

Annexes

Annexe 1	
NORME AFNOR Z 47-102 (1978).....	374
Annexe 2	
PRÉSENTATION DE L'EXPÉRIMENTATION.....	382
Annexe 3	
RÉSULTATS (1) : « LES MISES EN DOCUMENTS ».....	385
Annexe 4	
RÉSULTATS (2) : LES NOMS PROPRES DANS LES PRATIQUES DOCUMENTAIRES.....	386

Annexe 1 : norme AFNOR 1978

NORME FRANÇAISE ENREGISTRÉE	DOCUMENTATION PRINCIPES GÉNÉRAUX POUR L'INDEXATION DES DOCUMENTS	NF Z 47-102 Août 1978
SOMMAIRE		
		Page
	AVANT-PROPOS	2
1	OBJET ET DOMAINE D'APPLICATION	2
2	PROCESSUS ET FINALITÉ DE L'INDEXATION	2
3	RECONNAISSANCE DES CONCEPTS CONTENANT L'INFORMATION	3
3.1	APPRÉHENSION DU DOCUMENT	3
3.1.1	Document écrit	3
3.1.2	Autres documents	3
3.2	IDENTIFICATION DES CONCEPTS	4
3.3	SÉLECTION DES CONCEPTS	4
3.3.1	Étendue ou exhaustivité de l'indexation	4
3.3.2	Spécificité de l'indexation	5
4	REPRÉSENTATION DES CONCEPTS DANS LE LANGAGE DOCUMENTAIRE	5
5	CONTRÔLE DE LA QUALITÉ DE L'INDEXATION	6
	Annexe 1 : Schéma : l'indexation dans la chaîne documentaire	7
	Annexe 2 : Lexique	8
Enregistrée par décision du 1978-06-26 pour prendre effet le 1978-08-01		© AFNOR 1978 Droits de reproduction et de traduction réservés pour tous pays

Afnor 78224

NF Z 47-102 1^{er} Tirage 78-07

General methods for analyzing documents and determining their subjects. ☺
Indexierung von Dokumenten — Allgemeine Prinzipien.

AVANT-PROPOS

La présente norme a été établie à partir des travaux sur l'indexation effectués par le Comité technique « Documentation » de l'ISO (Organisation internationale de normalisation) et l'UNESCO.

1 OBJET ET DOMAINE D'APPLICATION

Cette norme a pour objet d'établir des méthodologies valables et cohérentes pour la description et la caractérisation des documents à l'aide de représentations de concepts (mots-clés, descripteurs, indices, etc.).

Ces principes sont conçus de telle manière qu'ils soient dans toute la mesure du possible indépendants de tout système documentaire. Ils constituent une base uniforme de règles ou de recommandations qui doivent permettre :

- une coopération plus facile entre les différents services d'information ;
- l'établissement de règles adaptées à des systèmes particuliers.

Ces recommandations doivent permettre une indexation adaptée à un système courant de recherche d'information dans un fonds documentaire. Mais chaque système doit aller plus loin dans le raffinement par l'élaboration de règles plus spécifiques adaptées aux besoins de ses utilisateurs, tout en restant compatibles avec ces principes généraux.

2 PROCESSUS ET FINALITÉ DE L'INDEXATION

— L'indexation est l'opération qui consiste à décrire et à caractériser un document à l'aide de représentations des concepts contenus dans ce document, c'est-à-dire à transcrire en langage documentaire les concepts après les avoir extraits du document par une analyse. La transcription en langage documentaire se fait grâce à des outils d'indexation tels que thésaurus, classification, etc.

Le présent texte ne concerne que l'indexation du contenu intellectuel du document et ne prend pas en compte les éléments caractérisant le support physique tels que le nombre de pages, le format, le type de document...

— La finalité de l'indexation est de permettre une recherche efficace des informations contenues dans un fonds de documents et d'indiquer rapidement, sous une forme concise, la teneur d'un document.

L'indexation conduit à l'enregistrement des concepts contenus dans un document, sous une forme organisée et facilement accessible c'est-à-dire à la confection d'outils de recherche documentaire tels que des index de livres ou de bulletins bibliographiques, des catalogues organisés de matières ou encore des fichiers manuels, mécaniques ou automatiques. La recherche des informations enregistrées, par exemple la sélection de documents répondant à une question s'opérera à partir de ces outils de recherche documentaire. Cette opération est possible par indexation de la question, opération analogue à l'indexation du fonds. L'indexation de la question permet d'accéder aux outils de recherche documentaire créés par l'indexation du fonds de documents.

