

Quels systèmes conceptuels peut-on préconiser pour les banques de données terminologiques en ligne?

La diffusion de ressources terminologiques au moyen de réseaux interactifs qui relient les centres de documentation présupposent l'accès aux archives de chacun de ces centres, qui emploient des systèmes d'organisation différents. La spécificité des banques de données terminologiques et la variété de systèmes de classifications en usage rendent très complexe la recherche en ligne d'informations terminologiques. En soulignant la nécessité d'impulser des études dans ce domaine d'une manière coopérative, cet article se propose d'examiner une solution directe qui permet de conserver l'organisation actuelle des archives existantes, et de parvenir à l'exploitation d'un système unique et complexe de connaissances, ouvert et multifonctionnel. Ce système devrait être solidement ancré dans une structure ontologique, et représenter une «cartographie» conceptuelle qui propose à l'utilisateur un parcours rationnel qui guide sa recherche. Nous indiquons en guise de conclusion quelques exemples de comparaisons et de corrélations parmi d'importants systèmes de classification.

Termes-clés : terminologie ; classification ; ontologie ; documentation.

1 Plurilinguisme et société de l'information

« **L**e plurilinguisme n'est pas, institutionnellement, un problème ; c'est un état de l'humanité... Le mouvement qui promet aujourd'hui en Europe de libérer et de faire revivre les identités ethniques, culturelles et linguistiques réprimées et opprimées est légitime et porté par le sentiment diffus de la fragilité de ce qui est libre et intègre. Nous devons cependant nous demander comment ce mouvement peut coexister avec la tendance (que les Américains appellent la globalisation) de faire du monde un réseau de communication qui dépasse et qui annule les frontières traditionnelles. Les communications que les réseaux informatiques d'aujourd'hui rendent (nous devrions dire, imposent) ininterrompues et instantanées, éliminant les délais et les erreurs produites par la diversité des langues et par l'obstacle de la traduction. Comment un tel mode de communication, orgueil de l'ingénierie des États-Unis, peut-il accepter les rythmes d'un humanisme respectueux des diversités culturelles et linguistiques, tel qu'il est exigé dans les organismes de la Communauté européenne, déjà très accaparés par la traduction de leurs propres textes dans les langues et dans les terminologies des États membres? » (Nencioni, 1997: 1-2).

Une réponse à cette question est fournie en partie par le Conseil de l'Union européenne dans sa décision

du 4 novembre 1996. Cette instance propose une politique innovatrice dans le secteur de la langue dont le but est la défense des différentes identités culturelles, leur valorisation et leur intégration pour la promotion du processus culturel, social, économique et productif dans la totalité de la Communauté européenne. La décision, qui vise l'adoption d'un programme communautaire pour la promotion de la diversité des langues dans la société de l'information, souligne dans ses prémisses l'importance des aspects culturels et linguistiques dans cette nouvelle société, importance qui trouve un écho jusqu'à la Conférence des ministres du G7 tenue à Bruxelles en février 1995.

Il ressort de la politique commune européenne le signe et la volonté d'imposer dans le circuit de la communication internationale le respect de l'identité historique et culturelle des différents pays et de valoriser ces ressources par le biais de la mondialisation.

Parmi les motivations qui ont présidé à la formulation de ce programme, retenons les suivantes :
 – l'avènement de la société de l'information pourrait permettre aux citoyens européens de mieux accéder à l'information et de saisir cette occasion unique de bénéficier de la richesse et de la diversité culturelle et linguistique de l'Europe ;
 – il faut s'efforcer de garantir à chaque citoyen européen l'égalité en matière de participation à la société de l'information, indépendamment de ses conditions sociales, culturelles, linguistiques ou géographiques ;

– il est essentiel de garantir à chaque citoyen européen le même accès à l'information; cette information doit être mise à sa disposition dans sa propre langue;

– les langues qui restent en dehors de la société de l'information risquent de se voir marginalisées de façon plus ou moins rapide. (Décision du Conseil de l'Union européenne du 4 novembre 1996, art 6, 10, 11, 12). Il est donc permis d'espérer que «le XXI^e siècle sera celui de la terminologie multilingue. L'essor du multilinguisme est un justificatif de taille pour inciter à la collaboration et au partage. S'il revêt un intérêt marqué, le multilinguisme devra constituer un des principaux axes de développement dans les années à venir, malgré ses coûts élevés et d'implantation et d'exploitation pour les banques de terminologie.» (Tremblay 1996: 149-150.)

2 Le patrimoine terminologique et les banques de données

Par langues de spécialité (ou technolectes) on entend le lexique qui permet d'exprimer les concepts et les connaissances propres aux domaines scientifique, technique ou professionnel, caractérisées par le fait qu'on s'adresse à une classe donnée de scientifiques et de professionnels qui partagent ces connaissances. Il existe une gamme assez large de langues spécialisées, représentée non seulement par le nombre de disciplines auxquelles ils se rapportent, mais aussi par le niveau de spécialisation scientifique ou technique. Plus l'unité terminologique a un degré élevé de spécialisation, moins elle est fréquente dans l'usage.

Il arrive souvent aussi que la spécialisation très poussée d'une discipline dépasse les limites de la

discipline même pour déborder dans des secteurs particuliers d'autres disciplines. Que l'on pense par exemple en cardiologie aux études de nature médicale et technologique qui concernent le rejet des valves cardiaques, en sénologie au diagnostic par l'imagerie ou en acoustique aux aspects physiques, biologiques et sociaux de la pollution sonore.

Le «terme» est l'unité de base des langues de spécialité. Pour qu'un mot soit défini comme terme, il doit se distinguer de l'ensemble d'autres éléments. La définition est une opération logique et linguistique destinée à fixer une limite ou une fin (définir), et le terme est le résultat de cette opération: «le terme désigne à la fois une "opération logique", et la "production d'une séquence en langue naturelle"; il désigne à la fois ces "opérations" et leurs "résultats"» (Rey 1979: 40).

Le langage partagé par la communauté à laquelle il est destiné est un instrument direct de communication, dans la mesure où il retranscrit, grâce aux termes propres qui le composent, ce qu'il est sensé définir. Cette condition est remplie si les termes sont sans ambiguïté et leur sens est univoque.

