

Bibliothèque nationale de France

## Service SRU de la BnF Documentation générale

**Bibliothèque nationale de France**  
**direction des Services et des réseaux**  
département des Métadonnées  
service Diffusion des métadonnées

Version : 24/01/19  
Auteur : Département des Métadonnées

# Le service SRU de la BnF

---

1. Présentation .....	2
2. Vue d'ensemble du service SRU de la BnF .....	2
3. Formuler une recherche.....	3
4. Réponse SRU et différents formats proposés.....	5
5. Exemples de recherche .....	7
6. Parcourir l'ensemble des résultats d'une requête et choisir son format.....	8
7. Annexes.....	9

## 1. Présentation

Le protocole SRU (Search/Retrieval via URL) est né de la volonté d'adapter le protocole d'échange Z39.50 aux standards technologiques du web (requêtes http et format XML).

Conçu en 2002 par le groupe « ZING » (« Z39.50 International Next Generation »), il a été développé et maintenu par la **Bibliothèque du Congrès** (Version 1.1 en 2004, Version 1.2 en 2007) :

<http://www.loc.gov/standards/sru>.

Depuis février 2013, le protocole est pris en charge par le consortium **OASIS** (Advancing open standards for the information society). La version 2.0, en cours d'élaboration, est hébergée sur le site de la Bibliothèque du Congrès : <http://www.loc.gov/standards/sru/sru-2-0.html>.

## 2. Vue d'ensemble du service SRU de la BnF

Le service SRU implémenté à la BnF est la [version 1.2 du SRU](#).

Il permet d'interroger le catalogue général de la BnF et de récupérer ses métadonnées dans différents formats structurés. Le contenu interrogeable est le même que celui du catalogue général de la BnF <http://catalogue.bnf.fr>, soit l'ensemble des données bibliographiques et d'autorité du catalogue général de la BnF.

L'URL de base du service est : <http://catalogue.bnf.fr/api/SRU>

Les 2 opérations implémentées par le service SRU sont :

- [Explain](#) : fournit la liste des critères interrogeables, modes d'interrogation, formats des données en sortie, valeurs par défaut, etc et renseigne sur le service (url de base du serveur, description générale, ...)  
<http://catalogue.bnf.fr/api/SRU?version=1.2&operation=Explain>
  - [SearchRetrieve](#) : permet de lancer une requête dans le catalogue général de la BnF (notices bibliographiques et d'autorité) et renvoie, encapsulées dans du XML, les notices du catalogue qui répondent à la requête. Cette opération rend obligatoire le paramètre **query** qui contient la requête proprement dite.
1. Formuler une recherche

## 3. Formuler une recherche

### 1) Structure de l'URL de requête

Un exemple de requête interrogeant le catalogue :

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.title all "Victor Hugo"](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.title all \)

- **Racine de l'URL** : <http://catalogue.bnf.fr/api/SRU>
- **version** : Paramètre obligatoire. Version du protocole SRU utilisé ("1.2")
- **operation** : Paramètre obligatoire. Valeurs autorisées :
  - "explain" : fournit la liste des paramètres disponibles : critères de recherche, formats disponibles en sortie, etc.
  - "searchRetrieve" : permet de lancer une requête dans le catalogue
- **query** : Paramètre obligatoire si "operation=searchRetrieve". Contient la requête proprement dite formulée selon le langage de requête [CQL](#) « Contextual Query Language ». Voir ci-dessous comment « Formuler la requête (query) : contextes, critères , ... ».

Un exemple plus complet permettant de parcourir les résultats et de choisir son format:

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.title all "Victor Hugo"&startRecord=1&maximumRecords=5&recordSchema=dublincore](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.title all \)