L'utilisation d'un langage documentaire pour ces deux opérations permet d'obtenir une coïncidence exacte du vocabulaire d'indexation des documents avec celui des questions auxquelles ces documents apportent une réponse.

Un schéma donné en annexe permet de situer l'indexation (son processus et ses buts) au sein de la chaîne documentaire.

Les principes d'indexation, objet de cette norme, sont valables aussi bien dans le cas où les opérations se déroulent manuellement que dans le cas où le système est plus évolué c'est-à-dire où il y a automatisation de la recherche documentaire.

Le processus de l'indexation se déroule en deux étapes essentielles :

- la reconnaissance des concepts contenant l'information dans les documents à indexer ;
- la représentation de ces concepts dans le langage documentaire.

3 RECONNAISSANCE DES CONCEPTS CONTENANT L'INFORMATION

La reconnaissance des concepts contenant l'information dans un document peut elle-même se subdiviser en trois phases, sans que la limite entre ces phases soit très nette, les opérations qu'elles comportent pouvant se superposer dans le temps.

Ces trois phases sont les suivantes :

- appréhension du contenu total du document ;
- identification des concepts représentant ce contenu ;
- sélection des concepts nécessaires à la recherche ultérieure.

3.1 APPRÉHENSION DU DOCUMENT

La prise de connaissance du contenu d'un document diffère selon le type de document à indexer.

3.1.1 Document écrit

C'est le cas le plus fréquent. En effet, les documents écrits (livres, articles de revues, ...) forment la plus grande partie des fonds des centres de documentation et des bibliothèques.

L'appréhension du document se fait par la lecture. Mais il est important de souligner que la lecture in extenso du document n'est pas indispensable, ceci pour des raisons économiques. L'indexeur ne doit cependant perdre aucune information utile contenue dans le document et il doit repérer les passages les plus riches en information tels que :

- le titre, le résumé ;
- l'introduction, les premières phrases des chapitres et paragraphes ;
- les graphiques, les tableaux de chiffres, les figures et leurs légendes ;
- la conclusion ;
- les notions mises en valeur par certains artifices typographiques (soulignage, caractères gras, ...).

Tous ces éléments du document représentent autant de points d'ancrage que l'indexeur devra prendre en compte pour identifier l'information.

Il est recommandé d'indexer à partir du document original, le titre ou le résumé ne contenant souvent qu'un aspect ou qu'une partie ou qu'une généralisation de l'information. Dans le cas d'une indexation à partir d'un titre ou d'un résumé, le problème de l'identification du contenu informatif du document est reporté sur l'analyste qui élabore le résumé ou sur l'auteur qui doit fournir un titre synthétisant au mieux ses travaux.

3.1.2 Autres documents

La démarche d'appréhension du document est différente dans le cas d'un document audiovisuel, visuel ou sonore, ou d'un objet à trois dimensions (une pièce de musée, une pièce archéologique, etc.). Les éléments à prendre en compte pour identifier l'information pourront être autres que ceux retenus dans le cas d'un document écrit.

S'il existe un document écrit, accompagnant le document ou l'objet, la démarche sera facilitée.

3.2 IDENTIFICATION DES CONCEPTS

Après la prise de connaissance du document, il est nécessaire que l'indexeur suive un schéma logique qui le guidera dans le choix des concepts décrivant le mieux l'information. Les principaux concepts à identifier peuvent être rangés dans un petit nombre de classes que chaque système documentaire peut préciser suivant le domaine couvert en distinguant nettement les catégories auxquelles appartiennent les concepts : s'agit-il d'un procédé ou d'un équipement ? d'une propriété ou d'un phénomène ?, etc.

L'indexeur identifie les concepts dans le document à partir d'un schéma adapté à son domaine.

Par exemple :

- pour les documents traitant de la thérapeutique médicamenteuse, l'indexeur devra rechercher le nom de la maladie, le nom du médicament, son mode d'administration, le résultat obtenu ;
- pour des documents traitant de la synthèse de composés chimiques l'indexeur devra identifier le composé de départ, le procédé de fabrication, les conditions opératoires, le produit obtenu,...

3.3 SÉLECTION DES CONCEPTS

Tous les concepts identifiés par l'indexeur comme représentatifs de l'information contenue dans le document, ne sont pas nécessairement retenus pour l'indexation. En effet, l'indexation doit essentiellement être adaptée à sa **fonction** propre dans le système où elle est utilisée. Elle peut servir pour un fonds donné de documents, à la confection d'un catalogue organisé de matières ou d'un bulletin bibliographique, à la recherche documentaire par des moyens mécaniques ou par ordinateur ou encore à la publication d'index, et elle peut alors être différente selon le type de produits qui doit en résulter.

