Module 6 - section 7 : Indexation

Claire SIBILLE - DE GRIMOÜARD - Marcel CAYA

Version du 29/09/09 27/12/2024

Table des matières

Objectifs 4

1. Introduction 6

1.1. Qu'est ce que l'indexation ? 6
1.1.1. L'indexation en archivistique 7
1.1.2. Un outil pour certains instruments de recherche 7

1.2. Indexation automatique et indexation manuelle 7

1.3. Le processus d'indexation 8

2. Choisir le vocabulaire et la méthode d'indexation 9

2.1. Principes 9
2.1.1. Diminution des formes existantes 10
2.1.2. Réduction des ambiguïtés du vocabulaire 10
2.1.3. Sélection du vocabulaire 11

2.2. L'indexation libre 13

2.3. Renvois et tableau méthodique 13
2.3.1. Les renvois 13
2.3.2. Le tableau méthodique 14

2.4. L'indexation contrôlée : thésaurus et listes d'autorité 16
2.4.1. Le thésaurus 17
2.4.2. Les listes d'autorité 19

3. Choisir le niveau d'indexation 22

3.1. Les noms communs 22
3.1.1. Un principe 23
3.1.2. Un problème de cohérence 24

3.2. Les noms propres 25
3.2.1. Les noms géographiques 25
3.2.2. Les noms de personne 27
3.2.3. Les autres noms propres 28

4. Mettre en forme matériellement un fichier 30

4.1. La rédaction d'une fiche manuelle 30

4.2. Le rangement des fiches 31

4.3. L'intégration dans une base de données 32

5. Conclusion 34

6. Évaluation des connaissances 36

Exercice : 6.1. Exercice 1 36

Exercice : 6.2. Exercice 2 36

Mentions légales **37**

Objectifs

Description du module :

Nous avons tous vu des archives définitives avant même de savoir ce qu'elles sont. L'image de l'album de photographies anciennes ou de vieux contrats d'hypothèque aux pages vieillies et fragiles vient immédiatement à l'esprit. Parce que nous avons tous visité un jour des bureaux de l'administration, nous sommes aussi familiers avec les grandes quantités de documents que les administrations modernes produisent et accumulent. Pourtant, au sein du grand public, peu connaissent la distinction entre les archives courantes et les archives définitives.

Les archivistes d'aujourd'hui travaillent à mieux faire connaître du grand public l'importance de bien gérer les archives courantes, ces archives d'aujourd'hui, pour mieux protéger éventuellement celles dont on aura besoin demain. Une fois désignées comme définitives, ces archives constituent souvent la seule trace mémorielle des faits, qu'ils soient simples reconnaissances de droits (ou de devoirs) ou le souvenir d'événements et de contextes particuliers.

Parce qu'elles ne représentent qu'une fraction de l'ensemble des documents qui ont existé, ces archives définitives deviennent le lien privilégié entre ce qui était et la représentation que nous voulons conserver des personnes, des organismes et des administrations. C'est pourquoi, afin d'en préserver l'authenticité, la signification et l'intégrité, elles doivent être traitées et conservées selon un ensemble de règles et de normes par des spécialistes qui savent comment assurer leur préservation à long terme tout en permettant leur repérage et leur utilisation dans les meilleures conditions possibles.

Le but du module est de :

- vous apprendre à reconnaître les archives définitives et l'ensemble des problèmes et des solutions liées aux traitements nécessaires à leur utilisation
- vous initier aux procédures d'évaluation, de collecte et de tri
- vous faire connaître les principaux instruments de recherche dans les archives
- vous familiariser, dans le respect des normes internationales, avec les pratiques de classement et de description des archives définitives
- vous initier plus particulièrement aux normes internationales de description (ISAD-G) et aux notices d'autorité (ISAAR) en archivistique pour les noms propres de personnes, collectivités, lieux et famille
- vous initier à l'indexation des descriptions des archives définitives
- vous faire connaître les normes de préparation d'instruments de recherche électronique en vue d'une diffusion sur Internet

L'apprenant doit être en mesure de :

- comprendre les particularités du traitement des archives définitives
- identifier les problèmes particuliers posés par la collecte et le tri des archives définitives
- classer et décrire un fonds d'archives (et ses parties) selon les normes internationales
- préparer des instruments de recherche sur papier
- comprendre, en attendant de pouvoir le réaliser, les exigences de leur diffusion sur formats électroniques

Positionnement :

L’étudiant devrait avoir déjà vu les cinq modules précédents, particulièrement les modules 4 et 5 sur la gestion des archives courantes et intermédiaires.

Conseils d'apprentissage :

Ce module est assez long. Nous conseillons de lire d'abord les pages principales, puis de revenir sur les encarts si l'on veut davantage d'informations.

La section 8 sur la publication électronique des instruments de recherche s’adresse plus particulièrement à des archivistes qui ont déjà acquis de l’expérience ; elle est cependant maintenue dans cette formation qui se veut de base afin d’avoir une information d’ensemble sur les archives définitives.

1. Introduction

Introduction

La préparation d'*instruments de recherche* sur papier crée souvent de grandes quantités de notices de *description* qu'il demeure toujours difficile d'utiliser avec rapidité et précision.

L'élaboration d'*index* permet de donner au chercheur des moyens supplémentaires de trouver les *documents* qui l'intéresse.

Même lorsque les instruments de recherche sont disponibles en version numérique, l'*indexation* permet d'accroître l'efficacité de la recherche en fournissant des clés de repérage supplémentaires.

Afin d'atteindre ses objectifs, cependant, la préparation d'index doit être faite selon les règles de l'art afin :

- de s'assurer que toutes les références pertinentes apparaissent
- et d'éviter qu'un trop grand nombre de références soient générées inutilement.

1.1. Qu'est ce que l'indexation ?

Introduction

L'Association française de normalisation (AFNOR) définit l'indexation comme :

« l'opération qui consiste à décrire et à caractériser un document à l'aide de représentations des concepts évoqués dans ce document, c'est-à-dire à transcrire en langage documentaire les concepts après les avoir extraits du document par une analyse » (1993).

Dans sa forme la plus simple, **l'index est une liste alphabétique** des **noms de personnes**, des **noms géographiques** et des **noms de matières** contenus dans un document assortie des **références correspondantes** destinées à les localiser.

L'index **ne doit pas être confondu** avec :

- le thésaurus, qui est un *« langage établi, a priori commun, normalisé, structuré, qui traduit en termes d'indexation ou de recherche tout concept devant entrer ou sortir d'un système documentaire donné »* (AFNOR, NZ 47-100) ,
- la liste d'autorité, qui est la liste alphabétique (sans hiérarchie) des termes que l'on peut utiliser pour l'indexation.

Pour plus de détails...
Pour en savoir plus : les normes ISO sur l'indexation et les thesaurus monolingues et multilingues ISO 999:1996 (novembre 1996) Information et documentation. Principes directeurs pour l'élaboration, la structure et la présentation des index ISO 2788 (1986) Principes directeurs pour l'établissement et le développement de thésaurus monolingues

Pour plus de détails...
ISO 5964 (1985) Principes directeurs pour l'établissement et le développement de thésaurus multilingues Un projet de révision des normes internationales sur les thésaurus monolingues et multilingues aborde la question de l'interopérabilité entre langages, intègre le multilinguisme en tant que variante linguistique d'un concept.

1.1.1. L'indexation en archivistique

L'indexation en archivistique porte :

- non pas sur le contenu des documents
- mais sur les notices descriptives.

« C'est une technique documentaire à deux degrés : en effet, seule l'analyse est faite à partir des documents eux-mêmes ; elle sert ensuite de support à la sélection des termes d'indexation. C'est pourquoi, quel que soit le niveau de finesse retenu (groupe de dossiers, dossier ou pièce), l'indexation repose sur une analyse intelligible, fiable et complète dont elle est le complément. »

(circulaire AD 94-8 du septembre 1994. Traitement des archives contemporaines. Indexation. Instruments de recherche).

