



Atelier

5/11/18

{BnF

# Bibliostratus

## Stratégie d'Alignement d'URLs pour la Transition bibliographique

« le passage à RDA nécessite des outils et un format de production plus adaptés à des **catalogues FRBRisés**. Le rôle du groupe Systèmes & Données est d'envisager quels pourraient être ces outils et de **préparer la migration des données des catalogues actuels vers une nouvelle structuration de l'information bibliographique**. »

<https://www.transition-bibliographique.fr/>  
<https://www.transition-bibliographique.fr/enjeux/position-francaise-rda/>  
<https://www.transition-bibliographique.fr/systemes-et-donnees/recommandations/>



Transition  
bibliographique

FRBRiser les catalogues et adopter  
de nouvelles règles de catalogage  
dérivées de RDA



# Bibliostratus

## C'est

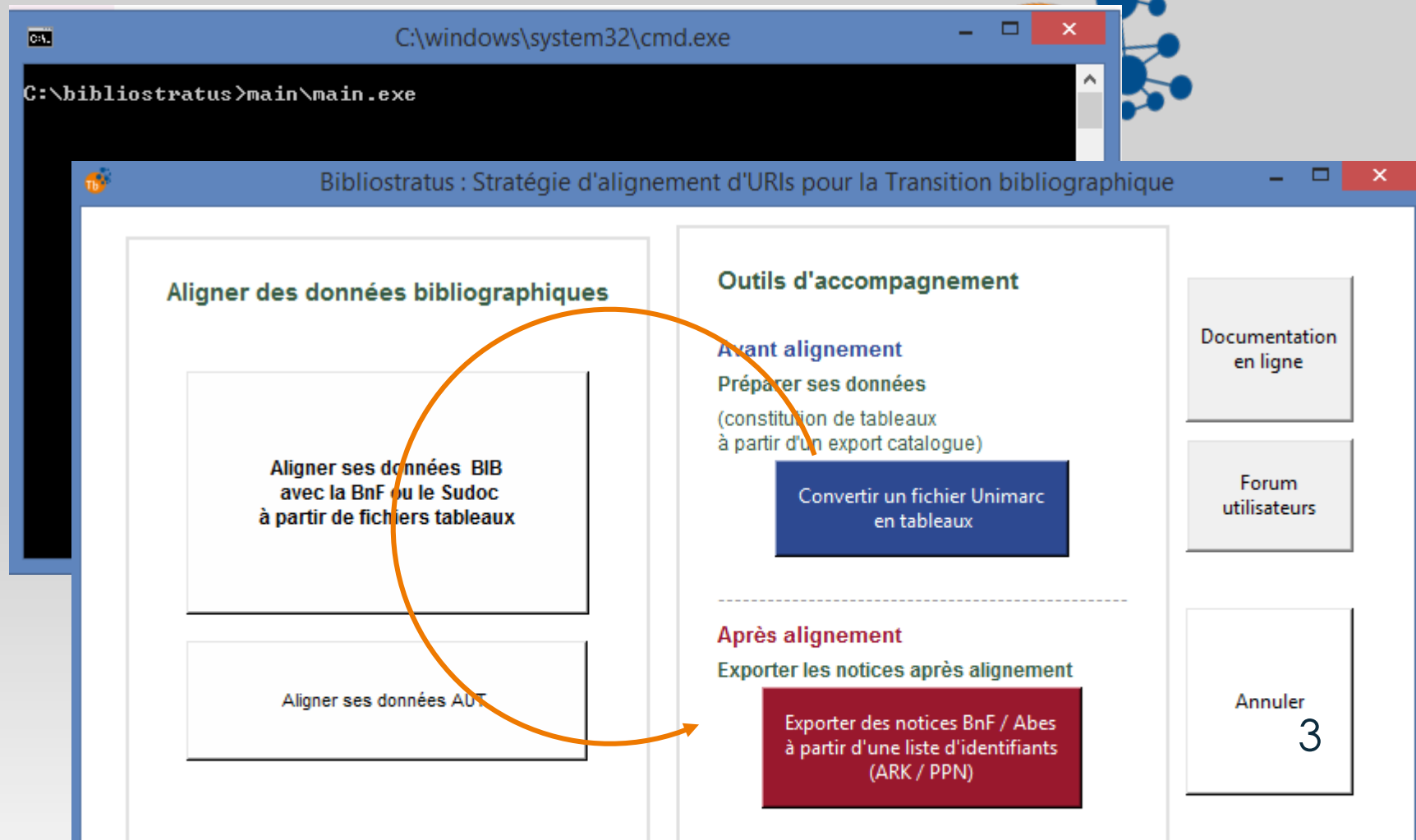
- **1 logiciel** développé dans le cadre du programme de Transition bibliographique
  - **3 modules de traitements** disponibles : transformation préalable des fichiers d'entrée si besoin ; alignement ; récupération de notices
  - des **mises à jour régulières** avec des améliorations au fil de l'eau :
    - documentation technique  
<https://github.com/Transition-bibliographique/bibliostratus>
    - page sur site TB  
<https://www.transition-bibliographique.fr/bibliostratus-alignement-donnees-catalogues/>
    - forum agorabib  
<http://www.agorabib.fr/topic/3317-bibliostratus-mettre-en-correspondance-ses-notices-avec-elles-de-la-bnf/>