L'adaptation de l'indexation à sa fonction et au domaine couvert par le document peut être obtenue par une variation des paramètres de l'indexation qui sont son **étendue** ou **exhaustivité** et sa **spécificité**.

3.3.1 Étendue ou exhaustivité de l'indexation

D'une façon générale, tous les concepts contenus dans le document et pouvant être pris en compte par le système documentaire doivent se retrouver dans l'indexation. Si un document traite de plusieurs sujets indépendants mais entrant dans le domaine couvert, il peut ou bien être indexé par plusieurs spécialistes ou bien être divisé en plusieurs parties indexées séparément.

Il est recommandé de se garder d'une optique trop étroite car l'indexation d'un document peut être destinée à plusieurs centres et ce cas se multipliera avec le développement des réseaux d'information. Ainsi dans un système scientifique et technique, il ne faut pas négliger l'indexation d'aspects non techniques tels que des notions économiques ou des facteurs sociaux, ou l'indexation d'aspects marginaux.

Dans tous les cas, le critère de sélection d'un concept doit être sa signification en tant qu'élément du contenu informatif du document. Ceci revient à se demander à quelles questions le document donnera une réponse pertinente et à s'assurer que les notions appropriées figurent dans l'indexation. Ce critère est fondamental car il résulte de la finalité même de l'indexation.

Il est essentiel que l'indexation couvre les besoins des usagers. Il est utile que l'indexeur connaisse le type de questions posées au système. Il ne faut pas négliger l'effet rétroactif des questions des utilisateurs sur l'indexation au fur et à mesure qu'apparaissent de nouveaux centres d'intérêt. Il faut quand même éviter que cette adaptation aux besoins de l'utilisateur n'aille jusqu'à la déformation de l'indexation.

Théoriquement, le nombre de descripteurs par document ne doit pas être limité arbitrairement mais doit seulement dépendre de la quantité d'information contenue dans le document. Une telle limitation nuirait à l'objectivité de l'indexation et aboutirait dans beaucoup de cas à une perte d'information importante pour les utilisateurs. Par contre, pour certains systèmes documentaires cette limitation peut être dictée par des considérations économiques. Dans ce cas, le choix de l'indexeur doit être fonction de l'importance relative du contenu informatif des concepts à retenir.

Il faut également tirer du document les concepts qui y sont implicitement contenus, et qui serviront à replacer les descripteurs dans un contexte approprié.

Exemple :

« Mesure du monoxyde de carbone émis par un moteur tournant au ralenti » implique la notion de « pollution atmosphérique ».

3.3.2 Spécificité de l'indexation

Elle est variable suivant les buts de l'indexation à l'intérieur d'un système documentaire. Le niveau d'indexation restera plus général pour un catalogue de bibliothèque que pour un outil de recherche documentaire destiné à une recherche rétrospective fine.

Dans le cas où l'indexation doit répondre à plusieurs objectifs, il est alors possible de couvrir tous les niveaux d'indexation du plus fin au plus général, permettant de répondre aux demandes très spécifiques ou se situant à un niveau plus général.

Il peut exister par ailleurs une relation entre le degré de spécificité de l'indexation et le degré de développement des concepts inclus par l'auteur dans un document. Une notion peu développée ou seulement citée par l'auteur si elle est retenue par l'indexeur, pourrait être indexée à un niveau de généralité élevé.

4 REPRÉSENTATION DES CONCEPTS DANS LE LANGAGE DOCUMENTAIRE

Il s'agit de formuler, en langage documentaire, à l'aide d'outils d'indexation, les concepts préalablement sélectionnés. Cette transcription se réalise aussi bien pour le fonds documentaire que pour la question. Elle conduit à l'organisation de l'information pour la rendre utilisable et accessible.

Il est donc recommandé de disposer d'outils d'indexation permettant de représenter l'information selon un certain nombre de règles spécifiques à chaque outil d'indexation. Les outils d'indexation les plus employés se divisent en deux groupes :

- outils de type « combinatoire » tels que thésaurus, lexique structuré, liste de sujets, etc. ;
- outils du type « catégoriel » où les concepts sont représentés par les indices ou les symboles de la classification utilisée.

