



HAL
open science

Encadrer la réingénierie d'un thesaurus : méthode, enjeux et impacts pour l'équipe d'un service de veille et documentation en entreprise

Lorraine Keller

► To cite this version:

Lorraine Keller. Encadrer la réingénierie d'un thesaurus : méthode, enjeux et impacts pour l'équipe d'un service de veille et documentation en entreprise. domain_shs.info.docu. 2013. mem_00945542

HAL Id: mem_00945542

https://memic.ccsd.cnrs.fr/mem_00945542v1

Submitted on 12 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



CONSERVATOIRE NATIONAL DES ARTS ET METIERS

Ecole Management et Société-Département CITS

INTD

MEMOIRE pour obtenir le
Titre professionnel "Chef de projet en ingénierie documentaire" INTD

RNCP niveau I

Présenté et soutenu par Loraine Keller

le 26 novembre 2013

Encadrer la réingénierie d'un thesaurus

**Méthode, enjeux et impacts pour l'équipe d'un service de
veille et documentation en entreprise**

Jury :

LAUMAILLE, Odile. Responsable du service Veille et documentation, BPCE SA (Responsable de stage)
LOPEZ-UROZ, Adriana. Responsable du Centre de ressources documentaires de l'INTD (Directeur de mémoire)

Promotion 43

*A mon père,
à sa générosité,
avec tout mon amour*

Remerciements

Je tiens à remercier en tout premier lieu Odile Laumaille, pour la confiance et le soutien sans faille qu'elle m'a témoigné durant toute cette mission et son stage préparatoire. Un grand merci également, et toute mon amitié, à l'ensemble de l'équipe du service Veille et documentation de BPCE, pour son accueil chaleureux et bienveillant, sa générosité et sa grande disponibilité. Cette expérience a été un bonheur grâce à vous toutes.

Merci à toute l'équipe de l'INTD, enseignants, intervenants et personnel de l'administration, pour son dévouement et son implication constants. Cette année de formation m'a apporté bien plus encore que je n'étais venue chercher : des compétences, bien sûr, une vocation et un métier, sans doute, mais avant tout de belles rencontres, enrichissantes et stimulantes. Je salue également l'ensemble de la promotion 2012/2013 : à vos côtés j'ai pris plaisir à apprendre et échanger chaque jour, dans la diversité et l'ouverture d'esprit. Un merci tout particulier à Adriana Lopez-Uroz, pour ses conseils avisés et son enthousiasme quant à mon sujet de mémoire.

Merci à celles qui ont bien voulu me consacrer du temps pour l'écriture de ce mémoire. A Michèle Hudon, d'abord, que j'ai eu le privilège de rencontrer à l'occasion de la conférence ISKO France 2013 à la FMSH. Outre son regard critique et son aide précieuse, j'ai découvert une personne aussi charmante qu'exigeante. Merci à Hélène Zysman et Sylvie Dalbin, qui ont partagé avec moi leur avis concernant la norme ISO 25964. Leur expérience « de l'intérieur » de sa conception a été un atout inestimable pour moi.

Enfin, un immense merci à l'homme qui partage ma vie, pour son soutien indéfectible et sa patience sans borne...

Notice

KELLER, Loraine. Encadrer la réingénierie d'un thesaurus : enjeux et impacts pour l'équipe d'un service de veille et documentation en entreprise. Mémoire professionnel INTD, Titre I, Chef de projet en ingénierie documentaire. Conservatoire national des arts et métiers - Institut national des Sciences et Techniques de la Documentation, 2013, 146 p.

RESUME

L'auteur de ce mémoire présente une mission de réingénierie de thesaurus documentaire spécialisé, réalisée pour le service Veille et documentation du groupe bancaire BPCE. Après avoir présenté la nouvelle norme ISO 25964 sur les thesauri et ses principaux apports, elle examine les spécificités d'un thesaurus spécialisé et sa place dans l'entreprise. L'objet central de ce mémoire est de fournir des clés méthodologiques pour la gestion de projets de ce type et d'analyser l'impact en termes de montée en compétences pour les documentalistes impliquées et pour le management. Pour conclure, l'auteur s'intéresse aux implications de la seconde partie de la norme, portant sur l'interopérabilité entre les vocabulaires dans le contexte du Web sémantique, et les opportunités d'applications en entreprise.

DESCRIPTEURS

Thesaurus ; Thésaurus ; Réingénierie ; Langage contrôlé ; Langage documentaire ; Vocabulaire contrôlé ; Norme ; Gestion des connaissances ; Système de gestion des connaissances ; Management de l'information ; Culture informationnelle ; Compétence informationnelle ; Interopérabilité ; Web sémantique ; Web de données ; Web 3.0 ; SKOS ; Ontologie ; Méthodologie ; Gestion de projet ; France

ENGLISH

This essay presents a mission of reengineering of a thesaurus for the information intelligence and documentation unit of the banking group BPCE. After introducing the new standard ISO 25964 about thesauri and its main contributions, it discusses the specificities of a specialized thesaurus in business environment. The main topic of this paper is to provide methodological advice for this type of project management and to analyze the impact on participants' skills and upon management of the team. To conclude, the author addresses the second part of the standard relating to interoperability between vocabularies in semantic Web context, and the involved opportunities for business applications.

Table des matières

Remerciements	3
Notice	4
Table des matières	5
Liste des tableaux et figures	8
Liste des sigles et acronymes utilisés	9
Introduction	10
PREMIERE PARTIE LE THESAURUS SPECIALISE UN OUTIL INDISSOCIABLE DE SON	
CONTEXTE DE MISE EN ŒUVRE	14
1 Le thesaurus : une préoccupation très actuelle	15
1.1 <i>Une mission portée par un « heureux évènement »</i>	15
1.1.1 Une nouvelle norme marquée par un contexte radicalement différent.....	15
1.1.2 Principaux objectifs de la nouvelle norme.....	16
1.1.3 Des avancées significatives	17
1.2 <i>Rappel des grands principes du thesaurus documentaire</i>	19
1.2.1 Quelques éléments de définition	19
1.2.2 Deux nouveautés introduites par la norme ISO 25964.....	20
1.2.3 Rappel des 3 types de relations sémantiques	21
1.3 <i>Un thesaurus, pour quoi faire ?</i>	25
1.3.1 Des fonctions traditionnelles	25
1.3.2 Un périmètre croissant et de nouveaux usages.....	26
1.3.3 Une utilité renforcée	27
2 Un thesaurus spécialisé : pour répondre à quels besoins ?	29
2.1 <i>Adaptation aux usages</i>	29
2.1.1 Décrire le fonds documentaire	29
2.1.2 Aider à la recherche	29
2.2 <i>Adaptation aux utilisateurs</i>	30
2.2.1 Bien connaître ses publics... ..	30
2.2.2 ... pour anticiper leurs besoins	30
2.3 <i>Le reflet des activités et de la culture d'une entreprise</i>	31
2.3.1 Un référentiel des activités de l'entreprise	31
2.3.2 Une matérialisation de la culture de l'entreprise.....	31
3 Production, utilisation et maintien d'un thesaurus spécialisé en entreprise : l'exemple de BPCE	33
3.1 <i>Petit historique du thesaurus de la CNCE</i>	33
3.1.1 Production du « Thesaurus Techlib »	33
3.1.2 Vie du thesaurus.....	35
3.2 <i>Activités de la Documentation générale de BPCE SA</i>	36
3.2.1 La constitution et la gestion du fonds documentaire.....	36
3.2.2 Au cœur de l'offre, les veilles documentaires thématiques	38
3.3 <i>Usages du thesaurus depuis la fusion</i>	40
3.3.1 Indexation	40
3.3.2 Recherche.....	40
3.3.3 Attribution des cotes d'ouvrages	41

DEUXIEME PARTIE MENER UN PROJET DE REINGENIERIE DE THESAURUS ELEMENTS DE METHODE	42
4 Repenser un thesaurus : une étape incontournable dans la vie d'un service de documentation.....	43
4.1 <i>Une mise à jour inhérente au cycle de vie d'un thesaurus.....</i>	43
4.1.1 Constat d'obsolescence	43
4.1.2 Obstacles rencontrés pour l'indexation.....	44
4.1.3 La longue gestation du projet	45
4.2 <i>Une évolution nécessaire suite à l'adoption d'un nouveau logiciel</i>	45
4.2.1 Mise en place du progiciel JLB-Net	45
4.2.2 Du « thesaurus Techlib » au lexique Mots clés de JLB	47
4.3 <i>Considérations préliminaires sur les implications du projet</i>	48
5 Choix méthodologiques pour la gestion du projet	50
5.1 <i>Analyse des risques et contraintes liés à la mission</i>	50
5.1.1 Présentation du contexte de réalisation.....	50
5.1.2 Durée et coût de l'intervention	52
5.2 <i>S'inspirer de SCRUM pour optimiser la gestion de projet</i>	53
5.2.1 Présentation de Scrum	53
5.2.2 Poser les bases de la réussite du projet	54
5.3 <i>S'assurer l'implication des membres de l'équipe</i>	56
5.3.1 Inscrire le projet dans les priorités du service.....	57
5.3.2 Inscrire le projet dans les tâches quotidiennes.....	57
5.4 <i>Un dialogue constant avec le management.....</i>	59
5.4.1 Pilotage	59
5.4.2 Reporting.....	60
6 Le chef de projet : un rôle de facilitateur	61
6.1 <i>Proposer une méthode de travail adaptée</i>	61
6.1.1 Mettre en place des outils pour optimiser le travail préliminaire.....	61
6.1.2 Animer des séances de travail en commun.....	63
6.2 <i>Organiser les principales étapes de la refonte du thesaurus</i>	64
6.2.1 Phase d'évaluation conceptuelle et de collecte des termes.....	64
6.2.2 Phase de hiérarchisation des concepts.....	66
6.2.3 Phase d'enrichissement du réseau sémantique	67
6.3 <i>Proposer un accompagnement constant et personnalisé</i>	68
6.3.1 S'adapter au besoin individuel.....	68
6.3.2 Identifier les inclinations personnelles.....	69
TROISIEME PARTIE QUELLES PERSPECTIVES POUR LE MANAGEMENT D'UN SERVICE DOCUMENTAIRE ?	71
7 Un nouveau thesaurus, et après ?	72
7.1 <i>Composition du thesaurus BPCE à l'issue du projet.....</i>	72
7.1.1 Un thesaurus spécialisé organisé autour des thématiques de veille.....	72
7.1.2 Description de la structure hiérarchique	73
7.1.3 Evaluation du RAF (reste à faire).....	74
7.2 <i>Outils mis en place pour l'exploitation du thesaurus</i>	74
7.2.1 Identification des besoins	74
7.2.2 Présentation adoptée pour le thesaurus « papier »	75
7.2.3 Outils logiciels pour l'exploitation du thesaurus	76

7.3	<i>Préconisations pour l'amélioration du thesaurus BPCE</i>	78
7.3.1	Analyse critique du livrable obtenu	78
7.3.2	Principaux axes de développement à exploiter	81
8	Consolider et valoriser : une organisation apprenante au cœur de l'entreprise	87
8.1	<i>Retour d'expérience : quels bénéfices pour l'équipe du service Documentation générale ?</i>	87
8.1.1	Renforcement de l'« intelligence informationnelle »	87
8.1.2	Compétences développés et capitalisés	88
8.2	<i>Quel ROI pour le management ?</i>	90
8.2.1	La mise en place d'une politique de gestion du thesaurus	90
8.2.2	Manager pour développer une organisation apprenante	92
8.3	<i>Impacts organisationnels : effets sur l'image de marque du service DG au sein de l'organisation BPCE</i>	94
8.3.1	Gains immédiats	94
8.3.2	Capitalisation à plus long terme	95
9	Les référentiels terminologiques aujourd'hui : évolution et perspectives pour l'entreprise	99
9.1	<i>ISO 25964 : une nouvelle norme pour encourager la transition vers le Web sémantique</i>	99
9.1.1	Une norme adaptée à l'évolution des pratiques	99
9.1.2	Encadrer et faciliter l'interopérabilité entre les vocabulaires	100
9.1.3	Panorama des principaux types de vocabulaires, caractéristiques et interopérabilité	103
9.2	<i>Les technologies du Web sémantique au service de l'entreprise</i>	105
9.2.1	Qu'entend-on par Web « sémantique » ?	105
9.2.2	Quelles perspectives pour l'entreprise	107
	Conclusion	110
	Bibliographie	113
	Annexes	123

Liste des tableaux et figures

Tableaux

<i>TABLEAU 1 : NOTATIONS EMPLOYEES POUR LES RELATIONS DANS LE THESAURUS BPCE</i>	22
<i>TABLEAU 2 : GRILLE DE RECENSEMENT DES MODIFICATIONS PAR DOMAINE</i>	63
<i>TABLEAU 3 : RECOMMANDATIONS POUR L'INTEROPERABILITE AVEC D'AUTRES LANGAGES</i>	104

Figures

<i>FIGURE I : SCHEMAS XML DU MODELE DE DONNEES STANDARD DE LA NORME ISO 25964 SUR LES THESAURUS</i>	18
<i>FIGURE II : METIERS DES UTILISATEURS DE LA DOCUMENTATION GENERALE (HISTOGRAMME DE REPARTITION)</i>	39
<i>FIGURE III : PRESENTATION SCHEMATIQUE DU THESAURUS BPCE</i>	75
<i>FIGURE IV : ESSAI D'APPLICATION DE FACETTES AU THESAURUS BPCE</i>	86
<i>FIGURE V : EXEMPLE D'APPLICATION DU MODELE "LIEN DIRECT" A 4 VOCABULAIRES</i>	101
<i>FIGURE VI : EXEMPLE D'APPLICATION DU MODELE "STRUCTURE PIVOT" A 4 VOCABULAIRES</i>	101
<i>FIGURE VII : LE « LAYER CAKE » DU WEB SEMANTIQUE – TIM BERNERS-LEE (SOURCE : AFNOR BIVI)</i>	106

Liste des sigles et acronymes utilisés

ADBS	Association des professionnels de l'information et de la documentation
BFBP	Banque Fédérale des Banques Populaires
BPCE	Banque Populaire Caisse d'Épargne
CNCC	Compagnie Nationale des Commissaires aux Comptes
CNCE	Caisse Nationale des Caisses d'Épargne
DG	Documentation générale (de BPCE SA)
DSI	Diffusion sélective d'informations
ISKO	<i>International Society for Knowledge Organization</i>
ISO	<i>International Organization for Standardization</i> (Organisation internationale de normalisation)
OPAC	<i>Online public access catalog</i> ou Catalogue de bibliothèque accessible en ligne
OWL	<i>Web ontology language</i> (langage de représentation des connaissances pour le Web)
RDF	<i>Resource description framework</i> (modèle de graphe pour la description des ressources et leurs métadonnées)
ROI	<i>Return on investment</i> (retour sur investissement)
RSS	<i>Really simple syndication</i> (format de données pour la syndication de contenus Web)
SIC	Sciences de l'information et de la communication
SIGB	Système intégré de gestion de bibliothèque
SKOS	<i>Simple knowledge organization system</i> (système simple d'organisation des connaissances)
SOC	Système d'organisation des connaissances
TAG	Thesaurus de l'Activité Gouvernementale (du Québec)
TAL	Traitement automatique des langues
URI	<i>Uniforme resource identifier</i> (identifiant uniforme de ressource)
W3C	World Wide Web Consortium (organisme de normalisation à but non lucratif pour les technologies du Web, fondé par Tim Berners-Lee en 1994)
XML	<i>eXtensible markup language</i> (langage de balisage extensible)

Introduction

Contexte de la mission

La mission présentée dans ce mémoire s'est déroulée du 20 juin au 6 septembre 2013, avec pour objectif la refonte du thesaurus¹ documentaire utilisé par le service de veille et documentation de BPCE SA. Issu de la fusion en juillet 2009 de la Caisse Nationale des Caisses d'Épargne (CNCE) et de la Banque Fédérale des Banques Populaires (BFBP), cet organe central regroupe quelques 1 200 employés, essentiellement cadres, en charge d'orienter la stratégie du groupe et de fournir les services supports nécessaires aux activités de ses différents actionnaires et filiales (Natixis, Crédit foncier, Banque palatine, Crédit coopératif...).

Le service « Documentation générale » du groupe BPCE s'est constitué, peu après la fusion, de la réunion d'une partie des documentalistes composant les services existant dans chacune des deux entités principales, et fut placé sous la direction de la personne occupant cette fonction pour la CNCE, depuis la création du service en 1990. A l'époque, l'une de ses priorités fut la mise en place d'un thesaurus spécialisé, adapté aux métiers et au périmètre des Caisses d'Épargne. C'est ce même thesaurus qui était encore utilisé jusque-là essentiellement par la Documentation générale de BPCE (DG).

Or, 20 ans d'évolution des domaines couverts par ce référentiel l'ayant déjà rendu de plus en plus délicat à utiliser pour l'indexation, le bouleversement de l'activité des utilisateurs, engendré par la création du groupe, a achevé de rendre évidente l'urgence de mettre à jour et d'enrichir ce thesaurus. Un troisième élément, néanmoins, fut déterminant dans le lancement de cette mission. Afin de mettre en commun les bases de données des fonds documentaires issus de la CNCE et de la BFBP, gérées par deux logiciels différents, la mise en place d'un nouveau logiciel a été décidée, suite à une analyse des besoins menée par Serda en 2010. La direction informatique opta alors pour JLB-Net, dont elle possédait déjà la licence. Bien que la complexité de la gestion de projet ait généré d'importants retards, ce nouveau SIGB² a enfin pu être déployé avec succès en janvier 2013. Le logiciel gérait désormais à l'indexation à la fois le thesaurus de la CNCE (issu de Basis Techlib) et la liste des mots-clés générés par le logiciel DIP en usage à la BFBP, les deux mêlés dans le champ « Mots clés ».

Pour achever la fusion des deux entités, il apparaissait dès lors incontournable de doter le groupe de son propre thesaurus, adapté à ses métiers, ses activités et surtout à sa nouvelle culture d'entreprise. Le besoin était d'autant plus pressant que la seconde phase du projet JLB, un portail d'interrogation de la base de données par les utilisateurs (OPAC³) accessible depuis le portail

¹ Nous avons choisi d'utiliser l'orthographe la plus proche de l'origine latine du terme, soit *thesaurus*, sans accent, et au pluriel *thesauri*, comme dans la norme (rédigée en anglais).

² Système intégré de gestion de bibliothèque

³ *Online public access catalog* ou Catalogue de bibliothèque accessible en ligne

intranet du groupe, était en cours de déploiement. Or, cette interface prévoyait la possibilité pour l'utilisateur de sélectionner lui-même ses mots-clés en affichant le thesaurus. Le positionnement et la crédibilité de la Documentation générale auprès de ses publics aura donc fort à gagner de sa capacité à présenter un vocabulaire cohérent, à jour et adapté aux besoins de ses utilisateurs.

Un premier projet de refonte avait été lancé fin 2011, encadré par l'une des documentalistes du service, dont la formation et l'expérience en ce domaine constituaient des ressources précieuses. Mais sa présence à mi-temps seulement, ainsi que la charge de travail de ses collaboratrices, toujours plus lourde – tandis que les ressources humaines du service ne cessaient de diminuer – l'ont empêchée de mener à bien sa mission. La responsable du service, elle-même diplômée de l'INTD, a alors eu l'idée de recourir à un élève du Titre 1 pour mener ce projet, dans le cadre de son stage de fin de formation. J'ai donc eu la chance d'intervenir pour cette mission en tant que chef de projet documentaire.

Structure et objectifs du mémoire

Ce mémoire n'a pas pour objet de revenir longuement sur ce qu'est un thesaurus. Plusieurs références bien meilleures pour cela sont présentées dans la bibliographie, à commencer par le très précieux *Guide pratique pour l'élaboration d'un thesaurus documentaire* de Michèle Hudon [9, HUDON], remis à jour en 2009, qui fut une référence incontournable tout au long de la mission présentée ici. Il nous semble bien plus éclairant, en regard du contexte de réalisation, de revenir dans un premier temps sur les caractéristiques et les usages d'un thesaurus spécialisé, tel qu'il peut être conçu et mis en œuvre à l'heure actuelle au sein d'un service documentaire en entreprise. Quelle est sa place dans le système d'information d'entreprise aujourd'hui ? Il est indispensable de répondre à cette question afin d'évaluer la pertinence et l'opportunité de mener un projet de refonte d'un tel outil, tout en évaluant avec lucidité les objectifs poursuivis et les enjeux. Nous nous appuyerons pour cela sur l'exemple observé au sein de BPCE SA, en suivant le développement, l'évolution et les usages successifs du thesaurus qui a fait l'objet de cette refonte.