# Bibliostratus

## Pour

- **préparer** les données de vos catalogues en identifiant les alignements potentiels avec celles de la BnF ou du SUDOC et IdRef
- **trouver** les identifiants ARK ou PPN à partir de l'analyse de lots de notices
- **aligner** des notices bibliographiques ou des notices d'autorités
- **recupérer** des données validées.





# Avant Bibliostratus

## Bien préparer ses données



- Analyser les données de son catalogue par requêtes pour identifier les chantiers
- Vérifier la présence ou non d'identifiants pérennes des notices bibliographiques et d'autorités
- Lister les règles appliquées au sein de l'établissement en terme de catalogage
  - Dérivation : règles de fusion / remplacement
  - Catalogage document en main : existe-t-il des règles locales ? Sont-elles plus spécifiques ou dérogent-elles au format standard ?
- Dans quels formats sont les données (UNIMARC, INTERMARC, MARC21, DublinCore, Onix) à l'importation et à l'exportation ?
- Quels types d'améliorations peuvent recevoir les données du catalogue ?



# Bibliostratus : le module **bleu**

- aide à la **préparation des données** pour le module d'alignement. Ce module n'est pas nécessaire si vous pouvez fournir un **fichier tabulé** extrait de votre SIGB **selon les critères requis** : les données attendues seront différentes selon le type d'alignement souhaité (bibliographique / autorité ou biblio pour alignement autorités) et selon le type de ressource à aligner (monographie, périodique, audio/vidéo etc.)
- **convertit** et génère des fichiers tabulés à partir de fichiers iso2709 ou XML en UTF8 sans [BOM](#).

Préparer

■ exemple

Conversion de fichiers de notices Unimarc en tableaux

En entrée

Sélectionner un fichier de notices Unimarc

UTF-8

ISO  
2709

Marc  
XML

Format de fichier

- ☒ iso2709 encodé UTF-8 [exemple](#)
- ☐ iso2709 encodé ISO-8859-1
- ☐ Marc XML encodé UTF-8

Type de notices

- ☒ bibliographiques
- ☐ autorités
- ☐ biblio - pour alignement Autorités

Je ne sais pas / Je ne comprends pas

En sortie

Préfixe des fichiers en sortie :

Le programme va générer plusieurs fichiers, par type de document, en fonction du processus d'alignement avec les données de la BnF et des métadonnées utilisées pour cela :

- code TEX : monographies
- code VID : audiovisuel (CD/DVD)
- code PER : périodiques
- autres non prévus pour le module d'alignement (blanc)

Les codes indiquent l'option d'alignement à choisir dans le module. S'il n'y a pas de code, le fichier ne peut être chargé tel quel dans le module blanc : il faut au moins reprendre les colonnes.

Pour répartir les notices en fichiers, le programme utilise les informations présentes dans les zones codées de chaque notice Unimarc