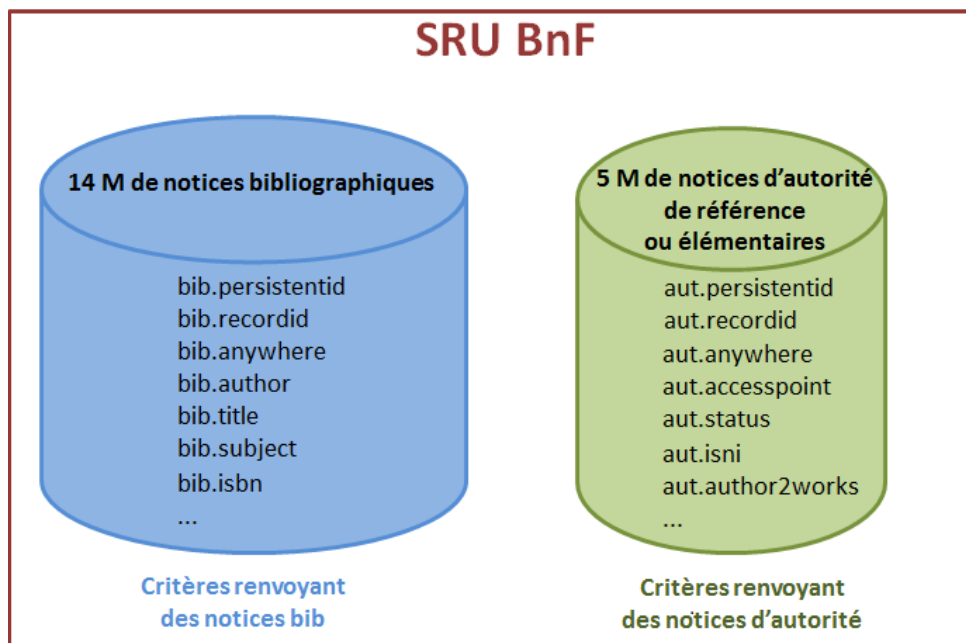
- **recordSchema** : optionnel. Définit le format dans lequel le serveur doit renvoyer les notices (ici "recordSchema=dublincore" (pour [Dublin Core](#), en XML)). Si non renseigné, c'est la valeur par défaut qui est prise : unimarcXchange (pour Unimarc encapsulé dans du MarcXchange).
- **maximumRecords** : optionnel (valeur par défaut = "10"). Indique le nombre de notices par page de résultats. Ce nombre doit être inférieur ou égal à "1000".
- **startRecord** : optionnel (valeur par défaut : "1"). Désigne la position du premier résultat renvoyé au sein de l'ensemble des résultats répondant à la recherche. Permet, combiné avec le paramètre "maximumRecords", de parcourir l'ensemble des résultats.

Voir les exemples sur « Parcourir l'ensemble des résultats d'une requête et choisir son format » plus loin dans le document.

### 2) Formuler la requête (query) : contextes, critères , ...

Le langage de requête utilisé pour formuler les recherches du SearchRetrieve est conforme à la norme SRU / CQL : [CQL Version 1.2](#)

Le service SRU de la BnF offre **l'accès à l'ensemble des données bibliographiques et d'autorité du catalogue général de la BnF**. (Voir Annexe 1 : Conditions de réutilisation des données de la BnF)



**Concernant l'accès aux notices d'autorité**, cet accès comprend toutes les notices d'autorité du catalogue général, à savoir : les notices d'autorité de référence (notices validées par la BnF) et les notices élémentaires (notices non retravaillées par un catalogueur ou un professionnel de la BnF).

La BnF attire votre attention sur le fait que, **par défaut, seules les notices d'autorité de référence sont interrogeables**. Ce choix de politique est le même que celui fait dans l'interface du catalogue général de la BnF où, par défaut, seules ces notices d'autorité de '*qualité*' (de référence) sont interrogeables. Nous rappelons également que via dans le serveur Z39.50 de la BnF, seules ces notices de référence sont accessibles.

Pour rechercher dans l'ensemble des notices d'autorité (de référence et élémentaires) du catalogue général de la BnF, l'index "**aut.status**" doit être utilisé et combiné aux autres critères portant sur les notices d'autorité avec pour valeur : (**aut.status any "validated sparse"**).

Pour rechercher dans les données du catalogue, **on peut combiner plusieurs clauses de recherche** par des AND, OR, ... . **Pour chaque clause, la recherche pour un critère donné se fait selon la relation** 'all', 'any', 'adj', ... suivant que l'on cherche 'tous les termes', 'un des termes' ou 'expression exacte' dans la notice.

**Deux contextes : 'bib' et 'aut'** permettent de préciser si l'on travaille avec des critères portant sur les notices bibliographiques du catalogue ou sur les notices d'autorité, permettant ainsi de clarifier l'usage du SRU et de combiner entre eux tous les critères d'un même contexte.

**Les critères** disponibles sont détaillés dans le **document "Tableau des critères"**.

On y trouve entre autres :

- **pour les notices bibliographiques** (critères préfixés par le contexte '*bib*') :  
bib.anywhere, bib.author, bib.title, bib.subject, bib.doctype, bib.recordtype, bib.persistentid, bib.recordid, bib.isbn, bib.issn, bib.ean, bib.ismn, bib.isrc, bib.publisher, bib.publicationdate, etc.
- **pour les autorités** (critères préfixés par le contexte '*aut*') :  
aut.anywhere, aut.accesspoint, aut.type, aut.status, aut.persistentid, bib.recordid, aut.isni, etc.

**Exemple** : Rechercher toutes les œuvres (Titre conventionnel, Titre uniforme textuel, Titre uniforme musical<sup>1</sup>) en rapport avec "Hamlet" de Shakespeare (qui comporte les mots "hamlet" et "shakespeare" dans le contenu de la notice) et qui sont des notices d'autorité de référence :

[query=\(aut.type any "tic tut tum"\) and \(aut.anywhere all "hamlet shakespeare"\)](#)

D'autres exemples de requêtes sont fournis plus loin dans le document "**Manuel et exemples de requêtes**".