Les termes d'indexation n'ont donc pas vocation à exprimer toutes les nuances contenues dans les analyses, ils n'ont qu'un rôle d'indicateur et sont obligatoirement réducteurs. Même si lors de la procédure d'indexation, ils sont directement associés à une cote ou à un groupe de *cotes*, ils renvoient d'abord à l'analyse telle qu'elle figure sur l'instrument de recherche et à laquelle le chercheur doit d'abord se référer pour valider sa recherche, préalablement à toute demande de communication.

1.1.2. Un outil pour certains instruments de recherche

Sans se substituer à l'analyse, l'indexation constitue donc un mode d'accès complémentaire au mode de classement choisi pour l'instrument de recherche.

Son intérêt est cependant lié à la présentation d'informations détaillées à indexer.

- Ainsi, l'indexation n'est pas pertinente pour les répertoires décrivant des documents sériels ou non nominatifs.

  Un index est inutile, par exemple, pour un répertoire des registres de l'enregistrement ou de l'état civil qui ne donne que le types d'actes et les dates extrêmes de chaque registre.
- De même, un index alphabétique ne s'impose pas pour un guide dans la mesure où les mots cités ne représentent qu'une faible partie du contenu des *fonds* eux-mêmes.

1.2. Indexation automatique et indexation manuelle

Les possibilités offertes par les technologies informatiques (prise en compte du texte intégral et du langage naturel, traitements lexico-sémantiques, etc.) peuvent être utilisées lorsque l'on souhaite obtenir des résultats rapides à coût réduit, si l'on veut reprendre par exemple de gros volumes de données. Le recours à ces méthodes ne constitue pas encore la recette miracle.

L'indexation "manuelle" à partir d'un vocabulaire structuré reste encore une bonne façon de guider l'utilisateur vers les analyses d'archives susceptibles de l'intéresser, en décrivant les documents par des termes normalisés non ambigus et indépendants du vocabulaire libre employé dans les analyses archivistiques.

L'indexation automatisée peut venir en complément d'une indexation structurée insuffisante, voire inexistante (cas des inventaires numérisés par reconnaissance optique de caractères).

Le plus important est finalement d'évaluer la rentabilité d'une indexation manuelle.

1.3. Le processus d'indexation

L'indexation se fait généralement à un seul niveau (pièce, dossier ou série organique) pour que le chercheur sache précisément sur quel niveau de détail il peut compter.

Exemple

S'il appelle un nom de commune, aura-t-il vraiment toutes les informations sur cette commune ?

Le processus d'indexation comprend deux étapes principales :

- à partir des notices descriptives, on identifie et on **sélectionne les concepts** qui seront représentés dans l'index, en accordant une attention particulière à l'intitulé, qui est la partie la plus significative de l'analyse ;
- la deuxième étape est celle de la **traduction en langage d'indexation**.

Un symbole, un terme ou une expression est alors choisi pour représenter de façon claire et concise chacun des concepts sélectionnés à l'étape de l'analyse.

Les termes d'indexation peuvent être contrôlés, se présentant dans une forme plus ou moins éloignée du langage naturel.

2. Choisir le vocabulaire et la méthode d'indexation

Introduction

La décision d'élaborer un index paraît simple en soi.

Les deux problèmes majeurs rencontrés dans le travail d'indexation sont :

- le choix du vocabulaire
- et le choix du niveau d'indexation.

Au-delà de la solution de ces problèmes, vous aurez aussi à choisir :

- entre l'indexation libre ou contrôlée,
- entre le recours à des renvois ou à un tableau méthodique.

Vous apporterez une réponse différente à toutes ces possibilités :

- selon le type d'instrument de recherche que vous avez à indexer,
- selon les types de recherche pour lesquels il sera utilisé et, évidemment,
- selon les moyens qui sont mis à votre disposition.

La **normalisation du vocabulaire** est indispensable à la cohérence de l'indexation. Cet effort de normalisation peut être guidé par des référentiels existants, dont beaucoup sont issus de l'activité d'autres secteurs patrimoniaux que les archives.

En France, les professionnels des bibliothèques notamment ont largement contribué à la production de normes AFNOR en travaillant sur le formalisme des noms de lieux géographiques, des noms de personnes et des noms de collectivités. Dans la plupart des pays, il existe des organismes qui officialisent les noms des lieux géographiques et les noms des collectivités. Pour les noms de personnes, il est important de suivre les données du fichier d'autorité.

Il faut également adapter le niveau de précision du vocabulaire d'indexation au niveau de description de l'analyse :

- éviter l'indexation détaillée au niveau d'un fonds ou d'un sous-fonds
- et, inversement, éviter d'indexer de façon trop générale des unités de description dont l'analyse est très précise.

Ainsi, les principes de normalisation et les outils dont on peut disposer permettent de surmonter les difficultés que pose le choix du vocabulaire, à condition bien sûr de faire preuve d'une grande rigueur et d'une grande discipline.

2.1. Principes

Introduction

On ne peut pas utiliser comme vocabulaire d'indexation les mots mêmes qu'on trouve dans un texte : le langage naturel a trop de formes et d'ambiguïtés pour pouvoir servir à une interrogation.

C'est pourquoi il faut le codifier :

- en éliminant tout d'abord les formes diverses d'un même terme,
- en sélectionnant ensuite les termes eux-mêmes.

2.1.1. Diminution des formes existantes

Quand on codifie, on doit respecter certaines règles :

- utiliser toujours un **substantif** : recrutement au lieu de recruter, donation au lieu de donner ;
- utiliser la forme française, y compris dans les noms propres quand une forme française existe : Londres et pas London, Venise et pas Venezia ;
- mettre les termes au singulier sauf dans deux cas :
  - quand le singulier n'est pas usité : travaux publics, archives…
  - quand le sens n'est pas le même au singulier et au pluriel : heure (temps)/heures (livre liturgique), mémoire (travail de recherche)/mémoires (souvenirs rédigés) ;
- quand plusieurs orthographes sont possibles, utiliser la plus généralement admise : clé plutôt que clef ;
- éviter la forme abrégée : "fac" pour faculté ;
- écrire les chiffres ordinaux en toutes lettres : "Premier Empire" et non 1er, "Troisième République" et non 3° ;
- développer les sigles sauf ceux qui sont universellement connus ; il faut penser qu'un sigle très courant dans notre profession ou dans les administrations avec lesquelles nous traitons n'est probablement pas connu du grand public : Unesco est connu de tous, mais «Conseil international des archives» devrait être préféré à ICA ou à CIA ;
- n'utiliser qu'une seule forme pour chaque nom propre.

Il s'agit d'harmoniser éventuellement l'orthographe mais surtout d'harmoniser la place des différents éléments du nom propre.

Exemple

Choisir "Chateaubriand, François-René de" et éliminer (Chateaubriand (F.-R. de), Chateaubriand (de), Chateaubriand (vicomte de), Chateaubriand (vicomte François-René de), Chateaubriand (de, vicomte) et autres... Ou choisir Labastide-Rouairoux et éliminer Bastide-Rouairoux (La).

2.1.2. Réduction des ambiguïtés du vocabulaire

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 :

Exemple

« source » signifiant cours d'eau naissant et base de recherche sera remplacé par « eau » ou « cours d'eau » et par « recherche historique » selon le sens.

- Il faut réserver les précisions entre parenthèses à des cas où aucun autre terme ne peut être utilisé de façon satisfaisante :

Exemple

ainsi kiwi (fruit) / kiwi (animal)

- On peut privilégier un sens par rapport à un autre quand l'un a un synonyme et l'autre pas :

Exemple

« 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 ».

- Il ne faut pas utiliser de terme dont la définition n'a jamais été précise ou a varié dans le temps :

Exemple

« province »

Vienne (département) / Vienne (Isère) / Vienne (Autriche)

ou Henri IV, roi de France / Henri IV, empereur germanique / Henri IV, roi d'Angleterre.