Le processus d'utilisation de ces outils d'indexation se définit comme la reconnaissance dans les vocabulaires contrôlés ou les classifications, des descripteurs ou indices permettant de représenter fidèlement les concepts sélectionnés dans les documents. L'utilisation des outils d'indexation est régie par des principes propres à chaque outil.

L'indexeur devra donc avant toute indexation bien connaître l'outil dont il dispose et les contraintes particulières qui lui sont imposées ainsi que les avantages dont il dispose. Ainsi, si l'outil d'indexation est un thésaurus, le nombre de descripteurs affectés à un document peut-être réduit sans perte d'information, les relations génériques et autres pouvant être entrées directement à partir du thésaurus.

L'utilisation d'un thésaurus implique une indexation par les descripteurs les plus spécifiques.

Certains systèmes documentaires utilisant des outils du type combinatoire admettent l'utilisation d'une pondération des descripteurs, d'un chaînage entre certains descripteurs, d'indicateurs de rôle, etc. Des règles particulières précises devront être formulées dans ce cas.

Si l'indexation se fait à l'aide d'outils de type catégoriel, les indices les plus spécifiques sont toujours situés dans un contexte plus large.

L'utilisation complémentaire des deux types d'outils d'indexation peut être bénéfique et faciliter ensuite la recherche des documents. Ainsi un brevet pourra être indexé par des descripteurs mais aussi de façon complémentaire par un indice de classification.

L'indexeur sera confronté au problème de formulation de concepts nouveaux ne pouvant être traduits dans le vocabulaire contrôlé ou dans la classification. Il faudra donc prévoir une procédure de mise à jour de l'outil d'indexation selon le système utilisé.

5 CONTRÔLE DE LA QUALITÉ DE L'INDEXATION

La qualité de l'indexation dépend de deux facteurs :

- la qualité de l'indexeur,
- la qualité de l'outil d'indexation.

Dans le cas idéal, pour un système donné, l'indexation d'un document devrait être identique quel que soit l'indexeur. Elle ne devrait pas non plus varier dans le temps pour un même indexeur si l'outil documentaire n'a pas été modifié. Une variation dans le temps ne peut être corrigée que par une rétroaction sur le système documentaire.

Pour que l'indexation soit cohérente, l'indexeur doit posséder un certain nombre de qualités dont la principale est l'impartialité. Il doit minimiser les effets, difficilement évitables, d'un jugement subjectif. Les difficultés d'obtention d'une bonne cohérence sont d'autant plus grandes que l'équipe d'indexation est plus nombreuse d'une part, et que les indexeurs sont plus éloignés les uns des autres (cas de l'indexation décentralisée), d'autre part. Dans ce cas un contrôle centralisé doit être mis en place.