Anne Guyon fait une comparaison très juste entre les termes employés dans un texte scientifique ou technique et les briques conceptuelles avec lesquelles une information se construit (Guyon 1993: 188). Pour comprendre l'information il convient de connaître les éléments de base, à savoir les termes.

Les glossaires, les dictionnaires, les thésaurus, les banques de données et autres sont de précieux recueils de terminologie, élaborés et gérés par des centres spécialisés pour parvenir à différentes fins institutionnelles. Ces recueils font aujourd'hui l'objet de recherches pour leur contenu conceptuel et terminologique. Les centres qui gèrent ces ressources,

classent et ordonnent le contenu de ces recueils selon des systèmes de classification consolidés ou originaux, de caractère général ou particulier, qui permettent la localisation et la consultation.

La constitution de réseaux interactifs, que relient les centres d'informations qui constituent les pôles du réseau, doit prévoir l'accès aux archives, organisées selon différents systèmes de classement. Les recherches d'information en ligne auront des résultats satisfaisants si on rationalise l'accès aux ressources terminologiques détenues par ces centres. Les collections qui ne seront pas accessibles par ce moyen et qui ne pourront être consultées perdront de leur utilité.

Pour récupérer les informations en ligne, la meilleure solution passerait par le respect des critères suivants:

- a) le partage d'un système de classification unique, qui comporte dans sa structure hiérarchique les informations les plus spécifiques correspondant au contenu de chaque archive;
- b) une liste des pôles avec les disponibilités des ressources dans un secteur particulier;

Ce genre de solution n'est cependant pas réaliste à cause de:

- La diversité des systèmes de classement en usage dont on a précédemment parlé;
- La spécificité du contenu d'une collection qui, la plupart du temps, a été établie avec une analyse conceptuelle rigoureuse qui de ce fait, devient difficilement intégrable dans un système collectif de caractère général;
- La multidisciplinarité qui caractérise généralement un patrimoine terminologique, ce qui rend difficile l'établissement d'un relevé correct.

L'utilisateur pourrait se faire guider de façon heureuse dans le processus rationnel de recherche si les

ressources terminologiques, ainsi que les ressources qui se réfèrent à un même thème, sont relevées avec une indication conceptuelle analytique qui distingue les différents contenus qui caractérisent les ressources elles-mêmes.

Le problème de la spécificité des banques de données terminologiques est mise en évidence par Louis-Claude Tremblay: «Les banques sont mises au point en fonction d'un mandat, d'une mission et d'objectifs bien précis, ce qui donne lieu à l'exploitation des champs de spécialité propres à la clientèle desservie.» (Tremblay 1996: 149). Si des échanges d'informations sont introduits et que la banque de données doit assumer des fonctions plus vastes et plus diverses que celles pour lesquelles elle était conçue, il en résulterait peut-être un alourdissement inutile pour l'utilisateur. Il faut néanmoins encourager une exploitation plus grande de ce patrimoine, en mettant en évidence les éléments de différenciation d'une banque de données qui puissent constituer «des avantages marqués en situation de vente, en situation de marketing». Tremblay note aussi que les clés d'accès au système d'information en réseau doivent être «à la fois significatives et précises. Ces clés sont des termes, des expressions justes dont la définition relève de l'activité terminologique. Une terminologie juste permet l'accès plus ciblé à cette terminologie». Et à ce sujet, l'auteur entrevoit de nouveaux thèmes de recherche, qui marqueront les futures orientations des banques de données terminologiques.

3 Une carte, guide du processus de recherche de l'utilisateur

Quelles sont les «expressions justes» qui permettent d'accéder avec

précision au système? Ce sont les termes qui représentent le contenu conceptuel de l'archive terminologique, le *concept-objet* en question et de façon plus spécifique les concepts corrélés à celui-ci (caractéristiques, procédés, actions, théories) qui révèlent des aspects particuliers de l'objet selon différents critères, sans nécessairement tenir compte de limites disciplinaires.

Un bon relevé sémantique d'un patrimoine terminologique doit mettre en évidence le concept-objet et ses facettes, les aspects traités et les critères qui ont présidé à sa création. Si ces aspects du problème ne sont pas suffisamment mis en lumière, le contenu de la banque de données peut être classé de façon incorrecte et par conséquent, donner des résultats de recherche décevants. Un lexique de la neige, qui par exemple, traite d'un lieu géographique caractéristique, avec ses phénomènes propres, et donc ses dénominations propres, se rapporte aussi aux aspects anthropologiques qui tiennent compte des traditions et de l'impact culturel que la neige a toujours suscité chez l'homme, aux aspects folkloriques, littéraires, psychologiques, sans oublier les divertissements, les sports associés, etc. Dans ce cas, une classification de la terminologie de la neige comme sous-classe de la météorologie ne serait pas représentative de tout le contenu du lexique.

Il en résulte qu'un concept-objet peut relever de n'importe quel domaine du savoir, et pour lui attribuer une identification adéquate pour les besoins de la classification, il faut donc identifier la fonctionnalité du concept-objet. Ainsi émerge la conception d'un système de références multidimensionnel, qui peut représenter l'objet sous différents points de vue et être indexé correctement pour le signalement du contenu sémantique du recueil.

L'activité de collaboration en vue de constituer un réseau d'informations terminologiques doit aussi viser l'étude de processus cognitifs, de méthodologies et d'outils, qui tout en valorisant les expériences acquises, permettent d'accéder rationnellement à des systèmes renforcés s'ils sont jugés satisfaisants, mais en introduisant des parcours cognitifs nouveaux et alternatifs.

Parmi les conclusions du séminaire «Compatibilité et intégration des systèmes de classement», on souligne à quel point l'infrastructure d'un système de documentation en réseau est valorisée par deux composantes essentielles, la télécommunication à l'échelle mondiale et les réseaux électroniques, mais on confirme aussi qu'en plus de cette dernière composante, il est indispensable de disposer d'une infrastructure qui mette en valeur l'organisation des connaissances et des informations (Recommandations 1995).

Une solution à ce problème de la réalisation d'une infrastructure de cette nature pourrait venir de l'analyse des systèmes de classification en usage, de l'étude de leur compatibilité, de leurs relations et de leur intégration éventuelle ainsi que de l'élaboration d'une cartographie conceptuelle. Celui qui cherche de l'information pourra faire appel à une telle cartographie en empruntant un parcours logique personnalisé, qui reflète son degré de spécialisation, ses connaissances et ses attentes en matière de documentation.

