

Règles pour l'indexation « sujet »

Plateforme Archives Documentation Numérisation (ADN), Maison des Sciences de l'Homme de Dijon, USR uB-CNRS3516. Auteur : Marie Royer, sous la direction d'Agnès Viola

Introduction

L'indexation est une méthode utilisée afin de retrouver efficacement un document. Pour ce faire, il est nécessaire de procéder au traitement de l'information textuelle contenue dans le document puis de structurer ces données dans des champs dédiés et de mettre ces informations à disposition sur un portail permettant une recherche.

En fonction de la nature du projet et des pratiques institutionnelles et/ou disciplinaires, voire des préconisations des tutelles, la nature de l'indexation peut être libre ou contrôlée. Dans le premier cas, cela relève du langage dit naturel, alors que dans le deuxième il existe des règles à suivre voire même des listes de vocabulaires contrôlés ou de thésaurus à l'exemple de RAMEAU ou Pactols.

Ce guide spécifie la façon dont doivent être saisies les valeurs de l'élément <subject> des fichiers EAD produits au sein de la plateforme ADN et qui se retrouvent en consultation sur le portail PANDOR dans la catégorie « mot(s) clé(s) » de l'indexation. Pour les besoins d'un projet, certains index sujet se voient appliquer un attribut et une liste d'autorité spécifique. Par exemple, pour la ressource Bulletin de l'OIV, les noms des vins ont fait l'objet d'un index basé sur une liste d'indication géographique protégée (IGP), fournie par l'Organisation Internationale du Vin.

Les règles éditées ci-après ne s'appliquent pas à ces cas particuliers.

Dans un contexte universitaire, en matière d'indexation sujet, l'usage du Répertoire d'autorité-matière encyclopédique et alphabétique unifié (RAMEAU) est privilégié, notamment pour faciliter les échanges avec des plateformes nationales comme ISIDORE. Le guide d'indexation RAMEAU est librement consultable en ligne à l'adresse suivante : <http://guiderameau.bnf.fr/>

En raison de la spécificité des fonds traités par la plateforme ADN, il arrive que des termes proposés par RAMEAU ne correspondent pas à la réalité de ce que l'on souhaite indexer.

Ex : « Libertaires » traduit dans RAMEAU par « Anarchistes ».

Dans les cas où aucune vedette matière ne répond au besoin d'indexation, il est employé l'indexation libre.

L'utilisation faite de RAMEAU au sein de la plateforme ADN ne prévoit pas l'usage des subdivisions géographiques et formes.

En effet, les corpus sont décrits en EAD, qui destine des balises spécifiques à cet usage à savoir :

<geogname> pour les termes géographiques

<genreform> pour les types de documents

Les aspects relatifs à l'indexation des types de documents et des termes géographiques ne sont pas traités ici.

Pour pallier les difficultés liées à l'emploi du langage naturel, d'autant plus par plusieurs personnes distinctes, les règles suivantes s'appliquent :

Règles d'indexation sujet libre

Les mots sujet

- Un descripteur ne doit représenter qu'un seul concept. Il faut songer à tous les pièges signalés qui peuvent se présenter pour la clarté de cette définition. Il ne faut pas retenir des mots qui peuvent avoir plusieurs sens.
Il faut réserver les précisions entre parenthèses à des cas où aucun autre terme ne peut être utilisé de façon satisfaisante
- Lors de l'indexation, toujours mettre en lettre capitale la première lettre du mot indexé.
Cas où la capitale s'impose :
 - Noms propres
 - Pour des institutions : État, Académie, Univers, Sénat, Cour des comptes, Université (mais pas l'université de Picardie), etc.
 - les points cardinaux : Nord, Est, Sud, Ouest
- Ne pas terminer le(s) mot(s) par des signes de ponctuation
- Indiquer les accents
- Les sigles se trouvent entre parenthèses après leur signification :
 - Institut universitaire de technologie (IUT)
- Ne pas utiliser d'abréviations
Ex : Philo → Philosophie
- Eviter d'utiliser les pronoms
- Privilégier le masculin pluriel sauf dans le cas où un mot n'a pas le même sens au singulier et au pluriel
Ex :

heure (temps)/heures (livre liturgique),
mémoire (travail de recherche)/mémoires (souvenirs rédigés)

- Ne pas mettre de concepts au pluriel :
 - Art, Cinéma, Jeunesse etc.
- Ne pas utiliser de verbe
- Un descripteur peut-être uniterme ou composé
Ex : Travailleur de nuit
- Ne pas prendre en compte (sauf cas exceptionnel) :
 - la forme des documents (bibliographie, interview, rapport, norme) qui font l'objet d'un index spécifique <genreform> en EAD
 - leur genre (conte, roman d'aventures...)
 - leur support physique (diapositive, disque...)
 - leur niveau d'utilisation (CP, collège...)
 - leur domaine d'utilisation (histoire, géologie...)
- Utiliser les formes françaises quand elles existent
Ex : Londres et non London
- Quand plusieurs orthographes sont possibles, utiliser la plus généralement admise, si connue
Ex : clé plutôt que clef
- Ecrire les chiffres ordinaux en toutes lettres
Ex : Premier Empire et non 1er, Troisième République et non 3°
- On peut privilégier un sens par rapport à un autre quand l'un a un synonyme et l'autre pas
Ex : « coutumes » a un sens précis dans le droit médiéval et il n'y a pas de terme équivalent ; par ailleurs, ce mot signifie aussi « traditions » ; On peut accepter son utilisation dans son sens juridique et l'interdire dans son autre sens parce que, dans ce cas, il peut être remplacé par « traditions ».
- Dans le choix entre les synonymes, il convient de privilégier le terme savant plus précis que le terme usuel ; il recouvre d'ailleurs généralement plusieurs "variantes" et son utilisation réduit ainsi notablement le nombre des descripteurs
Ex : employer « ovin » plutôt que « mouton » (ovin définit autant le mouton que la brebis, l'agneau ou le bélier alors que le mouton n'est qu'un ovin parmi d'autres), « cours d'eau » plutôt que « rivière » (dans les cours d'eau, il y a les rivières, les fleuves, les ruisseaux, les torrents... alors que la rivière n'est qu'une sorte de cours d'eau).

Les noms de personnes

Les noms de personnes sont indexés sous la forme Nom, Prénom

Ex :

Pagnol, Marcel

Chateaubriand, François-René de

Cas particulier des noms à particule

Le rejet ou non des particules des patronymes diffère selon des usages nationaux.

Pour plus de précisions, se reporter au guide du catalogueur de la BNF consultable à cette adresse : http://guideducatalogueur.bnf.fr/ABN/GPC.nsf/gpc_homepage?openform

A noter :

Cet index est à distinguer de l'auteur du document décrit qui fait l'objet d'un élément et d'un attribut spécifique dans les instruments de recherche encodés en EAD au sein de la plateforme ADN. Par exemple : <persname role="auteur">Duby, Georges</persname>