## 4. Réponse SRU et différents formats proposés

### 1) La réponse au SearchRetrieve

La réponse à une recherche par SearchRetrieve est renvoyée par le serveur SRU sous forme XML [conforme à la norme SRU 1.2](#).

Pour la requête suivante :

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author all "victor hugo"&recordSchema=dublincore](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author%20all%20victor%20hugo&recordSchema=dublincore),

la réponse comporte (voir le XML ci-dessous):

- la version du SRU utilisé ('1.2')
- le rappel de la requête exécutée ('*author all "victor hugo"*)
- le nombre total de résultats ('7761')
- puis l'ensemble des notices encapsulées dans l'élément <srw:records> avec :
  - Pour chaque notice (élément <srw:record>):
    - Les métadonnées globales : vocabulaire utilisé (ici 'dc') et format d'encapsulation ('xml')
    - Le contenu proprement dit de la notice (élément <srw:recordData>)
    - L'identifiant ark de cette notice (élément <srw:recordIdentifier>)
    - La position de cette notice dans la liste résultat (élément <srw:recordPosition>)
    - La date de création et de dernière modification de la notice dans le catalogue général de la BnF (éléments <ixm:attr name="CreationDate"> et <ixm:attr name="LastValidationDate"> dans <srw:extraRecordData>)

```
<?xml version="1.0" encoding="UTF-8"?>
<srw:searchRetrieveResponse xmlns:srw="http://www.loc.gov/zing/srw/" ...>
<srw:version>1.2</srw:version>
<srw:echoedSearchRetrieveRequest>
  <srw:version>1.2</srw:version>
  <srw:query>bib.author all "victor hugo"</srw:query>
</srw:echoedSearchRetrieveRequest>
<srw:numberOfRecords>7761</srw:numberOfRecords>
<srw:records>
  <srw:record>
    <srw:recordSchema>dc</srw:recordSchema>
    <srw:recordPacking>xml</srw:recordPacking>
    <srw:recordData>
      ...
    </srw:recordData>
    <srw:recordIdentifier>ark:/12148/cb371988266</srw:recordIdentifier>
```

<sup>1</sup> Les trigrammes et codes associés à ces valeurs sont donnés dans le tableau des critères en Annexe 2.

```

<srw:recordPosition>1</srw:recordPosition>
<srw:extraRecordData>
  <ixm:attr name="CreationDate">20001201</ixm:attr>
  <ixm:attr name="LastValidationDate">20001211</ixm:attr>
  <mn:score>16.686481</mn:score>
</srw:extraRecordData>
</srw:record>
<srw:record>...</srw:record>
<srw:record>...</srw:record>
</srw:records>
<srw:nextRecordPosition>4</srw:nextRecordPosition>
</srw:searchRetrieveResponse>

```

## 2) Formats de récupération des notices bibliographiques et d'autorité

Les formats proposés en sortie pour les notices sont l'[Unimarc](#) (avec ou sans analytique<sup>2</sup>) et l'[Intermarc](#) (avec ou sans analytique) pour les notices bibliographiques et d'autorité :

- ***unimarcXchange***
- ***unimarcXchange-anl***
- ***intermarcXchange***
- ***intermarcXchange-anl***

et le Dublin Core uniquement pour les notices bibliographiques :

- ***dublincore***

Pour choisir son format, il suffit de rajouter dans la requête le paramètre recordSchema et de préciser le format choisi comme ici "&recordSchema=dublincore" :

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author all "victor hugo" &recordSchema=dublincore](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author%20all%20victor%20hugo)

ou pour l'intermarcXchange :

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author all "victor hugo"&recordSchema=intermarcXchange](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.author%20all%20victor%20hugo)

Voir d'autres exemples sur « Choisir son format » plus loin dans le document.

- Le format Dublin Core employé est le Dublin Core simple (ou « non qualifié »), le même que celui des entrepôts OAI de la BnF.
- Pour les formats Unimarc et Intermarc, la réponse SRU étant au format XML, les notices en Unimarc ou Intermarc sont encapsulées dans du [MarcXchange 2.0](#) (généralisation du [MarcXML](#)) qui permet de représenter sous forme XML l'ensemble des champs de ces formats. On retrouve ainsi l'ensemble des zones de ces formats mais sous forme XML :
  - le label de la notice (zone 000) devient l'élément "leader",
  - les zones de contrôle (zones 00X) sont décrites dans des éléments "control field"
  - et les zones de données sont décrites dans des éléments "data field".

<sup>2</sup> Une notice analytique est le contenu détaillé du document, sous la forme d'une sous-notice encapsulée dans la notice principale. C'est particulièrement fréquent pour les enregistrements sonores (CD, cassettes), etc.  
Exemple : <http://catalogue.bnf.fr/ark:/12148/cb38257526m>

- Attention, les notices de spectacle n'ont pas de conversion en format Unimarc. La notice XML (le record) renvoyée comporte alors le message "erreur de traitement". De même pour les notices d'autorité Dewey que la BnF ne peut légalement diffuser. En unimarcXchange, pour éviter ce type de notices, il suffit de rajouter dans la query, la clause *not (bib.doctype="v")* ou *not (aut.type="dew")*.