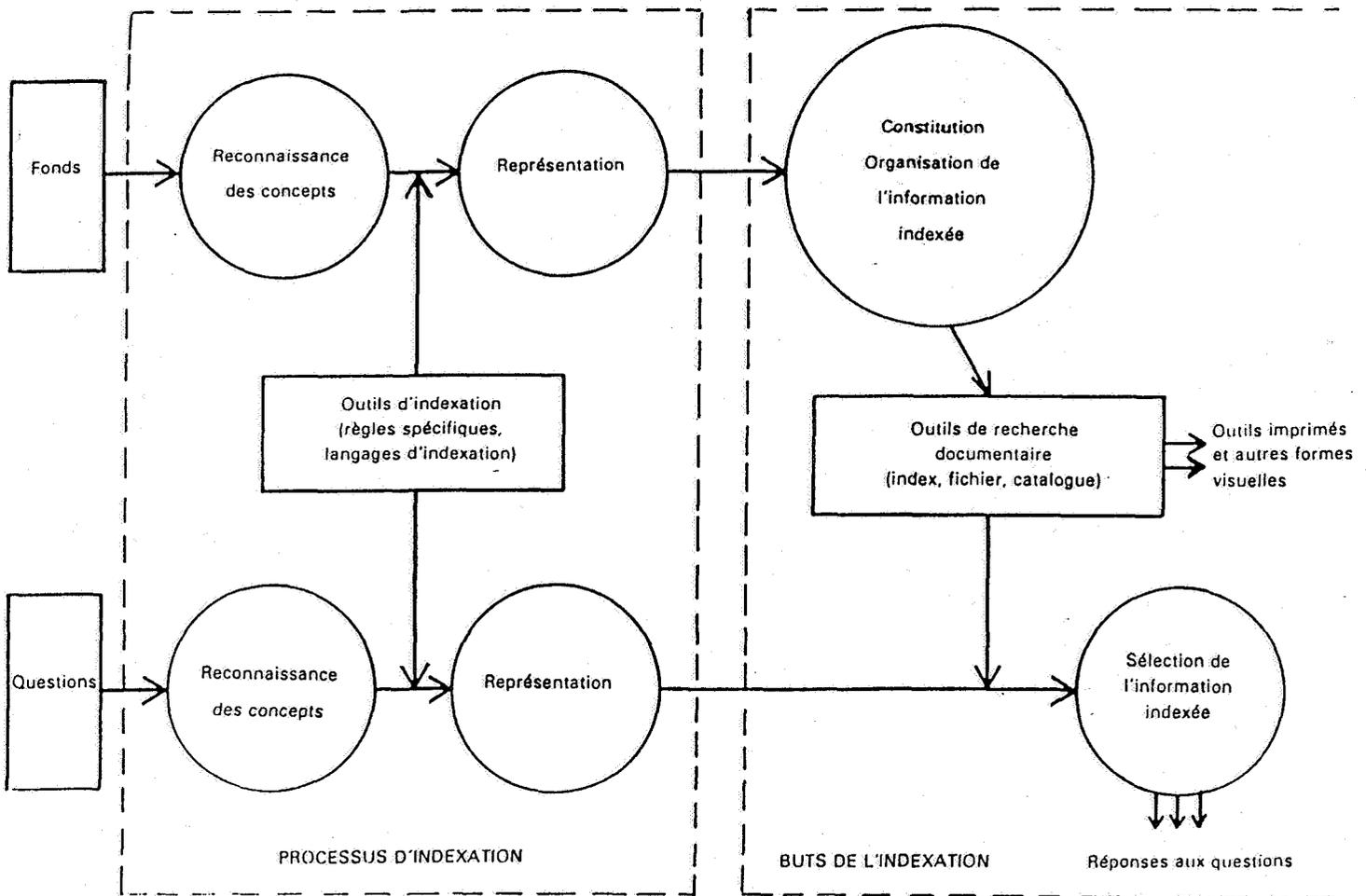
Il est nécessaire que l'indexeur soit un spécialiste de la branche dans laquelle il travaille et qu'il en connaisse le vocabulaire spécifique dans la langue des documents qu'il indexe, aussi bien que les règles et procédures spécifiques au système qu'il utilise.

La qualité de l'indexation peut être améliorée si l'indexeur connaît les types de questions posées. Il peut ainsi vérifier la qualité de son indexation en se mettant dans la situation de celui qui recherche une information. Il pourra par exemple vérifier si les descripteurs sélectionnés peuvent donner de fausses combinaisons créant du bruit à la recherche.

La qualité de l'indexation dépend aussi des qualités de l'outil documentaire utilisé. Sa qualité essentielle étant l'adaptation aux besoins nouveaux qui peuvent apparaître, et qui se traduira par une mise à jour périodique de l'outil, dont la fréquence sera fonction du système.

Le contrôle de la qualité de l'indexation peut se faire par une analyse des résultats de recherche documentaire, par exemple à l'aide du calcul des taux de rappel et de précision.

ANNEXE 1 : L'INDEXATION DANS LA CHAÎNE DOCUMENTAIRE



NF Z 47-102

ANNEXE 2

LEXIQUE

Les définitions ci-après sont données à titre provisoire en attendant la publication des normes de Vocabulaire de la documentation.

CLASSIFICATION

Structuration de notions en classes et subdivisions pour expliquer les relations sémantiques généralement hiérarchiques existant entre elles au moyen d'une notation.

CONCEPT (voir : NOTION)

DESCRIPTEUR

Mot ou groupe de mots retenus dans un thésaurus et choisis parmi un ensemble de termes équivalents pour représenter sans ambiguïté une notion contenue dans un document ou dans une demande de recherche documentaire.

DOCUMENT

Ensemble d'un support d'information, des données enregistrées sur ce support et de leur signification, servant à la consultation, l'étude, la preuve, etc.

IDENTIFICATEUR

Nom propre ou assimilé utilisé comme descripteur.

NOTE : Ce terme peut être un nom de projet, de personne ou de collectivité, de marque de fabrique, un nom géographique, une abréviation ou un sigle, etc.

INDEX

Ensemble ordonné de termes choisis et figurant dans un document avec une indication permettant de les y localiser.

