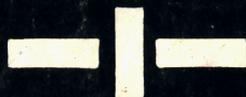


DOCUMENTATION  
ET INFORMATION

N. BELY / A. BORILLO / J. VIRBEL

N. SIOT-DECAUVILLE

**PROCÉDURES  
D'ANALYSE SÉMANTIQUE  
APPLIQUÉES  
A LA DOCUMENTATION  
SCIENTIFIQUE**



GAUTHIER-VILLARS

BIBLIOTHEQUE DU CERIST

PROCÉDURES  
D'ANALYSE SÉMANTIQUE  
APPLIQUÉES A LA  
DOCUMENTATION SCIENTIFIQUE

## OUVRAGES DE LA COLLECTION

### « DOCUMENTATION ET INFORMATION »

*Parus :*

- B. C. VICKERY.** — *La classification à facettes.* Guide pour la construction et l'utilisation des schémas spéciaux. Traduit de l'anglais par P. SALVAN.
- R. DUBUC.** — *La classification décimale universelle.* Manuel pratique d'utilisation.
- R.-C. CROS, J.-C. GARDIN, F. LÉVY.** — *L'automatisation des recherches documentaires.* Un modèle général : LE SYNTOL.
- L'organisation de la documentation scientifique.** Études par J.-C. GARDIN, E. DE GROLIER, F. LEVÉRY et l'Association nationale d'études pour la documentation automatique (A. N. E. D. A.).
- Z. DOBROWOLSKI.** — *Etude sur la construction des systèmes de classification.*
- C. LEGEARD.** — *Guide de recherches documentaires en démographie.*
- Économie générale d'une chaîne documentaire mécanisée.** Par F. ALOUCHE, N. BÉLY, R.-C. CROS, J.-C. GARDIN et J. PERRIAULT.
- R. CORMIER.** — *Les sources des statistiques actuelles.* Guide de documentation.
- R. DUBUC.** — *Exercices programmés sur la C. D. U.*
- Procédures d'analyse sémantique appliquées à la documentation scientifique.** Par N. BÉLY, A. BORILLO, N. SIOT-DECAUVILLE et J. VIRBEL.
- Classification médicale de la "National Library of Medicine".**  
Traduit par Dr G. NICOLE et M. NICOLE.

DOCUMENTATION ET INFORMATION

Collection dirigée par Paul Poindron,  
Directeur des études de l'Institut national  
des techniques de la documentation

1925

PROCÉDURES  
D'ANALYSE SÉMANTIQUE  
APPLIQUÉES A LA  
DOCUMENTATION SCIENTIFIQUE

par

N. BELY

A. BORILLO

N. SIOT-DECAUVILLE

J. VIRBEL

*CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE*

*PRÉFACE DE J.-C. GARDIN*

GAUTHIER-VILLARS ÉDITEUR

55, quai des Grands-Augustins, 75 - PARIS - VI<sup>e</sup>

1970

Procédures d'analyse sémantique appliquées à la documentation scientifique, par N. BELY, A. BORILLO, N. SIOT-DECAUVILLE, J. VIRBEL. Préface de J.-C. GARDIN. — Paris, Gauthier-Villars, 1970. — 16 cm, XVIII-242, fig. tableaux.

(Documentation et Information).



477j

© GAUTHIER-VILLARS, 1970

Toute reproduction, même partielle, de cet ouvrage est interdite. La copie ou reproduction, par quelque procédé que ce soit : photographie, microfilm, bande magnétique, disque ou autre, constitue une contre-façon passible des peines prévues par la loi du 11 mars 1957 sur la protection des droits d'auteurs.

## - TABLE DES MATIÈRES -

	Pages
Préface de J.C. Gardin . . . . .	XIII
<b>1e PARTIE : CADRE MÉTHODOLOGIQUE</b>	
<b>1. Fondements de la recherche . . . . .</b>	<b>3</b>
<b>2. Paramètres expérimentaux . . . . .</b>	<b>7</b>
2.1. <i>Le langage-cible . . . . .</i>	7
2.2. <i>Le domaine et le corpus . . . . .</i>	11
2.3. <i>Démarche générale . . . . .</i>	12
2.4. <i>Etapas de la recherche . . . . .</i>	14
<b>2e PARTIE : INDEXATION LEXICALE</b>	
<b>3. Démarche de l'indexation lexicale . . . . .</b>	<b>19</b>
3.1. <i>Morphologie des termes du langage naturel . . . . .</i>	19
3.2. <i>Les groupes de mots . . . . .</i>	20
3.2.1. <i>Définition des groupes de mots . . . . .</i>	20
3.2.2. <i>Reconnaissance des groupes de mots . . . . .</i>	21
3.3. <i>Statuts des termes LN quant à leur traduction . . . . .</i>	23
3.4. <i>Polysémies . . . . .</i>	25
3.4.1. <i>Définition . . . . .</i>	25
3.4.2. <i>Résolution des polysémies . . . . .</i>	26
3.4.2/1 <i>Polysémies de type "1" : "mots clef" . . . . .</i>	26
3.4.2/2 <i>Polysémies de type "2" : "catégories grammaticales multiples" . . . . .</i>	27
3.4.2/3 <i>Polysémies de type "3" : "analyse sémantique seule" . . . . .</i>	29