La seconde partie de ce mémoire est consacrée à une présentation détaillée du projet mené au sein du service Documentation générale de BPCE SA. Notre but est de fournir des pistes de réflexion et des clés pour la mise en œuvre d'un projet de réingénierie de thesaurus au sein d'une organisation. C'est pourquoi la méthode de gestion de projet et les adaptations apportées en cours de route seront particulièrement explicitées. Nous verrons en outre dans quels contextes doit être envisagée la réingénierie d'un thesaurus, et quel positionnement un chef de projet peut adopter pour aider son équipe à mener à bien une telle entreprise.

Enfin, l'objectif de la dernière partie sera d'apporter une analyse a posteriori des bénéfices d'un tel projet pour le management d'un service documentaire. Quelles étaient les attentes de la

responsable du service lorsqu'elle a conçu ce projet ? Comment évaluer, aujourd'hui, les résultats obtenus et l'impact du projet sur son équipe ? Enfin, de quelle montée en compétences témoignent les retours d'expérience apportés par chaque membre de l'équipe ? Toutes ces questions sont au cœur de ce mémoire, comme elles ont constitué la préoccupation centrale dans la gestion de ce projet. Nous ouvrirons pour terminer sur certains aspects mis en avant par la seconde partie de la norme ISO 25964 [2, ISO], qui n'ont pu être traités dans le contexte de cette mission, bien qu'ils nous semblent essentiels aujourd'hui. Il s'agit des questions d'interopérabilité avec les autres vocabulaires, dans le contexte de l'essor du Web sémantique. Ces technologies sémantiques – qui ont permis à Tim Berners-Lee, dès 2001, de rêver d'un tel changement de paradigme qu'on parle parfois de Web 3.0 – pénètrent de plus en plus, et avec grand profit, le monde de l'entreprise. Nous y résumerons les recommandations pour l'alignement des différents types de vocabulaires un thesaurus, avant de donner un aperçu des opportunités que présentent les technologies sémantiques actuelles pour les systèmes d'information en entreprise.

Nous nous appuyerons aussi souvent que nécessaire sur les éléments tels qu'ils sont présentés par la dernière norme ISO 25964 *Information et documentation : Thesaurus et interopérabilité avec d'autres vocabulaires* [1, ISO], parue en août 2011 (mars 2013 pour la seconde partie). Cette approche devrait permettre d'extraire de celle-ci les informations essentielles à quiconque souhaiterait mener un projet similaire – en particulier tant que la norme n'aura pu être traduite en français dans son intégralité. En attendant, on ne manquera pas de se référer au livre blanc (autre ressource qui nous fut indispensable) publié à l'initiative de l'AFNOR et reprenant les principaux points d'évolution par rapport à la norme ISO 2788 [4, DALBIN et al.]. Précisons que, le projet ayant porté sur un thesaurus monolingue, les thesauri multilingues ne seront pas évoqués ici. Enfin, les éléments donnés ici ont pour but de présenter les points qui nous ont semblés les plus pertinents, dans le contexte de cette mission, ou qui apportent une réelle évolution par rapport à la précédente norme. Ce n'est en aucun cas une étude exhaustive, mais nous espérons donner quelques clés permettant de gagner du temps dans la réflexion autour d'un thesaurus. Pour les exemples et plus de détail, nous renverrons directement au texte original.

Première partie

Le thesaurus spécialisé

Un outil indissociable
de son contexte de mise en œuvre

1 Le thesaurus : une préoccupation très actuelle

1.1 Une mission portée par un « heureux évènement »

Le travail sur le thesaurus mené lors de cette mission a bénéficié de la récente parution de la norme ISO 25964-1:2011: *Information et documentation - Thésaurus et interopérabilité avec d'autres vocabulaires – Partie 1 : thésaurus pour la recherche documentaire* (Août 2011) [1, ISO]. Celle-ci a représenté un véritable évènement dans le domaine de l'information-documentation, et particulièrement attendu lorsqu'on sait que la dernière édition de la norme ISO 2788, qu'elle remplace, datait de 1986.

1.1.1 Une nouvelle norme marquée par un contexte radicalement différent

Si le thesaurus en lui-même n'a que peu évolué, le monde, lui, s'est transformé depuis la première formalisation de ses principes par la norme ISO 2788, en 1974, époque où le papier était encore roi. Avec les nouvelles technologies de l'informatique, puis de l'Internet, les possibilités d'application des thesauri se sont démultipliées, et avec elles leur valeur pour les organismes qui les conçoivent et les utilisent. En conséquence, le nombre de logiciels de gestion de thesaurus, qu'ils soient embarqués dans des systèmes de gestion de bases de données ou autonomes (*standalone*), n'a cessé de croître depuis le début des années 1980. Ce contexte a amorcé le développement de langages informatiques pour les représenter, rendus bientôt indispensables par le déploiement de ces applications sur le Web. La question de l'interopérabilité est en parallèle devenue incontournable, au fur et à mesure que les ressources accédaient à une diffusion de plus en plus large, que les échanges se faisaient plus fréquents, que les données s'ouvraient.

En effet, les liens s'affirment de plus en plus étroits et pérennes entre l'outil conceptuel qu'est le thesaurus et l'informatique documentaire, qui permet leur exploitation toujours plus efficace [11 et 12, DALBIN]. La norme apporte des préconisations précises et essentielles sur ces deux points (Clauses 14 à 18). Elle vise, d'une part, à permettre aux logiciels documentaires d'atteindre une meilleure gestion des thesauri, et à garantir, d'autre part, leur déploiement optimal au sein de systèmes d'information de plus en plus complexes.

Pourtant, la norme ISO 25964-1 insiste en préambule sur l'intervention indispensable de l'être humain à chaque bout de la chaîne d'opérations :

« ISO 25964 opère la transition nécessaire afin d'être compatible avec le monde de la gestion électronique de l'information. Cependant, cette partie d'ISO 25964 retient l'hypothèse que l'intelligence humaine est habituellement impliquée dans la sélection

des termes d'indexation comme dans la sélection des termes de recherche. Si l'indexeur et le chercheur sont tous deux guidés vers le choix du même terme pour le même concept, alors les documents pertinents remonteront. C'est le principe essentiel qui sous-tend la conception de thesaurus, bien qu'un thesaurus puisse également être en jeu dans des situations où la machine fait les choix. »⁴

1.1.2 Principaux objectifs de la nouvelle norme

Au vu de ces bouleversements, il apparaît qu'un travail de longue haleine a dû être réalisé par l'équipe internationale de professionnels de l'information-documentation (Infodoc) réunis pour l'occasion, sous la direction de Stella G. Dextre Clarke⁵. Ce comité d'élaboration a par ailleurs réuni des membres issus de nombreux pays, dont certains ne sont que rarement mis à contribution dans les travaux de l'ISO. Cela n'a pu qu'enrichir sa réflexion, de même que le corpus de projets de référence étudiés, au cours d'un travail véritablement collaboratif. Il s'agissait, premièrement, de réunir leurs retours sur plus de 20 ans d'expérience, marqués par des évolutions radicales dans les équipements comme dans les manières de travailler.

De plus, cette mise à jour représentait une opportunité, qu'il était important de ne pas manquer, de clarifier les perspectives actuelles de la recherche et des initiatives portant sur les vocabulaires informatiques, en démontrant au passage l'apport indispensable des compétences documentaires dans ce domaine. Nous reviendrons plus longuement sur ce point, qui fait l'objet de la seconde partie de la norme⁶, parue en mars 2013, au cours de la dernière partie de ce mémoire.

Il est utile de préciser dès maintenant que nous utiliserons parfois le terme de « vocabulaire » pour se référer à différents types de langages informatiques traduisant des langages contrôlés utilisés pour l'indexation. Nous nous référons ici à la définition très large qu'en donne le W3C⁷ :

« Sur le Web sémantique, les vocabulaires définissent les concepts et les relations (aussi dénommées « termes ») utilisés pour décrire et représenter un domaine d'intérêt. Les vocabulaires sont utilisés pour classifier les termes qui peuvent être utilisés avec une application spécifique, caractériser les relations possibles, et définir les contraintes possibles dans l'utilisation de ces termes. »⁸

⁴ ISO 25964 [1, ISO], Introduction, p. VI. Traduction de l'auteur

⁵ Consultante en management de l'information, Oxford, UK

⁶ ISO 25964-2:2013 (Mars 2013) *Information et documentation - Thésaurus et interopérabilité avec d'autres vocabulaires - Partie 2 : Interopérabilité avec d'autres vocabulaires* [2, ISO]

⁷ World Wide Web Consortium, organisme de normalisation à but non lucratif (<http://www.w3.org/>)

⁸ Source : site du W3C, traduction de l'auteur. <http://www.w3.org/standards/semanticweb/ontology>

1.1.3 Des avancées significatives

Nous avons pu rencontrer Michèle Hudon, représentante du Québec au sein du comité d'élaboration, ainsi qu'Hélène Zysman, qui faisait partie du groupe de réflexion français, et échanger avec Sylvie Dalbin, qui le dirigeait. Leur aide nous a été précieuse pour identifier les principaux points positifs apportés par rapport aux normes ISO 2788 et 5964, ainsi que AFNOR NF Z47-100 et 47-101, mais aussi les pistes d'améliorations futures à envisager.

1.1.3.1 Un texte adapté aux pratiques du XXI^e siècle

Bien sûr, l'adaptation de la norme au contexte actuel de la gestion de l'information est unanimement saluée. Mettre enfin l'accent sur le développement de thesaurus dans un environnement numérique était une des ambitions principales. Cependant, il faudrait plutôt parler de « contextes », au pluriel, tant elle donne enfin la parole aux professionnels issus de tous types d'organismes où l'on gère de l'information, et affiche une volonté de sortir des milieux traditionnels que sont les bibliothèques et les centres de documentation. La variété des profils impliqués dans cette réflexion a ainsi permis la présentation de nombreux exemples, dédiés à des contextes et usages variés, pour mieux s'adresser à tous type de cas. Plus encore, Michèle Hudon a apprécié une norme qui se veut plus explicative, au-delà des recommandations, s'adressant ainsi plus largement aux différentes spécialités qui sont aujourd'hui impliquées dans la mise en œuvre de thesauri (informaticiens et programmeurs de systèmes d'information, terminologues, etc.). Elle propose notamment, en guise d'introduction, une liste de 64 définitions normalisées regroupant l'ensemble des termes développés dans la norme, formant un glossaire dédié au thesaurus particulièrement utile⁹.

1.1.3.2 Un outil puissant de communication avec les équipes techniques

La norme introduit notamment, comme outil essentiel pour les deux objectifs évoqués plus haut, un modèle de données XML permettant de faciliter l'échange de thesauri.¹⁰ Toutes les classes et relations proposés par la norme sont représentées dans ce schéma (Cf. FIGURE I) Ce modèle doit permettre un dialogue facilité entre les documentalistes en charge des thesauri et les informaticiens responsables de leur transposition dans les outils informatiques, et en particulier sur le Web. Hélène Zysman, qui intervient fréquemment dans des projets de grande ampleur en tant que spécialiste des langages documentaires pour la société Smile, a particulièrement insisté sur cet aspect. En effet, les clauses concernant les formats et protocoles d'échange de données, permettant l'intégration d'un thesaurus à un système informatisé, ont été l'objet d'un soin tout particulier – même si certains spécialistes du sujet semblent regretter qu'elle n'aille pas encore

⁹ La seconde partie de la norme reprend ces définitions et les complète, portant leur nombre à 89 [2, ISO, Clause 3]

¹⁰ Accessible à l'adresse URL : <http://www.niso.org/schemas/iso25964>

plus loin. Autre point d'amélioration notable par rapport aux anciennes normes, et qui constitue un alignement indispensable avec les usages actuels, la clause concernant les logiciels dédiés à l'élaboration et à la gestion des thesauri a été considérablement enrichie (Clause 14).

Enfin, les éléments les plus techniques concernant cet aspect font l'objet de la seconde partie de la norme dans son intégralité [2, ISO]. La question de l'interopérabilité entre les langages documentaires y est pour la première fois abordée en profondeur, posant les principes de base pour l'établissement de passerelles entre le thesaurus et d'autres vocabulaires, comme nous le verrons plus en détail dans la dernière partie de ce mémoire. Sylvie Dalbin regrette cependant « *le peu de réflexion et d'exploitation autour de la norme en raison de la confusion avec SKOS.* » Elle encourage une meilleure distinction entre le modèle de données de la norme, qui permet de construire et maintenir des vocabulaires, et SKOS, outil de diffusion et donc de leur utilisation. Pour elle : « *L'impact est assez dramatique (surtout en France d'ailleurs) sur le devenir des vocabulaires, SKOS n'intégrant aucune (méta)données de maintenance et de traçage dans le temps...* ». ¹¹

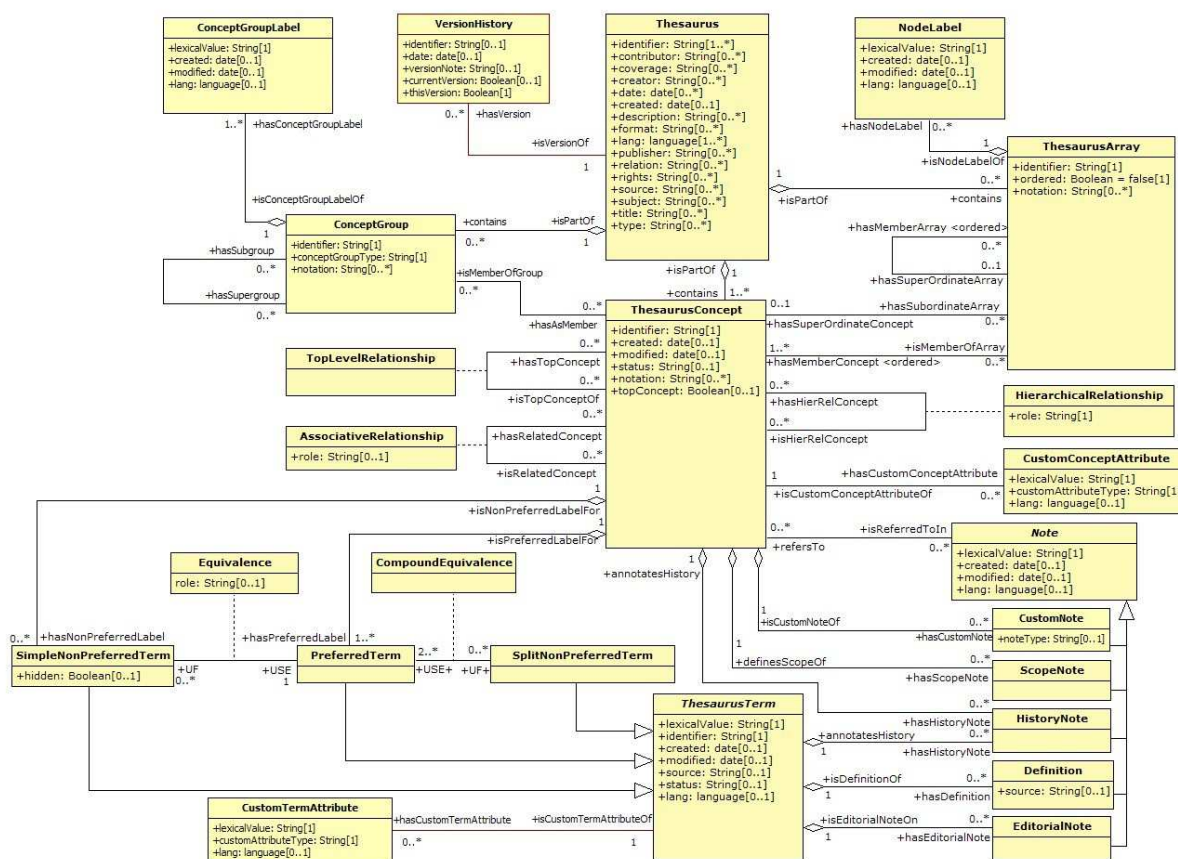


FIGURE I : SCHEMAS XML DU MODELE DE DONNEES STANDARD DE LA NORME ISO 25964 SUR LES THESAURUS

¹¹ Correspondance du 9 novembre 2013

1.2 Rappel des grands principes du thesaurus documentaire

Le thesaurus est un élément aussi essentiel aux métiers de la documentation qu'il échappe le plus souvent à l'appréhension du grand public – même si le Web de données et le travail collaboratif ont pu ces dernières années améliorer sa perception des pratiques d'indexation. Dans cette partie, nous développerons uniquement les points qui nous semblent essentiels pour ce mémoire. D'autre part, la grande spécificité du thesaurus, comparé à d'autres langages documentaires, étant la richesse des relations sémantiques qui y sont développées, il nous a paru intéressant d'en présenter un rappel ci-après.

1.2.1 Quelques éléments de définition

La mission menée à BPCE ayant une certaine dimension communication auprès de la direction des services généraux, dont le service documentation fait partie, il était important de faire comprendre en quoi consiste cet outil, mais aussi la complexité de sa conception. Voici la définition que nous avons proposée dans la note de cadrage présentée au début de la mission, afin d'en aborder tous les aspects de façon simple et concise :

Un thesaurus est une liste de termes contrôlés et normalisés, représentant chacun un concept unique d'un domaine de connaissances. Ces termes, appelés « descripteurs » (ou termes préférentiels), et leurs synonymes (ou « non-descripteurs ») sont organisés selon les relations qui structurent leur champ sémantique : relations hiérarchiques, d'équivalence (synonymie) ou d'association. Ainsi, un terme générique peut avoir un ou plusieurs synonymes, un ou plusieurs termes spécifiques – représentant des sous-concepts liés, et être associé à un ou plusieurs termes d'autres champs sémantiques.

Voici par ailleurs la définition proposée par la norme :

« Vocabulaire contrôlé et structuré dans lequel les concepts sont représentés par des termes, organisés de manière à rendre explicites les relations entre les concepts, et les termes préférentiels sont accompagnés d'entrées pour les synonymes et quasi-synonymes. »¹²

Il est important de retenir qu'il s'agit d'un langage artificiel, qui doit permettre une représentation formalisée et univoque des concepts présents dans le document à indexer. Il prend la forme d'une liste d'autorité organisée par des relations sémantiques, dans laquelle sont identifiés des descripteurs et des non descripteurs. Ceux-ci obéissent à des règles terminologiques propres et sont sélectionnés pour limiter au maximum tout risque d'ambiguïté.

¹² [1, ISO] 25964-1:2011, Clause 2.62, p.12 (traduction de l'auteur).

1.2.2 Deux nouveautés introduites par la norme ISO 25964

1.2.2.1 Une distinction claire entre concept et terme

Dans un souci d'interopérabilité et pour garantir l'évolutivité des thesauri, la nouvelle norme introduit une distinction formelle entre concepts¹³ et termes. On parlera alors de termes préférentiels et non-préférentiels, reliés aux concepts par une relation spécifique, et entre eux par une relation d'équivalence (qu'il est possible de spécifier). Chaque concept doit être représenté par au moins un terme préférentiel. Un thesaurus se définit dès lors comme un réseau de concepts, représentés par des termes et articulés selon trois grands types de relations. Les relations permettent de définir les concepts et de les situer dans l'univers sémantique décrit par le thesaurus. Cette distinction essentielle est l'un des points clés de la norme, indispensable au développement d'un schéma de données. Elle permet en outre les alignements nécessaires aux applications dans le cadre du Web sémantique, en permettant l'attribution à chaque concept d'un identifiant unique ou URI¹⁴ (Cf. 9.1.2).