Cette cartographie devra être fondée sur des concepts-objets généraux, abstraits et concrets, qui à leur tour comprendront, justement dans le but de rester généraux, des concepts plus spécifiques corrélés (*clusters*). Il est possible d'organiser pour les concepts en rapport avec chacun des clusters une structure réalisée selon un modèle commun,

qui identifie un réseau conceptuel primaire pour être mis en relation avec d'autres systèmes, universels et particuliers, et constituer ainsi un réseau très complexe, mais en gardant la possibilité de proposer un échafaudage (*une structuration*) ontologique. L'utilisateur qui interroge le système de documentation devra être en mesure de suivre sur la cartographie qu'on lui propose un parcours plus rationnel pour sa recherche. Il est par conséquent fondamental d'étudier la compatibilité des systèmes de classement et leurs corrélations pour définir un système de connaissances qui permette à l'utilisateur de visualiser automatiquement l'organisation du secteur en rapport avec l'objet de sa recherche et de trouver le centre d'information qui dispose de telles ressources.

4 Recommandations sur la compatibilité et l'intégration des systèmes de classement

L'intégration des systèmes conceptuels, universels et particuliers, est un sujet souvent débattu sur le plans de la théorie, de la méthodologie et des applications.

Nous venons de faire mention du séminaire de recherche, «Compatibilité et intégration des systèmes de classement», organisé en 1995 à Varsovie par l'ISKO (*International Society for Knowledge Organisation*) et la *Society for Professional Information*. Les recommandations, qui regroupent des conclusions partagées par des chercheurs de douze pays, revêtent une importance particulière dans le cadre de l'argument présenté dans cet article, tant pour l'objectif commun qui est d'organiser un système d'information à l'échelle planétaire,

que pour les suggestions qu'ils fournissent. Si l'on est conscient de la nécessité de partager l'infrastructure électronique et télématique pour organiser un système de documentation, il ne semble pas possible d'obtenir un tel consensus sur l'importance d'une infrastructure conceptuelle qui sert de support à la précédente, ni sur les efforts à consentir pour harmoniser et intégrer des systèmes de classification des connaissances.

Des recommandations sont plus particulièrement destinées:

- a) aux problèmes de recherche, de formation et d'échange d'informations;
- b) au développement de systèmes et de collaborations aux niveaux national et international.

En ce qui concerne le premier point, il convient d'encourager les études des principes et les méthodologies nécessaires pour établir les corrélations entre les systèmes de classement ayant des structures différentes. La *compatibilité* ne doit pas être poursuivie comme une fin en soi, mais pour les objectifs spécifiques à identifier au cas par cas, en évaluant les avantages associés à chaque solution.

Il est opportun de promouvoir la formation de spécialistes de l'information et de l'organisation de la connaissance, y compris de la représentation des connaissances, la classification et l'indexation, ainsi que la terminologie.

En ce qui concerne le développement de systèmes et de collaborations nationales et internationales, les recommandations mettent en évidence la nécessité d'accroître les corrélations entre les systèmes de classements existants et le rendement d'une base de connaissances, conceptuelle et terminologique, ouverte et multifonctionnelle, multilingue et intégrée, qui ne modifie d'aucune manière les différentes sources qui

l'alimentent. Un système ainsi conçu devra être en mesure de satisfaire non seulement les exigences de celui qui fait partie d'un tel réseau mais également celles de l'utilisateur final qui fait des recherches sur l'internet et à partir d'autres services en ligne, disponibles en différentes langues, autres que ceux utilisés pour les bases de données du réseau. En plus, le système pourra assumer le rôle d'outil d'indexation, y compris au niveau européen, de base conceptuelle pour des systèmes fondés sur la connaissance; de dictionnaire sous forme monolingue, bilingue ou plurilingue pour une utilisation commune; de dictionnaire /base de connaissances pour le traitement automatique du langage, comprenant l'usage de la langue naturelle et la traduction automatique; de classification/ontologie pour la normalisation des éléments donnés. Une base de connaissances ouverte à de tels développements sera alimentée par un effort coopératif qui vise à tirer les meilleurs résultats des informations provenant de différentes sources, mutuellement complémentaires et reliées.

La coopération internationale devra être attentive à définir et à maintenir des listes d'autorités de noms de personnes, d'organisations, etc.

Parmi les recommandations on compte aussi le développement de critères, de méthodologies, d'outils et de logiciels pour définir, gérer, convertir, harmoniser, restructurer et traduire des systèmes de classements monolingue et plurilingue, des systèmes de classification, des thésaurus et autres listes terminologiques comme les dictionnaires, en gardant comme perspective le contrôle de leur quantité.

5 Subjectivité et objectivité d'un système de classification

Une analyse rigoureuse menée par Ingetraut Dahlberg sur les principes qui dans l'histoire de la pensée philosophique, à partir de Platon, ont inspiré l'organisation de la connaissance, arrive à la conclusion que les classifications du savoir jusqu'à nos jours sont caractérisées par des principes subjectifs (Dahlberg 1995: 131). En particulier, ses études ont mis en lumière comment la division de champs de savoir (*Domaines de connaissances*) liés aux capacités humaines, telles que la mémoire, l'imagination et le raisonnement, était attribuée de façon hiérarchique à ces mêmes capacités, de telle sorte que celles-ci représentent les principales divisions du savoir. Dans la trilogie de Francis Bacon, par exemple l'*Histoire*, avec ses subdivisions, est mise en relation avec la mémoire, la *Poésie* avec l'imagination, la *Science* ou la *Philosophie* avec la raison ou l'intelligence.

Passant en revue l'histoire des systèmes de classification des bibliothèques, Dahlberg met l'accent sur l'influence que le schéma philosophique de Bacon a exercée sur le développement de nouveaux systèmes de classement pour ce qui concerne le regroupement et la séquence, mais aussi et d'une certaine manière la conception de ses classes principales. On en veut pour preuve les systèmes de W.T. Harris de 1870 et celui de M. Dewey de 1876, qui sont certainement inspirés du précédent. La classification de Dewey est parmi celles qui sont aujourd'hui encore les plus utilisées, et c'est également le modèle sur lequel Otlet et Lafontaine se sont appuyés pour développer la Classification décimale universelle.

