

# Classification par combinaison de facteurs

## Application à la documentation économique

Créé en 1957, REXECO est un centre d'information économique au service d'une quarantaine d'entreprises industrielles et financières. Il doit informer ses adhérents des contextes français et étrangers qui conditionnent leur activité. Il doit aussi les éclairer sur les perspectives d'avenir les plus probables, afin que ses adhérents puissent élaborer les stratégies les plus efficaces, les plus adaptées aux circonstances. Les soucis d'économies de moyens, de rapidité et d'actualité ne doivent pas abolir la préoccupation d'une information exhaustive et rigoureuse.

La mission de son service de documentation consiste à fournir les documents nécessaires à l'élaboration des études du centre et à renseigner les adhérents de REXECO, soit en répondant directement aux questions posées, soit en leur indiquant les sources qu'il convient de consulter pour leurs problèmes, généraux ou spécialisés. Son rôle est davantage de savoir ce qui existe en fait de documentation que de constituer sur place une sorte de dépôt légal des ouvrages d'économie dont la plupart seraient rarement utilisés. La bibliothèque de REXECO ne possède donc que les documents indispensables de façon quasi-permanente : statistiques françaises, étrangères et internationales, revues d'économie politique, de finances, ouvrages de méthodes statistiques et d'économétrie. Les grandes bibliothèques, dont la vocation est précisément d'acquérir et de conserver les ouvrages et périodiques intéressant leurs domaines respectifs, suppléent aux lacunes inévitables de la bibliothèque de REXECO, et celle-ci est heureuse de pouvoir rendre ici hommage à leur diligente coopération. Si le développement du fonds de documentation de REXECO est limité au strict indispensable, il faut, en revanche, multiplier au maximum les informations relatives aux ouvrages publiés dans le monde afin de retrouver dans les délais les plus rapides les références des publications parues sur n'importe quel sujet. Il est donc nécessaire de posséder un vaste catalogue qui puisse inclure, outre les documents détenus par REXECO, les titres cités dans les bibliographies, les catalogues d'autres bibliothèques, les ouvrages de références. La constitution d'un tel catalogue par les méthodes classiques est d'un prix de revient prohibitif, autant par la place occupée que par le travail d'établissement des fiches, et sa consultation en est malaisée.

La solution à ce double problème de stockage rapide et important des informations et de recherche accélérée à travers le fichier constitué a pu être trouvée dans le système FILMOREX qui utilise les techniques de la photographie et de l'électronique.

Les documents (sommaires, tables de matières, pages de titres, résumés) sont enregistrés photographiquement au moyen d'une caméra sur un film standard de 35 mm, en même temps que leurs différentes caractéristiques ou coordonnées, dégagées par une analyse préalable. Le film, une fois développé, est coupé au moyen d'un massicot, et l'on obtient ainsi des microfiches. Celles-ci sont divisées en deux zones, une zone « texte » sur laquelle est microphotographié le texte (ou le résumé du document), une zone « coordonnées » sur laquelle sont photographiées les coordonnées du document. Ces coordonnées sont représentées par des combinaisons de carrés, les uns noirs, les autres transparents, disposés en rangées parallèles qui correspondent à un code binaire. Chaque ligne de code peut comporter cinq chiffres et chaque microfiche peut contenir jusqu'à 14 lignes de code.

Pour la recherche et le regroupement des informations, les microfiches sont introduites dans le sélecteur FILMOREX. Le principe du sélecteur consiste à faire défiler les fiches dans un rayon lumineux excitant une cellule photo-électrique. Sur le trajet de ce rayon lumineux est placée une pellicule opaque dans laquelle on a perforé en code le critère recherché. Si une fiche porte le même critère en carrés noirs que celui qui est perforé dans la pellicule opaque, le rayon lumineux est interrompu, la cellule n'est plus excitée et agit sur un relais qui fait passer la fiche correspondante dans la case « sélection ».

Le sélecteur comporte, suivant les modèles, une, deux, trois ou quatre cellules photo-électriques, dont chacune est capable de lire une ligne entière de code, c'est-à-dire cinq chiffres à la fois, sur les fiches qui défilent à la vitesse de six cents par minute. Un système de commande permet d'actionner les différentes cellules, soit une à une, soit toutes ensemble, soit en combinaison.

Les microfiches triées par le sélecteur peuvent être utilisées de deux façons : elles peuvent être lues dans un lecteur de microfiches ou reproduites sur « photolisting », appareil qui en agrandit le texte photographiquement.

Le mécanisme est donc très simple. Le travail de dactylographie des fiches est supprimé dans la plupart des cas. Il suffit de photographier la première page du document, la table des matières d'un ouvrage, le résumé d'un article, ou une fiche bibliographique établie par un autre organisme, pour obtenir une fiche comprenant un grand nombre d'indications. Pour certaines informations de presse, la microfiche se substitue au document. La microfiche est faite en autant d'exemplaires que le document comporte de notions dominantes, chaque exemplaire étant automatiquement marqué de sa rubrique de classement. En classant ces microfiches dans leurs rubriques respectives, on constitue un préclassement qui évite de faire défiler, pour chaque recherche, tout le fichier dans le sélecteur.