Cette distinction permet en outre une avancée notable, saluée notamment par Michèle Hudon, dont c'est une des spécialités. Il s'agit de la réunion dans une même norme des thesauri monolingues et multilingues, qui faisaient jusqu'ici l'objet de deux normes séparées, de façon assez artificielle (la norme ISO 5964 de 1985 étant dédiée au thesaurus multilingue). Désormais l'ensemble de la norme ISO 25964 s'applique au thesaurus en général. Seule la Clause 9 *Equivalence across languages* (« Equivalence entre les langues », p. 50 à 57) est dédiée aux problématiques d'alignement conceptuel et terminologique entre langues naturelles. Ce type de relation d'équivalence interlinguistique vient ainsi enrichir la description sémantique d'un concept, au même titre que l'équivalence intralinguistique. Sylvie Dalbin a déclaré regretter, malgré tout, le manque de temps pour mieux structurer la différence entre terminologie linguistique et culturelle d'un côté, et approche conceptuelle de l'autre.

1.2.2.2 L'introduction des regroupements par facettes

Afin d'enrichir les moyens d'accès à un même concept, la norme ISO 25964 introduit les regroupements par facettes, qui peuvent éventuellement compléter un regroupement par domaine ou thématique. Une facette est définie ici comme « *un groupement de concepts appartenant à une même catégorie inhérente* »¹⁵. Sont ainsi donnés comme exemples de catégories de plus haut niveau pouvant servir de facettes :

¹³ Un concept étant défini comme une « unité de pensée » (*unit of thought*) [1, ISO, Clause 2.11, p.3] et un terme comme « un mot ou phrase utilisé pour représenter (*label*) un concept » [1, ISO, Clause 2.61, p. 12]

¹⁴ *Uniform resource identifier*

¹⁵ [1, ISO], Clause 2.20, p. 4 (traduction de l'auteur).

- Objets
- Matériaux
- Agents
- Actions
- Lieux
- Périodes

La Clause 11 *Facet analysis* (« Analyse par facettes », p. 68-69) présente les principes et avantages de ce type de regroupement, certes moins aisé à mettre en place que le traditionnel regroupement thématique, la notion de facette étant beaucoup plus abstraite (Cf. 7.3.2.4). Néanmoins, elle correspond tout particulièrement aux systèmes de recherche d'information actuels, dont elle enrichit considérablement les possibilités d'organisation et de filtrage. [4, DALBIN et al.] Ce type d'analyse, qui varie bien sûr en fonction du domaine d'application, convient en outre parfaitement aux thesauri très spécialisés, car ils enrichissent la navigation en liant les concepts selon d'autres aspects, et ils permettent d'améliorer la pertinence des résultats en augmentant le niveau de spécificité des requêtes. Pour y parvenir, les facettes peuvent elles-mêmes être subdivisées en sous-facettes pour enrichir les critères de filtrage. Par exemple, la facette *Acteurs* d'un domaine peut se diviser en *Acteurs institutionnels* et en *Acteurs privés*.

Ces facettes sont introduites par des « relais virtuels » (*node labels*), distincts des termes composant le thesaurus. Ils peuvent représenter deux types de relations :

- une caractéristique de division d'un concept de niveau supérieur : dans ce cas, tous les termes représentés en dessous sont des spécifiques du terme supérieur ;
- l'introduction d'une nouvelle facette : les termes qui suivent ce relai ne sont généralement pas des spécifiques du terme qui le précède.¹⁶

1.2.3 Rappel des 3 types de relations sémantiques

Nous nous contenterons dans cette partie de rappeler les grandes lignes de la structure théasaurale, en insistant s'il y a lieu sur les modifications proposées par la nouvelle norme. Pour le thesaurus de BPCE, une notation en français a été adoptée. Voici le rappel des sigles employés pour représenter les relations :

¹⁶ Voir l'exemple traduit de la norme exposé dans [4, DALBIN et al.], disponible en ligne.

TABLEAU 1 : NOTATIONS EMPLOYEES POUR LES RELATIONS DANS LE THESAURUS BPCE

Relation	Notation française	Symbole
Equivalence (intralinguistique)	EM (employer) EP (employé pour)	=
Hiérarchie	TG (terme générique) TS (terme spécifique)	< >
Association	TA (terme associé)	–

A ces notations principales s'ajoutent celles désignant les notes d'application (NA) et historiques (NH), qui peuvent être nécessaires pour expliciter un emploi spécifique d'un terme dans le contexte du thesaurus en question, ou pour relater une évolution dans la représentation d'un concept suite à une révision.

Un élément à garder en tête dans la définition des relations est la notion de paradigme vs syntagme¹⁷. Une relation paradigmatique est de nature conceptuelle, c'est-à-dire que l'on associera naturellement ces deux concepts. Au contraire, la relation syntagmatique est contextuelle, c'est-à-dire que l'on pourra la retrouver exprimée ponctuellement dans un document. La norme recommande de ne pas représenter ce second type de relation dans un thesaurus [1, ISO, Clause 4.3, p.16/17].

1.2.3.1 La relation d'équivalence

Dite aussi « relation de substitution » (AFNOR 1981), elle lie tous les termes d'une même langue qui expriment le même concept. Elle est définie dans la norme ISO 25964 comme « *la relation entre deux termes d'un thesaurus représentant chacun le même concept* »¹⁸. Elle recouvre trois types d'équivalents intralinguistiques (ou considérés comme tels) : la synonymie véritable, la quasi-synonymie et l'antonymie.

1.2.3.1.1 Synonymie véritable

Il peut d'agir de :

- la forme complète et son abréviation ou sigle ;
- les termes d'origines linguistiques ou culturelles différentes mais utilisés couramment pour désigner le même concept ;
- le nom populaire et le nom scientifique ;
- les variantes orthographiques ;
- les appellations ancienne et moderne.

¹⁷ Ce schéma est issu des travaux du linguiste danois Hjelmslev.

¹⁸ [1, ISO], Clause 2.18, p. 4 (traduction de l'auteur).

Les deux premiers types de synonymie sont ceux qui ont été le plus souvent rencontrés au cours de ce projet. Le secteur de la finance est particulièrement sujet à l'utilisation de termes anglo-saxons, bien souvent plus couramment que les termes francophones. De ce fait, si les termes français ont été privilégiés en principe, le terme anglo-saxon a plusieurs fois été défini comme descripteur dans un souci de meilleure adaptation à la culture des utilisateurs.

1.2.3.1.2 *Quasi-synonymie*

Elle concerne deux types de cas principalement :

- Des termes de niveaux hiérarchiques différents nécessaires à l'exhaustivité du thesaurus mais trop spécifiques pour être utilisés comme descripteurs. Un terme plus général leur sera alors préféré, dont ils seront définis comme employés pour.

Ex : Ours
EP *Ours blanc*
EP *Ours brun*

- Des concepts très proches difficiles à distinguer pour des non spécialistes.

1.2.3.1.3 *Antonymie*

Il s'agit dans ce cas des deux pôles d'un même axe sémantique.

Ex : Imposition
EP *Non-imposition*

Ces deux aspects d'un même concept n'ont pas besoin d'être deux descripteurs indépendants, mais ils sont importants à relier, dans la mesure où une ressource portant sur l'un fera fréquemment référence au second aspect du concept.

1.2.3.2 La relation hiérarchique

Ce type de relation, dite aussi « relation d'inclusion » associe un terme préférentiel générique (TG) à un ou plusieurs termes préférentiels spécifiques (TS). La norme ISO 25964 la définit comme « *la relation entre une paire de concepts dont la portée [limites sémantiques] de l'un est recouverte complètement par la portée de l'autre* »¹⁹. Michèle Hudon résume ainsi la fonction de ce type de relation, qui est à la base de toute la structure théssaurale :

« *La relation hiérarchique facilite la navigation verticale dans le thesaurus, permettant à l'utilisateur d'atteindre le degré de précision désiré dans l'indexation ou la recherche.* » [9, HUDON, p.57]

¹⁹ [1, ISO], Clause 2.23, p.5 (traduction de l'auteur).

On distingue trois types de relations hiérarchiques :

- la relation générique (de division à subdivision de même nature). Il s'agit du type de relation hiérarchique privilégié.

Ex : TGG Animal TSG Mammifère

- la relation partitive (de tout à partie)

Ex : TGP Système digestif TSP Estomac

- la relation d'instance (de classe générique à spécimen)

Ex : TGI Montagne TSI Alpes

Si la précision du type de relation hiérarchique reste facultative selon la norme ISO 25964, ces distinctions peuvent s'avérer utiles conceptuellement au cours de la phase de réflexion et de sélection des termes. C'est par ailleurs le genre de précision qui peut s'avérer nécessaire lorsqu'une évolution est envisagée vers un vocabulaire plus descriptif, de type ontologie.

Dans certains cas, et pour offrir plusieurs clés d'accès, on utilisera la polyhiérarchie pour lier un terme à des concepts de catégories différentes.

Ex : Caméra sous-marine TG Caméra à usage spécifique

Caméra sous-marine TG Matériel de plongée

Ex 2 : Crâne TGG Os

Crâne TGP Tête

Le modèle de données proposé par la norme ne fixe aucune limite quant au nombre de domaines auxquels un même concept peut être lié hiérarchiquement. Néanmoins, pour ce projet, il a été décidé de ne pas utiliser la polyhiérarchie, afin de ne pas compliquer la structure du thesaurus.

1.2.3.3 La relation associative

Ce dernier type de relation permet de lier des concepts généralement associés mentalement par le public utilisateur du thesaurus (à l'indexation ou à la recherche). La norme ISO 25964 la définit comme « *la relation entre une paire de concepts qui ne sont pas reliés hiérarchiquement, mais partagent une forte connexion sémantique* »²⁰. Elles sont donc non-hiérarchiques et peuvent lier des concepts appartenant à des domaines (ou micro-thesaurus) différents. On associera de préférence des termes de niveaux de spécificité comparables.

Une association conceptuelle étant par nature très subjective, il est important de toujours s'interroger sur son utilité et sa pertinence. Elle doit aider, selon les cas, à percevoir la signification

²⁰ [1, ISO], Clause 2.2, p.1 (traduction de l'auteur).

du descripteur et/ou appréhender son aire sémantique globale. A la recherche, ce type de relation constitue enfin une aide non négligeable pour orienter l'utilisateur vers des concepts liés à sa requête initiale qui devraient lui permettre de l'enrichir, augmentant ainsi ses chances d'accéder aux ressources attendues et pertinentes.

M. Hudon recense 13 principaux types de relations associatives [9, HUDON, p. 132-133] :

- la cause et l'effet : *Infection* et *Maladie*
- un tout et une composante essentielle : *Livre* et *Reliure*
- une action et son agent : *Jardinage* et *Jardinier*
- une action et son produit : *Jardinier* et *Jardin*
- une action et son objet : *Récolte* et *Cultures*
- une action et le lieu de son déroulement : *Enseignement* et *Ecole*
- une science et son objet : *Paléontologie* et *Fossile*
- un objet et sa propriété : *Poison* et *Toxicité*
- un objet et son application : *Ordinateur* et *Traitement des données*
- un objet et un de ses matériaux constitutifs : *Bijou* et *Pierre précieuse*

Mais elles peuvent être également utiles pour associer :

- des concepts de sens proche : *Bateau* et *Navire*
- des antonymes : *Tolérance* et *Préjugé*
- des concepts complémentaires : *Enseignement* et *Apprentissage*

En règle générale, une relation associative est réciproque, et de ce fait symétrique dans sa représentation au sein du thesaurus (c'est-à-dire signalée pour chaque terme de la relation).

1.3 Un thesaurus, pour quoi faire ?

On l'a vu, la structure et le fonctionnement d'un thesaurus restent essentiellement les mêmes d'une norme à l'autre [5, DEXTRE CLARKE]. Ce qui a certainement évolué depuis la dernière édition de la norme en 1986, en revanche, ce sont les objectifs poursuivis lors de la conception d'un thesaurus, ses utilisations initialement prévues, mais aussi les évolutions ultérieures, qui doivent aujourd'hui être systématiquement envisagées, tant les dernières années nous ont souvent confrontés à des projets rendus obsolètes avant même leur aboutissement.

1.3.1 Des fonctions traditionnelles

Comme on l'a évoqué plus haut (Cf.1.1.2), les fonctions essentielles du thesaurus évoquées par la norme ISO 25964 restent inchangées. Il s'agit de fournir un vocabulaire contrôlé et normalisé permettant :

- d'une part, d'indexer des documents dans une base de données en décrivant précisément leur contenu, grâce à un langage validé et partagé ;
- d'autre part, d'effectuer des recherches au sein de cette même base de données et d'assurer un rappel optimal des documents pertinents traitant d'un concept donné.

Le processus d'indexation peut être décrit à travers les deux étapes de l'opération de « *réécriture du contenu informationnel de document* » [17, COURBIERES, p. 157] :

- l'analyse documentaire, qui permet d'extraire les concepts contenus dans le document ;
- la représentation de l'information, à travers la traduction de ces concepts en langage documentaire.

Le thesaurus a donc pour fonction principale de garantir l'usage d'un même descripteur pour représenter un concept unique à chaque bout de cette chaîne d'opérations, quel que soit le nombre d'acteurs impliqués, qu'il s'agisse d'un documentaliste ou d'un utilisateur consultant la base. Ainsi, « *si [le professionnel de l'information-documentation] peut être vu comme une interface entre les utilisateurs et les documents d'un système documentaire, c'est notamment par le biais du ou des langages documentaires dont il dispose et qu'il utilise.* » [16, COURBIERES, p.131]

1.3.2 Un périmètre croissant et de nouveaux usages

1.3.2.1 Un usage généralisé par l'Internet

Le déploiement quasi systématique aujourd'hui des bases de données documentaires sur des réseaux utilisant le système Internet, qu'il s'agisse d'un intranet, d'un extranet ou encore du Web, a eu pour conséquence naturelle d'élargir l'usage des thesauri et autres langages contrôlés – sans que l'internaute en ait nécessairement conscience, cependant. Effectuant désormais le plus souvent ses recherches de façon autonome, du moins dans un premier temps, l'utilisateur se voit de plus en plus fréquemment offrir la possibilité d'accéder au(x) langage(s) d'indexation employé(s) par la base de données qu'il consulte, en totalité ou parfois dans une version simplifiée. Le thesaurus se positionne dès lors comme un guide au sein du fonds documentaire qu'il indexe, traduisant en filigrane la vision du monde (ou du domaine de spécialité) développée par l'organisme ayant composé et réuni ce fonds dans un but spécifique.

1.3.2.2 Des outils d'aide à la découverte

Une autre conséquence de cet accès privilégié au thesaurus est de permettre au public de l'utiliser comme un outil de découverte. La navigation par facette, souvent basée sur une taxonomie voire extraite d'un thesaurus, a popularisé cet emploi du thesaurus consistant à explorer sa hiérarchie à partir d'un sujet de recherche vague. En cela, le thesaurus est clairement devenu un outil de

sérendipité²¹, cette « sagacité accidentelle » qui nous permet de trouver ce que nous ne cherchions pas, mal ou sans le savoir. Doit-on dès lors envisager de concevoir dès le départ un thesaurus comme un outil de découverte d'un domaine ?

Certains usages le justifient certainement, et il me semble que le site portail du Québec en offre un parfait exemple avec le Thésaurus de l'Activité Gouvernementale (TAG)²². Il a été conçu comme « *un outil qui vous aide à trouver le mot juste permettant de faire une recherche efficace dans les sites Web de l'Administration québécoise* ». Et pour ce faire, le point de vue adopté est bien sûr celui de l'utilisateur : le citoyen cherchant à effectuer telle ou telle démarche auprès de l'administration. Le thesaurus a donc été construit à partir de facettes décrivant successivement :

- Les domaines couverts par l'administration
- Les processus dans lequel le citoyen souhaite s'inscrire
- Les documents qu'il cherche à obtenir
- La « clientèle » à laquelle il appartient (entreprise, groupe de personnes ou individu)
- L'intervenant auquel il souhaite s'adresser (organisme, service...)

La norme ISO 25964 évoque également cette nouvelle forme d'organisation du thesaurus par facettes (Cf. 1.2.2.2), qui n'exclut d'ailleurs pas une organisation thématique des termes, mais peut au contraire lui être combinée. Cet ajout d'une entrée par typologie est un apport considérable pour la recherche, puisqu'elle multiplie les voies d'accès et les associations possibles entre les concepts, qui peuvent ainsi déployer différents types de liens conceptuels.

1.3.3 Une utilité renforcée

1.3.3.1 Omniprésence des langages contrôlés

Les langages contrôlés représentent donc un outil central, pour les applications Internet comme on vient de le voir, mais aussi au sein des entreprises, en-dehors même des centres de documentation. En effet, de nombreuses applications devenues indispensables – pour la gestion de l'information, sa publication et le travail collaboratif (CMS, ERP, GED, etc.), le *knowledge management*, ou encore l'audit et la conformité – nécessitent des référentiels métiers, devenus parfois si imposants et complexes qu'ils peuvent être rapprochés d'un thesaurus, même s'ils n'ont pas toujours été conçus comme tels initialement. De nombreux projets portant sur l'implémentation ou la mise à jour de ces types d'outils font donc nécessairement intervenir des spécialistes des langages documentaires, pour répondre aux exigences de gestion de contenus

²¹ De l'anglais *serendipity*, néologisme créé en 1754 par Horace Walpole pour désigner les « découvertes inattendues, faites par accidents et sagacité ».

²² <http://www.thesaurus.gouv.qc.ca/tag/accueil.do> Voir aussi le très enrichissant retour d'expérience recueilli par Sylvie Dalbin sur son blog, *Descripteurs* [13, DALBIN] : <http://dossierdoc.typepad.com/descripteurs/2012/05/le-tag-thesaurus-activite-gouvernementale-quebec-temoignage.html>

désormais incontournables. Les consultant(e)s spécialisées dans ce domaine Sylvie Dalbin (Assistance et Techniques Documentaires) et Hélène Zysman (Smile) m'ont confirmé mener continuellement des missions portant sur les outils terminologiques d'organisation ou d'accès à la connaissance.

1.3.3.2 Des outils dédiés de plus en plus nombreux

Bien sûr, l'évolution des logiciels et des langages informatiques traduisant les vocabulaires a été un préalable indispensable à cette évolution du périmètre d'action, et partant des possibilités offertes dans la conception de thesaurus. Le standard SKOS, notamment, a largement permis et encouragé le mouvement de diffusion de ce langage (Cf. 9.1.3). Après s'être longtemps fait discrets, se contentant d'être assimilés à des modules embarqués indispensables aux logiciels de gestion de l'information, les outils dédiés aux référentiels font aujourd'hui de plus en plus parler d'eux. Le groupe « Taxonomies et Métadonnées » de l'ADBS leur consacre d'ailleurs une journée d'étude très attendue le 26 novembre 2013.²³

Les questions d'interopérabilité entre les langages est aujourd'hui au cœur des préoccupations en informatique documentaire, que le déploiement en reste interne à l'entreprise ou qu'ils soient exploités sur le Web. Ce n'est donc pas un hasard si la seconde partie de la norme [2, ISO] leur est plus particulièrement consacrée, s'ajoutant aux Clauses 14 à 18 qui clôturent sa première partie [1, ISO, p. 98-121]. Les outils autonomes (*standalone*) de création, gestion et maintenance de thesaurus se multiplient actuellement.²⁴ Signe de cet intérêt croissant et des enjeux pour la diffusion de l'information, plusieurs projets *open source*, libres et gratuits, se sont largement déployés ces dernières années : TemaTres (Argentine), ThManager (Espagne), The32W (Canada) principalement. Le Ministère de la Culture et de la Communication vient ainsi de développer son propre outil, baptisé GINCO (Gestion Informatisée de Nomenclatures Collaboratives et Ouvertes)²⁵, qui s'appuie sur la norme ISO 25964 et le standard SKOS.

Le thesaurus, comme les autres langages artificiels, est donc plus que jamais au centre des préoccupations. Penchons-nous maintenant plus avant sur les caractéristiques des thesaurus spécialisés, cette fois.

²³ <http://www.adbs.fr/du-thesaurus-aux-referentiels-terminologiques-134144.htm> (consulté le 13/11/2013). Les supports de la plupart des intervenants seront publiés sur le site.