Pour plus de détails sur ces formats, consulter les Annexes 3 et 4 « SRU et le format Dublin Core » et « SRU et les formats MARC ».

**Par défaut, le format de sortie est le format l'unimarcXchange.**

**L'encodage est l'UTF-8**

## 5. Exemples de recherche

On trouvera un grand nombre d'exemples de requêtes dans le **document "Manuel et exemples de requêtes"**.

Ci-dessous juste quelques exemples :

1. Les notices bibliographiques comportant tous les mots "marcel" et "proust" en Auteur et "recherche", "temps" et "perdu" en Titre :  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.author all "marcel proust"\) and \(bib.title all "recherche temps perdu"\)](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.author all )
2. L'ouvrage ayant pour ISBN 978-2-211-21125-3 :  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.isbn adj "978-2-211-21125-3"](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.isbn adj )
3. La notice bibliographique d'identifiant pérenne "ark:/12148/cb44496975d" :  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.persistentid any "ark:/12148/cb44496975d"\)](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.persistentid any )
4. Les Personnes ou Collectivités (notices d'autorité de référence) ayant les mots "victor" et "hugo" dans leurs formes retenues ou rejetées:  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(aut.type any "pep org"\) and \(aut.accesspoint all "victor hugo"\)](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(aut.type any )
5. Les notices d'autorité ayant pour ISNI 0000000121200982, 0000000121434842 ou 0000000121261347:  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=aut.isni any "0000000121200982 0000000121434842 0000000121261347"](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=aut.isni any )

## 6. Parcourir l'ensemble des résultats d'une requête et choisir son format

### 1) Parcourir l'ensemble des résultats d'une requête

Parcours des résultats 100 par 100 :

- Les 100 premiers résultats (de 1 à 100):  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.anywhere all "victor hugo"\)&maximumRecords=100](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.anywhere%20all%20%22victor%20hugo%22)&maximumRecords=100)
- Les 100 résultats suivants (de 101 à 200):  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.anywhere all "victor hugo"\)&startRecord=101&maximumRecords=100](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.anywhere%20all%20%22victor%20hugo%22)&startRecord=101&maximumRecords=100)

### 2) Choisir son format

Par défaut, le format de sortie est l'*unimarcXchange* (Unimarc encapsulé dans du MarcXchange).

- Les Fleurs du Mal ("ark:/12148/cb44496975d")
  - o en Unimarc (encapsulé dans du MarcXchange) : "**unimarcXchange**".  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.persistentid any "ark:/12148/cb44496975d"\)](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.persistentid%20any%20%22ark%3A%2F12148%2Fcb44496975d%22))
  - o ou au format Dublin Core : "**dublincore**".  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.persistentid any "ark:/12148/cb44496975d"\)&recordSchema=dublincore](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.persistentid%20any%20%22ark%3A%2F12148%2Fcb44496975d%22)&recordSchema=dublincore)
- La liste des œuvres de Victor Hugo (numéro de notice=11907966) en InterMarc (encapsulé dans du MarcXchange) : "**intermarcXchange**".  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(aut.author2work any "11907966"\)&recordSchema=intermarcXchange](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(aut.author2work%20any%20%2211907966%22)&recordSchema=intermarcXchange)
- Les notices de spectacles<sup>3</sup> « publiées » en Italie en InterMarc (encapsulé dans du MarcXchange) : "**intermarcXchange**".  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.doctype any "v"\) and \(bib.publicationplace any "it"\)&recordSchema=intermarcXchange](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.doctype%20any%20%22v%22)%20and%20(bib.publicationplace%20any%20%22it%22)&recordSchema=intermarcXchange)
- La notice de Piano concertos 2 & 4 de Sergei Rachmaninov ("ark:/12148/cb38225709g") en Unimarc avec analytiques (encapsulé dans du MarcXchange) : "**unimarcXchange-anl**".  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.persistentid any "ark:/12148/cb38225709g"\)&recordSchema=unimarcXchange-anl](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.persistentid%20any%20%22ark%3A%2F12148%2Fcb38225709g%22)&recordSchema=unimarcXchange-anl)
- Les grands interprètes de Wagner ("ark:/12148/cb38184184d") en InterMarc avec analytiques (encapsulé dans du MarcXchange) : "**intermarcXchange-anl**". Les analytiques sont embarquées sous forme arborescente dans les zones 986.  
[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=\(bib.persistentid any "ark:/12148/cb38184184d"\)&recordSchema=intermarcXchange-anl](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=(bib.persistentid%20any%20%22ark%3A%2F12148%2Fcb38184184d%22)&recordSchema=intermarcXchange-anl)

<sup>3</sup> Pour rappel, les notices de spectacle n'ont pas de conversion Unimarc.