- Il ne faut pas utiliser de mots qui ne diffèrent l'un de l'autre que par un ou des accents ou n'en autoriser qu'un des deux : cela pose souvent problème dans un système informatique car certaines tables ne prennent pas les accents en compte :

Exemple

Dans l'exemple « péché » : « pêche », on gardera pêche (aux poissons) parce que c'est un concept pour lequel on ne peut gère trouver de synonyme, tandis que la notion de péché peut être intégrée dans un descripteur religieux (« pratique religieuse » par exemple) ; la pêche fruit, elle, peut être indexée à fruit ou arboriculture.

2.1.3. Sélection du vocabulaire

Un concept ne doit être désigné que par un seul descripteur, nous l'avons déjà dit. Une langue est riche d'un immense vocabulaire et d'une infinité de nuances. Mais la prise en compte de cette richesse n'est pas compatible avec un système de description normalisé susceptible de répondre rapidement et précisément à toute demande, émanant de personnes très différentes dans leur culture, dans leur domaine de compétence, dans leur attente.

Il faut donc **réduire les termes d'indexation à un certain nombre de concepts**, suffisants pour pouvoir répondre à toutes les catégories de recherche.

- Il ne faut retenir qu'un mot ou une expression par concept, en rejetant donc tous les synonymes au sens très large

Exemple

on ne retiendra qu'un seul terme pour indexer le concept de climat (le thésaurus français a retenu "météorologie").

- Dans le choix entre les synonymes, il faut 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 :

Exemple

“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).

- Il faut éviter des expressions qui précisent les concepts mais au prix d’un accroissement des descripteurs.

La plupart de ces précisions peuvent être obtenues :

- en croisant des descripteurs (dans une base de données)
- ou en utilisant des descripteurs secondaires (sur une fiche manuelle),

Exemple

éviter de créer “histoire de France” (pourquoi pas celle de tous les pays ou de toutes les villes...) ; lors de l'interrogation d'un système informatisé, on croisera “histoire” et “France” ; dans un fichier manuel, on utilisera aussi les deux descripteurs, le croisement des deux étant indiqué dans le corps de la fiche ;

Dans certains cas, par contre, l’utilisation d’une expression s’impose :

- parce que le résultat du démembrement de l’expression serait incompréhensible : le croisement de “sécurité” et de “société” ne peut définir le concept de “sécurité sociale”
- parce que le croisement de deux descripteurs peut donner des résultats très différents :

Exemple

“philosophie” et “histoire” peuvent s'associer en “philosophie de l'histoire” ou en “histoire de la philosophie” ; il vaut donc mieux créer ces deux descripteurs.

Pour plus de détails... **Complément**

HUDON, Michèle. « Indexation et langages documentaires dans les milieux archivistiques à l'ère des nouvelles technologies de l'information »[Hudon]

ou, si vous pouvez vous le procurer : *HUDON, Michèle « Le Thésaurus : conception, élaboration, gestion »*.[Hudon]

2.2. L'indexation libre

Dans l'indexation libre, l'indexeur extrait les descripteurs des analyses ou les choisit librement sans vérification.

C'est le système le plus facile et rapide en amont, mais le résultat présente de très grandes difficultés à l'interrogation :

- atomisation du vocabulaire,
- plusieurs descripteurs pour un même concept,
- plusieurs concepts pour un même descripteur…

Attention

Dans un système informatique, toute forme devient un descripteur distinct. Toute faute de frappe, toute utilisation d'une forme proscrite, toute utilisation non codifiée de tirets, de différenciation ou non de l'article (La Bastide ou Labastide) et de position des différents éléments peut entraîner la création d'au moins huit descripteurs différents.

2.3. Renvois et tableau méthodique

Introduction

Les renvois ont été les premiers outils utilisés pour guider les chercheurs dans un fichier et aussi pour pallier certaines des imperfections de l'indexation libre.

Le tableau méthodique est venu par la suite organiser le vocabulaire au niveau au moins de l'interrogation, constituant le premier pas vers une indexation contrôlée.

2.3.1. Les renvois

Lorsqu'on se trouve face à un fichier déjà constitué selon le "principe" de l'indexation libre, on peut pallier certains défauts par des renvois, notamment dans le cas où on a utilisé plusieurs termes pour le même concept.

Exemple

ovin, voir aussi mouton

mouton, voir aussi ovin

Cela permet de rationaliser le fichier, de le rendre scientifiquement utilisable, mais cela ne constitue qu'une solution d'attente avant refonte du fichier.

Les renvois permettent de :

- relier entre elles des formes très différentes d'un même concept.

Exemple

Léningrad, voir Saint Pétersbourg

Beyle, Henri, voir Stendhal

ONU, voir Organisation des Nations Unies

- traduire des hiérarchies de noms communs

Exemple

production agricole

voir aussi : céréales, viticulture, maïs, colza, horticulture

- traduire des hiérarchies de noms de lieux

Exemple

Espagne

voir aussi : Andalousie, Aragon, Castille, Catalogne

voir aussi : Barcelone, Bilbao, Madrid, Salamanque, Valence

voir aussi : frontière franco-espagnole

- faire des renvois systématiques entre les communes et certaines circonscriptions locales.

Exemple

- Château-Verdun (baronnie de)

  voir aussi : Albiès, Aston, Aulos, Bouan, Cabannes (Les), Château-Verdun, Larcat, Pech, Sinsat
- Albiès

  voir aussi : Château-Verdun (baronnie de)

- indiquer les relations de continuité historique

Exemple

La commune de Lasserre a été distraite de celle de Tourtouse en 1888.

Il faut faire une fiche : Lasserre

Avant 1888 : voir Tourtouse

2.3.2. Le tableau méthodique

Tous les renvois donnés ci-dessus sont annulés par l'emploi d'un tableau méthodique (du moins sous forme d'instrument de recherche, pas en fichier).

Un tableau méthodique des mots-matières se réalise a posteriori, à partir d'un index ou d'un fichier existant. Il n'y a donc pas de tableau-type.

Il faut relever dans un fichier tous les mots-matières utilisés et les organiser de façon thématique, les thèmes découlant forcément du contenu des archives couvertes par l'index ou le fichier. Ceci afin d'offrir au chercheur la liste complète des termes d'indexation par domaine ; il connaîtra ainsi tous les mots utilisés dans le domaine de l'agriculture, de la vie politique, de la famille… et pourra utiliser le fichier de façon exhaustive et rapide.

Le tableau méthodique doit comprendre tous les mots-matières utilisés et seulement eux. Ces mots sont répartis au sein de rubriques thématiques et, à l'intérieur de ces rubriques, ils sont classés dans l'ordre alphabétique. Il n'y a pas de lien hiérarchique comme dans un thésaurus.

Un tableau méthodique peut être réalisé à partir de l'index d'un instrument de recherche papier et publié avec lui. Cela évitera à l'utilisateur de parcourir l'index tout entier. Son élaboration permet aussi de voir les éventuels défauts de l'index (utilisation de synonymes notamment) et de les corriger avant publication.

Tableau méthodique des mots-matières *Exemple*

administration	administration générale	conseillers d'État commissaires de la République médiateur de la République ministres parlementaires police préfets présidents de la République
	finances publiques	budget domaine enregistrement finances inspecteurs des finances
	fiscalité	douanes réforme fiscale
agriculture et pêche	économie rurale	comice agricole haras
	forêt	bois eaux et forêts
	production agricole	agriculture chanvre eaux de vie lin pêche plantes sucres vin
communications	messagerie	aéropostale postes

	transports	aéronautique aviation chemins de fer navires
économie	commerce	commerce économie marine marchande

Le tableau méthodique peut aussi s'appliquer aux noms géographiques : on regroupe dans un tableau les toponymes présents dans l'index ou le fichier par département, par pays...