²⁴ Sylvie Dalbin propose sur son blog *Descripteurs* [13, DALBIN] un très précieux panorama des logiciels disponibles : http://dossierdoc.typepad.com/descripteurs/2011/12/logiciel_de_thesaurus.html (dernière mise à jour : Juillet 2013).

²⁵ Plus d'informations ici : <http://www.culturecommunication.gouv.fr/Etudes-et-documentation/HADOC/Referentiels2/Les-vocabulaires-scientifiques-et-techniques/Le-projet-GINCO>

2 Un thesaurus spécialisé : pour répondre à quels besoins ?

Un thesaurus est toujours, par la force des choses, la traduction d'une certaine représentation du monde intrinsèquement liée aux activités, objectifs et contexte de son organisme de production. C'est pourquoi il s'avère bien souvent délicat de reprendre tel quel, même en partie, un thesaurus développé par d'autres pour l'appliquer à son propre projet et son environnement. C'est à coup sûr s'exposer à des arbitrages impossibles entre des points de vue inconciliables. Un thesaurus spécialisé – c'est-à-dire non généraliste – doit, pour être efficace, être conçu afin de s'adapter à trois paramètres également essentiels et intrinsèquement liés : les usages, les utilisateurs et les activités de l'organisme concerné, institution ou entreprise.

2.1 Adaptation aux usages

Un thesaurus a deux usages principaux, correspondant en miroir à ses deux fonctions principales : l'indexation et la recherche (Cf.1.3.1).

2.1.1 Décrire le fonds documentaire

L'un des premiers engagements que doit tenir tout langage documentaire est de décrire fidèlement le fonds documentaire qu'il doit permettre de représenter – ressources déjà réunies ou susceptibles d'être acquises à l'avenir par le centre de documentation. Bien sûr, le type d'analyse employé pour le traitement de l'information doit d'abord correspondre à chaque type de document. D'autre part, le plan de classement reflète la constitution du fonds en présentant de façon synthétique les différents domaines abordés. Enfin, l'outil permettant de décrire le plus précisément le contenu informationnel des documents est le langage d'indexation, dont le thesaurus constitue la forme la plus aboutie.

Le thesaurus doit donc être élaboré de façon à décrire au plus près le contenu d'un fonds documentaire donné. Pour un thesaurus spécialisé en entreprise, cela sous-entend la sélection des domaines concernant effectivement les activités de l'entreprise, au sein desquels sont rassemblés les concepts réellement susceptibles de décrire les ressources répondant aux besoins informationnels des collaborateurs de l'entreprise. C'est à cette condition qu'il pourra devenir un outil efficace et simple d'utilisation pour les documentalistes chargés de l'indexation et de la production de produits documentaires au cours du processus de traitement de l'information.

2.1.2 Aider à la recherche

Mis à disposition de l'utilisateur pour formuler son besoin d'information, le thesaurus est un outil précieux d'aide à la recherche dans la mesure où il aura permis, dans un premier temps, une

description fidèle et pragmatique des documents indexés dans la base de données. Il peut être dès lors défini comme un « *code commun aux deux instances de production et de réception [...] des produits documentaires élaborés dans le cadre de l'indexation.* » [16, COURBIERES, p.131]. Il est donc l'un des vecteurs principaux assistant le documentaliste dans sa fonction de médiation entre les documents, porteurs d'informations, et son ou ses publics, dans leur recherche d'information.

2.2 Adaptation aux utilisateurs

Le thesaurus doit permettre de représenter le fonds documentaire. Or, ce fonds est lui-même le reflet, plus ou moins idéalisé, de l'activité de l'organisation qui l'abrite et des profils d'utilisateurs qui constituent son public.

2.2.1 Bien connaître ses publics...

On a évoqué la fonction de médiation, qui est aujourd'hui au cœur de la conception des professions de l'information-documentation. Annette Beguin-Verbrugge en donne une définition, du point de vue pédagogique, qui nous paraît particulièrement pertinente ici :

« La médiation documentaire est une médiation qui n'implique pas le rapport personnel direct, mais qui implique cependant le recours au langage et la prise en compte du niveau de connaissance de l'utilisateur et de ses capacités d'abstraction. »²⁶

Ce sont bien les spécificités de l'utilisateur, son niveau d'expertise du domaine, mais aussi son niveau de maîtrise des techniques de recherche documentaire, qui doivent conditionner les pratiques d'indexation, considérées comme un outil privilégié de la médiation. Dans le contexte de l'entreprise, la connaissance du public visé passe prioritairement par la connaissance des métiers exercés et des activités menées pour le compte de l'organisation. Ce champ d'exercice détermine en effet les domaines et spécialités dans lesquels la valeur ajoutée de l'information fournie sera la plus élevée.

2.2.2 ... pour anticiper leurs besoins

Le travail de « facilitation de l'accès à l'information » [22, REGIMBEAU, p. 106] qui constitue l'objet des différents traitements documentaires ne peut être réalisé efficacement que si le documentaliste qui en a la charge connaît suffisamment ses utilisateurs pour être à même d'anticiper leurs besoins. Une parfaite connaissance de l'organisation dans laquelle il s'insère, des

²⁶ Lors du XIII^e Congrès national des Sciences de l'Information et de la Communication, 7-9 octobre 2002, Marseille. Cité dans [22, REGIMBEAU, p. 83]

métiers et des activités qui y sont exercés, mais aussi son expérience concrète des besoins, réunie au fur et à mesure des recherches effectuées, sont les principaux éléments dont il dispose pour ce travail.

Cette dimension anticipative s'exprime principalement dans deux activités essentielles d'un centre de ressources documentaires :

- lors du repérage des ressources potentiellement utiles à ses utilisateurs, et qui devront à ce titre intégrer le fonds documentaire ;
- lors de l'élaboration de produits documentaires décrivant ses ressources, notamment l'indexation, de façon à faire ressortir les éléments identifiés en amont comme nécessaires à l'utilisateur et le plus susceptibles de répondre à son besoin d'information.

2.3 Le reflet des activités et de la culture d'une entreprise

Chaque entreprise est unique, et s'il est difficile de plaquer un thesaurus existant sur les ressources documentaires d'une entreprise, c'est d'abord parce que ces ressources sont le reflet des activités au service desquelles elles ont été réunies, mais aussi – surtout ? – car elles expriment la culture de cette entreprise avec plus de justesse que n'importe quel document façonné *ad hoc* par son service de communication.

2.3.1 Un référentiel des activités de l'entreprise

On l'a vu, le langage documentaire est chargé, d'une part, de décrire fidèlement les ressources constituant le fonds documentaire réuni pour répondre aux besoins d'information des collaborateurs de l'entreprise, d'autre part, de correspondre aux métiers et activités des utilisateurs de la base de données pour mieux les aider dans leurs requêtes. Partant de cette double orientation – même si une adéquation totale à ces deux objectifs reste un idéal – il est aisé de considérer le thesaurus d'entreprise, spécialisé par nature, comme un référentiel des activités de l'entreprise, sous sa forme la plus évoluée.

2.3.2 Une matérialisation de la culture de l'entreprise

Les documentalistes chargés de réunir les documents utiles aux collaborateurs de l'entreprise pour laquelle ils œuvrent, *a fortiori* dans un service de veille, passent leurs journées à faire des choix. Ils choisissent les documents qui pourront être utiles à leur public, lui aussi au service de cette même entreprise. A travers ces choix, guidés par leur expérience et leur connaissance de l'organisation, ils réunissent un corpus qui dresse un portrait en ombre chinoise de la culture de celle-ci.

Un langage documentaire étant conçu comme le reflet du fonds documentaire qu'il permet d'indexer, il façonne nécessairement un langage qui peut être considéré comme une transcription, une représentation sémantique des éléments qui construisent la culture de l'entreprise au sein de laquelle il opère : ses métiers, ses préoccupations, et même ses valeurs. Il est même possible d'aller plus loin dans cette compréhension, en analysant les occurrences effectives de chaque descripteur dans la base de données, et d'en déduire l'importance de chaque concept au sein de ce tableau.

Au vu de ces aspects, l'apport que cet outil pourrait représenter pour l'organisation des différents outils de gestion de l'information présents dans l'entreprise, mais dont le contrôle échappe aux professionnels de l'information-documentation, paraît évident. Il leur appartient de faire connaître leur expertise et l'étendue de leur réflexion dans ce domaine au sein de l'entreprise, afin de fédérer et d'articuler les projets qui pourraient bénéficier de ce référentiel commun. Malheureusement, la complexité des organisations, le positionnement de ces professionnels, mais aussi la perception des fonctions documentaires dans l'entreprise représentent trop souvent autant de freins au déploiement de telles politiques unifiées de gouvernance de l'information.

3 Production, utilisation et maintien d'un thesaurus spécialisé en entreprise : l'exemple de BPCE

L'objectif du projet mené durant le stage était la refonte du thesaurus documentaire utilisé à l'indexation et à la recherche par les documentalistes du service Documentation générale de BPCE SA. Il est également disponible depuis peu à la recherche, directement par les collaborateurs de l'entreprise (utilisateurs), via le portail documentaire accessible depuis l'Intranet. Il s'agit dans le cas qui nous intéresse d'un thesaurus monolingue spécialisé, centré sur les activités banque, finance et assurance, complété par d'autres domaines nécessaires aux activités annexes menées dans le groupe. Nous l'analysons ici dans un double objectif : pour comprendre le langage qui a servi de base à cette mission bien sûr, mais aussi comme un exemple significatif de la conception et des usages d'outils d'indexation en entreprise.

3.1 Petit historique du thesaurus de la CNCE

Afin de mieux comprendre le travail de refonte qui sera détaillé dans la seconde partie, il nous semble intéressant de présenter dans un premier temps les conditions de production, de maintenance et d'utilisation du thesaurus utilisé jusqu'ici au sein du service.

3.1.1 Production du « Thesaurus Techlib²⁷ »

3.1.1.1 Le contexte : la création d'un service de documentation à l'organe central

Le thesaurus sur lequel a porté cette mission a été élaboré au cours de l'année 1990, à l'occasion de la création du service Documentation de la Caisse Nationale des Caisses d'Épargne. Plusieurs services se sont alors trouvés fusionnés, et leurs fonds documentaires respectifs réunis. Un logiciel de gestion de bibliothèque, BASIS Techlib, avait été acheté par la direction informatique pour informatiser le fonds ainsi constitué. La conception d'un langage d'indexation contrôlé adapté à l'entreprise devint donc l'une des priorités de la nouvelle responsable (aujourd'hui à la tête de la Documentation générale de BPCE).

3.1.1.2 Une méthode pragmatique

L'équipe constituant ce nouveau service se limitant à quatre documentalistes, et déjà fort sollicitées, une méthode rapide et efficace fut mise en œuvre. Dans un premier temps, des thesauri existants concernant les domaines d'activité de l'entreprise furent identifiés. Il s'agissait essentiellement de ceux de l'Association française des banques (AFB), de Delphes (base de

²⁷ Dans la suite de ce mémoire, nous parlerons du « thesaurus Techlib » pour faire référence au thesaurus qui a fait l'objet de cette refonte, en usage jusqu'ici à la DG de BPCE.

données économiques) et de la Commission des opérations de bourse (devenue l'Autorité des marchés financiers, AMF). Après avoir été redimensionnés et élagués, ils ont été enrichis des concepts et termes propres aux Caisses d'épargne, formant ainsi un « patchwork de thesauri ». Les principales préconisations de la norme ISO 2788 furent appliquées.

Les séances de travail se firent essentiellement en tête-à-tête entre la responsable du service et la spécialiste du domaine concerné. Des problématiques interpersonnelles importantes rendirent en effet le travail en équipe délicat. Néanmoins, l'efficacité de la méthode employée permit de concevoir un thesaurus stable et adapté à l'usage en quelques mois seulement. Sa création dans le logiciel BASIS demanda quelques mois supplémentaires, si bien que l'outil informatique était opérationnel fin 1991, soit un peu moins de deux ans après la création du service, en février/mars 1990.

3.1.1.3 Spécificité du thesaurus de la CNCE

On l'a vu, la conception du langage d'indexation et l'organisation du fonds physique de ressources documentaires ont dû être menés de front par les documentalistes de la CNCE. Pour des raisons pragmatiques, liées aussi bien à l'efficacité du projet qu'à l'adaptation aux besoins du service, les deux objectifs ont été menés conjointement. Un plan de classement a donc été créé pour organiser le fonds documentaire, qui a ensuite servi de matrice au thesaurus élaboré selon la méthode relatée plus haut. Ainsi, si les principes du thesaurus ont été appliqués dans l'esprit, la forme correspondait en réalité à celle d'un plan de classement agrémenté d'un très riche lexique hiérarchisé. Autre argument en ce sens, les relations associatives étaient matérialisées par un VA, « voir aussi », plus typique du plan de classement. Les termes proposés pour chaque domaine et sous-domaine du plan de classement permettaient donc, à la fois, d'identifier la meilleure cotation à attribuer à une ressource, et de constituer son indexat pour son référencement dans la base de données documentaire.

Néanmoins, cet outil hybride étant désigné par ses utilisatrices principales, les documentalistes de la CNCE, puis de la nouvelle structure créée lors de la fusion pour BPCE SA, comme un « thesaurus », nous continuerons de le désigner ainsi. Chaque outil ne doit-il pas avant tout être adapté à l'usage que l'on fait ? Or, nous avons présenté la logique ayant procédé à son élaboration, et il s'est révélé, durant de nombreuses années, correspondre parfaitement aux objectifs initiaux. On verra cependant dans la seconde partie que cette spécificité fut lourde de conséquences lors de la mission de refonte de ce dit thesaurus, plus de vingt ans plus tard.

3.1.2 Vie du thesaurus

3.1.2.1 Une maintenance au fil de l'eau

La maintenance de ce thesaurus a été menée au fil de l'eau, sans que le besoin de mettre en place une procédure formalisée se soit fait ressentir. Deux personnes, dont la responsable du service, étaient en charge des modifications. Les documentalistes, lorsqu'elles identifiaient un concept manquant lors de leurs activités d'indexation, proposaient l'ajout d'un descripteur correspondant. Si celui-ci était validé en réunion de service, il était introduit à la fois dans le thesaurus « papier » (inséré dans le plan de classement) et dans le module de gestion de thesaurus du logiciel Techlib. On procédait de même pour les changements de descripteurs ou autres adaptations nécessaires. Le thesaurus de la CNCE s'éloignait ainsi progressivement des référentiels qui lui avaient initialement servi de canevas, devenant peu à peu celui des activités propres à son entreprise.

Aucun suivi ne semble avoir été fait de ces opérations, si bien que, ce logiciel ne proposant pas de données statistiques concernant le thesaurus, il est impossible de connaître le volume des modifications, ajouts ou suppressions, opérés jusqu'en 2008. C'est à cette date en effet que le thesaurus a pratiquement cessé d'être mis à jour, notamment en raison d'un plan de sauvegarde de l'emploi ayant entraîné une réduction des effectifs de 60%. Néanmoins, ces nouveaux descripteurs ayant le plus souvent été ajoutés tels quels, sans le développement de leurs propres relations, on remarque dans ce thesaurus une proportion élevée de termes orphelins, qui donne une indication de la portée de ces modifications sur la structure thésaurale de l'ensemble du lexique. On précisera par ailleurs que le plan de classement, quant à lui n'a pas subi d'évolution durant cette période.

3.1.2.2 Evolution des usages du thesaurus à la CNCE

Initialement, le thesaurus informatisé sous le logiciel Techlib était utilisé uniquement par les documentalistes du service, pour l'indexation des ouvrages dans la base de données et la recherche au sein de ce fonds documentaire. En 2005, fut mis en place l'intranet de la CNCE, ouvrant l'accès à la base de données et à son thesaurus à tous les collaborateurs de l'organe central des Caisses d'épargne, seuls à pouvoir emprunter les ouvrages. D'autre part, un « club métier Documentation » devait permettre de fédérer les usages et les bonnes pratiques des différents documentalistes répartis dans le groupe, dont la plupart travaillaient seuls, disposant de ressources restreintes. La plupart de ces services disposaient déjà de leur propre système d'indexation lorsque le service documentation de l'organe central fut créé. Néanmoins le thesaurus développé par la CNCE leur fut communiqué, et certains ont pu s'en inspirer pour leurs projets internes.

La même année, une documentation juridique distincte fut créée, et la partie correspondante du fonds documentaire y fut transférée. Néanmoins, les notices concernant ces ouvrages n'ont pas toutes été supprimées de la base de données Techlib à l'époque. Cette partie du plan de classement comme du thesaurus sont également restées intactes.

3.1.2.3 Evolution du service documentation de la CNCE

Les articles n'étaient pas enregistrés dans cette base – contrairement aux pratiques de la documentation de la BFBP, qui utilisaient à l'époque le logiciel DIP. Des dossiers documentaires thématiques en format papier étaient constitués à l'intention des utilisateurs de la CNCE et des recherches rétrospectives pouvaient être effectuées à la demande. Leur redondance conduisit à une évolution vers des pratiques de veilles thématiques avec envois systématiques, qui devinrent peu à peu l'essentiel de l'activité.

Ainsi, en 1993, soit 3 ans après la création du service, y étaient traitées en moyenne 240 questions par mois, pour seulement 250 envois systématiques. En 2011, le service Documentation de BPCE qui lui a succédé ne traitait plus que 160 questions par mois. En contrepartie, la diffusion de ses produits de veille atteignait 6 000 envois mensuels. Cette évolution, née d'une demande spécifique de la direction commerciale pour le suivi des clientèles sectorielles et le marché de l'économie sociale, s'est si bien généralisée que la veille constitue aujourd'hui l'activité principale du service Documentation générale de BPCE.

3.2 Activités de la Documentation générale de BPCE SA

Un thesaurus étant avant tout conçu comme le reflet de ses usages, des besoins auxquels il répond, il est nécessaire de présenter d'abord son contexte d'utilisation, pour mieux comprendre ensuite les raisons qui ont poussé à sa refonte et les orientations qui ont été choisies.

3.2.1 La constitution et la gestion du fonds documentaire

Le fonds documentaire proposé par la DG et indexé dans la base de données du logiciel de gestion de bibliothèque, se compose essentiellement de deux types de ressources :

3.2.1.1 Les ouvrages

Rassemblé au fil des années, le fonds documentaires physique est constitué d'ouvrages de types monographies ou publications en série (revues de fonds). Ils traitent essentiellement :

- des domaines d'activité du groupe : la banque commerciale, les activités de marché et services financiers, l'assurance et l'immobilier en France ;
- de l'environnement de l'entreprise et autres informations d'ordre générique utiles à l'activité : institutions, politique, cadre législatif (très général), relations internationales,

sciences économiques, économie sociale et solidaire ou encore développement durable... ;

- des autres domaines nécessaires à l'activité de l'entreprise : ressources humaines et relations sociales, gestion commerciale et marketing, informatique et traitement de l'information, etc.

La base de données Techlib indexait, au moment de son intégration au logiciel JLB, de 5 000 à 6 000 notices. Un désherbage mené en 2012 a porté sur les ouvrages obsolètes, ainsi que sur les notices pour les faire correspondre à l'existant. Néanmoins, de nombreuses notices héritées de la base de données DIP, de la BFBP, indexant des articles de périodiques ont subsisté sans que les documents correspondant aient été conservés. Ces notices parasites ont rendu délicate l'analyse de l'emploi des termes du thesaurus.

3.2.1.2 Les études économiques sectorielles ou concurrentielles

Pour répondre aux besoins d'information spécifiques portant sur certains marchés ou secteurs d'activité, BPCE a négocié, via son service de négociation et achats, un contrat avec les fournisseurs d'études économiques DAFSA et Xerfi. Sur les 5 221 notices que compte la base de données en novembre 2013, 247 concernent ce type de ressources. Elles constituent une part importante des demandes utilisateurs.