Fichiers  
tabulés  
par type  
de doc

DÉFENSE  
DE  
FUMER

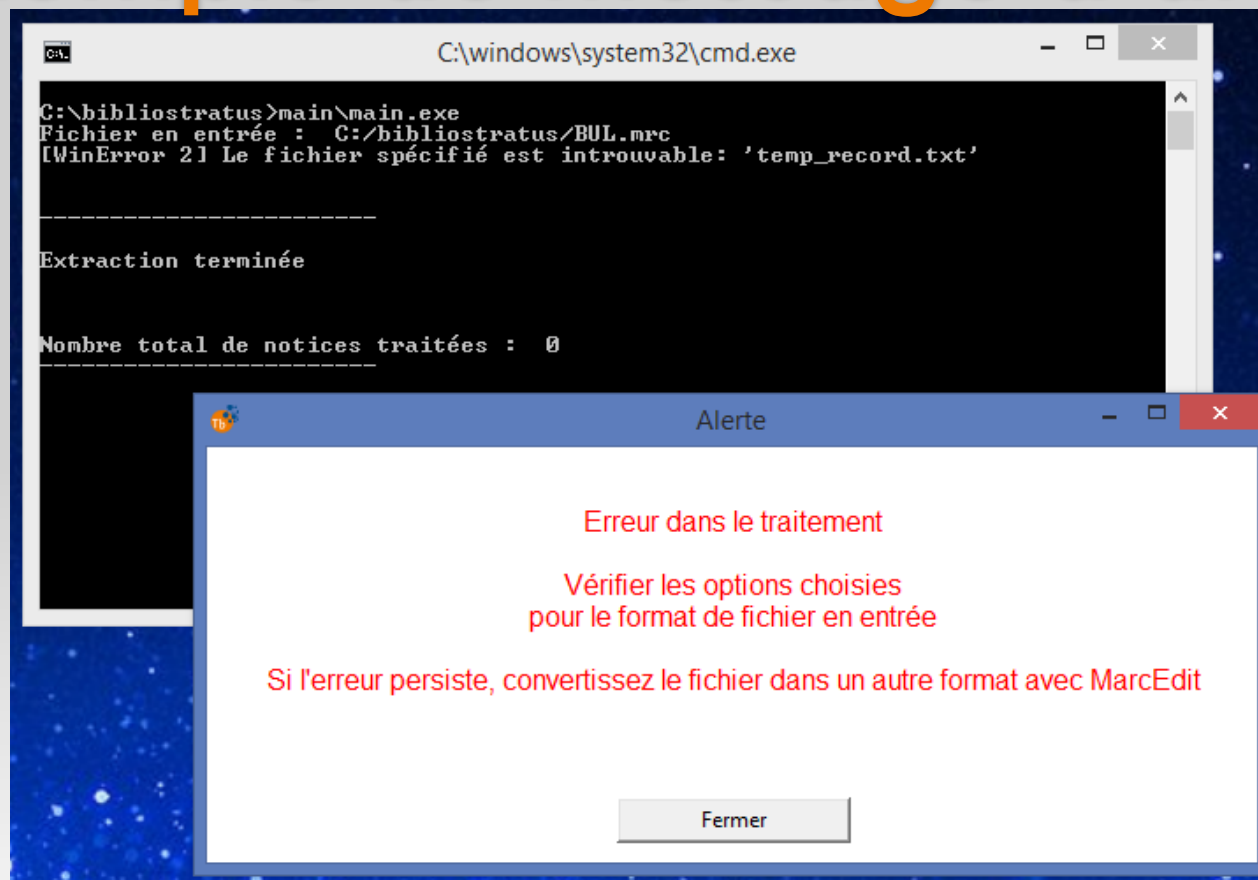


# Bibliostratus : le module **bleu**

## Exemple de message d'alerte

Préparer

ISO  
2709





# Bibliostratus : le module blanc

## Notices bibliographiques

Aligner

- **traite** en entrée un **fichier tabulé** contenant des informations bibliographiques, généralement :
  - des identifiants (ARK BnF, FRBNF, ISBN, EAN, etc.)
  - des chaînes de caractères (titre, auteur, date)
- **nettoie** et applique un ordre de priorité de traitement des données
- **génère des fichiers** avec des propositions d'**alignement**

Programme d'alignement de données bibliographiques avec la BnF et le Sudoc

En entrée :

Fichier contenant les notices :  Séparateur TAB, Encodage UTF-8

En sortie :

Nombre de fichiers

☒ 1 fichier

☐ Plusieurs fichiers (Pb / 0 / 1 / plusieurs ARK trouvés)

Type de documents

☒ [TEX] Monographies texte [exemple](#)

(Colonnes : Num Not | FRBNF | ARK | ISBN | EAN | Titre | Auteur | Date | Volume-Tome | Editeur)

☐ [VID] Audiovisuel (DVD)

(Num Not | FRBNF | ARK | EAN | N° commercial | Titre | Auteur | Date | Editeur)

☐ [AUD] Enregistrements sonores [exemple](#)