Tableau méthodique par noms de lieux ? *Exemple*

AFRIQUE

Afrique du Sud

Maroc

EUROPE

Croatie

- Zagreb

France

- Ain
  - Bourg-en-Bresse
- Allier
  - Vichy
- Alsace
- Bouches-du-Rhône
  - Aix-en-Provence,
  - Marseille
- Ile-de-France
- Roussillon

2.4. L'indexation contrôlée : thésaurus et listes d'autorité

Introduction

Les renvois, les tableaux méthodiques ont été les outils élaborés au fur et à mesure que la pratique de l'indexation prenait de l'ampleur. Ils sont réalisés *a posteriori*, ils organisent un vocabulaire déjà créé.

L'indexation contrôlée est plus complexe. Elle est l'aboutissement de la réflexion sur l'indexation, d'une longue pratique et de la prise en compte des diverses possibilités, mais aussi des contraintes offertes par les systèmes informatiques.

Dans le cadre de l'indexation contrôlée, le vocabulaire d'indexation est déjà défini a priori et son utilisation exclusive s'impose à l'indexeur.

Ce vocabulaire peut se présenter sous deux formes :

- la liste d'autorités, liste alphabétique de descripteurs et de non-descripteurs ;
- le thésaurus, liste structurée de descripteurs et de non-descripteurs reliés entre eux par des relations de hiérarchie, d'association ou d'équivalence.

ATTENTION !
On appelle descripteur un terme d'indexation autorisé et non-descripteur un terme interdit pour lequel on renvoie à un terme autorisé. **EXEMPLE :** ferme EM EXPLOITATION AGRICOLE (EM = employer. "Ferme" est un non-descripteur et on renvoie sur le descripteur "exploitation agricole").

2.4.1. Le thésaurus

Le thésaurus est évidemment l'outil le plus complexe à élaborer. Il s'applique généralement à un domaine précis.

Exemple

Le thésaurus des Archives de France s'applique à tous les fonds d'archives locales, publiques et privées, anciennes, modernes et contemporaines.

Il doit contenir tous les termes nécessaires à l'indexation dans le domaine qui est le sien. Il doit être complet dès sa création.

Il doit comporter un certain nombre de non-descripteurs, des termes qui ont été écartés mais qui, parce qu'il représentent des notions susceptibles de faire l'objet régulier d'interrogations, doivent être renvoyés sur le descripteur autorisé. C'est ce qu'on appelle la relation d'équivalence.

Exemple

Dans le thésaurus français, on a écarté "marée noire" au profit de "pollution de la mer", plus général, mais "marée noire" a été reconnue comme notion assez importante pour justifier qu'elle soit mentionnée, même si c'est pour la renvoyer sur "pollution de la mer".

Tous les termes d'un thésaurus doivent être mis en relation hiérarchique, c'est-à-dire organisés en chapitre, rubriques et sous-rubriques. Cela est très important pour obtenir une réponse complète. C'est là que réside la grande richesse du thésaurus.

Exemple

D'après le thésaurus W des Archives de France :

ÉQUIPEMENT	niveau hiérarchique n° 1
ENVIRONNEMENT	niveau hiérarchique n° 2
• PROTECTION DE LA NATURE	niveau hiérarchique n° 3

• ASSOCIATION DE DÉFENSE DE L'ENVIRONNEMENT • BIODIVERSITÉ • ESPÈCE PROTÉGÉE • FAUNE SAUVAGE • FLORE SAUVAGE • PARC NATUREL • RÉSERVE NATURELLE • SITE	niveau hiérarchique n° 4

Interroger à "site" donne des réponses. Mais interroger à "protection de la nature" donne aussi des réponses, de niveau plus général, sur les sites ; interroger "environnement" donne encore des réponses de niveau supérieur sur les sites. **Si on veut faire une recherche vraiment complète sur les sites, il faut interroger toute la hiérarchie.**

En sens inverse, la hiérarchie indique qu'une recherche sur la protection de la nature doit intégrer aussi des descripteurs de niveau inférieur.

Si le fichier est manuel, le lecteur ayant le thesaurus à sa disposition pourra de lui-même faire toutes les interrogations utiles.

Dans une base de données ayant intégré le thesaurus, s'il fait une interrogation à "site", le système demandera s'il veut aussi les termes en hiérarchie ; s'il fait une interrogation à "protection de la nature", les réponses seront les cotes indexées à "protection de la nature" mais aussi toutes celles indexées aux termes placés hiérarchiquement en dessous, dont "site".

Les relations hiérarchiques et les relations d'équivalence ne peuvent être établies qu'une seule fois pour un terme.

Un système informatique n'accepte pas qu'un terme soit présent dans plusieurs chaînes hiérarchiques, ni qu'un non-descripteur renvoie sur plusieurs descripteurs. On compense cette contrainte par des relations d'association.

Exemple

Dans le thésaurus des Archives de France, la plupart des descripteurs se rapportant aux établissements d'enseignement se trouvent dans une hiérarchie « Education » (chapitre 5).

Mais les universités populaires, elles, se trouvent dans une hiérarchie « Temps libre et sociabilité » (chapitre 11).

Puisqu'il n'est pas possible de faire dépendre « université populaire » des deux hiérarchies, on fait une association.

« Université populaire » est citée comme terme associé dans la hiérarchie « Enseignement » et « Université » est un terme associé à « Université populaire » dans la hiérarchie « Éducation ».

Pour plus de détails... ***Complément***

Exemple du Thésaurus des Archives de France (cf. Thésaurus des Archives de France)

Le thésaurus sert ainsi essentiellement pour les noms communs qui qualifient l'objet des documents, concepts qu'on peut ranger méthodiquement et hiérarchiser.

Pour plus de détails... *Complément*

Pour en savoir plus : un exemple de thésaurus comme outil d'aide à la recherche : le site des Archives départementales de la Vendée[1].

Les Archives départementales de la Vendée[2] ont élaboré en interne des référentiels destinés à guider l'internaute dans ses recherches, dans le cadre du développement, par un prestataire, d'une nouvelle interface de consultation et d'interrogation des descriptions archivistiques.

Le thésaurus Matières a été constitué a posteriori, à partir des termes figurant dans les instruments de recherche. L'architecture du thésaurus est proche du système Dewey, mais ce sont les renvois et les synonymies qui importent le plus au lecteur. Sur le site des Archives départementales, l'utilisateur peut donc sélectionner des termes à partir de plusieurs listes hiérarchisées : périodes historiques (certaines périodes sont découpées en séries d'années, par exemple, la Révolution), noms de lieux, typologie documentaire (forme littéraire, type de pièce, type de publication) et mots matières. Il est également possible d'interroger dans les mots de l'analyse. Toutefois, il ne peut y avoir de recherche qu'au même niveau, car il n'y a pas d'héritage de l'indexation.

2.4.2. Les listes d'autorité

Pour les noms propres, pour des actions et pour la typologie documentaire, on utilise aussi des **listes d'autorité** :

- il s'agit toujours d'un **vocabulaire contrôlé**,
  - avec descripteurs autorisés
  - et non-descripteurs interdits
- mais présentés dans un **ordre alphabétique, sans hiérarchie**.

Attention

En France, des listes « Actions », « Typologie documentaire » et « Contexte historique » constituent des annexes du Thésaurus pour la description et l'indexation des archives locales anciennes, modernes et contemporaines. Ce sont des listes fermées, les termes autorisés ayant été définis a priori. Les descripteurs figurant dans ces « Actions », « Typologie documentaire » et « Contexte historique » ne sont pas par eux-mêmes des termes d'indexation mais ils peuvent être associés à un ou plusieurs descripteurs du thésaurus-matières si le contexte documentaire l'exige.