	Pages
3.4.2/4 Polysémies de type "4" : "analyse syntaxique et sémantique . . . . .	31
3.4.2/5 Polysémies de type "5" : "descripteurs obligatoires . . . . .	34
3.5. <i>Inventaire des types d'outils nécessaires à l'indexation lexicale.</i>	35
<b>4. Description des outils linguistiques . . . . .</b>	<b>37</b>
4.1. <i>Le lexique documentaire . . . . .</i>	37
4.2. <i>Le dictionnaire automatique . . . . .</i>	38
4.2.1. Codes affectés aux entrées . . . . .	38
4.2.1/1 Codes grammaticaux . . . . .	38
A. Codes de catégories grammaticales . . . . .	38
B. Codes morphologiques . . . . .	39
a. Codes affectés aux verbes . . . . .	39
b. Codes affectés aux autres catégories variables . . . . .	40
4.2.1/2 Codes affectés aux groupes de mots . . . . .	42
4.2.1/3 Codes de polysémies . . . . .	43
4.2.2. Conventions pour la rédaction des articles du dictionnaire . . . . .	43
4.2.2/1 Transcription des signes diacritiques . . . . .	44
4.2.2/2 Codes de délimitation des éléments de l'article . . . . .	44
4.2.2/3 Règles pour la rédaction des articles . . . . .	45
4.2.2/4 Exemples d'articles . . . . .	45
A. Articles sans règle de résolution polysémique . . . . .	45
B. Articles avec règles de résolution polysémique . . . . .	45
4.3. <i>Les algorithmes d'analyse . . . . .</i>	46
4.3.1. Reconnaissance des groupes de mots . . . . .	46
4.3.2. Résolution des polysémies . . . . .	46
4.3.2/1 Type "2" . . . . .	46
4.3.2/2 Type "3" . . . . .	47
4.3.2/3 Type "4" . . . . .	48
<b>5. Le programme . . . . .</b>	<b>49</b>
5.1. <i>Les données . . . . .</i>	49
5.1.1. Le lexique . . . . .	49
5.1.2. Les désinences morphologiques . . . . .	49
5.1.2/1 Désinences de noms ou d'adjectifs . . . . .	49
5.1.2/2 Désinences de verbes . . . . .	50
5.1.3. Le dictionnaire . . . . .	50

	Pages
5.2. <i>Le traitement</i> . . . . .	50
5.2.1. Introduction des données . . . . .	50
5.2.1/1 Introduction du lexique . . . . .	50
5.2.1/2 Introduction des désinences . . . . .	51
5.2.1/3 Introduction du dictionnaire . . . . .	51
5.2.2. Traitement des résumés . . . . .	52
5.2.2/1 Première étape . . . . .	52
5.2.2/2 Seconde étape . . . . .	52
5.2.2/3 Troisième étape . . . . .	52
5.2.2/4 Quatrième étape . . . . .	53
5.2.2/5 Cinquième étape . . . . .	53
5.3. <i>Exploitation</i> . . . . .	54
<b>6. Résultats</b> . . . . .	55
6.1. <i>Reconnaissance des groupes de mots</i> . . . . .	55
6.2. <i>Résolution des polysémies</i> . . . . .	56
6.2.1. Résultats d'ensemble . . . . .	56
6.2.2. Algorithme "1" . . . . .	58
6.2.3. Algorithme "2" . . . . .	58
6.2.4. Algorithme "3" . . . . .	61
6.2.4/1 Contexte immédiat ("I") . . . . .	62
6.2.4/2 Contexte de la phrase ("P") . . . . .	63
6.2.4/3 Contexte du résumé ("R") . . . . .	63
6.2.5. Algorithme "4" . . . . .	65
6.2.5/1 Résolution des substantifs . . . . .	66
6.2.5/2 Résolution des adjectifs . . . . .	68
6.2.5/3 Les non-résolutions . . . . .	69
A. Structures non-reconnues . . . . .	70
B. Structures vides . . . . .	70
6.2.6. Algorithme "5" . . . . .	71
6.3. <i>Bilan d'ensemble</i> . . . . .	71