L'être humain et ses capacités ont donc inspiré la construction de classes intégrant les connaissances relatives aux sciences, aux arts et aux technologies.

Il serait souhaitable qu'un système de classification universelle soit fondé sur des principes ontologiques.

Un système universel fondé sur des objets généraux de l'être et sur l'usage d'un modèle de classement logique et systématique, qui permet de relever pour chaque objet les propriétés, la théorie et les méthodologies qui le caractérisent, relèverait des classes relatives aux objets simples, représentant une base objective aussi bien pour celui qui développe le système que pour celui qui l'utilise.

6 Information Coding Classification

L'ICC (*Information Coding Classification*) est fondée sur des principes ontologiques et sur la *théorie des niveaux intégratifs*.

Ce système de classification universelle, projeté et défini par Ingetraut Dahlberg, a été utilisé pour la codification des thèmes des banques de terminologie du CIRT (*Centro Italiano di Riferimento per la Terminologia tecnico-scientifica*) et c'est à cette expérience particulière que nous faisons référence ici.

Avant cette expérience, la volonté d'approfondir l'étude de la théorie avait été mise en valeur par les quelques résultats obtenus qui en confirment le bien-fondé. À la suite d'une expérimentation pratique relative à l'organisation systématique de termes relevant d'un secteur disciplinaire spécifique, des regroupements de termes avaient été constitués (clusters) sur la base de critères logiques qui ne tenaient pas compte d'un modèle préétabli: de tels

critères se sont révélés très similaires du modèle de catégories proposé par la chercheuse allemande, appelé «systématificateur».

La conviction de la validité pratique des principes qui ont inspiré l'ICC est également fondée sur l'utilité finale de la théorie. Les principes sur lesquels la classification est fondée permettent en effet de donner un ordre systématique à chaque champ du savoir et en particulier, à ces nouveaux thèmes, qui naissent de la recherche et qui se développent progressivement jusqu'à acquérir une connotation spécifique dans le domaine scientifique. De tels principes permettent d'organiser de façon systématique les informations relatives à un domaine particulier ou au contenu d'une banque de données spécialisée, et impliquent en premier lieu le chercheur qui, dans l'exercice de ses activités scientifiques, est directement impliqué dans le processus de l'évolution du savoir dans ce domaine spécifique.

6.1 Description du système

L'*Information Coding Classification* (ICC) sur la base de la théorie des niveaux intégratifs est représentée comme une matrice à double entrée (tableau 1).

Une entrée vise à relever les *objets* de la réalité avec toutes les unités de connaissance qui y sont associées. L'«objet» de la réalité avec ses prédicats, des concepts associés à celui-ci, est le champ de connaissance. L'autre entrée de la matrice relève en revanche du *mode d'analyser* l'objet.

Selon la théorie des niveaux intégratifs, on a mis en valeur les objets généraux correspondant à *neuf niveaux de l'être* et qui constituent ce qu'on appelle les «catégories d'êtres», et les *aspects* qui se rapportent à ces objets et qui correspondent à neuf «catégories de forme». Tous les concepts relatifs à un objet général,

0 GENERAL FORM	01 THEORIES, PRINCIPLES	02 OBJECT COMPONENT	03 ACTIVITY PROCESS	04 PROPERTY ATTRIBUTE	05 PERSONS OR CONT'D	06 INSTITUTION OR CONT'D	07 TECHNOLOGY & PRODUCTION	08 APPLICATION & DETERMINAT.	09 DISTRIBUTION & SYNTHESIS
1 FORM & STRUCTURE AREA	11 Logic	12 Mathematics	13 Statistics	14 Systemology	15 Organization Science	16 Metrology	17 Cybernetics, Control & Automation	18 Standardization	19 Testing and Monitoring
2 ENERGY & MATTER AREA	21 Mechanics	22 Physcis of Matter	23 General and Technical Physics	24 Electronics	25 Physical Chemistry	26 Pure Chemistry	27 Chemical Technology & Engineering	28 Energy Science and Technology	29 Electrical Engineering
3 COSMO & GEO-AREA	31 Astronomy & Astrophysics	32 Astronatics & Space Research	33 Basic Geosciences	34 Atmospheric Sciences & Technology	35 Hydrospheic & Oceanol. Sci. & Technology	36 Geological Sciences	37 Mining	38 Materials Science & Technology	39 Geography
4 BIO-AREA	41 Basic biological Sciences	42 Microbiology and Cultivaiton	43 Plant Biology and Cultivation	44 Animal Biology and Breeding	45 Veterinary Sciences	46 Agriculture & Horticulture	47 Forestry & Wood Sci. & Technology	48 Food Sciences and Technology	49 Ecology and Environment
5 HUMAN AREA	51 Human Biology	52 Health and Theoretical Medicine	53 Pathology and Practical Medicine	54 Clinical Medicine & Cure	55 Psychology	56 Education	57 Profession Sci., Labor., Leisure	58 Sport Science and Sports	59 Household and Home Life
6 SOCIO AREA	61 Sociology	62 State and Politics	63 Public Administration	64 Money and Finances	65 Social Aid, Social Politics	66 Law	67 Area Planning, Urbanism	68 Military Science and Technology	69 History Science and History
7 ECONOMICS & TECHNOLOGY AREA	71 General and National Economics	72 Business Economics	73 Technology in general	74 Mechanical & Precision Engineering	75 Building	76 Commodity Science & Technology	77 Vehicle Science and Technology	78 Transportation Technology & Services	79 Utilities and Service Economics
8 SCIENCE & INFORMATION AREA	81 Science of Science	82 Information Science	83 Informatics, computer science	84 Information in general	85 Communitat. Science	86 Mass Communication	87 Printing and Publishing	88 Communication Engineering	89 Semiotics
9 CULTURE AREA	91 Language and Linguistics	92 Literature and Philology	93 Music and Musicology	94 Fine Arts	95 Performing Art	96 Culture,Sciences, narrower sense	97 Philosophy	98 Religion and Secret Teachings	99 Christian Religion

Tableau 1 :
Information Coding Classification
Survey of Subject Groups © 1982, rev. 1992 I. Dahlberg, Frankfurt/Main, Germany

analysés selon les catégories de forme, constituent des regroupements de concepts associés qui forment des clusters; chaque cluster, à son tour, analysé selon les mêmes catégories, génère des clusters de concepts plus spécifiques. C'est ainsi que le processus peut être poursuivi à l'infini.