Pour plus de détails... *Complément*

Pour en savoir plus : exemple des Archives nationales d'outre-mer, Aix-en-Provence (France)

En France, dans le cadre de l'intégration des nouveaux instruments de recherche dans l'application IREL[3] (Instruments de Recherche En Ligne), une réflexion est en cours aux Archives nationales d'outre-mer sur l'élaboration de listes d'autorité pour les noms de personnes, les noms d'institutions, les noms de lieux et les mots matière, et sur l'établissement de notices d'autorité conformes à ISAAR(CPF) pour chaque personne ou institution citée dans les différents instruments de recherche. Il s'agit en effet de dégager une logique d'accès commune aux différents instruments de recherche pour guider l'internaute le plus efficacement dans sa recherche. Ce développement d'IREL se traduira dans la conception des formulaires d'interrogation qui seront offerts à ce dernier.

1. http://archives.vendee.fr/
2. http://archives.vendee.fr/
3. http://www.archivesnationales.culture.gouv.fr/caom/fr/index.html

Les règles d'écriture fixées par la norme AFNOR NF Z 47-200 (Liste d'autorités de matières, structure et règles d'emploi) pour les noms de personnes et les noms géographiques sont tout à fait valables pour les descriptions archivistiques ; elles peuvent aussi servir de guide dans d'autres pays.

Exemple

Pour les noms de communes et de lieux-dits, on se reportera en France aux publications de l'INSEE.

Pour les autres noms géographiques, on utilisera la forme la plus courante, en français si elle existe, et dans sa forme contemporaine.

Le plus important est de n'en utiliser qu'une, sauf exception due à l'histoire qu'il faut bien souligner par des renvois.

Il est impératif de ne pas avoir de descripteur homographe ; quand le cas se présente, il faut différencier par une précision entre parenthèse : par exemple, Villaret (commune de Lézat)/ Villaret (commune de Montbel).

Il faut également homogénéiser l'ordre des éléments d'un nom composé : ne pas mettre Mont Blanc et Cenis (Mont), ou Col du Tournalet et Puymorens (col du).

Pour le contexte historique, les prescriptions sont les mêmes. Il n'y a pas d'intitulé officiel des événements historiques. Il appartient à qui élabore une liste d'autorités de faire en sorte que

- chaque événement n'ait qu'un seul descripteur
- que ces descripteurs soient rédigés de façon homogène
- et qu'il n'y ait pas d'homographe.

Exemple

Il faut choisir entre Révolution française et Révolution de 1789, entre Première Guerre mondiale et Guerre de 1914-1918 ou Guerre 1914-1918 ; il ne faut pas choisir Guerre 1914-1918 d'une part et Seconde Guerre mondiale d'autre part ; il faut différencier des guerres, des batailles, des traités... qui portent le même nom par un élément ajouté, en général la date : par exemple Guerre d'Espagne (1808-1814) / Guerre d'Espagne (1936-1939).

D'une manière générale d'ailleurs, concernant un événement historique ou une période, il est utile d'inclure la(les) date(s) dans l'intitulé du descripteur.

Exemple

Traité des Pyrénées (1699), Directoire (1795-1799).

Les règles relatives aux noms d'organismes ou de personnes morales (que les normes bibliographiques appellent "collectivités") sont à rechercher dans la norme NF Z 44-060 Catalogage d'auteurs et d'anonymes. forme et structure des vedettes de collectivités-auteurs.

Pour les noms de personnes physiques, on peut se référer à la norme AFNOR NF Z 44-061 Forme et structure des vedettes noms de personne, des vedettes titres, des rubriques de classement et des titres forgés.

Outre les fichiers d'autorité de la Bibliothèque nationale de France, déjà cités, on se reportera, en France, à certaines des normes AFNOR (norme sur les noms de lieux, norme sur l'image fixe) qui contiennent également des listes d'autorité très utiles. Dans les autres pays, on trouvera une source qui puisse faire autorité.

Complément

Pour en savoir plus : les normes de l'Association française de normalisation (AFNOR) (cf. Normes de l'AFNOR)

Complément

Pour en savoir plus : l'indexation des noms de lieux, la norme AFNOR NF Z44-081 (cf. Indexation AFNOR)

3. Choisir le niveau d'indexation

Introduction

Le choix du niveau d'indexation, lui, reste un problème délicat. C'est pourquoi l'indexation devrait se faire toujours au même niveau de détail sinon, la réponse donnée n'est pas cohérente.

EXEMPLE :
Dans des dossiers d'élections politiques, on peut mentionner et indexer le nom des candidats ; si on ne le fait pas systématiquement, la réponse donnée est partielle, donc erronée : le chercheur pensera qu'il n'a à sa disposition qu'un dossier sur Monsieur X, alors qu'il y en a en fait plusieurs.

Trois situations se présentent :

- **L'index d'un *inventaire* ou d'un *répertoire***

  Il ne pose généralement que peu de problèmes, puisqu'il porte sur un ensemble limité, lui-même traité de façon homogène, qu'il est réalisé dans un temps limité et souvent par une même personne. Une harmonisation s'impose facilement.

- **Le fichier d'un fonds ouvert réalisé à partir d'instruments de recherche de même nature**

  Le fait que l'instrument de recherche couvre des documents d'archives qui présentent une certaine homogénéité demeure un atout. Par contre, l'ampleur du volume traité, l'étalement du travail dans le temps, l'intervention de plusieurs personnes rendent l'harmonisation plus aléatoire.

- **Un fichier général de service d'archives**

  Ici, les problèmes sont plus importants parce qu'un tel fichier est issu de l'indexation d'instruments de recherche de niveaux très divers, allant de l'inventaire sommaire au répertoire numérique, voire à l'état des fonds, et que les informations disponibles sont évidement de niveau très différents.

Les solutions sont différentes selon qu'il s'agit :

- de noms communs
- ou de noms propres.

3.1. Les noms communs

Introduction

Les noms communs touchent un large éventail de matières à divers niveaux.

Une simple accumulation de tous les noms désignant des objets produira souvent beaucoup de références inutiles sans assurance que toutes celles qui sont pertinentes seront incluses.

C'est pourquoi il convient de suivre un certain nombre de règles qui varient selon les types d'instruments de recherche que l'on veut indexer.

3.1.1. Un principe

Il faut prendre pour principe d'**indexer systématiquement tous les objets, principaux ou secondaires, mentionnés dans les unités archivistiques décrites** : si on a distingué un objet lors de l'analyse, c'est qu'il est digne d'intérêt. Cela satisfait à la fois au principe même de l'indexation qui est d'extraire toutes les notions susceptibles d'être interrogées (ce qui correspond bien aux objets des dossiers) et au souci de s'en tenir à un même niveau de détail.

On peut réduire l'indexation grâce à l'existence d'un thésaurus hiérarchisé.

Exemple

On peut indexer des dossiers spécifiques "boucherie", "boulangerie", "débit de boisson".

Ailleurs, un seul dossier concerne tous les commerces alimentaires : on peut l'indexer à "commerce alimentaire", même si dans l'analyse on a énuméré lesdits commerces.

Sans vocabulaire hiérarchisé, cela peut poser problème : le lecteur qui fait une recherche sur les débits de boisson n'ira pas chercher forcément à "commerce alimentaire" et n'aura donc pas toutes les références. Avec un vocabulaire hiérarchisé, ce problème disparaît. Mais cette solution n'est valable que si les descripteurs sont de niveau très proche (comme "commerce alimentaire" et "débit de boisson" qui sont de niveau 3 et 4 dans le thésaurus français.

Lorsque les objets sont multiples et variés, trouver un descripteur qui serve de facteur commun amène à un concept beaucoup trop général.

Exemple

Les dossiers d'un cabinet de préfet traitent de toutes les affaires intéressant le département (en France circonscription géographique et administrative) ; le seul facteur commun qu'on peut trouver est "administration générale", ce qui n'enrichit guère le fichier et revient quasiment à priver d'accès les archives du cabinet du préfet.

Dans ce cas, il faut indexer à tous les objets signalés dans les analyses (d'élection à inondation, de drogue à immigration, de fête à course cycliste…).