## 7. Annexes

### 1. Conditions de réutilisation des données de la BnF

Depuis le 1er janvier 2014, la BnF a placé ses métadonnées descriptives (données bibliographiques et d'autorité) sous la **Licence Ouverte** de l'État élaborée par la mission Etalab. L'utilisation de ces métadonnées est libre et gratuite sous réserve du maintien de la mention de leur source et de l'indication de leur date de récupération.

C'est dans ce cadre que le protocole SRU permet la récupération de l'ensemble des métadonnées exposées dans le Catalogue général de la BnF.

En savoir plus : <https://www.bnf.fr/fr/conditions-de-reutilisations-des-donnees-de-la-bnf>

### 2. Tableau des critères (index) de recherche

La liste des critères (ou index) de recherche est fournie dans le **document "Tableau des critères"**..

### 3. SRU et le format Dublin Core

Le format Dublin Core utilisé est le Dublin Core simple (ou « non qualifié »), le même que celui des entrepôts OAI de la BnF. Il n'est applicable qu'aux notices bibliographiques.

En savoir plus sur le Dublin Core et les entrepôts OAI de la BnF : <https://www.bnf.fr/fr/dublin-core> et <https://www.bnf.fr/fr/recuperer-des-notices-bibliographiques-en-dublin-core-oai-cat>.

Exemple : La notice (<ark:/12148/cb42000454s>) au format *dublincore* :

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.persistentid any "ark:/12148/cb42000454s"&recordSchema=dublincore](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.persistentid%20any%20%22ark%3A%2F12148%2Fcb42000454s%22%26recordSchema=dublincore)

Dans la réponse SRU, la notice au format Dublin Core est inscrite au sein de la balise *recordData* de la façon suivante :

```
<srw:record>
<srw:recordSchema>dc</srw:recordSchema>
<srw:recordPacking>xml</srw:recordPacking>
<srw:recordData>
  <oai_dc:dc
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
    <dc:identifiant>http://catalogue.bnf.fr/ark:/12148/cb42000454s</dc:identifiant>
    <dc:titre>La grotte Chauvet révélée / Guy Jouve</dc:titre>
    <dc:creator>Jouve, Guy (1945-....)</dc:creator>
    <dc:publisher>G. Jouve (Biviers)</dc:publisher>
    <dc:date>2009</dc:date>
    <dc:identifiant>ISBN 9782953021851</dc:identifiant>
    <dc:description>Code à barres commercial : EAN 9782953021851</dc:description>
    <dc:format>1 vol. (172 p.) ; 22 cm</dc:format>
    <dc:language>fre</dc:language>
    <dc:type xml:lang="fre">texte imprimé</dc:type>
    <dc:type xml:lang="eng">printed text</dc:type>
```

```

    <dc:type xml:lang="eng">text</dc:type>
    <dc:rights xml:lang="fre">Catalogue en ligne de la Bibliothèque nationale de
    France</dc:rights>
    <dc:rights xml:lang="eng">French National Library online Catalog</dc:rights>
  </oai_dc:dc>
</srw:recordData>
<srw:recordIdentifier>ark:/12148/cb42000454s</srw:recordIdentifier>
<srw:recordPosition>1</srw:recordPosition>
<srw:extraRecordData>
  <ixm:attr name="CreationDate">20090612</ixm:attr>
  <ixm:attr name="LastValidationDate">20091012</ixm:attr>
  <mn:score>2.1864283</mn:score>
</srw:extraRecordData>
</srw:record>

```

Exemple de la même notice ([ark:/12148/cb42000454s](http://catoai.bnf.fr/oai2/OAIHandler?verb=GetRecord&metadataPrefix=oai_dc&identifieur=oai:bnf.fr:gallica/ark:/12148/cb42000454s)) dans l'entrepôt OAI-CAT de la BnF :

[http://catoai.bnf.fr/oai2/OAIHandler?verb=GetRecord&metadataPrefix=oai\\_dc&identifieur=oai:bnf.fr:gallica/ark:/12148/cb42000454s](http://catoai.bnf.fr/oai2/OAIHandler?verb=GetRecord&metadataPrefix=oai_dc&identifieur=oai:bnf.fr:gallica/ark:/12148/cb42000454s)

Attention : toutes les notices du catalogue général ne sont pas présentes dans l'entrepôt OAI-CAT.

## 4. SRU et les formats MARC

### I) Structuration des éléments d'information : les formats MARC

Les notices des bibliothèques sont saisies dans des formats très structurés en zones et sous-zones : les formats **MARC**.

Plusieurs formats MARC coexistent, et la BnF fournit ses notices dans 2 de ces formats : Unimarc et Interarc. L'Interarc est utilisé essentiellement par la BnF, qui gère les évolutions de ce format. L'Unimarc est, lui, un format commun à de nombreuses bibliothèques dans le monde.