D'abord intégrées sous format papier au fonds documentaires, elles sont devenues peu à peu disponibles en version électronique. Cette année enfin, un accès direct aux bases de données de ces fournisseurs a été négocié pour les collaborateurs du groupe concernés par ces informations. L'accès se fait via l'espace intranet de la Documentation générale, qui gère et met en place les droits d'accès des utilisateurs. Aucune statistique précise n'est disponible à l'heure actuelle, mais l'équipe de la DG n'a pas remarqué une réelle diminution des demandes depuis l'installation de ce système. Il a simplement facilité l'accès aux études pour les collaborateurs du reste du groupe. De plus, beaucoup de demandes utilisateurs ne sont pas aussi spécifiques, portant plutôt sur « ce qu'on peut leur fournir » sur une thématique donnée, tous types de documents confondus.

3.2.1.3 La gestion des abonnements

A ces ressources présentes dans la base de données s'ajoutent bien sûr les périodiques dépouillés pour les veilles, et qui sont tenus à disposition des utilisateurs pour consultation. Les collections sont conservées sur un à deux ans, puis archivées. Outre les abonnements contractés pour le service, la Documentation générale est en charge de la gestion des abonnements souscrits pour les différents collaborateurs de l'entreprise qui peuvent y prétendre, soit environ 200 abonnements au total.

3.2.2 Au cœur de l'offre, les veilles documentaires thématiques

3.2.2.1 L'offre de veille

L'activité principale menée par les 5 documentalistes composant l'équipe de la DG est la production de produits de veille documentaire²⁸. L'essentiel de la presse généraliste, ainsi que les principales revues spécialisées dont les thématiques touchent les activités du groupe, sont analysées. Le service produit au total 15 veilles thématiques dont :

- 5 hebdomadaires, concernant les cœurs de métier ;
- 9 bimensuelles, dont 4 composées en majorité d'articles en texte intégral ;
- les rapports annuels des concurrents banque et assurance (durée de vie semestrielle) ;
- auxquels s'ajoutent 2 veilles règlementaires, diffusées quotidiennement.

Il s'agit d'une part de veilles « métiers », d'autre part de veilles « marchés » ou concurrentielles. Elles sont organisées autour de grandes thématiques s'adressant aux différentes entités du groupe. Il s'agit donc simultanément d'une activité de veille commerciale, concurrentielle, environnementale et réglementaire, stratégique ou encore créative. Dans chacune d'elles, les sous-chapitres évoluent en fonction de l'actualité. Les articles privilégiés sont assez généralistes, ou bien concernent directement le groupe ou ses concurrents. La grande variété des profils des lecteurs exige une sélection pouvant s'adresser à tous. Des articles plus pointus peuvent donc être adressés directement à un lecteur particulièrement concerné. A l'heure actuelle, et dans l'attente de la mise en place prochaine d'un module dédié à l'édition des veilles sous le logiciel JLB-Net, les veilles sont réalisées sous un format Word, diffusé aux abonnés par mail en pièce jointe. Enfin, les veilleuses-documentalistes réalisent également des panoramas à l'aide d'agrégateurs de presse.

Enfin, la veille concernant la réputation du groupe et de ses différents actionnaires ou filiales est gérée directement par la communication interne. Elle est automatisée et confiée à un fournisseur externe, envoyée quotidiennement aux instances de direction ainsi qu'à la Documentation générale. Les deux ou trois articles les plus pertinents sont enfin publiés sur l'intranet à destinations de tous les collaborateurs du groupe.

3.2.2.2 Publics et satisfaction estimée

Le positionnement de la Documentation générale la place comme interlocutrice privilégiée de la plupart des collaborateurs des différents organes et filiales, dont la majorité ne dispose pas d'une documentation en interne. Néanmoins, ses utilisateurs sont essentiellement les quelques 1 500

²⁸ Pour illustrer l'offre de veille documentaire à BPCE, se reporter au schéma de l'activité de veille en Annexe 7.

salariés, stagiaires et apprentis de BPCE SA (70% des utilisateurs²⁹). Elle centralise les demandes, répond à la plupart d'entre elles, et réoriente au besoin l'utilisateur si sa requête est trop pointue pour être traitée directement (notamment auprès de la documentation juridique de BPCE, ou de Natixis, dont la documentation interne dispose d'analyses sectorielles ou concurrentielles). D'autre part, elle met ses ressources à disposition des filiales et des actionnaires, via les documentalistes internes éventuellement.

L'enquête menée en 2011 a montré un taux de satisfaction très élevé quant aux services déjà proposés (91%), et le principal souhait évoqué a été un renforcement de la communication concernant les activités de la Documentation générale. Les principaux clients de ses produits de veille en 2011 étaient les services Risques et Finance (28,7%), Marketing et Distribution (16,8%), Systèmes d'information (15,2%) et Prestations de services et Conseil (14,1%), qui représentaient à eux quatre les trois quarts des utilisateurs. Il est probable que des proportions similaires soient observables pour les demandes de recherches. En outre, la Documentation générale s'adresse à une très grande variété de métiers, comme le montre la FIGURE II ci-dessous.

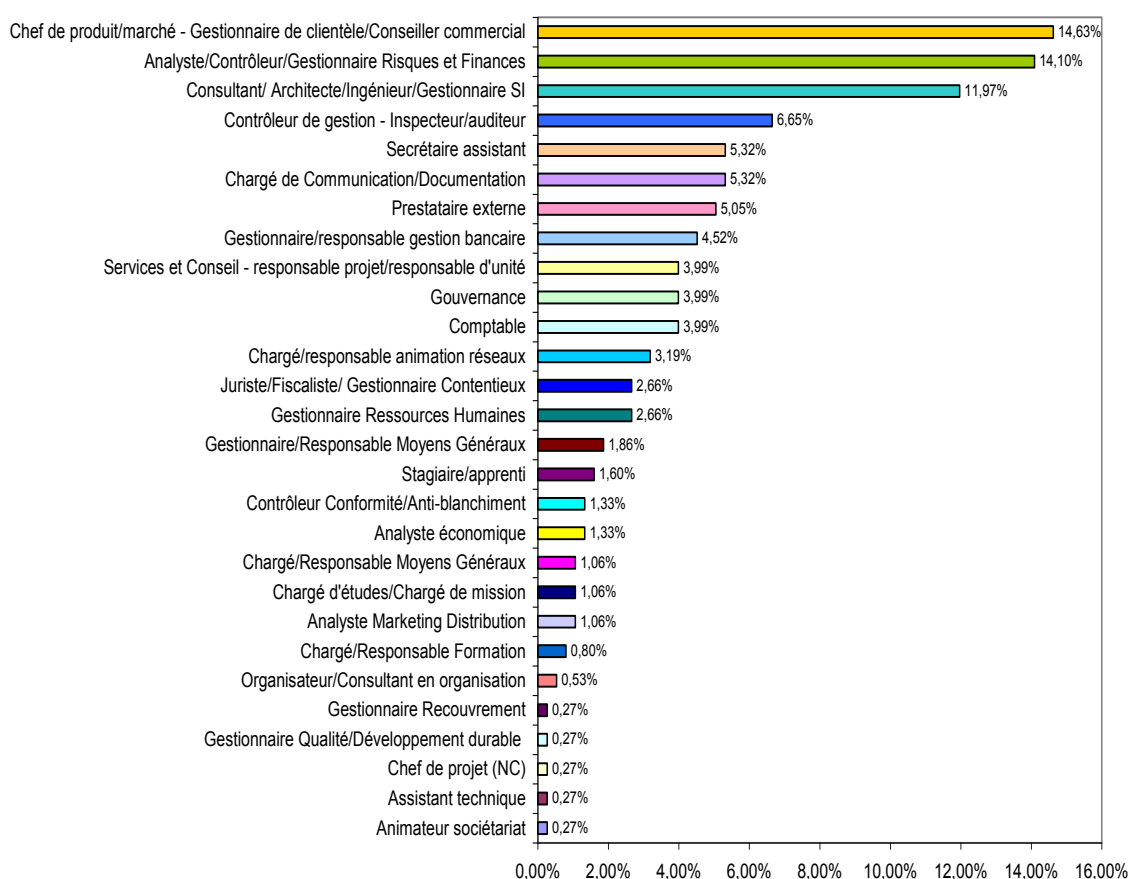


FIGURE II : METIERS DES UTILISATEURS DE LA DOCUMENTATION GENERALE (HISTOGRAMME DE REPARTITION)

²⁹ Les chiffres proposés dans cette partie ainsi que la Figure II sont issus de l'enquête qualité menée en interne par formulaires électroniques en juin/juillet 2011, soit deux ans après la fusion.

Pour mieux refléter les activités décrites ci-dessus, le service doit être renommé prochainement « Veille et documentation ». Cette initiative, portée par la responsable du service, est validée et soutenue par son N+1, le directeur du pôle Immobilier et Services généraux. Outre une meilleure adéquation avec la réalité des services apportés aux utilisateurs, une meilleure image du service et de ses produits auprès des collaborateurs de BPCE est attendue de cette modification.

3.3 Usages du thesaurus depuis la fusion

3.3.1 Indexation

L'usage principal du thesaurus jusqu'à présent est la description par les documentalistes des ouvrages de type monographies ou études économiques qui entrent dans le fonds documentaire. Les articles de fonds les plus pertinents pour l'activité, issus d'une liste de périodiques spécifiquement identifiés dans cette optique, ont également fait leur entrée dans la base de données depuis août 2013, pour y être également capitalisés.

Les pratiques d'indexation décrites par l'équipe de documentalistes sont classiques :

- Examen des mots du titre et du sous-titre, puis du résumé de quatrième de couverture
- Recueil des notions abordées dans le sommaire et éventuellement grâce à l'index
- Recherche de ces notions au sein du thesaurus dans JLB (champ « Mots clés ») ou identification de synonymes au besoin

Jusqu'à présent il n'y a pas de préconisation officielle encadrant des pratiques communes pour l'indexation. La profondeur d'indexation, notamment, est laissée à l'appréciation de chacune, en fonction du type d'ouvrage analysé et du nombre de notions importantes qui y sont traitées. Ainsi, l'une des documentalistes a évoqué les concepts nouveaux, qui peuvent nécessiter d'être mentionnées même s'ils ne sont pas centraux dans l'ouvrage en question, car permettant de répondre à des demandes utilisateurs émergentes. Néanmoins, cette démarche était rendue difficile par la non-utilisation des candidats-descripteurs. En effet, cette fonctionnalité étant absente du logiciel Techlib, cette pratique ne correspondait pas aux usages du service et n'avait pas encore été proposée.

3.3.2 Recherche

Depuis la fusion et dans l'attente de la mise en place du portail sous le nouveau progiciel JLB-Net Portail 2.0, les documentalistes étaient les seules à utiliser le thesaurus, tel qu'il apparaissait désormais dans JLB-Net Bibliothèques (Cf. 0) La mise en production du portail, entamée durant l'été 2013, met désormais la recherche dans la base de données en accès direct pour les utilisateurs du groupe. Cette évolution est encore trop récente pour disposer de quelques statistiques d'utilisation, mais il y a fort à parier que cette nouvelle ouverture, assortie à une

campagne de communication adaptée, devrait engendrer une augmentation importante du nombre de requêtes. Le portail permettant à l'utilisateur, dans le formulaire de recherche avancé, de sélectionner lui-même ses mots clés dans l'index, la consultation du thesaurus devrait également se généraliser.

3.3.3 Attribution des cotes d'ouvrages

L'une des premières fonctions attribuées au thesaurus lors de sa conception, et qui explique en grande partie son insertion sous forme de lexique au sein du plan de classement, était celle de guide pour le référencement des ouvrages au sein du fonds documentaire. En effet, certaines documentalistes n'étant pas spécialistes des domaines de la banque et de la finance, une mauvaise maîtrise des concepts traités dans les ouvrages pouvait conduire à des erreurs lors de l'identification de son domaine d'appartenance. Les descripteurs du thesaurus leur permettaient donc, en étant insérés dans les parties correspondantes du plan de classement, de les guider pour attribuer l'indice ou cote adéquate à la ressource qu'elles devaient enregistrer.

Cette utilisation du langage documentaire s'est poursuivie au sein de cette équipe, et a donc été prise en compte au cours des travaux de refonte du thesaurus, comme nous le verrons dans la seconde partie. Il est utile de préciser ici que les usagers, c'est-à-dire les collaborateurs de BPCE SA et plus largement du groupe, n'ont eu jusqu'ici qu'un usage très limité du thesaurus, le plus souvent accompagné par une documentaliste. C'est pourquoi leurs pratiques n'ont pu être abordées dans ce mémoire. L'ouverture de l'OPAC, on l'a évoqué, leur en donnera bientôt un accès direct et privilégié. Il sera bien sûr intéressant de faire une analyse de ces nouveaux usages, dès que les premières statistiques seront disponibles.

Deuxième partie

Mener un projet de réingénierie de thesaurus Éléments de méthode

4 Repenser un thesaurus : une étape incontournable dans la vie d'un service de documentation

Le langage documentaire employé par un service de documentation est donc à la base de l'ensemble de ses pratiques, dont l'objectif final est toujours le repérage de l'information (*information retrieval*). Rien d'étonnant dès lors à ce que le maintien et l'évolution de son thesaurus soit au centre de ses préoccupations.

4.1 Une mise à jour inhérente au cycle de vie d'un thesaurus

4.1.1 Constat d'obsolescence

Le thesaurus n'ayant plus été entretenu depuis 2008, et en l'absence de la fonction candidat-descripteur jusqu'à l'adoption de JLB (mais dont l'usage n'a pas encore été mis en place), presque aucun nouveau descripteur n'a été intégré en cinq ans. Or, le thesaurus est un outil qui doit, pour être utile et pertinent, rester en constante évolution. Afin de permettre une représentation correcte d'un domaine de connaissance, il doit en suivre les mutations. Cela signifie notamment intégrer les concepts émergents qui s'y imposent régulièrement, mais également suivre les évolutions linguistiques et sémantiques qui le touchent, et peuvent amener à privilégier différents descripteurs selon les périodes.

4.1.1.1 Des domaines en constante évolution

Ce constat est d'autant plus vrai pour des domaines d'activités aussi mouvants que la banque et la finance, touchés qui plus est durant les vingt années de vie de ce thesaurus par de profonds bouleversements réglementaires, institutionnels, structurels et même technologiques (moyens de paiement, etc.). L'introduction de nouveaux concepts ne pouvait suffire à sa mise à jour, et une restructuration en profondeur devenait inévitable. Par ailleurs, les domaines annexes touchant aux sciences sociales, au traitement de l'information, à la vie des entreprises ou encore à la politique économique avaient également connu des évolutions fondamentales, qu'il devenait indispensable d'intégrer. Même le domaine regroupant les secteurs d'activité nécessitait un rafraîchissement, tandis que tout restait à faire pour le domaine, désormais incontournable, du développement durable.

4.1.1.2 Une entreprise aux activités remodelées

D'autre part, suite aux restructurations provoquées par la fusion, l'activité de l'entreprise a, elle aussi, beaucoup évolué ces dernières années. Cette refonte était donc également l'occasion de purger les termes liés à l'activité de la CNCE : les plus de 250 descripteurs dédiés aux produits

de la Caisse d'épargne, notamment. En rafraichissant chacun des micro-thesauri, les documentalistes les ont naturellement fait évoluer en fonction, non seulement, de l'actualité du domaine, mais aussi, et surtout parfois, en s'appuyant sur les activités actuelles de leurs utilisateurs au sein du groupe.

4.1.2 Obstacles rencontrés pour l'indexation

Le progiciel de SIGB JLB-Net ne permettant pas la recherche en texte intégral, mais seulement sur les champs de la notice, la pertinence de l'analyse sémantique réalisée par les documentalistes pour constituer l'indexat d'un document revêt un caractère crucial pour la recherche d'information.

4.1.2.1 Difficultés à valoriser un thesaurus incomplet

A l'heure actuelle, les nombreux concepts absents du thesaurus sont autant d'obstacles à l'indexation exhaustive des documents analysés, et par conséquent dans l'identification de l'information pertinente pour répondre aux requêtes des utilisateurs. Ces manques, rencontrés quotidiennement par les documentalistes dans leurs activités, les obligent à des solutions de contournement, à des choix approximatifs, qui peuvent générer un risque de contresens et limiter également le repérage de l'information. On pourrait parler dans ce cas, plutôt que d'une expansion sémantique³⁰, d'une « dispersion sémantique ». En outre, la frustration liée à l'impossibilité de réaliser une description pertinente de certains ouvrages peut provoquer en réaction une tendance à la surindexation lorsque les descripteurs sont présents.

4.1.2.2 Une surreprésentation de mots-clés inutilisés

Lors du développement du thesaurus CNCE, détaillé plus haut (cf. 3.1.1.2), la méthode a consisté essentiellement à sélectionner des parties de thesauri existants, considérées comme utiles *a priori*, en les enrichissant éventuellement. Or, ce travail a été effectué conjointement avec la constitution du fonds documentaire auquel ce langage d'indexation devait ensuite être appliqué. Aucune évaluation n'a été menée par la suite pour évaluer l'adéquation effective du vocabulaire ainsi constitué. Des concepts identifiés comme manquants ont été ajoutés au gré des besoins (cf. 3.1.2.1), mais aucun nettoyage systématique des descripteurs superflus n'a été réalisé.

Ce constat entraîne un double inconvénient. A l'indexation, lors de la recherche des descripteurs adéquats par les documentalistes, ces termes peuvent ralentir le travail de façon inutile. Pire, avec l'ouverture de la recherche aux utilisateurs grâce au portail, la recherche pourra se faire

³⁰ « Opération consistant à étendre les termes de la question de l'utilisateur avec des termes proches en parcourant un dictionnaire ou un réseau sémantique. » (ADBS, « Vocabulaire de la doc », http://www.adbs.fr/expansion-semantique-17095.htm?RH=OUTILS_VOC)

directement par sélection dans l'index des mots clés de JLB, soit dans le thesaurus. Or, quoi de plus décourageant pour l'utilisateur que de sélectionner des termes qui ne renvoient à aucune notice, et ce de façon répétée ? Il devenait donc important, également pour le succès de cette nouvelle offre de service, de rafraîchir rapidement le thesaurus.

4.1.3 La longue gestation du projet

4.1.3.1 Une première tentative avortée

Devant toutes ces problématiques, une première mission de refonte du thesaurus avait été lancée en interne en mai 2011, soit avant même le lancement du projet concernant le nouveau logiciel documentaire. Mais l'absence de ressources suffisantes affectées exclusivement à la conduite de ce projet n'a pas permis alors de le mener à bien. Néanmoins, les éléments rassemblés à cette occasion, ainsi que le travail préliminaire effectué par l'une des documentalistes du service (et membre de l'équipe pour le présent projet) ont constitué une base de travail intéressante. En conséquence, les difficultés de gestion liées au non maintien d'un thesaurus à jour s'aggravant, la responsable du service a finalement choisi de faire appel à une aide extérieure pour la conduite du projet.

4.1.3.2 Principaux objectifs de la mission présentée

Le principal livrable attendu à l'issue du projet était un langage documentaire mis à jour, stabilisé et normalisé. Le corpus de base pour la refonte fut limité au thesaurus Techlib (soit le lexique hiérarchisé inséré dans le plan de classement). Le délai imparti, très court, ne permettait pas de fixer au préalable des objectifs précis en termes d'intégration des modifications au logiciel de gestion des ressources, JLB-Net. Le lexique des mots clés issus de DIP serait donc aligné ultérieurement au nouveau thesaurus, au fur et à mesure des traitements informatiques.

4.2 Une évolution nécessaire suite à l'adoption d'un nouveau logiciel

Après la fusion, un évènement majeur a impacté l'utilisation et l'évolution du thesaurus initialement conçu au sein de la CNCE : l'adoption d'un nouveau logiciel, grâce auquel ont été rassemblées les bases de données de la CNCE et de la BFBP, à l'image des deux équipes de documentalistes qui se trouvaient ainsi réunies.

4.2.1 Mise en place du progiciel JLB-Net

A l'issue de la fusion, le nouveau service Documentation générale de BPCE SA disposait de deux bases de données, sous deux logiciels distincts. Deux fonds aux caractéristiques spécifiques, indexés à l'aide de langages documentaires bien différents...