On peut indexer par groupe d'articles de même analyse.

Exemple

Un versement de matrices cadastrales ou de recensements de population peut comprendre 400 articles dont les analyses ne diffèrent que par le nom de la commune.

L'indexation matière ne se fera qu'une fois : les cotes extrêmes (345 W 1-500) représentent la référence, que ce soit sur une fiche manuelle ou dans un écran de saisie d'une base de données.

Cela allège l'indexation mais aussi les réponses (il n'y aura qu'une seule réponse pour les 500 dossiers).

Par contre, l'indexation noms de lieux se fera plutôt au niveau du dossier.

On peut faire une indexation par groupes d'articles même quand chaque article a son analyse bien individualisée. C'est le cas quand les objets de dossiers se répètent régulièrement.

Exemple

Prenons le cas de dossiers issus du contrôle de l'administration communale par la Préfecture, particulièrement du contrôle du budget, donc des investissements.

Nous avons 50 articles classés dans l'ordre alphabétique des communes, dans lesquels reviennent régulièrement, presque pour chaque commune, des sous-dossiers

- sur la voirie communale,
- sur les écoles,
- la mairie,
- la collecte et le traitement des ordures ménagères,
- l'assainissement de l'eau…

On peut indexer ces concepts une seule fois pour les 50 articles, en donnant toujours pour références les cote extrêmes de l'ensemble.

On indexera par contre au dossier le nom des communes et aussi des matières qui n'apparaissent qu'une ou quelques fois.

Exemple

VOIRIE COMMUNALE 345 W 35-85

travaux

contrôle budgétaire

1965-1972

(une seule fiche matière)

et

ALBIES 345 W 37

voirie communale

travaux, contrôle budgétaire

1967-1968

(une fiche pour chaque commune)

2.2. 3.1.2. Un problème de cohérence

Lorsque l'on constitue le fichier général d'un service, notamment sous forme d'une base de données informatique, on doit faire face à un gros problème de cohérence parce que les données proviennent des instruments de recherche existants dans le service et que, dans tout service et surtout dans des services anciens, ayant connu différentes pratiques archivistiques, il existe des instruments de recherche réalisés à des niveaux de description différents.

Pour qu'un tel fichier, une telle base de données soit scientifiquement utilisable, il faut essayer d'harmoniser le niveau de réponse ; il faut donc

- limiter – sans perdre de substance – l'indexation issue d'une description à la pièce,
- compenser la faiblesse de l'information fournie par un répertoire issu d'une description succincte au dossier.

Pour plus de détails... *Complément*

Harmoniser, au sein d'un fichier général ou d'une base de données, une indexation issue de descriptions de niveaux différents. (cf. indexation)

3.2. Les noms propres

Introduction

Parce qu'ils reflètent aussi bien le contenu que le contexte des documents, les noms propres utilisés dans les analyses sont des clés essentielles pour l'identification des documents.

Ils ne peuvent cependant fournir des accès utiles pour la recherche qu'à condition que l'indexation suive des règles rigoureuses qui en garantissent la précision et la pertinence.

Les noms de personnes, de familles, de personnes morales et de lieux géographiques sont l'objet de normes bien établies qui nous guident dans l'établissement d'un fichier d'autorité, un outil essentiel pour bien contrôler toutes les variations possible de noms propres.

Pour plus de détails...
On consultera • la norme ISAAR (CPF) (2ème version) à l'adresse http://www.ica.org/sites/default/files/ISAAR2FR.pdf • les Règles canadiennes pour la description des documents d'archives canadiennes à l'adresse http://www.cdncouncilarchives.ca/f-archdesrules.html.

3.2.1. Les noms géographiques

Introduction

Les noms géographiques constituent un accès très utile au chercheur.

Ils doivent cependant être énoncés de façon à la fois assez précise et suffisamment exhaustive pour fournir des repères efficaces.

Aux problèmes de la modification des noms de lieux à travers les âges s'ajoutent plusieurs autres, notamment ceux de l'utilisation du même vocable pour des lieux différents, ceux des niveaux (localité, région etc) et la nécessité de présenter ces noms avec une indication de l'aire géographique ou du pays visé.

3.2.1.1. À quel niveau indexer ?

Les noms géographiques sont innombrables et il faut se donner une règle.

Cette règle varie forcément avec la compétence et les attributions du service ou du centre d'archives.

- Dans des archives communales française, par exemple, on indexe généralement au nom des quartiers, voire au nom des rues. C'est un niveau pertinent. On ne peut pas se contenter du nom de la commune : tous les documents seraient indexés à ce nom.
- Dans des archives départementales ou régionales, le niveau pertinent est celui de la commune. On n'indexe pas à un lieu-dit, sauf si ce lieu-dit est particulièrement important localement : un site touristique ou archéologique, une station balnéaire ou de sports d'hiver, le lieu d'implantation d'un organisme particulier (un établissement religieux, scolaire, un complexe industriel…). Ces lieux-dits doivent faire obligatoirement l'objet d'un renvoi vers la commune dont ils dépendent – et vice versa – puisque c'est la commune qui est prise comme niveau ordinaire d'indexation.
- Dans des Archives nationales, on alterne entre circonscriptions locales d'un certain niveau (les départements en France) et communes d'une certaine importance, selon la nature des dossiers.

Partout, on prendra en compte, **au sens très local** du terme, les noms de "régions" ou "provinces", les noms de "pays" les cours d'eau ou les plans d'eau (lacs, étangs…), les montagnes et autres éléments importants du paysage.

Ces règles permettent normalement de prendre en compte tous les noms géographiques contenus dans les analyses d'un répertoire numérique détaillé, puisque ce sont les mêmes qui président à la localisation de l'objet des dossiers lors de la rédaction des analyses.

Attention

Elles permettent par contre de réduire dans l'indexation le nombre des toponymes contenus dans un inventaire où souvent, notamment dans le cas d'archives anciennes, les lieux-dits sont cités.

Ces lieux-dits peuvent bien sûr figurer dans l'index propre de l'inventaire, puisqu'ils existent dans les analyses, mais on ne les fait pas figurer dans le fichier général d'un service d'archives ; on indexe leur occurrence au nom de la commune.

3.2.1.2. Des noms à ne pas prendre en compte

Des noms de lieu "balises"

Pour respecter l'homogénéité de l'indexation, il y a un erreur à ne pas commettre : c'est de prendre en compte des noms propres qui ne sont cités que comme "balises", comme limites, ou comme localisation d'une institution.

Ils ne sont pas utilisés pour eux-mêmes, ils "encadrent" d'autres lieux qui, eux, ne sont pas cités et leur prise en compte aboutirait à une indexation très incomplète, donc induirait en erreur.

Exemple

Dans l'analyse suivante :

F/12/356 à F/12/372 Ministère de l'Industrie. – Usines, autorisations de construction (classement alphabétique par département).

F/12/356 Ain à Calvados.

Il ne faut pas indexer Ain ou Calvados qui ne sont cités que parce qu'ils représentent le premier et le dernier dossier de la liasse ; d'autres départements font l'objet de dossiers au même titre que ces deux-là.

Des noms de lieu "limites"

Pour les mêmes raisons, on n'indexera pas aux noms de lieux "limites".

Exemple

On ne choisira pas les termes Saint-Lary et Castillon pour un dossier concernant les travaux sur la route de Castillon à Saint-Lary : si les travaux se font sur toute la route, l'indexation est incomplète puisque d'autres communes sont concernées et peut-être que les travaux ne se font ni sur la commune de Castillon ni sur celle de Saint-Lary mais sur une autre non mentionnée.

Dans ce cas, l'indexation "lieu" doit se faire au nom de la route : "route départementale n° 15" ou, éventuellement "route de Castillon à Saint-Lary" si le descripteur a été créé, ce qui n'est pas la même chose que le nom des deux communes.

L'idéal bien sûr, c'est de faire des renvois (ou une association dans une base de données) entre cette route et toutes les communes qu'elle traverse.