Le matériel FILMOREX est conçu avec un système de codification par ligne comportant chacune vingt cases, dix cases noires et dix cases blanches, permettant ainsi 180.000 combinaisons environ (\*). Chaque code figure au bas de cartes (dites cartes vocabulaires) sur lesquelles on indique en vedette la signification attribuée et que l'on classe (alphabétiquement pour les sujets, auteurs, lieux..., chronologiquement pour les dates). Pour toute opération d'enregistrement, on puise dans ce fichier vocabulaire et on assemble les cartes codées correspondant au document à entrer.

C'est l'élaboration de ce fichier ou codification qui constitue la partie délicate du travail préparatoire à la mise en service de FILMOREX. Notamment lorsqu'il s'agit d'un langage aussi souple que celui des sciences sociales, il convient de faire appel à des spécialistes de ces disciplines plus qu'à des techniciens des systèmes de classification.

A REXECO, la mise en code des notions économiques, juridiques, sociologiques, démographiques, nécessaire pour que tout document puisse être d'abord dépouillé, puis retrouvé sans risque d'erreurs, a nécessité la collaboration des économistes du centre qui ont méthodiquement étudié et classé 4.000 notions environ. Au cours de ce long et minutieux travail, ils ont redécouvert quelques principes classiques de codification, qui méritent d'être rappelés ici.

Les notions étudiées sont presque toujours définies par certains types de coordonnées, tels l'auteur, la nature du document, le lieu, le temps, le sujet. Si la plupart de ces paramètres n'ont posé que des problèmes mineurs, il est évident que la codification des notions de sujets a été beaucoup plus difficile. Et en premier lieu leur choix n'a été résolu qu'après de délicates analyses des termes qui ont permis notamment le rapprochement des synonymes, rapprochement indispensable pour éviter l'émiettement des sujets à la recherche.

Aux vrais synonymes qui sont rares, on a ajouté les mots de sens voisin qui peuvent être assimilés à des synonymes. Ainsi le « crédit à la consommation » et la « vente à tempérament » ont une signification un peu différente mais portent le même code, car l'expérience a prouvé que les publications relatives à ces sujets sont indistinctement demandées sous l'un ou l'autre terme. Les antonymes sont également considérés comme synonymes car les notions de sens contraire intéressent une même recherche. Exemples : « Guerre = Paix », « Equilibre global = Inflation = Déflation », « Contingentement = Libération des échanges », « Dévaluation = Réévaluation ». Par contre, on a attribué des codes différents à certaines notions de sens voisin que l'on ne désire pas regrouper. Sur les fiches ayant pour vedettes ces notions la mention du second code est reportée, accompagnée d'un signe équivalent à « différent de » pour indiquer au lecteur qui souhaite trouver

---

(\*) Réduites de moitié du fait de la possibilité de division du code en deux parties (radical et terminaison).

des documents sur ces deux sujets voisins qu'il doit demander à la sélection les deux codes correspondants. Exemples : Crédit  $\neq$  Banque, Banque  $\neq$  Crédit.

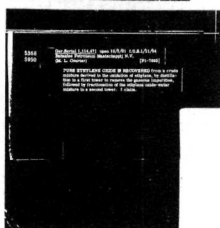
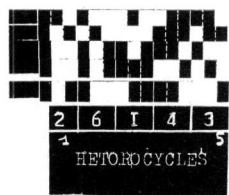
La détermination du sens des notions à coder a nécessité une analyse très précise de chaque terme et a conduit de ce fait à une étude des relations des notions entre elles. Ces rapports n'apparaissent pas toujours de façon évidente. S'il est certain que l'acier et le fer sont des métaux ferreux qui ont d'étroites liaisons entre eux, il est infiniment plus délicat de discerner les relations existant entre des notions abstraites telles que les sentiments sociaux ou les phénomènes de groupe et de les classer. Or, le système FILMOREX a prévu, grâce à la division de son code en deux parties (radical et terminaison), la possibilité d'exprimer une hiérarchie, mais une seule. Dans l'exemple précité des métaux ferreux, acier, fer, sidérurgie auront le même radical et des terminaisons différentes afin que l'on puisse obtenir à volonté, soit les documents enregistrés sur le seul sujet « acier », soit l'ensemble des documents se rapportant aux métaux ferreux, quelle qu'ait été la désinence du code à l'entrée.

On a aussi largement utilisé cette possibilité de hiérarchisation pour la codification de la chronologie, car si, à l'enregistrement, on situe les documents dans des limites de temps précises, les recherches portent, au contraire, sur des indications de dates beaucoup plus larges : « la prochaine décennie », « l'après-guerre », « l'entre-deux-guerres », etc...