**Par exemple**, pour l'ouvrage d'Audrey Becker, « *Les relations diplomatiques romano-barbares en Occident au Ve siècle : acteurs, fonctions, modalités* », paru en 2013 ([ark:/12148/cb436440194](http://catoai.bnf.fr/oai2/OAIHandler?verb=GetRecord&metadataPrefix=oai_dc&identifieur=oai:bnf.fr:gallica/ark:/12148/cb436440194)) :

- **les informations sur le titre** sont stockées :
  - **en zone 245 en Interarc** : ([notice en Interarc](#))
    - sous-zone \$a : Mention principale du titre
    - sous-zone \$e : complément du titre
 La zone de titre en Interarc contient :
 

245 1. \$a Les |relations diplomatiques romano-barbares en Occident au Ve siècle \$d Texte imprimé \$e acteurs, fonctions, modalités \$f Audrey Becker
  - **en zone 200 en Unimarc** : ([notice en Unimarc](#))
    - Mêmes sous-zones (on retrouve le titre en 200\$a et \$e)
    - La zone de titre en Unimarc contient :
 

200 1. \$a Les relations diplomatiques romano-barbares en Occident au Ve siècle \$b Texte imprimé \$e acteurs, fonctions, modalités \$f Audrey Becker
- **l'auteur principal d'un ouvrage** est mentionné :
  - **en zone 100 en Interarc** (avec sous-zones distinctes pour indiquer son nom, son prénom et ses dates de naissance/décès)
  - **en zone 700 en Unimarc**

### II) Distinction Notices bibliographiques / Notices d'autorité – Documentation sur les formats

Selon qu'on est sur une notice bibliographique (décrivant un document) ou sur une notice d'autorité (décrivant un auteur, un sujet, un lieu -- tout ce qui n'est pas un document), la même zone ne signifiera

pas la même chose.

D'où le besoin de documenter :

- le format Intermarc
  - pour les notices bibliographiques
  - pour les notices d'autorité
  
- le format Unimarc
  - pour les notices bibliographiques
  - pour les notices d'autorité
  - Le format Unimarc de diffusion des métadonnées descriptives de la BnF

### III) Unimarc et Intermarc en MarcXchange

Pour les formats Unimarc et Intermarc, la réponse SRU à l'opération searchRetrieve étant au format XML, les notices en Unimarc ou Intermarc sont encapsulées dans du [MarcXchange 2.0](#) (généralisation du [MarcXML](#)) qui permet de représenter sous forme XML l'ensemble des champs de ces formats.

Le format MarcXchange est issu de la généralisation du format [MarcXML](#)<sup>4</sup> à l'ensemble des formats MARC, le MarcXML étant conçu initialement uniquement pour le format MARC 21. Le MarcXchange a donc pour objectif de fournir une alternative à la norme ISO 2709 *Information et documentation – Format pour l'échange d'information* afin de permettre l'échange de notices de n'importe quel format MARC dans un environnement XML. Le travail a été mené dans le cadre de l'ISO : la norme ISO 25577 *Information et documentation – MarcXchange* [publiée en 2008](#) et révisé en 2013 : <https://www.iso.org/fr/standard/62878.html>. L'agence de maintenance est la Library of Congress : <http://www.loc.gov/standards/iso25577>.

En MarcXchange :

a) chaque élément "record" encapsulant une notice comporte 3 attributs<sup>5</sup> :

1. id : donnant l'identifiant de la notice (l'ark de la notice),
2. format : indiquant le format MARC utilisé dans la notice (INTERMARC, UNIMARC),
3. type : précisant s'il s'agit d'une notice bibliographique ou d'autorité (Bibliographic, Authority).

b). chaque zone du format MARC devient un élément comportant en attributs l'étiquette et les indicateurs :

- le label de la notice (zone 000) devient l'élément "leader",
- les zones de contrôle (zones 00X) sont décrites dans les éléments "control field"
- et les zones de données sont décrites dans les éléments "data field".

Pour ces éléments "data field", chaque sous-zone devient un sous-élément avec le code de sous-zone en attribut, alors que pour les éléments "leader" et "control field" les données sont traitées comme des chaînes de caractères.

Ainsi, pour l'ouvrage d'identifiant 34345163 de Victor Hugo, la zone de Titre (en Unimarc) sera représentée par :

```
<mx:datafield ind1="1" ind2=" " tag="200">
  <mx:subfield code="a">Les Misérables</mx:subfield>
  <mx:subfield code="b">Texte imprimé</mx:subfield>
  <mx:subfield code="f">Victor Hugo</mx:subfield>
</mx:datafield>
```

<sup>4</sup> Un exemple en marcxml via le SRU de la LOC :

<http://lx2.loc.gov:210/lcdb?version=1.1&operation=searchRetrieve&query=dinosaur&maximumRecords=5&recordSchema=marcxml>

<sup>5</sup> Les attributs format et type, on était ajouté par rapport au MARCXML.

```
</mxc:datafield>
```

Exemple : La notice d'identifiant [34345163](https://catalogue.bnf.fr/ark:/12148/cb34345163n) au format Unimarc encapsulé dans le marcXchange (*unimarcXchange*) (format par défaut):