Des noms de lieu localisation d'une institution

Il ne faut pas indexer au nom d'un lieu géographique qui n'est que la localisation d'une institution.

Exemple

Les registres de la justice de paix de Castillon ne doivent pas être indexés à “Castillon” mais au “canton de Castillon”, ressort de ladite justice.

La baronnie de Château-Verdun ne doit pas être indexée à “Château-Verdun” mais à “baronnie de Château-Verdun”.

Des renvois doivent être faits, bien sûr ou des hiérarchies créées entre l’institution et les communes qui en composent le ressort.

On n'indexera pas non plus des minutes de notaire au nom de la commune siège de l'étude : les minutes d'une étude couvrent de nombreuses communes en milieu rural ; même dans le cas d'études urbaines, les minutes couvrent généralement les villages des alentours.

3.2.1.3. L'absence des noms de lieu pertinents

Le nom de lieu pertinent peut ne pas apparaître dans les analyses.

Exemple

Écoles. – Travaux (classement dans l’ordre alphabétique des communes).

Des dossiers classés dans l’ordre alphabétique des communes existent mais le répertoire est trop succinct (cas fréquent dans les répertoires quelque peu anciens).

L’idéal est de restituer les noms de communes dans l’indexation, faute de quoi ces dossiers sont privés d’accès et ne seront sans doute jamais consultés.

Exemple

Il n’y a pas de dossier au nom des communes mais les documents comportent un élément important et identifiable sur chaque commune du département : c’est le cas par exemple des résultats d’élections, des résultats d’enquêtes statistiques (population, agriculture…).

On peut créer un descripteur général du genre “toutes communes” (il s’agit implicitement du département ressort du service d’archives) ou “communes de l’Ain”, en faisant une fois pour toutes un renvoi du nom de chaque commune vers ce descripteur.

3.2.2. Les noms de personnes

Introduction

L'utilisation des noms de personnes constitue l'un des accès les plus utiles au chercheur. L'archiviste doit cependant prendre un certain nombre de précautions pour s'assurer de la précision et de l'exhaustivité de ses choix d'accès. Les problèmes varient aussi selon la période couverte par les instruments de recherche.

3.2.2.1. Dans les archives modernes et contemporaines (postérieures à 1800)

Les noms de personnes sont souvent cités comme intitulés de dossiers nominatifs : dossiers de personnel, d’étrangers, de nomades, de personnes surveillées…

On peut aussi trouver des répertoires d’archives de juridictions portant le nom des parties dans les analyses de dossiers de procédure (mais c’est rare pour la période moderne et inexistant pour la période contemporaine puisque les documents de justice ne sont généralement pas communicables et que les noms ne peuvent donc être cités dans les instruments de recherche).

La question peut être réglée par les inventaires-index qui présentent les analyses de pièces ou de dossiers dans l’ordre alphabétique des noms propres de personnes.

On peut également faire des exceptions, selon les besoins du service.

Exemple

En France, on pourra choisir d'indexer

- dans les dossiers de personnel, pour des personnages qui sont plus susceptibles que d'autres de faire l'objet de recherches
    - membres du corps préfectoral
    - dignitaires ecclésiastiques
    - parlementaires
    - maires des communes les plus importantes
- dans les archives judiciaires pour une affaire particulièrement célèbre, même localement.

On indexe les personnes dont le nom figure dans les analyses

- non pas comme intitulé individuel parmi des centaines d'autres ,
- mais en raison de leur action ou de leur présence particulière :
    - un président de la République
    - un chef d'État étranger en visite
    - un personnage connu venu en touriste…

3.2.2.2. Dans les archives anciennes

Il existe, en France notamment, de nombreux inventaires analytiques d'archives de cours de justice ou d'institutions ecclésiastiques d'Ancien Régime où l'analyse à la pièce, à la page du registre, fait apparaître une grande quantité de noms propres.

Faire un index de ces inventaires ne pose pas de problème car le traitement est homogène (généralement, tous les noms des parties prenantes d'un acte sont cités).

Mais l'intégrer dans un fichier du service d'archives risque de déséquilibrer celui-ci.

On peut choisir une solution intermédiaire en ramenant l'indexation au nom de famille (par exemple : "Ferrand, famille de Prades"), avec des exceptions pour les personnages connus, dignitaires, membres de grandes familles seigneuriales…

3.2.3. Les autres noms propres

Si d'autres noms propres apparaissent dans les analyses, on les indexe. Par contre, il est utile de créer des descripteurs spécifiques pour des édifices notables qui font l'objet de nombreuses références. Pour les autres, on en restera à une double indexation, commune et type d'édifice (Castelnau / hôtel de ville).

Il convient d'indexer tous les noms d'organismes qui figurent dans les analyses. Dans ce cas, il ne faut pas confondre l'indexation avec un fichier de producteurs. Un organisme, une administration ne sera indexée qu'en tant qu'objet et non pas en tant que *producteur* d'archives.

Exemple

On indexera à "Direction départementales de l'Équipement" des dossiers sur le fonctionnement de cette administration, l'organisation de ses services, son personnel, ses locaux, mais pas les dossiers produits par elle.

Pour les axes de communication, c'est au nom de la route, de la voie de chemin de fer ou de tout autre axe qu'il faut indexer et non pas au nom des communes "terminus".

On doit indexer tous les événements historiques mentionnés dans les analyses. On peut faire aussi une indexation sur des périodes, d'après les dates des documents, à condition que ces "périodes" ne soient pas trop larges. On peut également hiérarchiser des événements ou des événements et une période (des batailles et des traités par rapport à une guerre).

4. Mettre en forme matériellement un fichier

Introduction

Après avoir fait le choix du vocabulaire et du niveau d'indexation, il faut passer :

- à la rédaction et au rangement
- ou à la saisie.

4.1. La rédaction d'une fiche manuelle

Lorsqu'on élabore un fichier, il faut rédiger **une fiche pour chaque entrée** : une analyse dont on extrait **deux** descripteurs noms communs et **trois** descripteurs noms propres donnera lieu à la rédaction de **cinq fiches**.

Une fiche doit porter quatre éléments :

- le descripteur en haut à gauche,
- la ou les cotes du ou des articles en haut à droite,
- un texte résumant ce qui, dans l'analyse, est en rapport avec le descripteur,
- les dates extrêmes des documents du ou des articles.

Exemple

TOULOUSE 23 W 54

Ravitaillement pendant l'Occupation

1942-1944

et

OCCUPATION 23 W 54

Ravitaillement de Toulouse

1942-1944

et

RAVITAILLEMENT 23 W 54

Ravitaillement de Toulouse

1942-1944

La rédaction de la partie texte d'une fiche est la partie la plus délicate. Dans la fiche manuelle, on ne recopie pas l'analyse toute entière. Il ne faut mettre que ce qui est en rapport avec le descripteur et le mettre de façon la plus résumée possible tout en restant compréhensible. La même analyse donne donc lieu sur les fiches à des notices différentes.

Exemple

Soit l'analyse suivante :

347 W 5 Construction de pistes forestières en forêt communale d'Esplas-de-Sérou.

1989

On rédige les deux fiches suivantes (indexées selon le thesaurus W) :

1. ESPLAS-DE-SEROU 347 W 5

   Forêt communale

   aménagement de pistes forestières

   1989

2. aménagement forestier 347 W 5

   forêt communale d'Esplas-de-Sérou

   aménagement de pistes forestières

   1989

Les fiches de renvoi ne portent pas de cote : il s'agit de relations de principe entre les descripteurs.

Exemple

Un renvoi d'équivalence :

ARAUX

voir ERP

Araux est un lieu-dit et Erp la commune dont ce lieu-dit dépend.

Un renvoi de hiérarchie :

RÉVOLUTION FRANÇAISE

voir aussi Prise de la Bastille, bataille de Valmy (1792)

Tous ces termes sont autorisés. On met "voir aussi" parce qu'il y a aussi une indexation à "Révolution française".