Chaque cluster de concepts est un objet de la réalité et représente un champ de connaissance. L'ICC est articulé de telle sorte qu'il permet d'ordonner n'importe lequel de ces champs grâce aux catégories de forme (modèle systématificateur).

6.2 Caractéristiques du système

Le choix de l'ICC est dicté par certains avantages importants du système :

1) la possibilité de caractériser le champ de connaissance, c'est-à-dire le thème de la banque de données terminologiques, à partir de *niveaux primitifs de catégories* et non d'une liste de disciplines. Cette qualité du système permet de relever un thème selon des points de vue différents et de les représenter par des notations différentes. Comme nous l'avons déjà indiqué, le sujet d'un fonds terminologique, même hautement spécifique, n'intéresse pas qu'une seule discipline définie traditionnellement par des divisions consolidées du savoir, mais manifeste plutôt, grâce à sa haute spécificité, une individualité propre exprimée par différents concepts afférents à plusieurs disciplines, concepts qui peuvent être regroupés selon divers aspects et critères, mais qui trouvent dans cet argument leur intégration et leur unicité. Par exemple, la culture de céréales, classée sous la cote 465, *Production végétale* (sous-classe de 46 Agriculture et horticulture; 4 domaine de la biologie); 43 *Biologie végétale et culture de plantes* (sous-classe de 4, domaine de

la biologie); 412 *Génétique* (sous-classe de 41, Sciences biologiques élémentaires; 4, domaine Biologie); 485, *Aliments végétaux* (sous-classe de 48, Sciences alimentaires et technologie; 4, domaine Biologie).

2) L'ICC comporte uniquement des classes qui définissent les champs de connaissance. Ces classes se réfèrent à des concepts généraux qui peuvent être mis en relation avec des concepts généraux d'autres systèmes. Cette caractéristique est un outil efficace pour « cartographier » des concepts hiérarchiquement plus élevés qui appartiennent à d'autres systèmes de classement.

3) Les dénominations des classes sont solidement ancrées à une définition (en langue allemande). Un classement systématique est construit sur cette base, et une répartition dans les bonnes classes est ainsi assurée.

4) Les principes de classement sur lesquels la classification est fondée permettent de représenter des concepts composés, et en particulier de relever les deux composantes d'une combinaison et leur relation syntaxique (par exemple Objet+activité/application). La relation syntaxique entre un sujet et le concept qui en forme le prédicat est ainsi expliquée.

Par exemple, « 351-428 » indique *Microbiologie de l'eau* (351 correspond à *Principes généraux de l'hydrographie et de l'océanographie*, 428 correspond à *Applications de la microbiologie*). Respectant les catégories de forme, la combinaison se réfère au niveau « 1 » pour indiquer les principes généraux du sujet (dans ce cas précis 351, où 35 est Hydrographie et océanographie) et au niveau « 8 », pour indiquer l'application (dans ce cas 428, où 42 est Microbiologie et culture). Ce genre de procédure est très utile pour incorporer de nouveaux sujets.

7 Comparabilité des systèmes

Le travail de mise en relation de l'*Information Coding Classification* et de la Classification décimale universelle a été traité par Ingetraut Dahlberg (Dahlberg 1982: 138). La table 2 est un exemple de correspondance entre les deux systèmes selon la notation de l'ICC, et la table 3 selon celle de la CDU. Dans cette dernière on relèvera le fait significatif que, pour une classe de la Classification décimale universelle qui ne correspond pas totalement à une classe propre de l'autre système, plusieurs codes de l'*Information Coding Classification*, reliés entre eux par des signes particuliers (+, / ou ;), sont marqués. Dans le tableau de comparaison ces signes ont une signification précise pour indiquer des relations entre la classe d'un système et plusieurs classes d'un autre système.

Des correspondances analogues auraient pu être indiquées d'après l'ordre alphabétique.

Il est opportun de faire remarquer que les relations entre l'*Information Coding Classification* et la Classification décimale universelle s'étendent à des niveaux plus particuliers que ceux qui ont été publiés. Les corrélations entre les classes ont pu être établies grâce à l'aide de la définition de chaque classe. L'actualité de ce travail est liée à sa future mise à jour et à la gestion automatique.

Dans la phase de projet du CIRT, on a présenté le problème de la multiplicité de structures des bases de connaissances. La nécessité de conserver de telles structures a suscité le projet de comparer les différents systèmes dans le but de pouvoir les gérer de façon intégrée. Dans une telle optique, l'ISRD, qui a organisé et qui gère la structure centrale de référence du CIRT, est en train