4.2.1.1 Un choix imposé par la DSI

Lorsque BPCE fut créée en juillet 2009, l'immeuble déjà occupé par la CNCE, au 50 rue Pierre Mendès-France (Paris 13), fut choisi pour accueillir le siège de l'entité BPCE SA, organe central du nouveau groupe. Le service documentation de la CNCE y conserva donc ses locaux, devint en janvier 2010 la « Documentation générale » de BPCE SA et accueillit les ressources du service documentation de la BFBP. La responsable du service à la CNCE, depuis sa création en 1990, a également conservé cette responsabilité pour BPCE. La réflexion menée à l'époque a conduit à privilégier la pérennité des pratiques et des outils de cette structure.

Une étude des besoins poussée menée par une consultante du cabinet Serda en 2010 avait doté le service Documentation générale d'un cahier des charges fonctionnel très précis, pour aider à la sélection du nouveau logiciel³¹. Celui-ci devait, outre les fonctions de gestion des ressources classiques, intégrer des fonctionnalités permettant d'optimiser la production des veilles, en proposant la création de profils et une gestion fine des périodicités d'envois, très variables. Le projet prévoyait en outre la création d'un portail OPAC accessible depuis l'intranet à l'ensemble des collaborateurs du groupe. La direction des Systèmes informatiques, en charge de la gestion des projets, a identifié un logiciel déjà présent dans l'organisme pouvant répondre en partie aux spécifications requises, JLB-Net Bibliothèques. Celui-ci se trouvait déjà en usage à la Documentation juridique, qui avait été séparée de la Documentation CNCE en 2005. Ce progiciel de SIGB classique, développé et distribué par la société JLB Informatique depuis 1981, ne couvrait pourtant pas l'ensemble des besoins exprimés.

4.2.1.2 Une gestion de projet complexe

Malgré son inadéquation, ce choix fut imposé et le projet mené avec l'aide d'une responsable projet interne à la DG assurant la maîtrise d'ouvrage, transfuge de la Documentation juridique et donc familiarisée avec l'outil. Le budget estimé n'ayant pas été obtenu pour 2011, le projet fut retardé et une partie seulement fut obtenue pour 2012. Mais les difficultés de dialogue avec les chefs de projet informatique, et surtout les nombreuses améliorations demandées au prestataire pour répondre au cahier des charges ralentirent considérablement l'installation de l'outil. L'application de base, JLB-Net Bibliothèques, pour la gestion des ressources, fut enfin installée en janvier 2013. Le logiciel de veille, malgré une refonte totale pour répondre à la demande, ne fut pas validé lors de sa phase de test au printemps 2013. Enfin, l'application JLB-Net Portail 2.0 (OPAC), dont la mise en production a débuté à l'été 2013, ne donnait pas encore pleine satisfaction et était encore en cours de développement en novembre.

³¹ Pour en savoir plus à ce sujet, consulter l'article de Christophe Duteil, « BPCE prépare la fusion de ses logiciels documentaires », paru dans Archimag – Guide pratique : Choisir ses logiciels, n°43, août 2011, p.28

4.2.2 Du « thesaurus Techlib » au lexique Mots clés de JLB

Une partie du projet d'installation du nouveau logiciel consistait bien évidemment à récupérer les bases de données existantes pour les fusionner dans la nouvelle base dans JLB-Net. Cela signifiait, par la même occasion, récupérer les langages documentaires utilisés pour indexer chacun des fonds documentaires.

4.2.2.1 Deux techniques d'indexation différentes

Le langage documentaire conçu au sein de la CNCE de la façon exposée plus haut (cf. 3.1.1) avait ensuite été intégré au module de gestion de thesaurus proposé par le progiciel BASIS Techlib (Open Text). Au service documentation de la BFBP, créé dans les années 1980, le progiciel de gestion documentaire utilisé était DIP Maker v.4, commercialisé par DIP Systèmes (devenu e-Dip en 2006). Le langage d'indexation choisi à l'époque était une simple liste de mots clés, sans aucune relation entre eux. Les deux fonds documentaires avaient donc été indexés avec des langages tout à fait différents, même si, portant sur les mêmes domaines d'activité, l'environnement sémantique était similaire. Mais l'absence de relations pour l'un signifiait nécessairement des pratiques d'indexation en décalage, puisqu'on ne pouvait dans ce cas compter sur les techniques d'autopostage, c'est-à-dire le :

*« Procédé permettant d'effectuer automatiquement une indexation complémentaire d'un document ou d'une question par tous les descripteurs appartenant à la même branche de l'arborescence du thesaurus que le descripteur le plus spécifique utilisé lors de l'indexation. L'autopostage **générique** (vers un niveau supérieur) peut être effectué lors de l'indexation et lors de la recherche. L'autopostage **spécifique** (vers un niveau inférieur) s'effectue lors de la recherche. »³²*

Ce double mode d'indexation devait évidemment engendrer des complications non négligeables une fois ces langages réunis au sein d'une même base de données.

4.2.2.2 Problématique de ce lexique hybride

Le progiciel Techlib possède un module intégré de gestion de thesaurus, mais l'ensemble des termes, descripteurs comme non-descripteurs, apparaît dans les applications d'indexation et de recherche dans une même fenêtre, à partir du champ intitulé « Mots clés ». Deux modes d'affichage sont proposés, par ordre alphabétique dans les deux cas : lexique ou thesaurus (faisant alors apparaître les relations). Les termes définis comme non préférentiels sont automatiquement remplacés par les descripteurs équivalents lorsqu'ils sont sélectionnés.

³² Définition de l'ADBS, « Vocabulaire de la doc » : <http://www.adbs.fr/autopostage-16226.htm>

Le premier problème engendré par la fusion des bases de données fut bien sûr l'apparition de nombreux doublons dans la liste des mots clés. La politique adoptée était de privilégier les termes issus du thesaurus Techlib dans le choix des descripteurs. Mais les deux langages apparaissant simultanément dans la liste des mots clés, on imagine le risque de ne pas utiliser chaque fois la même entrée, et le silence ainsi potentiellement généré. Heureusement, les mots clés issus de Dip étaient orthographiés avec un tiret entre les mots, limitant les risques lorsqu'il s'agissait de termes composés. Par exemple, le terme *Agence bancaire* issu du thesaurus (26 notices) avait pour doublons le terme *Agence banque* (47 notices) et le mot clé issu de Dip *agence-banque* (287 notices), soit une dispersion considérable. La campagne de suppression des doublons stricts menée au cours de la mission a concerné environ 240 mots clés, portant sur plus de 11 300 notices, grâce à la fonction de traitement par lot (sachant qu'une partie avait déjà été traitée lors des phases de test). Mais il subsiste de nombreux doublons dans les concepts pour lesquels des descripteurs différents avaient été choisis, et qu'il reste à identifier.

Pour donner une meilleure idée de la composition de ce lexique, voici les statistiques relevées en début de mission :

- Mots clés : 5038
- Génériques : 359
- Spécifiques : 1128
- Employés : 811
- Employés pour : 1129
- Termes associés : 844

Ces chiffres permettent de comprendre la proportion de termes du lexique n'appartenant pas à un langage de type thesaurus, ou du moins sans relation (2 422, soit près de la moitié³³). Le lexique Dip comportait environ 1 800 mots clés lors de son intégration à JLB-Net.

4.3 Considérations préliminaires sur les implications du projet

Si l'on considère la portée organisationnelle d'un thesaurus (Cf. 2.3.2) – potentielle ou réelle selon la place occupée stratégiquement par le service en charge du thesaurus dans l'entreprise – on comprend que tout projet portant sur la conception ou la refonte d'un thesaurus ne saurait être anodin. Plusieurs points nous semblent, à ce stade, devoir être pris en compte pour ce type de projets :

³³ Si l'on exclut le chiffre portant sur les termes employés, qui n'est pas significatif en soi, ainsi que les termes associés

- La constitution de l'équipe chargée de l'architecture du thesaurus et du choix des concepts, puis des descripteurs : leur niveau de connaissance de l'entreprise et de ses différents métiers doit être optimal, pour garantir la qualité et l'efficacité du référentiel ;
- La communication autour du projet auprès des différentes pôles de l'entreprise qui peuvent être concernés par sa mise en œuvre : direction dont dépend le service chargé du projet (*a minima*), autres services liés à la recherche, la gestion ou la diffusion de l'information, à tous les niveaux, et dans tous les organes du groupe éventuellement.

Mais pour l'heure, examinons la façon dont a été conduite cette mission concrètement, en essayant d'en dégager des éléments de méthode réutilisables.

5 Choix méthodologiques pour la gestion du projet

La norme ISO 25964 consacre sa Clause 13, *Managing thesaurus construction and maintenance* [1, ISO, p. 88-98], aux différents aspects de la gestion d'un projet de conception de thesaurus, de sa planification à la publication du produit final et sa mise à jour. En compléments de ces recommandations très générales, nous présentons ici les pistes qui ont été choisies pour un projet de refonte de thesaurus, dans le contexte spécifique de ce service de veille et documentation.

5.1 Analyse des risques et contraintes liés à la mission

De nombreux aspects doivent être analysés et pris en compte, afin de déterminer comment mener au mieux un projet aussi exigeant que la refonte d'un langage documentaire, à commencer par le contexte et les ressources disponibles, et d'identifier les contraintes organisationnelles et les risques qui en découlent. Au cours de cette mission, une première phase de cadrage a permis de minutieusement monter un plan d'action, conçu pour être le plus efficace possible en fonction de ces paramètres.

5.1.1 Présentation du contexte de réalisation

5.1.1.1 Un positionnement à l'image de la complexité organisationnelle du groupe

L'organe central du Groupe BPCE, BPCE SA, se caractérise par un positionnement profondément complexe. En effet, il se trouve à la fois au service de ses actionnaires principaux, qui se trouvent être deux réseaux bancaires concurrents, et de ses nombreuses filiales, tout en se voyant déléguée la responsabilité de définir le positionnement stratégique et l'orientation de l'ensemble de cette construction.

5.1.1.1.1 *La Documentation générale BPCE SA au cœur du réseau documentaire du groupe*

L'organisation des différentes structures fournissant de l'information dans l'ensemble du groupe BPCE constitue un système tout aussi complexe, dont l'articulation et les interactions ne sont pas aisées à appréhender. Différents organes du groupe disposent de leur propre service de documentation, composé parfois d'un seul documentaliste, tandis que certaines personnes peuvent assurer ce type de fonctions au sein d'autres services disséminés. La Documentation Générale de l'organe central de BPCE se trouve au centre de ce système. Il s'agit de la continuité du service fondé en 1990 au sein de la CNCE, évoqué précédemment. Des collaboratrices appartenant à la documentation de la BFBP, mais aussi des salariés des deux organes non issus de la documentation – mais qui devaient être réaffectés suite à la fusion – s'y sont peu à peu agrégées. Suite à plusieurs départs non remplacés, l'équipe se limite aujourd'hui à huit personnes, dont cinq veilleuses-documentalistes (y compris la responsable), correspondant à seulement 3,2

ETP, une responsable projets et deux assistantes. La DG centralise l'ensemble des ressources et les met à disposition de toute l'entreprise. L'autre principal service de documentation exerce au sein d'une filiale et se compose d'une douzaine de personnes. Produisant essentiellement des analyses sectorielles, elle se considère d'abord au service de ses propres utilisateurs, ce qui a pu rendre la négociation des ressources communes délicate.

5.1.1.1.2 Positionnement dans l'organigramme

La DG fait partie de la direction Immobilier et Services généraux (DISG), elle-même intégrée au pôle Opérations, au même titre que les directions des Systèmes informatiques et des Programmes. Ce positionnement dans l'organigramme pose de nombreux problèmes, à commencer par son éloignement des instances de direction, pénalisante lorsqu'il s'agit de négocier le budget ou de réaliser un projet. Les problématiques documentaires, et même les caractéristiques du métier de documentaliste, se trouvent par ailleurs bien éloignées de celles des autres services dépendants de la DISG. La responsable de la DG, qui a pu comparer par le passé les mérites de différents positionnements pour son service, y voit néanmoins certains avantages. Les Services généraux ont en effet en commun une vocation de fournisseur de service à destination de l'ensemble du groupe, ce qui dispense notamment de toute considération de confiscation des ressources au profit d'une seule direction.

5.1.1.2 Composition de l'équipe projet

Le commanditaire du livrable, qui a également assuré le pilotage du projet, était la responsable du service. L'ensemble des veilleuses-documentalistes du service se sont vues impliquées comme membres de l'équipe projet, en tant que spécialistes sur leurs domaines de veilles thématiques respectives, y compris la responsable, qui assumait ainsi un double rôle dans le projet. Une intervention ponctuelle a été également confiée à la responsable de projets interne, en tant qu'experte juridique, pour rationaliser la partie correspondante du thesaurus en fonction des besoins préalablement évalués.

5.1.1.2.1 Répartition initiale des tâches

La charge de travail attribuée à chacune était très variable, dépendant essentiellement des domaines de veille dont elle a la responsabilité. Ainsi l'une des veilleuse-documentaliste n'avait qu'un sous-domaine à traiter, l'immobilier, tandis que la responsable projets, qui n'était pas incluse initialement dans l'équipe projet, est intervenue pour la refonte du domaine juridique, qui ne concerne que partiellement le fonds documentaire de la DG. Néanmoins, les quatre autres membres se retrouvaient toutes face à une tâche conséquente, puisqu'elles devaient traiter trois à cinq thématiques. La responsable du service devait en outre traiter les domaines appartenant au cœur de ce thesaurus spécialisé, banque, finance et assurance, qui demandaient un travail de recherche très important. Enfin, la documentaliste la plus expérimentée dans la conception de

thesaurus (à qui avait été confiée la première mission, cf. 4.1.3) s'est vue confiés les domaines plus généralistes, ne faisant pas l'objet d'une veille spécifique, en plus de son propre domaine d'expertise, le développement durable, qu'il restait à créer intégralement. Son aide a également été précieuse sur les aspects théoriques liés à la conception de thesaurus et pour l'articulation des différents micro-thesauri.

5.1.1.2.2 Un niveau de compétences documentaires hétérogène

Sur les six documentalistes ayant participé au projet, trois avaient une formation initiale dans ce domaine. La responsable du service est titulaire du diplôme supérieur de l'INTD, après une maîtrise d'Economie. Elle a fait toute sa carrière dans la banque et a créé deux services de documentation, à la Société Générale puis à la Caisse d'Épargne. La seconde, après une maîtrise d'Histoire, a passé le Certificat d'aptitude aux fonctions de bibliothécaire, avant de suivre un DESS en Information et Documentation à l'IEP de Paris³⁴. La troisième a suivi un cursus de Techniques de l'information, option Documentation, dans un IUT et a fait toute sa carrière dans des centres de documentation (enseignement supérieur, ingénierie, puis 13 ans à la BFBP puis BPCE).

Parmi les deux autres membres de l'équipe, l'une est titulaire d'une licence AES et documentaliste depuis trois ans, après avoir exercé diverses fonctions à la Caisse d'Épargne (formatrice et chargée d'études entre autres). La seconde a suivi une formation initiale en secrétariat. Elles ont suivi plusieurs formations à l'ADBS pour découvrir les divers aspects du métier. La responsable projet, enfin, après une formation juridique (droit des affaires) a exercé toute sa carrière dans la documentation (cabinets d'avocats, responsable de la documentation juridique à la CNCE puis BPCE).

5.1.2 Durée et coût de l'intervention

5.1.2.1 Un délai de réalisation court et des contraintes fortes

Le projet s'est déroulé du 20 juin au 6 septembre 2013. Ce délai très court, d'environ 11 semaines, a imposé une organisation au plus juste, ainsi qu'une implication forte de l'ensemble des membres de l'équipe projet. De surcroît, la mission s'est déroulée durant la période des congés d'été, rendant plus prégnantes encore les contraintes organisationnelles. Pour les acteurs impliqués, les tâches affectées dans le cadre de ce projet se sont ajoutées à celles inhérentes à leurs activités quotidiennes, ajoutant une contrainte supplémentaire quant à leur disponibilité. De surcroît, parmi les cinq principales participantes du projet, seules deux travaillent à temps plein, deux aux trois et quatre cinquièmes, et la dernière à mi-temps.

³⁴ Son mémoire a d'ailleurs porté sur ce sujet : *ALEXIS au pays des thesaurus*, 1987 (notice disponible à l'adresse <http://www.sudoc.fr/096728418>)

Les principales phases du projet peuvent être distribuées selon la « méthode PDCA » :

- PLAN Phase de cadrage/planification : du 20 au 28 juin
- DO Phase de réalisation : du 1er juillet au 23 août + bilan à mi projet
- CHECK Phase d'actions correctives et évaluation du reste à faire : du 26 au 30 août
- ACT Phase de capitalisation et bilan du projet : du 02 au 06 septembre

5.1.2.2 Un budget restreint

Ce paramètre était une condition importante de la faisabilité du projet, le service fonctionnant actuellement avec des ressources très restreintes, et en sous-effectif depuis des récents départs non remplacés. Le principal coût additionnel aux frais de fonctionnement habituels du service initialement prévu était la rémunération de la stagiaire/chef de projet durant les deux mois et demi de sa mission. L'autre coût notable est l'investissement dans la norme ISO 25964, volumes 1 et 2, commandée auprès de l'AFNOR en version papier, pour un montant de 366,95 €. Michèle Hudon regrette d'ailleurs ce prix, qu'elle a qualifié d'exorbitant, d'autant que le texte n'est disponible qu'en anglais, divers organismes nationaux ayant choisi de ne pas le traduire.

5.2 S'inspirer de SCRUM pour optimiser la gestion de projet

Devant de telles contraintes pour la gestion de projet, en particulier concernant la faible disponibilité des membres de l'équipe du fait des congés et des rythmes de travail disparates, il semblait nécessaire de mettre au point une méthode quelque peu originale.

5.2.1 Présentation de Scrum

5.2.1.1 Les méthodes Agiles

Les méthodes Agiles sont nées dans les années 1990 pour optimiser la gestion de projets, notamment dans l'ingénierie logicielle. Un manifeste, publié en 2001, résume ainsi les principes fondamentaux qui guident cette démarche :

Manifeste pour le développement Agile de logiciels³⁵

Nous découvrons comment mieux développer des logiciels par la pratique et en aidant les autres à le faire. Ces expériences nous ont amenés à valoriser :

- ***Les individus et leurs interactions*** plus que les processus et les outils
- ***Des logiciels opérationnels*** plus qu'une documentation exhaustive
- ***La collaboration avec les clients*** plus que la négociation contractuelle
- ***L'adaptation au changement*** plus que le suivi d'un plan

³⁵ Manifeste et ses « 12 principes sous-jacents » disponibles à l'adresse : <http://agilemanifesto.org/iso/fr/>

Nous reconnaissons la valeur des seconds éléments, mais privilégions les premiers.

Quelles que soient les différentes méthodes issues de cette communauté de pratiques, la structure centrale est un processus qui privilégie un cycle de développement itératif, incrémental et adaptatif. Les vertus de ces principes et l'efficacité observée de ces méthodes se sont révélés tels, qu'elles se diffusent aujourd'hui en s'adaptant à tous les niveaux de l'entreprise, s'imposant peu à peu comme des techniques transversales de gestion de projet.

5.2.1.2 Les grands principes de SCRUM

Parmi les méthodes Agiles, Scrum, élaborée depuis 1996, est l'une des plus répandues actuellement. Il s'agit d'une méthode de planification incrémentale, permettant de maîtriser l'atteinte d'objectifs à court terme, grâce à la livraison du produit par tranches successives. En programmation, ces phases de test récurrentes, assorties d'aller-retour fréquents entre l'équipe de réalisation et le commanditaire, permettent une grande adaptabilité et un niveau de qualité final très élevé, favorisant par ailleurs l'innovation. Inspirée du rugby, la philosophie de Scrum repose sur l'implication et l'effort collectif des membres de l'équipe, pour garantir la réalisation d'un objectif commun et tangible. L'autre aspect indissociable est celui de la gestion de la qualité, puisque chaque étape du cycle de développement (ou *sprint*) constitue un affinement itératif des résultats de la précédente.

5.2.2 Poser les bases de la réussite du projet

Si cette méthode ne pouvait s'appliquer de façon littérale à ce projet, elle n'en constituait pas moins une approche pertinente, proposant des outils susceptibles d'apporter un avantage décisif dans l'atteinte des objectifs fixés, compte-tenu du délai très court.