### 3e PARTIE : INDEXATION SYNTAXIQUE

<b>7. Exposé de méthode</b> . . . . .	77
7.1. <i>Nature des relations logiques</i> . . . . .	77
7.2. <i>Format des relations logiques</i> . . . . .	78
7.3. <i>Passage du langage naturel à la représentation documentaire</i> . . . . .	80
7.3.1. Méthode de reconnaissance et d'extraction des relations logiques . . . . .	82
7.3.2. Utilisation du réseau notionnel . . . . .	84

	Pages
<b>8. Analyse de l'énoncé en langage naturel</b> . . . . .	87
8.1. <i>Type d'analyse</i> . . . . .	87
8.2. <i>Exploitation de l'énoncé</i> . . . . .	88
8.3. <i>Méthode d'analyse</i> . . . . .	90
8.3.1. La reconnaissance des catégories grammaticales . . . . .	90
8.3.2. La caractérisation de la fonction . . . . .	92
8.4. <i>Résultats de l'analyse</i> . . . . .	98
<b>9. Extraction des relations logiques</b> . . . . .	101
9.1. <i>Extraction des relations consécutives et comparatives</i> . . . . .	101
9.1.1. Les mots-outils . . . . .	101
9.1.2. Les schémas syntaxiques . . . . .	103
9.1.3. Règles de construction des syntagmes . . . . .	105
9.1.4. Recherches des suppléants et des coordonnés . . . . .	106
9.1.5. Vérification sémantique . . . . .	108
9.2. <i>Extraction des relations associatives</i> . . . . .	109
9.2.1. Construction des syntagmes . . . . .	110
9.2.2. Restrictions . . . . .	111
9.2.3. Vérification sémantique . . . . .	111
<b>10. Synthèse des résultats</b> . . . . .	115
10.1. <i>Types de résultats</i> . . . . .	115
10.1.1. Des syntagmes complets . . . . .	115
10.1.2. Des syntagmes incomplets . . . . .	115
10.1.3. Des descripteurs isolés ou isolats . . . . .	116
10.2. <i>Traitement des interprétations multiples</i> . . . . .	116
10.3. <i>Opération de raccordement</i> . . . . .	117
10.4. <i>Les opérations de développement</i> . . . . .	118
<b>11. Programme</b> . . . . .	119
11.1. <i>Organisation générale</i> . . . . .	119
11.2. <i>Introduction des données permanentes</i> . . . . .	120
11.2.1. Les codes grammaticaux et fonctionnels . . . . .	121
a) <i>Forme externe</i> . . . . .	121
b) <i>Forme interne</i> . . . . .	122
11.2.2. Règles d'analyse grammaticale . . . . .	122
a) <i>Forme externe</i> . . . . .	122
b) <i>Forme interne</i> . . . . .	122

	Pages
11.2.3. Schémas . . . . .	122
a) Forme externe . . . . .	122
b) Forme interne . . . . .	123
11.2.4. Coordonnés et suppléants . . . . .	124
a) Forme externe . . . . .	124
b) Forme interne . . . . .	124
11.2.5. Règles associatives . . . . .	124
a) Forme externe . . . . .	124
b) Forme interne . . . . .	125
11.3. <i>Traitement des résumés</i> . . . . .	125
11.3.1. Lecture des résumés . . . . .	125
Traitement des parenthèses . . . . .	125
11.3.2. Interprétation syntaxique d'une phrase (organigramme 2) . . . . .	126
Format des résultats partiels (Voir Annexe 21B) . . . . .	126
11.3.3. Recherche des syntagmes à l'aide des schémas synta- xiques (organigramme 3) . . . . .	128
11.3.3.1. Orientation de l'exploration . . . . .	128
11.3.3.2. Recherche des suppléants et des coordon- nés . . . . .	128
11.3.3.3. Médiation . . . . .	130
11.3.3.4. Format des résultats partiels (Voir Annexe 21B) . . . . .	130
11.3.3.5. Interprétations multiples . . . . .	131
11.4. <i>Recherche des syntagmes associatifs</i> (organigramme 4) . . . . .	131
11.4.1. Exploration de la phrase . . . . .	131
11.4.2. Format des résultats partiels (Voir Annexe 21B) . . . . .	133
11.4.3. Les isolats . . . . .	134
11.5. <i>Les raccordements</i> . . . . .	134
11.6. <i>Impression des résultats finals</i> . . . . .	134
11.6.1. Condensation à l'intérieur d'un résumé (programme 6) . . . . .	134
11.6.2. Format des résultats finals (Voir Annexe 21C) (programme 7) . . . . .	134
11.7. <i>Exploitation</i> . . . . .	135

