

Taxinomies, thésaurus et ontologies - Article de R. TEXIER

1 Les vocabulaires contrôlés

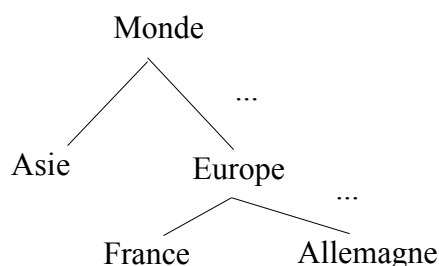
Lorsqu'on souhaite constituer un fond documentaire ou une base de connaissance, qu'elle soit informatisée ou non, on se pose très rapidement la question de son organisation : comment structurer, comment classifier tout ce contenu de façon pérenne et claire pour tous les utilisateurs afin qu'ils puissent y retrouver facilement les informations dont ils ont besoin ?

Très tôt, les premiers « documentalistes », ou bibliothécaires, ont adressé le problème en classant leur fond par rubriques, sous-rubriques, ... définies avec soin en évitant les termes confus ou imprécis et en respectant une logique stricte issue du domaine ou de la discipline couverte par le fond documentaire à organiser. On obtient ainsi des terminologies normalisées (du moins, localement) permettant d'indexer sans ambiguïté n'importe quel document, le rendant ainsi rapidement accessible aux lecteurs.

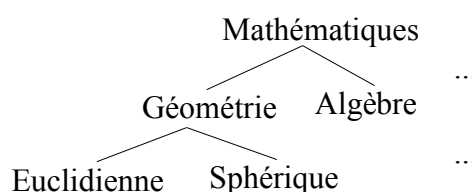
Ces terminologies et les relations entre les termes qu'elles contiennent constituent ce qu'on appelle des *vocabulaires contrôlés*. Cette approche « de bon sens » a fait ses preuves, et les vocabulaires contrôlés sont d'usage courant aujourd'hui, dans tous les domaines. Ils ont même connu un nouvel essor, sous des formes de plus en plus sophistiquées, avec la généralisation de l'informatique. De plus, l'explosion des autoroutes de l'information qui génèrent et diffusent des masses toujours plus importantes de données et de connaissances, rend le besoin d'organisation et de structuration des contenus toujours plus aigu.

2 Les taxinomies

La taxinomie est la forme la plus simple, la plus intuitive et la plus utilisée des vocabulaires contrôlés : elle se présente sous la forme d'une hiérarchie de termes généralement organisés dans le sens de la spécialisation (du général au particulier) bien que d'autres organisations soient possible (imbrication, composition, ...). On y représente qu'un seul type de relation à la fois : par exemple, si vous classez selon une règle de spécialisation, vous ne pouvez changer en cours de route et classer selon la taille.



Relation de division



Relation de spécialisation

3 Les thésaurus

Le thésaurus est une taxinomie sur laquelle on ajoute des relations « horizontales » en plus de la relation verticale qui structure la taxinomie originale. Ces relations expriment une connexité entre sujets par des liens du type « relatif à », « voir aussi » ou « synonyme de », ... et permettent d'élargir le champ d'étude ou de recherche.

Les thésaurus sont aussi utilisés pour construire des taxinomies multilingues en ajoutant des relations du type « se traduit en L par... ». Cette approche n'est valable que si on est sûr que la modélisation choisie est universelle (indépendante de la langue et de la culture) ce qui est rarement le cas en dehors des domaines techniques et scientifiques.

4 Les Ontologies

Le terme ontologie appartient d'abord à la philosophie, et plus particulièrement à la métaphysique. C'est la discipline qui traite de l'être en tant qu'être. Rien à voir donc, à l'origine, avec les vocabulaires contrôlés. C'est donc par un curieux détournement (peut-être inspiré par les *catégories* d'Aristote) que ce terme en est venu à désigner, dans le monde des sciences de l'information, la forme la plus sophistiquée des vocabulaires contrôlés.

Dans une ontologie, non seulement on normalise un ensemble de termes dans un lexique, mais on formalise et explicite aussi les relations qui les lient. On définit donc ainsi une grammaire (et donc un langage formel) sur ces termes et relations.

Exemple partiel :

<i>Représentant de</i>	
chat	mammifère
chien	mammifère
souris	mammifère

<i>Mange</i>	
chat	souris
chien	croquettes
souris	fromage

Le tableau de gauche peut aussi être représenté par une hiérarchie car cette relation est typique des taxinomies.

Ainsi, puisque les relations sont explicitées, désignées, on peut effectuer des recherches à leur sujet : « quels sont les mammifères ? », « qui mange souris ? », « que mange le chien ? », ... De plus, par chaînage, on peut vérifier des assertions du type « Les chats sont des mammifères qui mangent des souris ».

Les ontologies ne sont pas sans rappeler les *réseaux sémantiques*.

Les taxinomies et les thésaurus peuvent être considérés comme des ontologies dont les relations seraient restées implicites car « évidentes » pour le lecteur (mais pas pour l'ordinateur).

5 L'approche Easy KM™

Easy KM d'Elikya est une solution permettant de deployer un système de partage des connaissances rapidement au sein d'une organisation et avec une facilité d'accès et de prise en main maximum par les utilisateurs.

La taxinomie est la forme de vocabulaire contrôlé la plus connue, la plus intuitive et la mieux maîtrisée par le public. C'est donc celle qui est nativement supportée par Easy KM via le principe des « arbres de classement ». C'est aussi celle qui permet le mieux la définition de *profils* pour décrire les besoins et les centres d'intérêt – ou d'expertise – de chaque utilisateur.

Mais Easy KM va plus loin, de manière à concilier l'intuitivité des taxinomies avec le pouvoir d'expression des ontologies.

Les experts d'une organisation utilisant Easy KM définissent et font vivre autant de taxinomies différentes que nécessite le domaine d'activité ou d'investigation de leur « communauté ». Les documents ou éléments de connaissances apportés par les membres de l'organisation viennent ensuite peupler ces taxinomies comme autant de feuilles – ou éléments terminaux – de ces hiérarchies.

Cependant, un élément de connaissance donné peut figurer simultanément sur autant de taxinomies différentes qu'il vérifie de critères pertinents et même en plusieurs positions au sein d'une même taxinomie (propriétés cumulées).

Cette propriété offerte par Easy KM ouvre sur une puissance d'expression comparable à celle des ontologies, la complexité en moins.

Exemple :

Imaginons que nous ayons construit deux taxinomies, une sur la classifications des animaux (mammifères, reptiles, ... et sous catégories), disons A, et une autre sur les régimes alimentaires, disons B. Si on dispose d'un documents sur les chats, on a la possibilité de le classer simultanément parmi les mammifères (A) et les mangeurs de rongeurs, les mangeurs de petits oiseaux voire aussi dans les buveurs de lait (propriétés cumulées dans B). Et ainsi de suite dans autant de taxinomies pertinentes parmi celles dont on dispose (par exemple une sur le marché des animaux domestiques, ...).

Ainsi, lors d'une recherche qui nous retournerait ce document sur les chats, on aurait la possibilité, en un seul clic, d'ouvrir nos investigations sur l'ensemble des mammifères ou l'ensemble des buveurs de lait, ...

Si on regarde l'exemple donné en 4 sur les ontologies, on se rend compte que l'on peut répondre aux mêmes questions et vérifier les mêmes assertions.

Tout en étant simple d'utilisation, et avec un apprentissage minimum, Easy KM permet donc de mettre en œuvre des systèmes puissants et expressifs au service du management du savoir.