4.2. Le rangement des fiches

Le rangement se fait dans l'ordre alphabétique

- des formes choisies
- mais aussi des formes rejetées.

Exemple

On retient la forme suivante : SULLY, Maximilien de Béthune, duc de

Il faut faire aussi la fiche suivante :

BETHUNE, Maximilien de, duc de Sully

voir SULLY, Maximilien de Béthune, duc de

Quand plusieurs personnes portent le même nom :

- on les classe dans l'ordre alphabétique des prénoms
- si elles ont les mêmes noms et prénom, on les classe dans l'ordre alphabétique des qualités
- en dernier recours dans l'ordre chronologique de leurs dates de vie.

Les noms composés sont classés après les noms simples (Hardouin, Nicolas, puis Hardouin-Mansart, Jules).

Les particules de, du, des ne sont pas prises en compte dans l'ordre alphabétique ("Chateaubriand, de") et non "de Chateaubriand"). Le ou La sont considérés, eux, comme partie intégrante du nom et sont classés à leur place (Le Nain, le Nôtre).

Pour les noms géographiques, on ne tient pas compte des articles, qui sont rejetés après le nom (Seine, la).

Attention

Ce fut longtemps le cas des articles composant le nom de certaines communes : Mans (Le), Rochelle (La). On considère maintenant que l'article est une composante pleine et entière du nom de la commune et on dit "ville de le Mans" et non plus "ville du Mans".

Les noms géographiques qui ont une composante nom commun sont classés à leur composante nom propre : Blanc (mont), Tournalet (col du).

Comme pour les noms de personnes, dans les cas de lieux homographes, on classe dans l'ordre alphabétique de l'élément différenciant :

- dans l'ordre des départements ou des pays pour des communes de même nom,
- dans l'ordre des communes pour les lieux-dits…

Pareillement, on classe les noms composés après les noms simples.

Pour plus de détails... **Complément**

Au Canada, les méthodes de mise en forme d'un fichier d'autorité sont aussi expliquées dans deux ouvrages complémentaires à la présentation des Règles de description pour les documents d'archives. Voir :

- *Louise Gagnon-Arguin, "Une introduction au contrôle d'autorité pour le traitement des noms propres en archivistique"*[Gagnon-Arguin]
- *Elizabeth Black, "Le contrôle d'autorité : Un manuel destiné aux archivistes"*[Black]

4.3. L'intégration dans une base de données

Pour intégrer un fichier d'autorité dans une base de données, on remplit simplement des champs d'indexation dans des champs de saisie supplémentaires créés à cette fin. L'indexation est rattachée à la cote et fournit en réponse l'analyse entière de l'article.

Si on a affaire à un système gérant un vocabulaire organisé et n'acceptant que des termes créés a priori, la saisie de l'indexation ne pose aucun problème : tous les termes interdits sont automatiquement rejetés, les descripteurs ne sont acceptés que dans le champ qui leur a été attribué (un terme de typologie ne pourra être indexé par erreur en mot-matière…).

Les rattachements hiérarchiques et d'association doivent être faits a priori, c'est-à-dire lors de la création du descripteur, pas au moment de l'indexation. Ainsi, on ne se soucie pas des renvois lors de l'indexation elle-même.

Si on utilise un logiciel qui autorise l'indexation libre, la procédure d'indexation entraîne, ou devrait entraîner des contrôles permanents sur le vocabulaire et sur les renvois avant saisie.

Certaines bases permettent d'enrichir l'indexation par des notices complétant les descripteurs noms propres : éléments biographiques d'un personnage, historique du territoire d'une commune, d'une administration…

5. Conclusion

Pendant longtemps, l'indexation a été une opération menée à l'issue de l'achèvement d'un classement et avant l'impression de l'instrument de recherche sur papier. L'index était conçu seulement comme une annexe d'inventaire ou de répertoire. Il était généralement fait par une seule personne, dans un laps de temps assez restreint et sur un volume d'analyses lui aussi restreint. Il se présentait sous une forme aboutie (manuelle ou dactylographiée), relue avant impression. On y détectait donc facilement les anomalies et on les corrigeait.

Avec le développement des instruments de recherche synthétiques et la volonté croissante de disposer d'une information multiple à l'échelle d'un centre d'archives, l'indexation est devenue une opération qui vise le plus souvent la constitution de fichiers ouverts et de plus en plus volumineux, opération qui prend souvent la forme d'un travail collectif. Les défauts sont multipliés et les fichiers deviennent irrécupérables.

Ces remarques valent tout autant pour la préparation d'instruments de recherche qu'on voudra diffuser électroniquement. Bien que la prochaine section s'adresse plutôt à des usagers qui ont acquis une certaine expérience, le débutant voudra quand même en prendre connaissance pour en tenir compte éventuellement dans son travail. De moins en moins d'instruments de recherche sont publiés sur papier parce que de telles éditions coûtent très cher et prennent beaucoup de temps à réaliser. Un nombre croissant de centres préfèrent recourir à l'édition électronique sur Internet qui se fabriquent de façon plus automatiques, sont moins coûteuses et qui peuvent se corriger plus facilement, en cas d'erreur.

Pour plus de détails... *Complément*

Pour en savoir plus : le web sémantique et les vocabulaires contrôlés

L'accès et l'interconnexion sémantique des collections patrimoniales est un des axes de recherche en cours. Les thèmes, sujets, concepts représentés dans les ressources sont comparables et constituent autant de points d'interconnexion pour rendre ces collections interopérables. Mais ces collections sont généralement indexées, chacune avec ses langages spécifiques. Ces derniers constituent souvent des outils hétérogènes, conçus et élaborés selon des règles d'indexation qui leur sont propres. L'hétérogénéité réside autant dans le type d'outil et son mode d'élaboration et de représentation (thésaurus, liste d'autorité) que dans les variations présentes dans la couverture sémantique de concepts comparables. Les environnements linguistiques hétérogènes se surajoutent également à cette complexité. Comment permettre à un même public d'accéder à l'ensemble des documents de deux ou plusieurs collections indexées chacune avec un vocabulaire distinct en utilisant indifféremment les termes de l'un ou de l'autre de ces vocabulaires ? Comment utiliser les termes d'un langage de représentation spécifique pour chercher dans les collections qui n'ont pas de lien direct avec ce langage ? Pour atteindre ces objectifs, les termes en question doivent être alignés et les vocabulaires concernés doivent être rendus interopérables.

Deux orientations majeures se distinguent dans les travaux en cours. L'une a trait à l'établissement manuel des équivalences entre les termes des vocabulaires différents utilisés pour les collections visées à interconnecter. L'autre, plus récente, utilise les techniques du web sémantique, partant du principe que les vocabulaires contrôlés sont de véritables systèmes d'organisation de connaissances et correspondent donc au type d'artefacts dont la conception constitue à la fois le but et le moyen de la vision du *Web sémantique*.

En d’autres termes, les techniques du Web Sémantique peuvent donner une nouvelle vie aux outils traditionnels que sont les vocabulaires contrôlés en les exploitant au moyen de technologies standards développés pour les réseaux. Les vocabulaires contrôlés se voient ainsi utilisés dans des contextes élargis de mise en relation de ressources.

6. Évaluation des connaissances

Introduction

Dans cette section, de très nombreux exemples d'indexation vous ont été donnés. Vous pouvez essayer de les relire tout en essayant de les refaire, cela vous aidera fortement pour les exercices qui suivent.

Exercice : 6.1. Exercice 1

Cliquez sur le lien pour accéder à l'exercice.

6.1. Exercice 1 (cf. sec7_6.1.pdf)

Exercice : 6.2. Exercice 2

Cliquez sur le lien pour accéder à l'exercice.

6.2. Exercice 2 (cf. sec7_6.2.pdf)

Mentions légales

©AIAF - PIAF