

- **Cours /élément de module : M 2.1.1 - Metadata et Description de documents**
- **Niveau: 1^{ère} année - Cycle Ingénieur des connaissances et des données**
- **Professeur : M. Amine SENNOUNI**
- **Année universitaire: 2020-2021**
- **Partie 2/2 du cours: La description des documents par le biais des métadonnées : normes et formats internationaux (documents physiques, pages web et documents numériques)**

AXE 1: Description des documents physiques (papier et
électroniques)
(Format MARC-UNIMARC)

Approche classique

ISBD (ER)

- Prévoit **deux modes de traitement** en fonction du type de ressource électronique (sur support / accès en ligne)
- toute ressource électronique en ligne est réputée « publiée » (Postulat)
- sont exclus : « jouets programmés, calculateurs et autres objets programmés »
- **notion de bibliothèque numérique *absente*** de l'ISBD(ER) : pas de distinction entre fonds numériques et reste des fonds, un seul catalogue.

Approche classique

Formats MARC

Nés du besoin d'un schéma **permettant à des bases de données bibliographiques de communiquer entre elles**, ce que ne permet pas l'ISBD (ER)

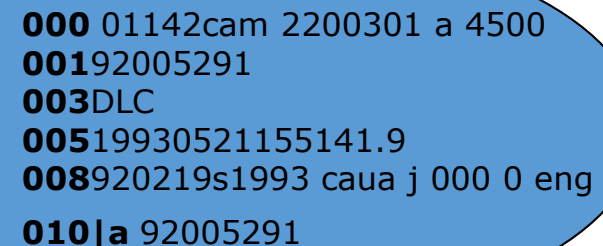
- Reposent tous sur la norme ISO 2709 "Format pour l'échange d'information"
 - norme adaptée à l'information bibliographique car elle **permet de gérer des zones** (et sous-zones)
 - de longueur variable
 - facultatives
 - répétables
 - définit la **structure** d'une notice
 - Label (zone fixe de 24 caractères)
 - Répertoire
 - Structure générique des zones

Approche classique

MARC 21

- Le format MARC 21 (Machine-Readable Cataloging ou Normes de Catalogage Lisibles par Machine) a été développé dans cette optique ; il met en forme les règles préconisées par l'ISBD de telle manière que les ordinateurs puissent les interpréter. Il fournit des champs pour stocker le titre de la ressource, l'individu ou l'organisme auteur, etc.
- Certaines données MARC 21 ne correspondent pas à une description bibliographique. Ces données permettent à différents systèmes bibliographiques d'échanger l'enregistrement.
- (Idem UNIMARC)

Bloc d'identification



```
000 01142cam 2200301 a 4500
00192005291
003DLC
00519930521155141.9
008920219s1993 caua j 000 0 eng
010|a 92005291
```

Approche classique

UNIMARC

- créé par l'**IFLA**
- dans le cadre du programme du **CBU** (Contrôle bibliographique universel)
- pour sortir de la « babélisation » des formats nationaux (**USMARC** : format national aux USA, **CAN/MARC** : format national au Canada, **MARC 21** : fusion d'USMARC et de CAN/MARC, reconnue par l'IFLA comme format d'échange, **INTERMARC** utilisé par la Bibliothèque nationale de France etc...)
- première édition : **1977**
- format « pivot » des **échanges** d'information bibliographique
- adopté ensuite comme format de **saisie** par des bibliothèques

Approche classique

UNIMARC

Spécificités pour les ressources électroniques :

- 215 : Description matérielle : champ **omis** pour les ressources en ligne, remplacé par 230
- 230 : Caractéristiques de la ressource électronique (**obligatoire** pour les ressources en ligne)
- 336 : Note sur le **type** de la ressource électronique
- 337 : Note sur la **configuration requise**
- 856 : Localisation électronique et accès (**URL**)

Qu'est ce que l'UNIMARC

- Définit une structure pour **l'échange sur support électronique, le traitement par logiciel** de données bibliographiques
- Les données elles-mêmes référentiels sont définies par des externes :
 - **règles de catalogage**
 - **systèmes d'indexation**

Naissance d'UNIMARC

- Nécessité de concevoir une interface pour **reformat**er les informations provenant de **source hétérogène**
- Problème d'incompatibilité – **Format universel d'échange** qui accepte les notices créées à partir de n'importe quel format MARC
- **Exportation:** Conversion des notices de son format local vers le format d'échange
- **Importation:** Conversion des notices reçues du format d'échange vers le format local
- **Risque** de perte de données

Objectifs d'UNIMARC

Caractéristiques:

- Respect de la norme ISO 2709
- Traite tous les médias
- Structure cohérente en blocs logiques
- Construit sur une architecture particulière qui regroupe:
 - les informations qui appartiennent à la description du document,
 - les informations qui relèvent de la responsabilité
 - les informations qui procèdent à l'analyse du sujet.