(Num Not | FRBNF | ARK | EAN | N° commercial | Titre | Auteur | Date | Editeur)

☐ [PER] Périodiques [exemple](#)

(Num Not | FRBNF | ARK | ISSN | Titre | Auteur | Date | Lieu de publication)

☐ [CAR] Cartes

(Num Not | FRBNF | ARK | ISBN | EAN | Titre | Auteur | Date | Editeur | Echelle)

Aligner de préférence :

☒ Avec la BnF (et à défaut avec le Sudoc)

☐ Avec le Sudoc (et à défaut avec la BnF)

Aligner les notices BIB

Fichiers résultats

7

Exemple de message d'alerte

```
C:\windows\system32\cmd.exe
C:\un-probl%C3%A8me-dencodage--passez-en-xml-avec-marccedit
Consultez le fichier TO-ALERT-notices_pb_encodage.txt

Extraction terminée

texte-monographie : 862
texte-périodiques : 676
texte-collection : 2
ressource électronique-monographie : 14

Nombre total de notices traitées : 1555

Exception : Tkinter.TclError:
Traceback (most recent call last):
  File "C:\un-probl%C3%A8me-dencodage--passez-en-xml-avec-marccedit\init.py", line 1699, in <lambda>
    file "noticesbib2arkBnF.py", line 1831, in <lambda>
    file "noticesbib2arkBnF.py", line 1577, in launch
    File "noticesbib2arkBnF.py", line 1569, in perimpr
AssertionError
```

Alerte

Le fichier contient un nombre insuffisant de colonnes.

Vérifiez les options indiquées dans le formulaire ou le format du fichier en entrée

Format attendu - Ligne 1 :

Num Not : BIUT438517

FRBNF :

ARK :

ISSN :

Titre : Activité scientifique 1993-1994

Auteur : Corto, Claudius Bopaul





# Bibliostratus : le module blanc

## Notices d'autorités

Aligner

Aligner ses données d'autorité avec la BnF ou IdRef

En entrée

Fichier contenant  
les notices d'autorité à aligner

Sélectionner un fichier  
Séparateur TAB, Encodage UTF-8

☒ Mon fichier a des en-têtes de colonne

Type de données en entrée

☒ [PÉRS] Liste de notices Personnes [exemple](#)

(N° Notice AUT | FRBNF | ARK | ISNI | Nom | Prénom | Date de début | Date de fin)

☐ [ORG] Liste de notices Organisations

(N° Notice AUT | FRBNF | ARK | ISNI | Nom | Prénom | Date de début | Date de fin)

☐ Liste de notices bibliographiques [exemple](#)

(N° Notice AUT | N° notice BIB | ARK Bib | FRBNF Bib | Titre | Date de publication | ISNI | Nom | Prénom | Dates auteur)

Aligner de préférence :

☒ Avec la BnF (et à défaut avec le Sudoc)

☐ Avec IdRef (et à défaut avec la BnF) - uniquement à partir d'AUT

☒ Relancer sur isni.org en cas d'absence de réponse

En sortie

☒ 1 fichier

☐ Plusieurs fichiers  
(Pb / 0 / 1 / plusieurs ARK trouvés)

☐ Récupérer les métadonnées BnF

Préfixe des fichiers en sortie

Aligner  
les autorités

Fichiers  
résultats

Annuler  
8

■ **traite** en entrée un **fichier tabulé** contenant des données d'autorités issues de notices d'autorités ou bibliographiques





# Bibliostratus : le module blanc

## Analyse et sélection

- **vérifier** les fichiers résultats générés : 1 ou plusieurs fichiers selon l'option choisie et un rapport statistique
- **exemple**

Nb ARK trouvés Nb notices concernées

|    |     |
|----|-----|
| 0  | 755 |
| 2  | 10  |
| 1  | 92  |
| 8  | 1   |
| 5  | 1   |
| 3  | 2   |
| 19 | 1   |

