure une zone, une sous-zone ou même la notice entière au moment de l'import.

Exemple : ne pas importer l'indexation issue d'un système que l'on ne souhaite pas utiliser.

Exemple à la Bpi, nous utilisons uniquement l'indexation Rameau, nous filtrons donc les indexations établies sur la base de noms géographiques BnF : exclure de l'import les zones 607 qui comporte la valeur AFNOR NF Z 44-081 en \$2.

Cela devrait aussi permettre de filtrer des notices identifiées comme incomplètes par les agences bibliographiques.

Dans le format de diffusion de la BnF, cette information est donnée par la valeur 3 en position 17 du label :

“3 sous-niveau 3. La notice est incomplète et peut être complétée par la suite par l'agence de catalogage”

Diapo 29 :

Le paramétrage des imports doit se faire en cohérence avec les nouvelles pratiques de description établies avec les catalogueurs.

Il est important de protéger les valeurs ajoutées qui sont propres à l'établissement afin d'encourager les catalogueurs à saisir ces informations.

Dans le SIGB de la Bpi, la protection peut se faire au niveau du paramétrage de la zone où l'on peut déterminer que la zone est protégée.

Pour pouvoir intervenir plus finement, la protection peut se faire également au moment de l'import.

Une autre option de protection de zone permet à la fois d'importer et de conserver une zone. Nous

l'utilisons par exemple pour les zones d'indexation matière : cela nous permet de conserver notre indexation matière tout en important celle de la BnF si celle-ci est différente.