	Pages
<b>12. Résultats de la phase syntaxique d'indexation</b> . . . . .	137
12.1. <i>Evaluation de l'analyse syntaxique</i> . . . . .	137
12.1.1. Déroulement de l'analyse . . . . .	138
a) Arrêts dûs à des altérations dans les données . . . . .	138
b) Arrêts dûs aux règles . . . . .	138
12.1.2. Résultats de l'analyse . . . . .	140
12.2. <i>Résultats de la recherche des relations consécutives et comparatives</i> . . . . .	142
12.2.1. Evaluation quantitative des résultats . . . . .	142
a) Les mots-outils . . . . .	142
b) Les mises en relations . . . . .	143
12.2.2. Evaluation de la procédure de recherche des syntagmes . . . . .	144
a) Les syntagmes complets . . . . .	144
b) Les syntagmes incomplets . . . . .	146
c) Absence de mise en relation . . . . .	147
12.2.3. Examen des résultats en fonction de leur valeur informative . . . . .	150
12.3. <i>Résultats de la recherche des relations associatives</i> . . . . .	153
12.3.1. Evaluation qualitative des résultats . . . . .	153
12.3.2. Evaluation de la procédure de recherche des syntagmes associatifs . . . . .	154
12.3.2/1 Les mises en relation incorrectes . . . . .	154
12.3.2/2 Les isolats . . . . .	155
12.3.3. Examen des résultats en fonction de leur valeur informative . . . . .	158
12.4. <i>Valeur des opérations de condensation et de raccordement</i> . . . . .	159
<b>Conclusion</b> . . . . .	161
<b>Annexes</b> . . . . .	165

## TABLE DES ANNEXES

	Pages
ANNEXE 1 : Catégories grammaticales . . . . .	167
ANNEXE 2 : Codes morphologiques des catégories variables (autres que le verbe). . . . .	169
ANNEXE 3 : Codes morphologiques des verbes . . . . .	172
ANNEXE 4 : Reconnaissance des groupes de mots (GpM) . . . . .	175
ANNEXE 5 : Résolution des polysémies type "2" : homographies substantif-verbe (N.V) et substantif-participe présent (N.T.) . . . . .	176
ANNEXE 6 : Résolution des polysémies type "2" : homographies adjectif-verbe (A.V.) . . . . .	177
ANNEXE 7 : Résolution des polysémies type "2" : homographies substantif-adjectif (N.A.) et substantif-verbe- adjectif, (N.V.A.). . . . .	178
ANNEXE 8 : Résolution des polysémies type "2" : homographies substantif-participe passé (N.B.) . . . . .	179
ANNEXE 9 : Résolution des polysémies type "4" : analyse syntaxique . . . . .	180
ANNEXE 10 : Echantillon du dictionnaire automatique . . . . .	181
ANNEXE 11 : Echantillon du lexique . . . . .	184
ANNEXE 12 : Echantillon des règles d'analyse syntaxique . . . . .	187
ANNEXE 13 : Liste des mots outils . . . . .	193
ANNEXE 14 : Schémas consécutifs (exemples) . . . . .	198
ANNEXE 15 : Schémas comparatifs (exemples) . . . . .	201
ANNEXE 16 : Table des coordonnés et des suppléants . . . . .	204

	Pages
ANNEXE 17 : Recherche des coordonnés et des suppléants . . . . .	207
ANNEXE 18 : Règles associatives . . . . .	208
ANNEXE 19 : Echantillon du réseau notionnel . . . . .	210
ANNEXE 20 : Echantillon du listing de la sortie lexicale . . . . .	213
ANNEXE 21 : Echantillon d'analyse syntaxique . . . . .	225
21-A : Texte soumis à l'analyse syntaxique . . . . .	226
21-B : Résultat intermédiaire . . . . .	234
21-C : Résultat final de l'analyse syntaxique . . . . .	240