INDEXATION

Représentation par les éléments d'un langage documentaire ou naturel, des notions résultant de l'analyse du contenu d'un document en vue d'en faciliter la recherche.

MOT CLÉ

Mot choisi dans le titre ou le texte d'un document, caractérisant son contenu et permettant la recherche de ce document.

NOTION

Toute unité de pensée.

THÉSAURUS

Vocabulaire contrôlé et dynamique de termes (descripteurs et non-descripteurs) obéissant à des règles terminologiques propres et reliés entre eux par des relations sémantiques.

NOTE : Ce vocabulaire est employé pour représenter sans ambiguïté dans un langage artificiel des notions exprimées en langage naturel et contenues dans des documents ou dans des demandes de recherche documentaire.

Annexe 2 : présentation de l'expérimentation

1- Objectif de l'expérimentation

L'expérimentation menée avait pour objectif d'apporter des éléments de réponse aux deux questions suivantes :

- une même source peut-elle faire l'objet de « mises en document » différentes ?
- s'il y a une diversité de mises en documents, à quel niveau et sous quelle forme se manifeste-t-elle ?

Pour appréhender ces deux questions, nous avons procédé à l'analyse comparative du traitement documentaire d'une même source.

Par traitement documentaire, on entend :

- la sélection des unités textuelles à indexer ;
- l'affectation de descripteurs aux unités textuelles retenues ;
- l'objectif d'utilisation assigné aux unités textuelles sélectionnées et indexées.

2- Méthode

2.1- Choix de la source

Pour observer le mécanisme de sélection documentaire, nous avons choisi comme source un périodique généraliste, supposé dense et varié : le quotidien *Le Monde*. L'édition retenue a été celle du 1er décembre 1994, date qui convenait aux participants sollicités et qui présentait l'avantage de comporter un supplément, lui toujours spécialisé¹.

Ne nous étant pas livrée à une analyse fine des publications du *Monde*, il est difficile d'évaluer la représentativité de l'édition retenue et par suite de l'expérimentation menée.

À titre indicatif, signalons quelques caractéristiques du numéro traité :

- numéro de 24 pages dont deux pages « non informatives » (publicité, p. 5 ; annonces immobilières p. 17) ;
- numéro comportant 92 « textes » dont :
 - 53 articles signés
 - 7 encadrés de rappel²
 - 16 dépêches identifiables (AFP ou Reuter)
 - 16 articles non signés
 - 5 hors-textes
 - 1 rectificatif

¹. Le supplément du jour concernait les arts et spectacles (ancienne formule du journal).

². Nous avons distingué article et encadré de rappel car les documentalistes n'ont pas toujours sélectionné les deux.

2.2- Choix des participants

Le type de source retenu a orienté le choix des participants : ce sont essentiellement les centres de documentation qui se livrent au dépouillement de la presse quotidienne.

Le nombre de participants a arbitrairement été arrêté à dix, une fois obtenue la diversité souhaitée :

- *diversité de statut* :
6 organismes relèvent du secteur privé, 4 du secteur public ;
- *diversité des couvertures documentaires* :
3 organismes documentaires généralistes, 7 spécialisés ;
- *diversité des langages documentaires* :
2 thésaurus, 1 liste de vedettes matières, 2 plans de classification, 6 lexiques de classement ;
- *diversité de l'utilisation des articles du Monde* :
alimentation de bases de données bibliographiques, alimentation de dossiers de presse, fabrication de revues de presse, diffusion sélective d'information³.

La liste des dix participants est la suivante :

	Statut	Couverture documentaire	Langage documentaire	Utilisation des articles
Capital	Privé Presse	Micro-économie Entreprises	Lexique de classement	Alimentation dossiers / DSI
CNDP Centre national de documentation pédagogique	Public Éducation nationale	Communication	Lexique de classement	Revue de presse bimensuelle
Cour des Comptes	Public Juridiction financière	Financements publics	Plan de classification	Revue de presse quotidienne Alimentation dossiers
La Croix	Privé Presse	Généraliste	Lexique de classement	Revue de presse quotidienne Alimentation dossiers
Documenta- tion française	Public Service du Premier ministre	Information politique Actualité	Thésaurus	Alimentation de base de données
Le Figaro (documentation économique)	Privé Presse	Économie Entreprises	Lexique de classement	Alimentation dossiers / DSI Revue de presse quotidienne
FNSP Fondation nationale des sciences politiques	Public Ensei- gnement supérieur	Sciences politiques Actualité	Plan de classification Liste de vedettes matières (Rameau)	Alimentation dossiers Alimentation de catalogue
InfoMatin	Privé Presse	Généraliste	Lexique de classement	Alimentation dossiers

³. Abrégée DSI dans le tableau ci-dessous.