WARNING: This is not so much a correlation table but rather the attempt to provide a systematic listing of the subjects contained in the bibliography with indication of their UDC Numbers, whenever these occurred. The non-indented subject headings/keywords are identical with ICC classes, the indented ones are subsumptions of such classes, mostly without the three-digit class number and heading which may, however, be found in Annex 2.						
				275	technical gases	661.1/.6/.8/.9
					organic materials	661.7
					colour industries; paints	667
					tensides; surface active agents	
				278	cosmetics-perfumery; cosmetics technol.	
				28	ENERGY SCIENCE AND TECHNOLOGY	
ICC		UDC		282	solar energy	
12	MATHEMATICS	51		283	steam power	621.1
122	algebraic groups; division algebra				water power, hydraulic energy;	
13	STATISTICS	31			hydroenergetics	621.1
132	mathematical probability				pneumatic energy	621.5
133	sampling			284	atomic energy	621.039
15	ORGANIZATION SCIENCE AND TECHNOL.				nuclear technol.; nuclear fission	621/621.0
154	Operations research				reactor; reactor names; thermo-nuclear	
155	Planning			285	plants;	
157	Rationalisation			286	accelerator engg.	
16	METROLOGY	389			explosives	662.1/.5
17	CYBERNETICS			287	fuels	662.6/.9
173	control; control systems			288	energy conservation	
175	automatic control engineering;				general economics of energy	620.9
	automation facilities	681.5		29	ELECTRICAL ENGINEERING	621.3
				293	Electrical power engineering	621.31
19	TESTING AND CHECKING				electric energy transportation;	
192	quality control; nondestructive testing				electrical energetics	
198	materials testing	620		294	integrated circuits, production engg.	
21/4	PHYSICS	53			power switching technology	
21	MECHANICS	531/534		297	electric motor	
213	aerodynamics			298	illumination engineering; lighting	628
217	ballistics				air-conditionning; refrigeration technol.	
22	PHYSICS OF MATTER			31	ASTRONOMY, ASTROPHYSICS	52
221	quantum field theory			313	radioastronomy	
222	high energy physics			33	BASIC GEOSCIENCES	55
223	radiation physics			332	sediments; structural geology	
227	nuclear sciences			334	Geodesy and surveying	528
228	solid state physics; crystallography	548		335	Geophysics; exploration geophysics	550.3
229	thin films				geothermal physics and technology	
23	GENERAL AND TECHNICAL PHYSICS			338	Palaeontology; palaeobotany, palinology	56
232	wire explosions; high temperature			339	Historical geology	
	physics; vacuum physics			34	METEOROLOGY, ATMOSPHERIC SCIENCES	551.5
235	thermal physics			344	air pollution	
236	optics, applied, technical; colors			346	bioclimatology	
237	acoustics, technical	681.8/.9		35	HYDROSPHERIC AND OCEANOLOGICAL	
24	ELECTRONICS				SCIENCES	551.49
242	laser			354	hydrology	556
243	semiconductor diode; transitors			355	Oceanographic, Maritime sciences	
244	voltaic cell				coastal environments;	
246	electronic tube			356	Marine technology	
25	PHYSICAL CHEMISTRY			357	Hydraulic engineering	626/627
255	radiochemistry; isotopes; radiation				hydraulic, water construction works;	
	chemistry				waterways	
26	PURE CHEMISTRY	54		358	Water economy	
263	Analytic chemistry			36	GEOLOGICAL SCIENCES	
264	uranium			363	soil mechanics	
265	alkaloid chemistry; organic reactions;			365	Mineralogy	551/553
	organic compounds			366	Soil sciences;	631/632
267	polymers				peat	
27	CHEMICAL TECHNOL. AND ENGINEERING	66		37	MINING	622
273	Chemical process technology			373	metal extraction	
	liquefaction technique	621.5		376	petroleum technology; petroleum	
274	chemicals (inorganic materials);				mineral oil processing	665.6/.7

Tableau 2: Corrélations entre ICC et CFU (Source: ICIB 1, p. 133)

UDC		ICC			
			35	Public Administration	63
1	PHILOSOPHY	97		Military Art and Science	68
11	Metaphysics. Fundamental problems	972	36	Social Welfare and Relief.	65
122/129	Special metaphysics	972	368	Insurance	656
13	Philosophy of mind and spirit. Metaphysics of mental life	9725	37	Education. Teaching. Training	56
14	Philosophical systems and points of view. Weltanschauung	976/8	379.8	Leisure	578
159.9	Psychology	55	389	Metrology. Weights and Measures	16
16	Logic; Theory of knowledge. Methodology of logic	11; 973	4	(vacant; linguistics transferred to 80)	
17	Moral philosophy. Ethics	975			
2	RELIGION. THEOLOGY	98/99	50	MATHEMATICS AND NATURAL SCIENCES	12+21/49
21	Natural theology	981	51	Generalities. Nature and conservation	49
22	Holy Scripture. Holy Bible	992	52	Mathematics	12
23	Dogmatic theology	994	528	Astronomy. Astrophysics. Space Research	31; 32
24	Practical theology	993	53	Geodesy	334
25	Pastoral theology	993	54	Physics	21/24
26	The Christian Church in general	996	548	Chemistry	25/26
27	General History of the Christian Church	999	549	Crystallography	228
28	Christian sects and denominations	998	55	Mineralogy	365
29	Non-christian Religions	984/8	550.3	Geology	33
3	SOCIAL SCIENCES	6	551.49	Geophysics	335
30	Sociology. Sociography	61	551.5	Hydrology	35
31	Statistics	13	56	Meteorology	34
32	Politics	62	57	Palaeontology	338
33	Political Economy. Economics. Trade. Commerce	71/72	58	Biological Sciences	41/44; 51
34	Law. Jurisprudence	66	59	Botany	43
				Zoology	44

Tableau 3: Corrélations entre CDU et ICC (Source: ICIB 1, p. 138)

d'étudier la définition d'un métasystème pour la gestion intégrée de l'ICC avec d'autres classifications utilisées pour classer les patrimoines terminologiques qui relèvent du CIRT.

Le classement international de normes (ICS) est un système très utilisé dans le domaine terminologique. Ce genre de système, que l'on doit à l'Iso, est conçu pour classer tous les différents secteurs de nature technique qui ont la particularité d'avoir été normés, parmi lesquels est comprise la classe 01 relative à « Généralités. Terminologie. Normalisation. Documentation ».

L'ICC et l'ICS ont fait l'objet d'études. Il ressort des analyses comparatives des divergences de nature théorique, méthodologique et applicative.

En ce qui concerne l'aspect théorique, le classement international de normes (ICS) est fondé sur des *principes inductifs* de caractère expérimental, tandis que

l'International Coding Classification (ICC) est fondé sur des *principes déductifs* qui font référence à des catégories primitives, relatives à des objets concrets ou abstraits (Meo-Evoli, Negrini et Farnesi 1998).

La diverse nature des deux systèmes influe fortement sur les aspects méthodologiques et applicatifs des systèmes proprement dits.

Du point de vue méthodologique, l'organisation inductive présente une souplesse de structure qui privilégie avant tout les niveaux hiérarchiques inférieurs. Ces nouveaux ajouts pourraient cependant entraîner la modification de la structure générale du classement dès lors que les concepts hiérarchiquement plus élevés, définis préalablement, n'incluraient pas correctement les nouvelles unités conceptuelles introduites.