5.2.2.1 Eléments de SCRUM intégrés au projet

Tout en conservant le rôle de chef de projet, qui nous semblait indispensable dans un contexte où les membres de l'équipe projet n'étaient pas déchargés de leurs tâches quotidiennes, un mode d'organisation inspiré de Scrum a été imaginé.

5.2.2.1.1 Les cycles de travail ou « runs »

En fonction des disponibilités de chaque membre (ou « experte », terme que nous avons choisi d'appliquer dans ce contexte), selon ses congés mais aussi son rythme de travail, des cycles de travail ont été planifiés. Ces cycles, appelés « *runs* » pour conserver la métaphore sportive, étaient dédiés chacun à une thématique précise, assignée à une experte. Au besoin, pour certains sujets transverses, le chef de projet pouvait solliciter le support d'une ou deux contributrice(s) à une experte, qui jouait alors le rôle de leader et toujours décisionnaire sur ses domaines.

Si dans la méthode Scrum originale les *sprints* doivent toujours avoir une durée fixe (2 à 4 semaines en général), le contexte de ce projet nous a conduites à une adaptation en fonction des contraintes des différentes expertes. La durée pour chaque *run* fut fixée à :

- 1 semaine (5 jours) pour les collaboratrices à temps pleins
- 1 semaine et demie (6 jours) pour celle aux 4/5^{ème}
- 3 semaines (7 à 9 jours) pour celles aux 3/5^{ème} ou à mi-temps

Ces durées ont été conçues pour s'adapter aux tâches quotidiennes rapportées au temps de travail, en prenant en compte la charge supplémentaire prévue pour chacune dans le cadre du projet, selon ses domaines d'intervention.

5.2.2.1.2 *Les scrums ou « mêlées quotidiennes »*

Le second aspect inspiré de la méthode Scrum, qu'il nous a paru profitable d'adopter pour ce projet, est précisément celui qui lui donne son nom, en faisant référence aux mêlées de rugby. Il s'agit d'une réunion quotidienne de l'équipe projet qui ouvre chaque journée de travail. Chacun y exprime son « reste à faire » (RAF) et les obstacles ou freins qu'il a pu rencontrer. Le chef de projet (ou *scrum master*) peut alors réévaluer les délais pour chaque tâche et affecter des ressources supplémentaires sur celles qui le nécessitent.

Ce mode de fonctionnement apporte plusieurs avantages majeurs pour la gestion du projet :

- une plus grande adaptabilité du planning, garantissant la maîtrise des délais ;
- un repérage en temps réel des points de blocage, mais aussi des points de convergence entre deux domaines d'expertise, nécessitant une réflexion conjointe ;
- un point de focus quotidien sur le projet, permettant d'entretenir la motivation des membres de l'équipe et leur concentration sur les tâches liées au projet.

5.2.2.2 Adaptations nécessaires au cours du projet

Les éléments empruntés devaient permettre :

- une organisation facilitée, en parallèle de tâches quotidiennes déjà lourdes ;
- d'encourager un niveau élevé de participation et d'implication dans le projet, à la hauteur des ambitions en termes de qualité fixés initialement ;
- d'introduire une dimension ludique pour le lancement du projet.

5.2.2.2.1 *Des runs à géométrie variable*

L'organisation générale sous forme de *runs* dédiés chacun à une partie du thesaurus a été appliquée tout au long du projet. Cependant, la longueur de ces *runs* n'a cessé d'être adaptée, en fonction de la rapidité d'avancement de chaque experte, mais aussi de la complexité de chaque domaine, impossible à anticiper lors de la phase de cadrage. Il a également été possible de diviser un *run*, initialement prévu pour traiter un domaine beaucoup trop large, en deux afin de conserver la motivation de l'experte, tout en évaluant plus finement l'avancement réel. Un tableau d'avancement (tenu par le chef de projet) permet de retracer a posteriori le planning de réalisation effectif. (Cf. Annexe 1)

5.2.2.2.2 *Des « mêlées » tributaires du calendrier des présences*

Les mêlées quotidiennes, prévues au départ dans le but d'encourager et de maintenir la motivation de l'équipe, se sont espacées au fil des départs en vacances, et en fonction des effectifs présents, pour finalement disparaître. Le but avait été amplement atteint, puisque l'implication de l'ensemble de l'équipe s'est très vite révélée, comme nous le verrons, excellente – au-delà même du niveau espéré. Il devenait inutile de maintenir artificiellement un moment qui, par ailleurs, pouvait se révéler inhibant en présence de la responsable du service. Ces échanges se sont donc poursuivis de façon moins formelle, en tête-à-tête ou petits groupes d'expertes dont la réflexion sur leurs domaines se trouvaient en convergence à ce moment.

5.2.2.3 Bilan de la méthodologie projet

Cette méthode de gestion de projet a rempli ses principaux objectifs, à savoir :

- La réalisation incrémentale du livrable final a permis d'évaluer constamment le niveau de qualité global et d'assurer une homogénéité suffisante ;
- Les mises en commun régulières des produits de *run* ont favorisé leur affinement itératif (bien que l'équipe n'ait jamais été au complet durant la mission) ;
- Le délai a pu être globalement respecté, tout en atteignant les exigences du commanditaire quant au produit final.

5.3 S'assurer l'implication des membres de l'équipe

Durant la phase de cadrage, il est vite apparu que le levier essentiel pour espérer la réussite du projet, malgré les délais et nombreuses contraintes, consistait à obtenir l'adhésion sans faille de l'ensemble des membres de l'équipe. Seul ce paramètre pouvait éviter d'essuyer un échec similaire à celui rencontré lors de la première tentative de refonte du thesaurus, deux ans plus tôt.

5.3.1 Inscrire le projet dans les priorités du service

5.3.1.1 Implication du commanditaire

A l'image de Scrum, où le commanditaire (*product owner*) est présent en permanence au plus près de l'équipe projet pour dicter ses priorités, cette mission a bénéficié d'une implication totale de la responsable de la DG. Ce rôle de supervision n'était pas simple à exercer en parallèle des nombreuses activités liées à sa fonction, et surtout en tant que participante active au projet, dont elle constituait même une ressource en sollicitation constante. Néanmoins, son intervention en tant que décisionnaire finale était nécessaire chaque fois qu'une question surgissait, qui ne pouvait être réglée par consensus. Il pouvait s'agir de la délimitation de deux domaines, de la définition du domaine d'appartenance d'un concept, ou encore de décisions d'ordre structurel, comme l'alignement avec le plan de classement.

Ainsi, il est très vite apparu que pour mener à bien ce type de projet dans un temps aussi limité, il est indispensable de désigner une personne pour jouer rôle de « gardienne de l'esprit du thesaurus ». Que cette personne soit, selon le contexte, la plus expérimentée sur le sujet ou bien la plus au fait des pratiques des utilisateurs de l'outil, sa tâche est d'insuffler une vision à l'équipe en charge du projet. Pour le chef de projet, c'est elle qui doit définir les objectifs à remplir pour le livrable et les critères de satisfaction attendus. Bien sûr, quand le projet s'étale sur une plus longue période, un mode de décision plus collectif est à privilégier, car il peut permettre d'étendre la réflexion et d'atteindre un niveau de qualité supérieur.

5.3.1.2 Positionnement et autonomie du chef de projet

L'implication et les attentes fortes de la responsable du service ont permis d'initier la mission dans les meilleures conditions en offrant au chef de projet un positionnement solide au sein du service. Son intervention était attendue et le soutien dont bénéficiait son action, l'autorité qui lui était conférée pour la mise en place et la gestion du projet ont été affirmées à plusieurs reprises. De fait toute latitude nous ont été laissées pour imaginer un mode de gestion de projet adéquat, puis pour l'organisation du déroulement des tâches avec chaque experte.

5.3.2 Inscrire le projet dans les tâches quotidiennes

Nous avons exposé plus haut le principe de répartition des responsabilités pour la refonte du thesaurus et la méthode mise en place pour la planification des tâches. Une organisation bien pensée était en effet incontournable dans le cadre d'une équipe projet déjà très occupée par la vie normale du service.

5.3.2.1 Des tâches quotidiennes nombreuses et variées

Nous avons présenté dans la première partie les principales activités menées au sein de la DG. Les veilles thématiques en constituent la tâche la plus lourde, mobilisant la responsable du service et les quatre autres veilleuses-documentalistes une bonne partie de leur temps. Ainsi, la matinée est généralement dédiée au dépouillement des périodiques. Chacune est responsable de l'analyse d'un ou plusieurs quotidiens, ainsi que d'un certain nombre de titres plus spécifiquement liés à ses domaines de veille. Elle identifie ce qui lui est utile directement et recense pour ses collègues les articles traitant de leurs domaines. La confection des produits de veille est l'autre aspect très gourmand en temps de cette activité. La mise en place du module de veille de JLB-Net, dont la première livraison n'a pas satisfait les besoins, représentera une aide précieuse sur ce point.

En parallèle de cette activité de fond, chacune est amenée bien sûr à répondre à des demandes utilisateurs. Certains se déplacent à la documentation, il faut alors les recevoir et les interroger pour cerner leur besoin, mais la plupart adressent désormais leurs demandes par email. Il y a ensuite bien sûr la gestion des prêts (trois semaines) et des retours. Une assistante est en charge des commandes d'ouvrages, qu'ils soient destinés à intégrer le fonds documentaire ou à l'usage d'un collaborateur qui en a fait la demande (prêt « longue durée »). Dans le premier cas, la documentaliste qui traite ce domaine procède à sa réception à son indexation dans JLB-Net Bibliothèque et lui attribue une cote avant de l'intégrer au fonds ou de prévenir l'utilisateur qui en avait fait la réservation.

L'adresse mail du service est enfin le réceptacle de tous types de questions touchant à la diffusion de l'information au sein de l'entreprise :

- accès aux études économiques en ligne via l'intranet ;
- base de données Norma regroupant les communications à caractères réglementaire et normatif sur l'intranet, mise en place en 2012 et dont la direction de la conformité assure la modération ;
- questions touchant aux espaces métiers de l'intranet (outil également géré par la communication interne), etc.

Les documentalistes répondent à ces questions, en réorientant au besoin l'utilisateur vers le bon interlocuteur. Enfin, une réunion de service à lieu en principe tous les jeudis après-midi, afin d'échanger sur l'activité de la semaine et exposer toutes les questions soulevées, proposer des idées...

5.3.2.2 ... qui sont autant d'occasions de réfléchir au thesaurus

Imposer le travail sur le thesaurus comme une priorité à inclure dans ces journées déjà bien chargées, en particulier pour celles présentes à mi-temps ou aux trois cinquièmes, représentait

donc une gageure. Il était essentiel de leur donner envie d'y participer activement et de se dégager du temps pour les phases de réflexion et de recherches indispensables. En effet, sur une durée aussi courte, chaque experte devait avancer seule sur sa thématique en cours, parallèlement à ses collègues. Obtenir un niveau d'implication élevé était le seul moyen de garantir une autonomisation des membres de l'équipe suffisante pour l'efficacité du projet. Un premier levier utilisé a été de jouer sur l'aspect divertissant que pouvait revêtir le travail, via la méthode de gestion de projet proposée (« *scrum like* »), mais aussi par la mise en valeur d'une activité ponctuelle sortant de l'ordinaire (intérêt du travail en mode projet).

La stratégie adoptée a été d'encourager dès le départ la réflexion sur le thesaurus dans tous les aspects du travail quotidien. Peu après le lancement du projet, une session de formation a permis de mettre à jour les connaissances théoriques sur le thesaurus. Cela a permis de jeter un regard nouveau sur l'outil utilisé jour après jour, à la lumière d'une nouvelle compréhension. Chacune a ainsi été amenée à s'interroger sur ses pratiques d'indexation. Enfin, au gré du *run* en cours, le travail sur la veille thématique correspondante devenait l'occasion de réfléchir sur les concepts nouveaux d'un domaine, sur les termes employés actuellement et leur adéquation avec ceux présent dans le thesaurus Techlib, etc. Bref, peu à peu, la réflexion sur le thesaurus a infusé l'ensemble des activités menées par l'équipe, les deux s'enrichissant mutuellement, les encourageant à creuser des connaissances trop souvent effleurées.

5.4 Un dialogue constant avec le management

5.4.1 Pilotage

Une réunion hebdomadaire s'est tenue entre le chef de projet et le pilote (excepté durant les vacances d'OL). Fixées au vendredi matin, l'objectif de ces réunions de pilotage était le suivi des activités menées chaque semaine, l'identification des obstacles rencontrés ou des risques potentiels repérés, enfin la décision des actions correctives ou préventives à mettre en place. Chacune de ces réunions a fait l'objet d'un compte rendu envoyé au pilote, afin de valider le contenu des points évoqués et d'acter les décisions prises, échéances fixées, etc. Ces échanges ont peu à peu pris la relève, bien que plus espacés, des *scrums* en permettant de contrôler ce qui fonctionnait et d'identifier rapidement les points de blocage pour y remédier. Ils étaient enfin l'occasion d'évoquer les éventuelles problématiques interpersonnelles qui ne manquent pas de surgir au cours d'un travail en groupe, et pour le chef de projet d'obtenir les clés de compréhension qui pouvaient lui échapper. Cela permettait ainsi à OL d'adapter son discours et son comportement pour renforcer au mieux l'autorité de son chef de projet, lorsque cela s'avérait nécessaire.

5.4.2 Reporting

Compte-tenu du court terme du projet, de son périmètre limité et de la fréquence des instances de pilotage, un outil de reporting allégé paraissait suffisant. Les mêlées quotidiennes devaient permettre en outre de décharger les membres de l'équipe de cette contrainte, puisque le feedback exprimé à ces occasions était ensuite formalisé directement par le chef de projet. De plus, la réunion de service était l'occasion de faire un point hebdomadaire en équipe (malheureusement jamais au complet). Y était présentés les derniers *runs* achevés, les travaux en cours et rappelés les objectifs à court terme. Ces moments étaient aussi l'occasion de maintenir la concentration sur le projet et une visibilité pour chacune des membres de l'état d'avancement. Par ailleurs, un espace dédié au projet a été créé sur le serveur commun du service, afin de permettre l'accès à toutes aux ressources proposées pour faciliter le projet. Un fichier intitulé « Outils projet » réunissait en particulier les outils de suivi alimentés par le chef de projet, soumis à la validation de l'équipe :

- Tableau global d'avancement, récapitulant la répartition des *runs* et mis à jour en temps réel (Cf. Annexe 1) ;
- Un tableau de reporting quotidien retraçant les activités menées par chacune, les observations ou problèmes rencontrés et les solutions apportées ;
- Les comptes rendus des réunions de service hebdomadaires ;
- Un suivi du dé-doublonnage du thesaurus JLB mené par le chef de projet.

6 Le chef de projet : un rôle de facilitateur

Le rôle d'un chef de projet, qu'il soit interne ou externe, sur une mission de refonte de thesaurus est d'abord celle d'un facilitateur. Il doit encourager la réflexion et garantir un niveau de qualité élevé pour le livrable final. Pour y parvenir, sa mission comporte trois aspects essentiels : offrir des moyens adaptés, encadrer un mode de réflexion efficace et, avant tout, se montrer adaptable.

6.1 Proposer une méthode de travail adaptée

En fonction du contexte de réalisation, des spécificités de l'équipe projet réunie et de l'expérience de ses membres, certaines méthodes de travail pourront se révéler plus efficaces que d'autres. Il est donc important de réfléchir en amont aux outils qui peuvent être mis à disposition pour faciliter la réflexion et la réalisation des différentes tâches. Sans oublier de vérifier leur pertinence en cours de route.

6.1.1 Mettre en place des outils pour optimiser le travail préliminaire

6.1.1.1 Constitution d'une bibliothèque de thesauri existants

C'est une constante dans les normes sur la conception de thesaurus, il est toujours préconisé de s'inspirer de langages existants [1, ISO, Clause 13.1.3.3, p. 89]. A quoi bon réinventer la roue ? C'est la méthode qui avait été appliquée à l'origine lors de la création du thesaurus Techlib à la CNCE (Cf. 3.1.1.2). Le thesaurus que nous avons à refondre en était donc l'émanation. Même si celui-ci constituait notre source essentielle pour le choix des termes, une liste de thesauri existants portant sur les domaines à traiter a été compilée et mise à disposition des membres de l'équipe dès le début des travaux. Tous étaient disponibles en ligne, et ceux qui le pouvaient ont été téléchargés et capitalisés sur l'espace commun dédié au projet. Une ressource particulièrement utile nous a aidée à repérer les thesaurus francophones disponibles en ligne. Il s'agit du « thésauro-annuaire³⁶ » alimenté par Bruno Richardot sur son site intitulé *Tard Bourrichon*.

Nous listerons ici uniquement les thesauri et autres langages documentaires qui ont été le plus utilisés par les expertes. Pour la liste complète des ressources sélectionnées et proposées, ainsi que les adresses URL, se reporter à l'Annexe 6.

- EuroVoc, thesaurus multilingue de l'Union européenne
- Delphes (thesaurus matière) pour le domaine économique

³⁶ Disponible à l'adresse <http://www.tard-bourrichon.fr/documents/THESAURUSonline/tab/Thesauro-annuaire.html> (dernière mise à jour le 30/09/2013).

- Classification JEL (*Journal of Economic Literature* de l'American Economic Association)
- Thesaurus AISS de la Sécurité sociale
- Thesaurus de l'Activité Gouvernementale du Québec

D'autre part, nous avons pris contact avec la Compagnie Nationale des Commissaires aux Comptes (CNCC) afin de commander le *Thésaurus Audit Comptabilité* édité par cette structure (mars 2010), qui n'est pas disponible en ligne³⁷. Du fait des vacances, celui-ci n'a été réceptionné par l'experte concernée qu'en toute fin de mission.

6.1.1.2 Autres sources exploitées

Outre ces ressources propres aux langages d'indexation proposées à l'équipe, chaque experte était libre de puiser parmi les ressources à sa disposition pour trouver les informations qui lui semblaient nécessaires pour mieux appréhender ses domaines. Bien sûr, comme on l'a évoqué plus haut, les veilles déjà réalisées étaient une mine pour identifier les principaux concepts du domaine traités dans l'actualité. Mais d'autres sources étaient indispensables pour valider certaines définitions ou pour mieux comprendre les relations existant entre deux notions. En voici les principales, données par les membres de l'équipe :

- le Code monétaire et financier ;
- des ouvrages de référence du domaine ;
- des dictionnaires et lexiques spécialisés (Vernimmen sur la finance d'entreprise et la gestion bancaire, Dictionnaire technique de la Bourse et des produits financiers, dictionnaire d'économie, etc.) ;
- des manuels de formation ;
- des études Xerfi.

En ce qui concerne les ressources en ligne, un usage extensif et généralisé a été fait de l'encyclopédie Wikipédia, pour mieux cerner certains concepts (après s'être assuré de la qualité de l'article : exhaustivité, sources présentées, etc.). Les sites portails de certains ministères se sont révélés utiles également (Ecologie et Développement durable, LégiFrance), mais aussi des sites plus spécifiques à chaque domaine (collectivités territoriales, Crédit Immobilier, Novethic).

Enfin, les demandes des utilisateurs pouvaient orienter sur des points à d'intérêt à creuser ne figurant pas encore dans le thesaurus.

³⁷ Plus d'informations dans le communiqué de presse : http://www.cncc.fr/download/thesaurus_2010-02-communiqu%C3%A9_de_presse.pdf

6.1.1.3 Grille de recensement des modifications proposées

Afin de garder une trace de tous les changements proposés et validés, un outil très simple a été imaginé. Un fichier Excel a été créé pour chaque experte, comprenant autant de feuilles qu'elle avait de thématiques à traiter. Il devait dans un premier temps permettre en un coup d'œil de comprendre l'environnement sémantique du terme à ajouter ou modifier (Cf. TABLEAU 2 ci-dessous).