Analyser

Sélectionner

Problème d'accès  
URL Type de probl  
<http://catalogue.>  
<http://ca>

Fichiers  
résultats

Fichiers  
ARK/PPN

|    | A          | B     | C                                                | D                            | E                                             | F     | G    | H   | I               | J              | K    | L            | M                  |
|----|------------|-------|--------------------------------------------------|------------------------------|-----------------------------------------------|-------|------|-----|-----------------|----------------|------|--------------|--------------------|
| 1  | NumNot     | nbARK | ark trouvé                                       | Méthode                      | ark initial                                   | FRBNF | ISBN | EAN | Titre           | auteur         | date | Tome/Volum   | editeur            |
| 53 | BIUT181149 | 2     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Cartonnages     | dafsa          | 1995 |              | Groupe             |
| 54 | BIUT181150 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Liqueurs et a   | dafsa          | 1995 |              | Groupe             |
| 55 | BIUT181151 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Aliments pou    | dafsa          | 1995 |              | Groupe             |
| 56 | BIUT181154 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Extraction d'   | dafsa          | 1995 |              | Groupe             |
|    | BIUT181156 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Production et   | dafsa          | 1995 |              | Groupe             |
|    | BIUT182334 | 0     |                                                  |                              |                                               |       |      |     | Le refus de c   | jean-christi   | 1967 |              | Librairie          |
|    | BIUT195512 | 0     |                                                  |                              |                                               |       |      |     | Les effets du   | jacques kays   | 1960 |              | LGDJ               |
|    | PPN0485600 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel georg   | 1952 | Tome II      | Librairie générale |
|    | PPN0485597 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel rene j  | 1952 | Tome I       | Librairie générale |
|    | BIUT197778 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel emile   | 1953 | Tomes XII et | Librairie générale |
|    | PPN0485628 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel paul    | 1963 | Tome XIV     | Librairie générale |
| 64 | BIUT199099 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel georg   | 1957 | Tome V       | Librairie générale |
| 65 | BIUT199101 | 0     |                                                  |                              |                                               |       |      |     | Traité pratique | marcel paul    | 1954 | Tomes VI et  | Librairie générale |
| 66 | BIUT199180 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Rapport sur l   | france des co  | 1968 |              | la Documentatio    |
| 67 | BIUT199393 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Influence de    | paul durand    | 1961 |              | Librairie générale |
| 68 | BIUT203511 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Calcul écono    | centre d'eco   | 1968 | I            | Presses universi   |
| 69 | BIUT203976 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre BNF contenu dans titre initial, Titre-A |       |      |     | Etude des éco   | commerce pe    | 1996 |              | COPEF              |
| 70 | BIUT204108 | 1     | <a href="http://ark:/12148/cl">ark:/12148/cl</a> | Titre-Auteur-Date + contrôle | Titre 200\$a, Titre-Auteur-Date               |       |      |     | Droit admini    | valérie marcel | 1950 |              | Sirey              |



# Bibliostratus : le module **rouge**

- **récupérer des notices** à partir des identifiants ARK ou PPN/IdRef obtenus et qui ont été analysés et vérifiés
- **charger** dans son système local.

Bibliostratus : Récupérer les notices complètes BnF / Sudoc / IdRef à partir d'une liste de n° de notices

Récupérer

Sélectionner un fichier contenant une liste de n° de notices ARK BnF ou PPN Abes (un numéro par ligne)

Fichiers ARK/PPN

☒ N° de notices bibliographiques [exemple](#)

☐ N° de notices d'autorités [exemple](#)

-----

☒ Mon fichier a des en-têtes de colonne

Récupérer aussi les notices d'autorité liées

☐ auteurs ☐ sujets ☐ oeuvres

Notices à récupérer en :

☒ Unimarc

☐ Interimarc

☐ [ARK BnF] Unimarc avec notices analytiques

☐ [ARK BnF] Interimarc avec notices analytiques

Format du fichier :

☒ iso2709

☐ Marc XML

Préfixe fichier(s) en sortie

OK

Documentation en ligne

Forum utilisateurs

Annuler

Notices Marc

{ BnF

  
sudoc





# Bibliostratus : questions, remarques, échanges



Transition  
bibliographique

FRBRiser les catalogues et adopter  
de nouvelles règles de catalogage  
dérivées de RDA

## À vous de jouer...



Illustrations

<https://gallica.bnf.fr/ark:/12148/btv1b550041484>

<https://gallica.bnf.fr/ark:/12148/btv1b53033620q>

<https://gallica.bnf.fr/ark:/12148/btv1b53071102q>

<https://gallica.bnf.fr/ark:/12148/btv1b59048262>

<https://gallica.bnf.fr/ark:/12148/btv1b590241274>

<https://gallica.bnf.fr/ark:/12148/btv1b69145500>

<https://gallica.bnf.fr/ark:/12148/btv1b50437883>

<https://gallica.bnf.fr/ark:/12148/btv1b530995413>

<https://gallica.bnf.fr/ark:/12148/btv1b9037324>

<https://gallica.bnf.fr/ark:/12148/btv1b6931374>



# Pour les établissements du réseau SUDOC

- Quelques précisions...