Du point de vue de l'application, les avantages et les limites de la classification de nature inductive sont réciproquement inverses par rapport à ceux de la classification de nature

déductive. Si on retient et on fusionne les aspects positifs de chacun des deux classements on pourra obtenir un système complexe fondé sur des principes rigoureux de structures qui caractérisent l'ICC, et utiliser la classification articulée de concepts spécifiques et nouvellement introduits de l'ICS, ou d'autres classifications comparées à l'ICC.

De telles motivations expliquent l'effort d'intégrer des ressources cognitives de systèmes déjà utilisés, en laissant à chaque système ses propres caractéristiques et son autonomie. Par le truchement du réseau de relations comparatives, les classements seront reliés entre eux et constitueront un système global unique. Le réseau relationnel, qui met en évidence de différentes façons les équivalences conceptuelles (correspondance, inclusion, quasi-équivalence) et les différentes structures, représente la cartographie qui doit guider le processus de recherche de l'utilisateur. Avec cette procédure, la cartographie articulée et complexe réunit des

nœuds de structure qui mettent en relief une typologie d'équivalence, et se fondent sur une solide structure ontologique.

Une recherche réalisée par l'Institut d'études sur la recherche et la documentation scientifique du CNR a défini le modèle CoReC

(*Comparazione e Relazione di Classificazione*) pour représenter de façon formelle des classifications et des relations de comparaison (Meo-Evoli et Negrini 1998). L'application a permis de définir une table de comparaison entre ICC et ICS pour la classe 01 («Généralités.

Terminologie. Normalisation. Documentation» (tableaux 4 et 5)). Pour de plus amples renseignements sur la problématique, les solutions adoptées au niveau formel et le développement du projet, le lecteur est invité à se reporter à Meo-Evoli et Negrini (1998).

Icc code	Icc description	relation	Ics code	Ics description
18	STANDARDIZATION	is included in	01	GENERALITIES. TERMINOLOGY. STANDARDIZATION. DOCUMENTATION
826	Documentation	is included in	01	GENERALITIES. TERMINOLOGY. STANDARDIZATION. DOCUMENTATION
915	Terminology	corresponds to is included in	01.020 01	Terminology (principles and coordination) GENERALITIES. TERMINOLOGY. STANDARDIZATION. DOCUMENTATION
F	FORM DIVISIONS	is included in	01	GENERALITIES. TERMINOLOGY. STANDARDIZATION. DOCUMENTATION
F3	DICTIONARIES, TERMINOLOGIES	corresponds to	01.040	Vocabularies
F3/S12	MATHEMATICS	is included in	01.040.07	Mathematics. Natural Sciences (Vocabulaires)
F3/S168:F3/S746	Precision instrument for measurement technology	is about	01.040.17	Metrology and measurement. Physical phenomena (Vocabularies)
F3/S18	STANDARDIZATION	is included in	01.040.01	Generalities. Terminology. Standardization. Documentation (Vocabularies)
F3/S193	Testing technology	corresponds to	01.040.19	Testing (Vocabularies)
F3/S21	MECHANICS	is included in	01.040.07	Mathematics. Natural Sciences (Vocabularies)
F3/S22	PHYSICS OF MATTER	is included in	01.040.07	Mathematics. Natural Sciences (Vocabularies)
F3/S226:F3/S738	Technology in physics of liquids	is about	01.040.23	Fluid systems and components for general use (Vocabularies)
F3/S226:F3/S744	Tank & containers of liquids	is about	01.040.23	Fluid systems and components for general use (Vocabularies)
F3/S226:F3/S788	Mechanical handling of liquids	is about is about	01.040.23 01.040.21	Fluid systems and components for general use (Vocabularies) Mechanical systems and components for general use (Vocabularies)
F3/S23	GENERAL & TECHNICAL PHYSICS	is included in	01.040.07	Mathematics. Natural Sciences (Vocabularies)
F3/S24	ELECTRONICS	is included in corresponds to	01.040.07 01.040.31	Mathematics. Natural Sciences (Vocabularies) Electronics (Vocabularies)
F3/S25	PHYSICAL CHEMISTRY	is included in	01.040.07	Mathematics. Natural Sciences (Vocabularies)
F3/S26	PURE CHEMISTRY	is included in	01.040.07	Mathematics. Natural Sciences (Vocabularies)

Tableau 4 : Corrélations entre ICC et ICS

Ics code	Ics description	relations	Icc code	Icc description
01	GENERALITIES. TERMINOLOGY. STANDARDIZATION DOCUMENTATION	includes includes includes includes	915 18 826 F	Terminology STANDARDIZATION Documentation FORM DIVISIONS
01.020	Terminology (principles and coordination)	corresponds to	915	Terminology
01.040	Vocabularies	corresponds to	F3	DICTIONARIES, TERMINOLOGIES
01.040.01	Generalities. Terminology. Standardization. Documentation (Vocabularies)	includes includes includes	F3/S915 F3/S18 F3/S826	Terminology STANDARDIZATION Documentation
01.040.03	Sociology. Services. Company Organization and Management Administration. Transport (Vocabularies)	includes includes includes includes includes	F3/S61 F3/S79 F3/S72 F3/S63 F3/S78	SOCIOLOGY SERVICE ECONOMICS MANAGEMENT OF ENTERPRISES PUBLIC ADMINISTRATION TRANSPORT TECHNOL. & SERV.
01.040.07	Mathematics. Natural Sciences (Vocabularies)	includes includes includes includes includes includes includes includes includes includes includes includes includes includes	F3/S12 F3/S21 F3/S23 F3/S24 F3/S22 F3/S25 F3/S27 F3/S31 F3/S34 F3/S35 F3/S33 F3/S41 F3/S43 F3/S44	MATHEMATICS MECHANICS GENERAL & TECHNICAL PHYSICS ELECTRONICS PHYSICS OF MATTER PHYSICAL CHEMISTRY CHEMICAL TECHNOL. & ENG. ASTRONOMY AND ASTROPHYSICS. ATMOSPHER. SCI. METEOROLOGY HYDROSPH. & OCEANOLOG. SCI. BASIC GEOSCIENCES BASIC BIOLOGICAL SCIENCES PLANT BIOLOGY AND CULTIVAT. ANIMAL BIOLOGY AND BREEDING
01.040.11	Health care technology (Vocabularies)	is about	F3/S527	Medical physics & biomed. technol.
01.040.13	Environment and health protection. Safety (Vocabularies)	is about	F3/S49	ECOLOGY AND ENVIRONM. SCI.
01.040.17	Metrology and measurement. Physical phenomena (Vocabularies)	is about corresponds to	F3/S168:F3/S74 F3/S16	Precision instruments for measur. technology METROLOGY
01.040.19	Testing (Vocabularies)	corresponds to	F3/S193	Testing technology
01.040.21	Mechanical systems and components for general use (Vocabularies)	is about is about	F3/S741:F3/S21 F3/S226:F3/S78	Principles of mechanical engineering in Applied mechanics Mechanical handling of liquids
01.040.23	Fluid systems and components for general use (Vocabularies)	is about is about is about	F3/S226:F3/S78 F3/S226:F3/S74 F3/S226:F3/S73	Mechanical handling of liquids Tanks & containers of liquids Technology in physics of liquids