Le Monde	Privé Presse	Généraliste	Thésaurus	Archivage du journal Alimentation dossiers
Télérama	Privé Presse	Culture Médias Société	Lexique de classement	Alimentation dossiers Revue de presse quotidienne

3- Déroulement de l'expérimentation

3.1- Calendrier

L'expérimentation s'est déroulée approximativement d'octobre 1994 à octobre 1995 :

Septembre 1994

Octobre 1994

Novembre 1994

Mars 1995

Août-septembre 1995

Contacts et accord des participants

Mise au point des consignes et du questionnaire

Envoi des consignes et du questionnaire

Recueil des réponses et entretiens

Dépouillement et analyse des résultats

3.2- Consignes et questionnaire

Les consignes de travail étaient les suivantes :

- reporter le travail normalement effectué dans le cadre du service sur l'exemplaire du *Monde* fourni, en indiquant : le cochage, l'indexation et l'utilisation envisagée ;
- indiquer le temps réel passé au dépouillement du journal ;
- suivre le traitement documentaire du numéro : indiquer les modifications apportées à la sélection et/ou à l'indexation et/ou à l'utilisation au cours des quatre mois suivants.

Le questionnaire joint aux consignes portait sur :

- la couverture documentaire ;
- les périodiques dépouillés au même titre que *Le Monde* ;
- le mode de stockage du journal (texte intégral, coupures de presse ; supports imprimé, électronique, magnétique) ;
- outils et méthodes documentaires employés ;
- une rubrique de commentaire libre sur le dépouillement réalisé.

Annexe 3 : les mises en documents

Cette annexe présente, sous une forme chiffrée, le résultat issu de l'analyse du cochage effectué ; elle met en regard le nombre d'articles retenus (colonne « Nbr. art. ») et le nombre d'unités documentaires constituées à partir de ceux-ci (colonne « Nbr. doc. »).

Nom de l'organisme	Mise en document		Observations		
	Nbr.art	Nbrdoc	Temps passé	Stockage du Monde	Autres
Capital	5	3	10 mn	Texte intégral imprimé (1 an) Coupures	Le Monde par Minitel
CNDP	2	1	10 mn	Texte intégral imprimé (8 jours)	
Cour des comptes	6	6	15 mn	Texte intégral imprimé ("longtemps") Coupures CD-Rom	Journal de référence
La Croix	41	42	2 heures	Texte intégral imprimé ("longtemps") Coupures	Journal de référence
Documentation française	22	13	1h30	Coupures de presse (microfiches)	Journal de référence
Le Figaro-Éco	6	4	15 mn	Texte intégral imprimé (1 an) Coupures	
FNSP	78	85	4 heures	Texte intégral imprimé (8 jours) Coupures	Journal de référence Pôle associé Sciences Politiques
Info Matin	44	37	45 mn	Texte intégral imprimé (1 an)	Le Monde par l'Européenne de Données Abont. AFP
Le Monde	89	89			
Télérama	30	28	30 mn	Texte intégral imprimé (6 mois)	Le Monde par l'Européenne de Données

Annexe 4 : les noms propres dans les pratiques documentaires

La place des noms propres dans les pratiques documentaires peut être appréhendée à trois niveaux :

- au niveau global des grilles d'indexation utilisées ;
- au niveau médian de la distribution des termes d'indexation ;
- au niveau local de la nature linguistique des descripteurs matières.

1- Présence de la catégorie « nom propre » dans les grilles d'indexation

Organisme documentaire	Grille d'indexation
La Croix	Comprend deux champs d'indexation réservés aux noms propres : - lieu géographique de l'action - personnes citées
Le Monde	Constituée de quatre champs d'indexation imposant l'emploi de noms propres : - territoires étrangers - territoires français - personnes physiques, personnes morales - auteurs dont l'œuvre est citée
Documentation française	Comprend un champ obligatoire pour les descripteurs géographiques
FNSP	Le plan de classification impose l'entrée par un nom géographique ⁴
Cour des comptes	Adopte le plan de classification de la FNSP

4. « Chaque fois que cela est possible on classe d'abord sous le nom d'un pays, ou sous celui d'un ensemble géographique et seulement si aucune localisation n'est possible dans les rubriques générales (9.1 et 9.3 à 9.7). Dans ces cas, le chiffre 9 prend la place du nom de pays ».