Format UNIMARC

Le **format** définit :

- les noms des champs (**étiquettes**), des sous-champs (**codes de sous-champs**), et leur correspondance avec les catégories définies par les normes internationales de catalogage
- la valeur des **indicateurs** de traitement
- les **règles de syntaxe** (répétabilité, caractère obligatoire ou non, ...)
- la structure et le **contenu** des champs de **données codées**

UNIMARC: structure

- Les informations constituent un **enregistrement** UNIMARC sont réparties en **champs** ou zones qui prennent place dans **10 blocs** (de 0 à 9) et sont précédées d'un **label de notice**.
- Chacun de ces 10 blocs peut comprendre **100 champs** (200 à 299 pour le bloc 2 par exemple).
- Tous les champs possibles ne sont pas utilisés. Seulement **167** sur les **1000** possibles sont définis

UNIMARC: structure

Une structure en blocs fonctionnels :

- 0. Bloc d 'identification (contient ISBD 8)
- 1. Bloc d 'information codée
- 2. Bloc d 'information descriptive (ISBD 1 à 6)
- 3. Bloc des notes (ISBD 7)
- 4. Bloc des notices bibliographiques liées (liens ou accès)
- 5. Bloc des variantes de titre (accès)
- 6. Bloc de l 'indexation par sujets (accès)

UNIMARC: structure

Une structure en blocs fonctionnels :

- 7. Bloc de la responsabilité intellectuelle (accès)
- 8. Bloc de données à usage international
- 9. Bloc de données à usage national

Un format très structuré, ce qui le rend très avantageux en tant que format de traitement

UNIMARC: structure

Les 10 blocs sont les suivants :

0XX Bloc des numéros d'identification

1XX Bloc des informations codées

2XX Bloc des informations

descriptives **3XX** Bloc des notes

4XX Bloc des liens avec d'autres notices

5XX Bloc des titres associés

6XX Bloc de l'indexation matière

7XX Bloc de la responsabilité

intellectuelle **8XX** Bloc des données

internationales **9XX** Bloc des données

locales

UNIMARC: Bloc/Zone

La zone est découpée en:

- Etiquette
- Indicateurs
- Sous-champs

Etiquette

Une **étiquette** de 3 chiffres qui identifie la zone qui lui est associée de 001 à 999

Le premier chiffre correspond au bloc auquel appartient le champ.

Exemple

s: 010

ISB

N

200 Titre et mention de responsabilité

225 Zone de la collection

300 Zone des notes

Indicateurs

- Deux **indicateurs de traitement** qui peuvent prendre l'une des valeurs : **#, 0, 1, 2, 3.**
- Le **#** est une valeur neutre, les autres valeurs sont significatives et permettent d'indiquer au système qu'il doit appliquer un traitement particulier au contenu du champ.

200 \$a Je craque..... \$b texte imprimé \$f Guy

- **#** Le premier indicateur du champ 200 prendra la valeur **1** si on souhaite que **le titre propre soit *indexé*** ou la valeur **0** si on ne le souhaite pas.

Sous-champs

Un **sous-champ** introduit par un ensemble de deux caractères:

Un **séparateur** représenté par **\$** suivi d'un **identificateur** représenté par une lettre ou un chiffre. Ces deux caractères forment **le code de sous-champ**.

200 1# **\$a** Je craque **\$b** texte imprimé **\$f** Guy
Bedos

010 ## **\$a** 2-7876-5433-9 **\$b** br. **\$d** 250 F

Exemple

┌ Étiquette ┌ Code de sous-champs ┐
200 1# Caligula \$f Albert Camus
 \$a Indicateurs └ Contenu du sous-champ ┘

200 1# \$a ≠NSB≠ Les ≠NSE≠ contes du chat perché \$f Marcel
Aymé

Exercices

200 1# \$a ≠NSB≠ Les ≠NSE≠ contes du chat perché \$f Marcel
Aymé

210 ## \$a Paris \$c Presses universitaires de France \$d

1998 010 ## \$a 2-7879- 7643-8 \$d 250 F

Exemple

- Le signe □ encadrant les abréviations **NSB** (non sorting character beginning) et **NSE** (non-sorting character end) signifie que le titre sera classé à C et non à L.
- Il sert donc à **éliminer les mots vides**