Tableau 5: Corrélations entre ICS et ICC

8 Considérations en guise de conclusion

Le parcours constructif qui a été proposé n'est pas dépourvu de difficultés de réalisation. Les travaux

mentionnés sont le résultat de recherches conduites de façon individuelle, privées du support constructif dérivé d'une collaboration active, de l'appui nécessaire pour être étendues à la comparaison avec d'autres systèmes consolidés ou

approfondies à des niveaux hiérarchiques plus particuliers.

L'auteur a voulu promouvoir une initiative qui, grâce à la contribution de quelques savants du domaine, se proposait d'étudier l'intégration de différentes classifications en usage

après de centres de terminologie en Europe, pour obtenir un système harmonisé de langages documentaires consultables sur l'internet.

Un problème que nous n'avons pas souhaité évoquer ici est celui des langues. Si les notations d'une classification permettent de communiquer par dessus les barrières linguistiques, il est de toute évidence qu'un système d'information en réseau télématique ne pourra pas ne pas mettre à la disposition de l'utilisateur final les structures de classification explicites au moyen des termes. Ces termes, classés selon des méthodologies diverses, représentent la cartographie permettant à l'utilisateur d'effectuer un parcours de recherche. Les systèmes de classification pour rechercher des informations dans les banques de données terminologiques en ligne ne sauraient se satisfaire d'une présentation dans la seule langue anglaise. Ce serait en contradiction avec le rôle des banques de données terminologiques. L'accès doit être prévu dans les différentes langues des communautés linguistiques qui apportent leur propre contribution au réseau.

En regard des nombreuses mesures visant à promouvoir l'activité à contenu technologique innovateur, qui prévoient le support et l'utilisation de réseaux télématiques pour la diffusion de l'information, il serait nécessaire de réaliser des études et des recherches dans le secteur de la théorie et de la classification conceptuelle, secteur qui jusqu'à présent n'a pas enregistré de progrès significatif.

*Giliola Negrini,
CNR,
Istituto di studi sulla ricerca e
documentazione scientifica,
Rome.*

Traduction: J. Humbley.

Bibliographie

- Dahlberg (Ingetraut), 1974: *Grundlagen universaler Wissensordnung*, Munich, Saur Verlag Dokumentation, n° 18.
- Dahlberg (Ingetraut), 1982: «*Information Coding Classification – Principles, Structure and Application Possibilities*» dans *Classification systems and Thesauri*, 1950-1982, International Classification and Indexing Bibliography, ICIB 1, Frankfurt, Indeks Verlag, (107-132).
- Dahlberg (Ingetraut). 1982: «Annex 4» dans: *Classification systems and Thesauri*, 1950-1982, International Classification and Indexing Bibliography, ICIB 1, Frankfurt, Indeks Verlag, (138-139).
- Dahlberg (Ingetraut). 1995: «Historical paradigms in the philosophy of classification», dans G. Negrini (éd.), *Categorie, oggetti e strutture della conoscenza*, Atti del seminario, Roma, 1-2 dicembre 1994, CNR-ISRDS, Roma, CNR, (Note di bibliografia e di documentazione scientifica, 60) (121-132).
- Dahlberg (Ingetraut), 1996: «The Compatibility Guidelines – a re-evaluation» dans *Compatibility and Integration of Order Systems*. Proceedings of Research Seminar, ISKO and TIP, Varsovie, sept. 13-15, 1995 (32-45).
- Dahlberg (Ingetraut), 1996: «Library Catalogs in the Internet: Switching for Future Subject Access»: dans Green (R.) (ed.): *Knowledge Organization and Change*, 4th International ISKO Conference, Washington, 15-18 July, Frankfurt, INDEKS Verlag. (155-164).
- Guyon (Anne), 1993: «L'information scientifique et technique» dans *Actes de la première université d'automne en terminologie*, Rennes 2, 21-26 Septembre 1993, Paris, La Maison du Dictionnaire.
- International Standards Organization, 1997: *International Classification for Standards (ICS)*, Iso Catalogue 1997 (33-39);
- Isko, 1995 «Recommendations of the Research Seminar on Compatibility and Integration of Order Systems». organisé par ISKO and TIP, Varsovie, sept. 13-15, 1995, dans *Knowledge Organization*, 22, 1995, nos 3-4 (174-175)
- Meo-Evoli (L.) et Negrini (G.), 1998: *CoReC: a formal model for relations between two classification systems*, CNR/ISRDS/RT 8/98.
- Meo-Evoli (L.), Negrini (G.) et Farnesi (T.), 1998: «ICC and ICS: Comparison and Relations between two Systems based on Different Principles» dans *Structures and Relations in Knowledge Organization*, Proceedings of the 5th ISKO Conference, Lille, 25-29 August (sous presse).
- Nencioni (G.) 1997: *Plurilinguismo in Europa*, dans La Crusca per voi, n°15.
- Rey (Alain) [1979] 1992, *La Terminologie: noms et notions*, Paris, Presses universitaires de France (Que sais-je?)
- Tremblay (Louis-Claude), 1996: «Les avenues de collaboration entre les banques de terminologie», *Terminologies nouvelles*, n°15.
- Unione Europea, *Decisione del Consiglio dell'U.E. del 4 novembre 1996, riguardante l'adozione di un programma pluriennale per la promozione della diversità linguistica della Comunità nella società dell'informazione (SIML)*.