[http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.recordid any "34345163"](http://catalogue.bnf.fr/api/SRU?version=1.2&operation=searchRetrieve&query=bib.recordid%20any%2034345163)

Dans la réponse SRU, la notice en marcXchange est inscrite au sein de la balise *recordData* de la façon suivante :

```
<srw:record>
<srw:recordSchema>marcXchange</srw:recordSchema>
<srw:recordPacking>xml</srw:recordPacking>
<srw:recordData>
  <mxc:record ... id="ark:/12148/cb34345163n" format="UNIMARC" type="Bibliographic">
    <mxc:record format="UNIMARC" id="ark:/12148/cb34345163n" type="Bibliographic">
      <mxc:leader>      nam 22      3 450 </mxc:leader>
      <mxc:controlfield tag="001">FRBNF343451630000001</mxc:controlfield>
      <mxc:controlfield
tag="003">http://catalogue.bnf.fr/ark:/12148/cb34345163n</mxc:controlfield>
      <mxc:datafield ind1=" " ind2=" " tag="039">
        <mxc:subfield code="o">OPL</mxc:subfield>
        <mxc:subfield code="a">021204084</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1=" " ind2=" " tag="100">
        <mxc:subfield code="a">19960410d1863      m  y0frey50      ba</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1="0" ind2=" " tag="101">
        <mxc:subfield code="a">fre</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1=" " ind2=" " tag="102">
        <mxc:subfield code="a">FR</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1=" " ind2=" " tag="105">
        <mxc:subfield code="a">|||z 00|||</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1=" " ind2=" " tag="106">
        <mxc:subfield code="a">r</mxc:subfield>
      </mxc:datafield>
      <mxc:datafield ind1="1" ind2=" " tag="200">
        <mxc:subfield code="a">Les Misérables</mxc:subfield>
        <mxc:subfield code="b">Texte imprimé</mxc:subfield>
        <mxc:subfield code="f">Victor Hugo</mxc:subfield>
      </mxc:datafield>
      ....
      ....
      ....
    </mxc:record>
  </srw:recordData>
  <srw:recordIdentifier>ark:/12148/cb34345163n</srw:recordIdentifier>
  <srw:recordPosition>1</srw:recordPosition>
  <srw:extraRecordData>
    <ixm:attr name="CreationDate">19960410</ixm:attr>
    <ixm:attr name="LastValidationDate">20160128</ixm:attr>
    <mn:score>2.1864283</mn:score>
  </srw:extraRecordData>
</srw:record>
```

#### IV) Récupérer les éléments d'information avec XPath

Si vous voulez récupérer ces informations au moyen de chemins [XPath](#)<sup>6</sup> (que ce soit en XSL, PHP, Python ou autre), voici quelques exemples.

Attention à certains éléments d'information qui sont répétables, comme l'indexation Sujet.

Le contexte est la balise <record> qui encadre l'ensemble des informations.

---

<sup>6</sup> Syntaxe permettant d'exprimer un chemin dans un arbre XML, jusqu'à l'élément d'information voulu.



Type de notice	Élément d'information	Intermarc	Unimarc	Commentaires
<b>biblio-graphique</b>	Numéro de notice	@Numero	@Numero	Le numéro de notice (qu'on peut interroger dans le SRU avec le critère de recherche RecordId), est un attribut de la balise <record>
<b>biblio-graphique</b>	Titre Complément du titre	mxo:datafield[@tag='245']/mxo:subfield[@code='a'] mxo:datafield[@tag='245']/mxo:subfield[@code='e']	mxo:datafield[@tag='200']/mxo:subfield[@code='a'] mxo:datafield[@tag='200']/mxo:subfield[@code='e']	Intermarc : Zones 245\$a et 245\$e Unimarc : Zones 200\$a et 200\$e
<b>biblio-graphique</b>	Nom de famille de l'auteur principal Prénom Identifiant	mxo:datafield[@tag='100']/mxo:subfield[@code='a'] mxo:datafield[@tag='100']/mxo:subfield[@code='m'] mxo:datafield[@tag='100']/mxo:subfield[@code='3']	mxo:datafield[@tag='700']/mxo:subfield[@code='a'] mxo:datafield[@tag='700']/mxo:subfield[@code='b'] mxo:datafield[@tag='700']/mxo:subfield[@code='3']	Attention : ces zones servent à identifier des auteurs "Personnes physiques". Des auteurs "Organisation" (association, administration, etc.) sont en zone 110 (Intermarc) ou 710 (Unimarc) Le 100\$3 (Intermarc) ou 700\$3 (Unimarc) permet ensuite de relancer une recherche dans le SRU si on veut récupérer d'autres informations sur cet auteur
<b>biblio-graphique</b>	Nom, prénom et identifiant des co-auteurs	mxo:datafield[@tag='700']/mxo:subfield[@code='a'] mxo:datafield[@tag='700']/mxo:subfield[@code='m'] mxo:datafield[@tag='700']/mxo:subfield[@code='3']	mxo:datafield[@tag='701']/mxo:subfield[@code='a'] mxo:datafield[@tag='701']/mxo:subfield[@code='b'] mxo:datafield[@tag='701']/mxo:subfield[@code='3']	
<b>biblio-graphique</b>	ISBN	mxo:datafield[@tag='020']/mxo:subfield[@code='a']	mxo:datafield[@tag='010']/mxo:subfield[@code='a']	
<b>biblio-graphique</b>	EAN	mxo:datafield[@tag='038']/mxo:subfield[@code='a']	mxo:datafield[@tag='073']/mxo:subfield[@code='a']	