200 1# \$a #NSB# La #NSE# fontaine des gazelles

Format Unimarc:

2XX Bloc des informations
descriptives

Zone 1: Titre et mention de responsabilité

Données ISBD	Sous Zone UNIMARC
Titre propre	\$a
Autre titre propre du même auteur	\$a
Type de document	\$b
Titre propre d'un autre auteur	\$c
Titre parallèle	\$d
Sous titre - complément du titre	\$e
1 ^{ère} mention de responsabilité	\$f
MR suivante	\$g
N° de partie / Titre de partie	\$h / \$i
Numérotation du volume dans une zone imbriquée	\$v
Langue du titre parallèle	\$z

Zone 1: Titre et mention de responsabilité

1- Titre propre [type de document] / mention de responsabilité

La rose des Tudors [texte imprimé]/ **Dominique Fernandez.** – Ed. rev.
Et augm. – Paris : Julliard, 1989. – 139 p. couv. Ill. en coul. ; 20 cm.

200 1 # \$a #NSB # La #NSE# rose des Tudors **\$b** Texte imprimé
\$f Dominique Fernandez

200 1# \$a # NSB # La #NSE# rose des Tudors ----- Larose des
Tudors

200 1# \$a #NSB # L' #NSE# Allemagne ----- L' Allemagne

Zone 1: Titre et mention de responsabilité

1- Titre propre [type de document] / mention de responsabilité

La guirlande de Campra [Musique imprimée]/ par Arthur Honegger,
Daniel Lesur...[et al]

**200 1# \$a #NSB# La #NSE# Guirlande de Campra \$b Musique imprimé
\$f par Arthur Honegger, Daniel Lesur...[et al]**

Plusieurs noms dans une **même mention de responsabilité**, on ne répète pas
le
\$f

Zone 1: Titre et mention de responsabilité

1- Titre propre [type de document] / mention de responsabilité

Piano transcription [Enr.sonore]/ Beethoven, Liszt, comp. ; Glenn Gould, piano

200 1 # \$a Piano transcription **\$b** Enr.sonore **\$f** Beethoven, Liszt, comp **\$g** Glenn Gould, piano

Plusieurs noms dans une **même mention de responsabilité**, on ne répète pas le **\$f**. la mention de responsabilité suivante est introduite par **\$g**

Zone 1: Titre et mention de responsabilité

**3- Titre propre [type de document] = titre parallèle =
titre parallèle / mention de responsabilité**

Biblioteca [texte imprimé] = Libraries = Bibliothèques / Ernest Malaga

**200 1# \$a Biblioteca \$b Texte imprimé \$d = Libraries \$d = Bibliothèques
\$f Ernest Malaga \$z eng \$z fre**

Zone 1: Titre et mention de responsabilité

4- Titre complément du titre/ première mention
responsabilité; mention de responsabilité; troisième
mention de responsabilité e

Le trésor des Pharaons [texte imprimé]: la joaillerie égyptienne de la
période dynastique/ Cyril Aldred , trad. de l'anglais par Catherine David,
photogr.
D'Albert Shoucair. – Paris : Tallandier, cop. 1979. – 64 p.

200 1# \$a #NSB#Le #NSE# trésor des Pharaons \$b Texte imprimé \$e
la joaillerie égyptienne de la période dynastique \$f Cyril Aldred \$g
trad. de l'anglais par Catherine David \$g photogr. d'Albert Shoucair

Zone 1: Titre et mention de responsabilité

5- Titre propre; autre titre du même auteur; autre titre du même auteur / première mention de responsabilité; deuxième mention de responsabilité

Lettres adressées à des petites filles [texte imprimé] suivi de Alice racontée aux petits enfants, et de trois poèmes sans queue ni tête/ Lewis Carroll, texte français par Henri Parisot. – Paris : Flammarion, 1975. – 250p.

200 1# \$a Lettres adressées à des petites filles **\$b** Texte imprimé **\$a** **#NSB#** suivi de **#NSE#** Alice racontée aux petits enfants **\$a** **#NSB#** et de **#NSE#** Trois poèmes sans queue ni tête **\$f** Lewis Carroll **\$g** texte français par Henri Parisot

Exercice 1

La Fontaine [texte imprimé/ Roger Duchêne

Titre de couv. : « Jean de la Fontaine »

Corrigé 1

La Fontaine [texte imprimé/ Roger Duchêne

Titre de couv. : « Jean de la Fontaine »

200 1# \$a La Fontaine \$b texte imprimé \$fRoger Duchêne

On n'a pas enlevé l'article--Note

Zone2: Edition

- **Occurrence** FR
- **Valeur des indicateurs**
 - 1^{er} indicateur = # blanc
 - 2^{èm}_e indicateur = # blanc

Données ISBD	Sous Zone UNIMARC
Mention d'édition	\$a
Mention parallèle d'édition	\$d
Mention de responsabilité relative à l'édition	\$f
Mention de responsabilité suivante	\$g
Autre mention d'édition	\$b

Zone2: Edition

1- Mention d'édition

La rose des Tudors [texte imprimé]/ Dominique Fernandez. – **Ed. rev et augm.** – Paris : Julliard, 1989

205 ## \$a Ed. rev et augm.