L'unique hiérarchie autorisée par le système pose le problème du choix de l'échelon que l'on prend pour radical. Dans le cas des Impôts, si l'on décide de placer « Impôt » au sommet de la hiérarchie, on pourra très bien grouper sous ce radical l'impôt sur le revenu, l'impôt sur la dépense, l'impôt sur le capital, etc..., mais on ne pourra pas situer les différentes catégories d'impôt sur le revenu (impôt sur le revenu des personnes physiques, impôt sur les sociétés) par rapport à l'impôt sur le revenu. On préfère donc élire pour radicaux « Impôt sur le revenu », « Impôt sur la dépense », « Impôt sur le capital ». Cette solution interdit évidemment d'obtenir à la recherche, si l'on demande le seul code « Impôt », tous les documents enregistrés sous les radicaux « Impôts sur le revenu », « Impôt sur le capital », « Impôt sur la dépense ». Cette difficulté peut être résolue par la possibilité qu'offre FILMOREX de conjuguer plusieurs notions à l'enregistrement. Il suffit de conjuguer obligatoirement le code « Impôt » chaque fois que l'on enregistre un document sur une des catégories d'impôts.

Toutefois, ainsi qu'il a été dit plus haut, les rapports hiérarchiques entre les notions sont loin de s'imposer à l'esprit de façon absolue. En effet, alors que dans les branches techniques l'organisation est détaillée, dans les sciences humaines l'organisation objective est faible et autorise les contestations à l'égard de toute hiérarchie proposée. Aussi la conjugaison des notions offerte par FILMOREX est-elle la raison essentielle de la supériorité de ce système sur les classifications basées sur la hiérarchie.

**MICROFICHE D'UNE ETUDE THEORIQUE DE REXECO  
SUR LES MARGES BENEFICIAIRES DES ENTREPRISES EN ALLEMAGNE**



	1 - Allemagne	Paramètre de Lieu
zône	2 - Rexeco	» Auteur
code	3 - Description	» Point de vue
	4 - Théorie	» Point de vue
	5 - 1951/60	» Date
	6 - 1961	» Date
	7 - 1962	» Date
	8 - Bénéfices	» Sujet
	9 - Entreprise	» Sujet
	10 - Prix de revient	» Sujet
	11 - Distribution	» Sujet
zône texte	12 - Coût de la main- d'œuvre	» Sujet

La juxtaposition des notions est infinie, car si la conjugaison simultanée est bornée à 14 notions, le jeu des combinaisons successives permet les capacités d'enregistrement les plus étendues et les plus variées. Lieux, dates, points de vue s'ajoutent aux sujets selon les nécessités dégagées par l'analyse des publications sans l'alourdissement des charges que créerait le même travail dans un système classique de catalogue. Par ailleurs, la conjugaison des notions a autorisé la décomposition de ces dernières en leurs éléments constitutifs. C'est ainsi que le terme « Agriculteurs » exprimant l'association de deux notions : les travailleurs dans l'agriculture, n'a pas de code particulier, mais est traduit par la conjugaison des codes « Travailleurs » et « Agriculture ». A l'enregistrement, on introduit ces codes ; à la sélection, ou bien l'on désire avoir la notion complète et l'on demande les deux codes, ou bien l'on effectue une recherche soit sur l'agriculture, soit sur la population active, et dans chacun des deux cas le document est sélectionné. La dispersion des sujets à la recherche est ainsi souvent évitée, si bien que, chaque fois que l'on a pu décomposer les notions à coder, on a adopté la solution de la conjugaison des éléments. En voici quelques exemples :

Dollar = Monnaie + Etats-Unis.

Budgets économiques = Comptes nationaux + Prévision.

Emprunts publics = Emission + Titres d'état.

La description des problèmes de la codification et des solutions que l'on s'est efforcé d'y apporter pourrait proscrire l'emploi de systèmes mécanisés, si l'on ne retenait que l'importance de l'investissement initial immobilisant pour une période de temps relativement longue des spécialistes. Mais deux ans de fonctionnement du système FILMOREX permettent d'apprécier l'efficacité de cet investissement et de justifier ce dernier. Au cours de ce laps de temps, on a constaté que l'enregistrement et la recherche des documents

ont pu être effectués par différentes personnes qui connaissaient les principes de la codification mais n'avaient pas été obligées de retenir les notions conservées, le langage de la machine ayant la même souplesse que celui des économistes. Par ailleurs, il a été vérifié que dans les concepts des sciences humaines, l'organisation hiérarchique est un cadre trop rigide pour la complexité des sujets et la diversité des éclairages que l'on peut souhaiter, ce qui explique la difficulté des classements dits méthodiques ou par matières, alors que la conjonction des notions, avec ses combinaisons en nombre pratiquement illimité est la procédure suivie le plus souvent par l'esprit humain. Facilitant les dépouillements détaillés des publications, qu'interdirait toute méthode classique d'indexation, la conjugaison multiple des notions permet de retrouver une quantité d'informations à la fois volumineuse et précise.

Marie-Jeanne MAKSUD,

Bibliothécaire de REXECO,

(Centre de recherches pour l'expansion de l'Economie).