<b>bibliographique</b>	Indexation Sujet : libellé et précisions	<pre> mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='a'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='x'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='y'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='z'] </pre>	<pre> mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='a'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='x'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='y'] mxc:datafield[starts-with(@tag, '6')]/mxc:subfield[@code='z'] </pre>	<p>Plusieurs zones sont utilisées pour indexer un document :</p> <p><b>600</b> : indexation à un nom de personne (comme sujet du document)</p> <p><b>606</b> : indexation sujet utilisant le vocabulaire contrôlé Rameau (606\$a : terme principal, \$x = subdivision thématique -- le \$y et le \$z servent à donner des précisions géographiques et chronologies -- <a href="#">exemple</a>)</p> <p>La zone <b>607</b> sert à une indexation proprement géographique (<a href="#">même exemple</a>)</p> <p>Le chemin XPath donné en exemple permet de récupérer les zones dont le nom (@tag) commence par "6"</p>
<b>autorité</b>	Nom de personne	<pre> mxc:datafield[@tag='100']/mxc:subfield[@code='a'] (nom) mxc:datafield[@tag='100']/mxc:subfield[@code='m'] (prénom) mxc:datafield[@tag='100']/mxc:subfield[@code='d'] (dates) </pre>	<pre> mxc:datafield[@tag='100']/mxc:subfield[@code='a'] (nom) mxc:datafield[@tag='100']/mxc:subfield[@code='b'] (prénom) mxc:datafield[@tag='100']/mxc:subfield[@code='f'] (dates) </pre>	
<b>autorité</b>	Autres noms (pseudonymes, etc.)	<pre> mxc:datafield[@tag='400']/mxc:subfield[@code='a'] mxc:datafield[@tag='400']/mxc:subfield[@code='m'] </pre>	<pre> mxc:datafield[@tag='400']/mxc:subfield[@code='a'] mxc:datafield[@tag='400']/mxc:subfield[@code='b'] </pre>	Voir par exemple la notice de Boris Vian <a href="#">en Intermarc</a> ou <a href="#">en Unimarc</a>



autorité	Dates de naissance et de mort	<pre>substring(mxc:controlfield[@tag='008'], 29 , 4) (naissance) substring(mxc:controlfield[@tag='008'], 39 , 4) (mort)</pre>	<pre>substring(mxc:controlfield[@tag='103'], 2 , 4) (naissance) substring(mxc:controlfield[@tag='103'], 12 , 4) (mort)</pre>	<p>Les dates de naissance et de mort sont présentes en 2 endroits :</p> <p>en 100\$d (intermarc) ou 100\$f (unimarc), c'est du texte libre (on va donc y trouver "4 av. J.-C.")</p> <p>Si on cherche des dates sur 4 chiffres, il vaut mieux aller les chercher dans des zones strictement structurées : c'est le cas de la zone <b>008</b>, ou les caractères <b>29</b> à <b>32</b> (numérotation à partir de 1 en XPath) contiennent la date de naissance, et <b>39</b> à <b>42</b> la date de mort (même chose pour la zone 103 en Unimarc, positions différentes)</p>
autorité	Concept Rameau : forme retenue et formes rejetées	<pre>mxc:datafield[starts-with(@tag, '16')]/mxc:subfield[@code='a']  mxc:datafield[starts-with(@tag, '46')]/mxc:subfield[@code='a']</pre>	<pre>mxc:datafield[starts-with(@tag, '2')]/mxc:subfield[@code='a']  mxc:datafield[starts-with(@tag, '4')]/mxc:subfield[@code='a']</pre>	<p>Le vocabulaire Rameau est <b>composé de plusieurs sous-ensembles</b>.</p> <p>Selon le sous-ensemble auquel appartient le concept trouvé, le libellé sera dans une zone différente -- mais commençant toujours par "16" :</p> <ul style="list-style-type: none"> <li>• 166 pour les noms communs</li> <li>• 167 pour les noms géographiques</li> <li>• etc.</li> </ul> <p><a href="#">Liste complète</a></p> <ul style="list-style-type: none"> <li>•</li> </ul> <p>Les formes alternatives seront en 46X (le dernière caractère X étant le même que pour le libellé principal retenu)</p>