205 ## \$a 3° éd. entièrement refondue

205 ## \$a [Nouv. Éd.]

Adresse bibliographique

- **Zone 210**
- **Occurrence F NR**
- **Valeur des indicateurs**
 - 1^{er} indicateur = # blanc
 - 2^{èm}
e indicateur = # blanc

Zone 4: Adresse

Données ISBD	Sous Zone UNIMARC
Lieu de publication	\$a
Lieu suivant de publication	\$a
Adresse de l'éditeur	\$b
Nom de l'éditeur	\$c
Date de publication	\$d
Lieu de fabrication	\$e
Lieu suivant de fabrication	\$e
Adresse du fabricant	\$f
Nom du fabricant	\$g
Date de fabrication	\$h

Zone 4: Adresse

5- Lieu de publication: nom de l'éditeur, date (Dpt-Lieu d'impression: nom de l'imprimeur)

Paris: éd. PUF; Paris, 1976et Montrond (25- Alençon : Impr.
Jacqueset Montrond

210 ## \$a Paris **\$c** PUF **\$d** 1987 **\$e** 25- Alençon **\$g**Impr. Jacques
et Montrond

Zone 5 de la collation 215

- **Zone 215**
- **Occurrence F R**
- **Valeur des indicateurs**
 - 1^{er} indicateur = # blanc
 - 2^{èm}
e indicateur = #
blanc

Zone de la Collation

Données ISBD	Sous Zone UNIMARC
Indication du type de document et importance matérielle	\$a
Autres caractéristiques matérielles	\$c
Format	\$d
Matériel d'accompagnement	\$e

Zone de la Collation

1- Importance matérielle : mention d'illustration ; format

350 p. : ill., couv. Ill. en coul. ; 25 cm

215 ## \$a 350 p. **\$c** ill., couv. Ill. en coul. **\$d** 25 cm

2- Type de présentation (Importance matérielle): mention d'illustration ; format + matériel d'accompagnement

3 vol. (230, 120, 350 p.) : ill., cartes ; 21 cm.+ 6 transparents

215 ## \$a 3 vol. (230, 120, 350 p.) **\$c** ill., cartes **\$d** 21 cm **\$e** 6 transparents

Zone 6 de la collection et de la monographie en plusieurs volumes = 225

- **Zone 225**
- **Occurrence F R**
- **Valeur des indicateurs**
 - 1^{er} indicateur:** forme du titre en relation avec le titre saisi comme point d'accès dans la zone **410** et en **461**
 - 0= différent de la forme saisie en zone de liens
 - 1= pas de forme d'autorité établie en zone de liens
 - 2= semblable à la forme d'autorité saisie en zone de liens
 - 2^{ème} indicateur:** # blanc

Zone 6 de la collection et de la monographie en plusieurs volumes = 215

Données ISBD	Sous Zone UNIMARC
Titre de la collection ou de la monographie en plusieurs volumes	\$a
Titre parallèle	\$d
Complément du titre	\$e
Mention de responsabilité	\$f
Numéro de la sous collection	\$h
Titre de la sous collection	\$i
N° du volume à l'intérieur de la collection ou de la sous collection	\$v
ISSN de la collection	\$x

Zone de la Collection

1- (Titre propre de la collection, ISSN; numérotation dans la collection)

(Les cahiers pédagogiques, ISSN 0108-141X ; 3)

**225 2# \$a#NSB# Les #NSE# cahiers pédagogiques
\$x0108- 141X \$v 3**

L'indexation des données se fait à partir des zones indexables de la notice de collection à laquelle est liée la notice de la monographie que l'on catalogue

HTML : Evolution

- HTML 1.0
 - Texte de base, images, liens hypertextes
- HTML 2.0
 - Formulaires de saisie
- HTML 3.2
 - Justification (gauche, centre, droite)
 - Tableaux
 - Équations mathématiques
- HTML 4.0
 - Feuilles de style

Figures
Imagemap
...