2- Distribution des termes d'indexation

Abréviations utilisées dans le tableau :

- Nbre de I : rappel du nombre global d'indexations réalisées ;
- Nbre de G : indication du nombre de noms de lieux géographiques (pays, villes, régions, etc.) utilisés pour indexer ;
- Nbre de PP : indication du nombre de noms de personnes physiques utilisés pour indexer ;
- Nbre de PM : indication du nombre de noms de personnes morales (entreprises, manifestations culturelles, institutions, partis politiques, etc.) utilisés pour indexer ;
- Nbre de M : indication du nombre de « descripteurs matières » utilisés pour indexer ;
- % : part des noms propres dans les indexations.

Tableau de distribution

Organisme	Nbr I	NbrG	Nbr PP	Nbr PM	NbrM	%
Capital	3		2	1		100%
CNDP	1			1		100%
Cour des comptes	6					
La Croix	42	14	9	3	16	62%
Doc. française	13					50%
Le Figaro-Éco	4	1		2		100%
FNSP	85					94%
Info Matin	37	15	8	2	12	67,5%
Le Monde	89				11	88%
Télérama	28	7	5	3	5	75%

Commentaire du tableau

La distribution proposée ci-dessus est incomplète et donc approximative : la grille d'analyse retenue ne permet pas en effet de rendre compte complètement de tous les cas de pratiques documentaires ; certaines évoluent dans le temps, d'autres sont trop complexes pour être rendues ici dans le détail :

- *exemple de pratique complexe* : La Documentation française dispose de règles de composition de descripteurs fondées sur le

mélange nom propre-nom commun⁵ (exemple : « PS - parti politique position ») ; elle travaille en outre sur deux niveaux d'analyse (détaillé et général) partiellement redondant : nous avons essayé d'estimer plus que de comptabiliser la proportion des noms propres ;

- *exemple d'évolution de l'indexation dans le temps* : élimination d'articles une fois diffusés par revue de presse (cas de Télérama, de la Cour des comptes).

3- Nature linguistique des descripteurs thématiques

En examinant la composition des descripteurs thématiques (colonne « nbr M » dans le tableau ci-dessus), on peut établir une distinction entre les descripteurs thématiques comportant uniquement des « noms communs » et les descripteurs thématiques comportant à la fois des « noms communs » et des « noms propres ».

Abréviations utilisées dans le tableau :

- Nbre M : indication du nombre d'indexations thématiques ;
- M NC : descripteurs thématiques constitués uniquement de noms communs ;
- M NC-NP : descripteurs thématiques constitués à la fois de noms communs et de noms propres.

Nom	NbrM	M NC	M NC et NP
La Croix	16 ⁶	- institutions - corruption - sida / conférence - avortement - enseignement supérieur - chômage / piste	- organisation internationale / OCDE - élections présidentielles 1995 / Droite / Séguin - ONU - organisation internationale / OTAN - drogue / France - cinéma / festival / Nantes - organisation internationale / OMC
Info Matin	12	- élections présidentielles / primaires - politique / financement - drogue / dépenalisation - IVG / commandos - immigration - cinéma / festivals - universités	- télévision / Arte - presse / Actuel - organisation internationale / OCDE - GATT - transport maritime / Achille Lauro

⁵. On distingue ces règles de composition « nom commun-nom propre » des entrées de lexiques de classement constituées, de façon figée, d'un nom propre et d'un nom commun.

⁶. Trois de ces seize indexations sont identiques.

Le Monde	11	<ul style="list-style-type: none"> - analyse, chiffre ; emploi transitoire ; jeunesse / chômage, octobre, indice mensuel, 1994. 	<ul style="list-style-type: none"> - Bosnie, affrontement ethnique, États-Unis, comportement, politique étrangère, déclaration. - Russie, prise de position politique, constitution, souveraineté. - Koweït, adoption parlementaire, projet de loi, femme, université, interdiction. - Grande-Bretagne, projet, budget, 1995, 1996. - sommet politique du gouvernement, dialogue nord-sud, pays en voie de développement, signature, projet, Afrique, Russie. - Venezuela, décès, prison, évasion, chiffre. - Éthiopie, adoption parlementaire, État fédéraliste, liste. - Burundi, attentat, décès - Nigeria, décès, date, élection législative, 1995. - Zimbabwe, accord, force d'interposition, Angola.
Télérama	5	<ul style="list-style-type: none"> - institution - drogue - sida - enseignement - organisation internationale 	