# Outils et méthodologie



Transition  
bibliographique

ERBRiser les catalogues et adopter  
de nouvelles règles de catalogage  
dérivées de RDA

- Préparer une migration

[http://documentation.abes.fr/sudoc/manuels/administration/prepa\\_migration/index.html#PreparerSaMigration](http://documentation.abes.fr/sudoc/manuels/administration/prepa_migration/index.html#PreparerSaMigration)

Taux de recouvrement avant exemplarisation : Identifiants (PPN, ISBN, ISSN...) ; Dates/Auteurs/Titres

[http://documentation.abes.fr/sudoc/manuels/echanges/exemplarisation\\_automatique/prepafichierettest.html](http://documentation.abes.fr/sudoc/manuels/echanges/exemplarisation_automatique/prepafichierettest.html)

- Webservices

<http://www.abes.fr/Espace-Pro-Acces-direct-a/API-et-Web-Services>

- Programme d'alignement des notices autorités personnes physiques \*

<https://fr.slideshare.net/abesweb/jabes2017-qualinca-alignements>

...

\* #jabes17 :: session parallèle : Le signalement augmenté : des algorithmes au service des catalogues...



Administration et outils

[Lien profond vers l'OPAC](#)

[Gestion des bibliothèques](#)

[Gestion des utilisateurs Sudoc](#)

[Gestion des utilisateurs Sudoc-PS](#)

[Statistiques](#)

[Webservices Sudoc](#)

[Le Sudoc en RDF](#)

[IdRef en RDF](#)

[Recommandations pour la  
migration de données](#)



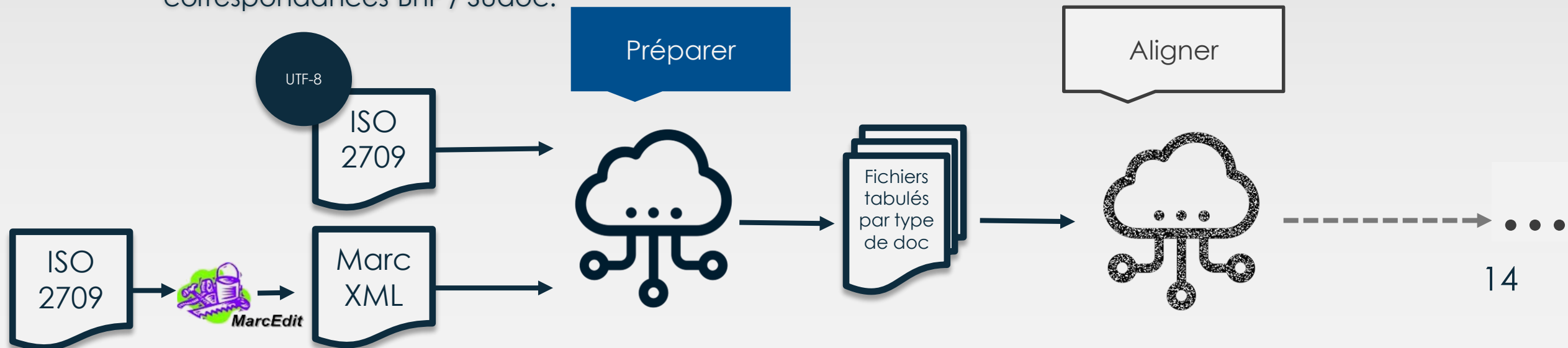
# Tests sur des notices locales



Transition  
bibliographique

FRBRiser les catalogues et adopter  
de nouvelles règles de catalogage  
dérivées de RDA

- Récupération de notices bibliographiques locales **non traitées par les outils ABES** de taux de recouvrement
- Fichiers ISO 2709 fournis par des établissements
- UTF8 ou non
- Résultats divers selon les notices fournies : fichiers tabulés par type de documents avec correspondances BnF / Sudoc.



# Exploiter les résultats



Transition  
bibliographique

FRBRiser les catalogues et adopter  
de nouvelles règles de catalogage  
dérivées de RDA

- Vérification des identifiants proposés.
- Récupération des notices BnF dans la base d'appui Sudoc / la base de production
- Exemplarisation