TABLEAU 2 : GRILLE DE RECENSEMENT DES MODIFICATIONS PAR DOMAINE

TERME A SUPPRIMER	TERME A MODIFIER	TERME A AJOUTER	TG	TS	EM	EP	TA	NA

Ces fichiers ont été réunis dans l'espace commun, de sorte que chaque membre de l'équipe pouvait ajouter ses suggestions à une autre experte, au fur et à mesure de ses propres travaux. Ainsi, quand elle rencontrait un concept intéressant à intégrer mais n'appartenant pas directement à l'un de ses domaines, elle pouvait le proposer, en précisant éventuellement une relation d'association. Ils constituaient donc également un outil de communication entre les membres de l'équipe, destiné à favoriser les échanges et le travail en commun, pour pallier au manque de temps pour se réunir.

6.1.2 Animer des séances de travail

6.1.2.1 Des séances en tête-à-tête régulières et productives

Proposées initialement comme une assistance pour les expertes, les séances de travail en tête à tête avec le chef de projet se sont vite révélées constituer des moments indispensables. Ces rendez-vous, programmés pour chaque *run*, étaient en effet l'occasion pour chacune :

- d'exposer et de valider le travail préparatoire réalisé jusque-là ;
- de confronter les points de vue et d'effectuer des recherches complémentaires pour affiner un concept ;
- de réfléchir aux relations appropriées entre les termes ;
- de restructurer l'ensemble du domaine en fonction des besoins identifiés ensemble, etc.

La compétence de l'experte sur son domaine se trouvait donc, lors de ces séances, complétée par le savoir-faire du chef de projet concernant la conception de thesaurus. Les discussions qui en résultaient permettaient en outre de réorienter le travail restant en fonction des orientations décidées, optimisant ainsi son efficacité.

6.1.2.2 Des temps de travail en commun indispensables

Outre ces séances à deux, certains points ont pu nécessiter de réunir des expertes traitant de domaines concomitants, ou dont les thématiques de veille concernent un autre domaine. Ainsi, l'experte en charge de la veille « Gestion de patrimoine » a été amenée à travailler étroitement avec celle chargée du sous-domaine « Activité bancaire », dans lequel la première s'insère.

D'autre part, à la fin de chaque *run*, les domaines ou sous-domaines qui venaient d'être traités étaient systématiquement mis en commun pour avis complémentaire du reste de l'équipe. Ce regard extérieur devait permettre, à la fois, de repérer des erreurs ou des manques, et d'identifier des concepts insérés en double – en l'absence de polyhiérarchie.

6.2 Organiser les principales étapes de la refonte du thesaurus

6.2.1 Phase d'évaluation conceptuelle et de collecte des termes

Chaque *run* débutait nécessairement par une phase de réflexion préliminaire, ayant pour but d'évaluer l'existant et de définir les manques.

6.2.1.1 Evaluation du thesaurus Techlib

Les problèmes posés à l'indexation par le thesaurus Techlib, ou plutôt la liste des mots-clés hybride qui en résulte dans JLB-Net, ont été exposés plus haut. (Cf. 4.1.2 et 0) Mais l'objectif du projet étant une refonte du thesaurus existant, une première phase essentielle devait consister à évaluer cette source. Pour en conserver la majorité de concepts utiles, il s'agissait d'identifier les parties nécessaires et d'éliminer les parties inutiles.

Pour cela, deux étapes principales étaient menées par les expertes lors de leurs travaux préliminaires, en début de *run* :

- l'identification des termes obsolètes au sein du thesaurus papier, en s'appuyant sur leur expertise du domaine ;
- la vérification de l'utilisation des termes dans JLB-Net.

6.2.1.2 Défrichage des termes inutilisés

Cette deuxième étape avait pour but d'identifier les termes non utilisés, ou très peu, à l'indexation jusqu'ici, comme pouvant légitimement être listés comme candidats à la suppression. Parfois, un examen plus attentif pouvait cependant s'avérer utile. En effet, la base de données DIP de la BFBP indexait essentiellement des articles. Ceux-ci couvrant souvent des domaines pointus, ils pouvaient demander un niveau d'indexation plus profond. Or, ces descripteurs ne sont pas nécessairement applicables au fonds documentaire de BPCE. Dans tous les cas, la suppression

des termes inutilisés est une étape indispensable pour garantir une meilleure description du fonds documentaire existant, et ainsi une adéquation satisfaisante avec l'information proposée et nécessaire.

6.2.1.3 Collecte des concepts pertinents et choix des descripteurs

Une fois réalisée cette phase importante d'évaluation et de défrichage, l'experte disposait enfin d'une vision plus claire et objective de la couverture de son domaine par le thesaurus Techlib. Après avoir repéré les concepts à conserver, c'est-à-dire pertinents et effectivement utiles, la seconde étape consistait à valider les descripteurs utilisés pour les décrire.

6.2.1.3.1 Sélection des termes à conserver ou à modifier

A ce stade, les termes effectivement utilisés à l'indexation correspondaient donc à des concepts pertinents. Cependant, le terme choisi à l'époque pour le représenter devait à son tour être évalué. Cette fois, c'est l'expérience du vocabulaire employé dans les sources traitant du domaine qui permettait à l'experte de s'orienter. Parfois, le terme défini à l'époque pour traduire un concept nouveau n'avait pas pris, ou était passé de mode. Dans ce cas, il suffisait parfois d'invertir une relation d'équivalence. Lorsqu'il était nécessaire de définir un nouveau terme préférentiel, l'ancien terme devait néanmoins être conservé comme non descripteur (éventuellement avec une note historique), afin de réindexer par la suite les notices concernées.

Cette étape était également l'occasion de vérifier si les descripteurs correspondaient aux principaux critères de sélection préconisés par la norme [1, ISO, Clause 6, p.21-37] :

- nom ou groupe nominal de forme singulière (à moins que le sens n'en soit affecté) ;
- non ambiguïté des termes ;
- sigle développé comme descripteur, sauf exception (usage commun, sens oublié) ;
- orthographe la plus répandue (selon une source reconnue)

Le critère de non-ambiguïté des termes est en particulier une aide importante pour choisir le terme préférentiel. Loin d'être un critère universel, il doit orienter vers le terme dont la compréhension sera la plus évidente pour les utilisateurs constituant le public cible du thesaurus.

6.2.1.3.2 Recherche de nouveaux concepts ou de nouveaux descripteurs

La dernière étape de ce travail sur les concepts et les descripteurs consistait enfin à compléter ceux qui avaient été conservés par ceux identifiés comme manquants. Les différentes sources présentées plus haut (Cf. 6.1.1), et bien sûr les veilles archivées et les connaissances des expertes sur leurs domaines ont ici été mises à contribution. De nombreuses recherches ont été nécessaires pour mieux cerner ces concepts à insérer dans le thesaurus, pour lesquels toutes les relations sémantiques étaient à créer.

La norme propose d'autres pistes dans le choix des termes qui se sont avérées utiles pour traiter ces nouveaux concepts :

- préférer des termes établis au jargon ou aux néologismes (sauf si leur usage est déjà répandu ou pour des concepts émergents) ;
- le choix des termes d'emprunts est réservé à ceux dont l'usage est déjà bien établi.

Ces deux recommandations s'appliquent particulièrement aux domaines qui nous intéressent. En effet, la finance comporte plusieurs notions complexes pour lesquelles le terme anglo-saxon est le plus couramment utilisé. Cela s'exprime à l'indexation, puisque les ouvrages les utilisent, mais surtout à la recherche, puisque les utilisateurs les emploieront de préférence dans leurs requêtes, que celles-ci soient faites aux documentalistes ou, bientôt, en direct via le portail. La complexité de ces concepts impose de s'en tenir au terme le plus courant, pour ne pas induire des erreurs de traduction qui pourraient perturber l'accès à l'information. Néanmoins, une vérification systématique a été réalisée auprès de la base de données France Terme³⁸, pour obtenir la traduction officielle des termes d'emprunt si elle existe, et ainsi la définir en non descripteur. Ainsi, si le terme français parvient à s'imposer, il suffira d'inverser la relation d'équivalence.

6.2.2 Phase de hiérarchisation des concepts

6.2.2.1 Reconstruction et consolidation de la structure hiérarchique

Le thesaurus Techlib en version papier correspondait à un index thématique et alphabétique. Le module de gestion de thesaurus de JLB ne propose quant à lui qu'un affichage alphabétique des termes, complétés de leurs relations. Pour offrir aux expertes une vision plus juste de la hiérarchie de leurs domaines, nous avons procédé à une réorganisation de chaque domaine pour obtenir une présentation thématique et hiérarchisée de l'existant. Celle-ci présentait deux avantages principaux :

- plus de lisibilité pour la suite de la réflexion ;
- identification des concepts orphelins facilitée ;
- élimination systématiques des incohérences (relation associative non symétrique, etc.).

Après avoir retravaillé cette version, en fonction du résultat de la première phase exposée plus haut, en séance de travail avec l'experte, on pouvait procéder à l'insertion des nouveaux concepts et à la construction de leurs propres relations au sein de cet ensemble « nettoyé ».

³⁸ Site du ministère de la Culture, consacré aux termes recommandés au *Journal officiel* de la République française : <http://www.culture.fr/franceterme>

6.2.2.2 Ajustement de la profondeur d'indexation

Une autre problématique posée à l'indexation par le thesaurus Techlib concerne la profondeur d'indexation. En effet, du fait de son mode de conception, la réflexion portant sur l'adéquation entre les niveaux de spécificités dans la hiérarchie et les besoins réels pour la description du fonds documentaire a été très limitée. Aujourd'hui, avec plus de vingt ans d'expérience de mise en œuvre de ce langage contrôlé, un ajustement s'avérerait nécessaire. Cette démarche est d'autant plus urgente avec la mise en place du portail OPAC, car elle permet de limiter le bruit et le silence engendrés par une profondeur d'indexation inadéquate.

Là encore, l'examen des notices indexées dans JLB-Net permet d'identifier les termes surexploités, qui indiquent qu'il faut augmenter le niveau de profondeur d'indexation. En effet, cela indique qu'il s'agit d'un descripteur trop générique, sous lequel sont regroupés des documents qui pourraient être décrits plus efficacement grâce à des spécificités de ce concept. À l'inverse, les termes spécifiques sous-exploités indiquent que la profondeur d'indexation est trop élevée et que le terme générique est peut-être suffisant ici. Bien sûr, cela reste à valider en fonction de l'expérience du domaine de l'experte. Elle peut estimer qu'il s'agit d'un concept peu connu jusque-là, mais qui connaît une émergence qui justifie de rester dans le thesaurus. Une autre solution est de définir une relation d'équivalence (de type quasi-synonymie), qui permette de faire apparaître le terme (notamment à la recherche) tout en renvoyant en pratique à un concept plus large et plus utile.

6.2.3 Phase d'enrichissement du réseau sémantique

6.2.3.1 Repérage des associations pertinentes

Le thesaurus Techlib, tel qu'implémenté dans JLB-Net, comportait 844 relations associatives (soit 422 relations réciproques). Lorsque deux concepts ayant été conservés étaient associés par une relation jugée pertinente, c'est-à-dire dans la majorité des cas, celle-ci est restée inchangée. Pour les concepts ajoutés, quelques associations ont pu l'être également directement par l'experte du domaine.

Pour la validation de ces associations, nous nous sommes appuyées sur la Clause 10.3, *The associative relationship*, de la norme [1, ISO, p.63-66]. L'objectif essentiel ici est d'apporter une valeur ajoutée à la recherche, en suggérant des termes additionnels ou alternatifs à la requête initiale. Le critère d'évaluation de la pertinence d'une association n'est donc pas que la relation soit évidente (auquel cas le lien serait inutile), mais qu'elle est suffisamment naturelle pour être proposée systématiquement à quiconque recherche l'un des termes concernés. La norme explique ainsi que l'un des indices est qu'on ne puisse définir l'un des termes sans avoir recours au second.

Pour ce projet, nous avons fait le choix de n'associer que des termes appartenant à des sous-domaines différents. Le but était de ne pas compliquer inutilement la structure de ce langage contrôlé qui devait rester limité, tout en encourageant un travail plus poussé sur la hiérarchisation des concepts.

6.2.3.2 Mise en commun du travail finalisé

La dernière étape de chacun des *rums* consistait à soumettre le travail effectué aux autres membres de l'équipe projet au cours des réunions de service. Un domaine ou sous-domaine n'était présenté au groupe qu'une fois qu'il avait été entièrement refondu et sa structure stabilisée, avec validation finale de l'experte concernée. Cette mise en commun avait trois objectifs essentiels pour la qualité finale du livrable :

- correction des erreurs formelles (fautes de frappe, présentation...)
- identification de concepts déjà présents dans un autre domaine (doublons) ;
- proposition de concepts manquants par les expertes de domaines proches ;
- proposition de relations associatives avec des termes présents dans les domaines de chacune.

Seule étape où les membres de l'équipe pouvaient réellement se pencher sur le travail réalisé par leurs collègues, ces moments constituaient des temps importants dans la gestion de ce projet. Ils étaient l'occasion de faire des points réguliers et concrets sur des aspects parfois mal compris de la conception de thesaurus ou sur les objectifs spécifiques et la politique adoptée pour le projet. Ils étaient surtout l'occasion d'échanger toutes ensemble sur le thesaurus et la méthode de travail, de partager des conseils, des sources pertinentes, des astuces pratiques... C'est lors de ces échanges que le thesaurus a le plus gagné en qualité et en homogénéité, que son réseau sémantique s'est le plus enrichi. Malheureusement, du fait des contraintes évoquées plus haut, elles n'ont jamais pu réunir l'équipe au complet.

6.3 Proposer un accompagnement constant et personnalisé

Si ces grandes étapes de travail se sont révélées plus ou moins incontournables, le niveau d'accompagnement par le chef de projet a varié selon les membres de l'équipe, en fonction de leurs besoins et attentes respectives. Dans tous les cas, un rythme d'au moins une séance de travail par semaine avec chacune s'est rapidement mis en place.

6.3.1 S'adapter au besoin individuel

Dès le début du projet, les réactions des expertes ont été très contrastées vis-à-vis de l'accompagnement qui leur était proposé. Les séances de travail n'étaient volontairement pas programmées à l'avance, afin d'en laisser l'initiative à chacune, en fonction du besoin ressenti.

6.3.1.1 Des profils distincts

On peut distinguer initialement trois profils :

- **Accompagnement permanent** : les séances de travail sont perçues comme indispensables pour avancer. Le travail préparatoire est concrétisé par une réflexion en commun, quotidienne ou presque. Des objectifs pour la suite du *run* sont établis en fin de séance, pour préparer la suivante.
- **Accompagnement ponctuel** : les séances de travail sont demandées chaque fois qu'une aide semble nécessaire, ou que suffisamment d'éléments ont été rassemblés pour souhaiter les valider ensemble.
- **Accompagnement de finalisation** : la séance de travail vient clore un *run*. La nouvelle hiérarchie est reconstruite selon le nouveau modèle et contrôlée ensemble. Si quelques modifications restantes sont identifiées, elles sont effectuées immédiatement.

6.3.1.2 Une demande d'accompagnement croissante

Néanmoins, ces trois profils ne sont pas à proprement parler un indice du niveau d'autonomie des expertes. En effet, l'efficacité de l'accompagnement constant mené avec le premier profil a très vite fait ses preuves, incitant les autres profils à demander des séances de travail plus régulières. Les échanges que permettaient ces séances, notamment avec une personne non-experte du domaine, se révélaient particulièrement fructueux et enrichissants. Ils étaient donc appréciés de toutes, quelle que soit l'expertise en conception de thesaurus par ailleurs.

D'autre part, le troisième profil n'a concerné réellement que la chef de service (OL), dont le niveau de formation et l'expérience justifiaient une telle autonomie. Son emploi du temps chargé et le nombre de ses domaines d'expertise lui imposaient de s'organiser au plus juste. De plus, les domaines sous sa responsabilité (banque, finance, assurance) demandaient énormément de recherches en amont pour être mieux appréhendés et actualisés : travail solitaire s'il en est !

6.3.2 Identifier les inclinations personnelles

En fonction de sa formation, de son expérience, de ses habitudes ou encore de sa personnalité, chaque individu développe sa propre méthode de travail, avec laquelle le chef de projet doit composer pour être efficace.

6.3.2.1 Des démarches individuelles

Dès le lancement du projet, une fois cadrés la méthodologie et le planning, chaque experte a été laissée libre de mener ses tâches à sa manière. Au préalable, une présentation des grands principes du thesaurus avait permis une remise à niveau sur le sujet, et les outils proposés à l'équipe avaient été détaillés. La répartition des domaines à traiter en *runs* aux échéances fixées

à l'avance (mais non rigides) garantissait une bonne visibilité, devant permettre une grande autonomie pour articuler tâches quotidiennes personnelles et tâches liées au projet. En schématisant, deux démarches principales ont pu être adoptées spontanément par les expertes :

- **L'approche « thesauro-centrée »** : partir d'une analyse fine de la présentation de son domaine dans le thesaurus Techlib, puis évaluer son adéquation à la réalité du terrain, telle que perçue à travers des veilles et des recherches complémentaires.
- **L'approche « domaine-centrée »** : commencer par se réappropriier le domaine de connaissances par des recherches plus poussées qu'à l'habitude (ouvrages de référence et Web), pour mieux évaluer l'existant et le travail à accomplir.

Il est intéressant de noter que, même si les tendances personnelles influencent forcément le choix de l'une ou l'autre approche, celui-ci n'était pas systématique. En effet, le niveau de connaissance effectif d'un domaine – ou du moins son appréhension personnelle – était un facteur non négligeable, qui pouvait selon les *runs* encourager des démarches différentes de la part d'une même experte. D'autre part, l'ampleur des domaines ou sous-domaines à retraiter pouvait varier considérablement, encourageant à tester différentes approches.

Quelle que soit la démarche adoptée, il revenait au chef de projet d'équilibrer le processus de réflexion lors des séances de travail, en incitant l'experte à compléter son approche initiale par la seconde. En effet, les deux sont bien sûr indissociables et complémentaires pour mener à bien la refonte d'un thesaurus. Pour cela, une capacité d'écoute et d'observation sont indispensables pour cerner les différents profils et s'y adapter, comme on le verra plus loin.

6.3.2.2 Une appétence commune pour le travail collaboratif

L'équipe d'un centre de veille et documentation n'a bien souvent que peu d'occasions de travailler ensemble. Chacun y a ses propres responsabilités et les tâches sont si nombreuses ! Un projet comme celui-ci est donc l'opportunité de travailler autrement et d'échanger. C'était sans doute l'une des attentes fortes de cette équipe, bien qu'implicite.

Et en effet, le grand regret exprimé par l'ensemble de l'équipe lors du bilan de projet est de n'avoir pu, en raison des contraintes de temps et d'organisation, bénéficier de plus de moments dédiés au travail en équipe. En effet, malgré les étapes de mises en commun régulières, la place de la réflexion transversale s'est avérée beaucoup plus limitée qu'il ne l'aurait fallu. Un plus haut niveau de partage aurait été bénéfique, à la fois, à la qualité du nouveau thesaurus et pour les acquis qu'en ont retirés les acteurs du projet. Ce paramètre est donc à prendre en compte dans l'évaluation du temps nécessaire à la conduite d'un projet de refonte, parallèlement à l'évaluation des tâches en elles-mêmes. Il participe des enjeux pour le management, que nous aborderons dans la prochaine partie.

Troisième partie

Quelles perspectives pour le management d'un service documentaire ?

7 Un nouveau thesaurus, et après ?

Trois questions semblent essentielles, lorsque l'on vient de mener un projet de refonte de thesaurus. Il s'agit évidemment d'évaluer le livrable obtenu, puis d'en estimer les conséquences en termes de maintenance et d'évolution, de faire le point, enfin, sur les outils permettant de l'exploiter et le valoriser.

7.1 Composition du thesaurus BPCE à l'issue du projet

Nous avons vu la méthode appliquée pour aboutir au thesaurus BPCE et ses principaux objectifs. Il est temps d'examiner plus en détail sa composition et sa structure, fruit d'un compromis entre l'existant et les objectifs qualitatifs du commanditaire et du chef de projet. Pour illustrer cette partie, on pourra se reporter aux Annexes 2 et 3 de ce mémoire (table des matières et exemple d'un domaine du thesaurus BPCE).

7.1.1 Un thesaurus spécialisé organisé autour des thématiques de veille

A l'issue de la refonte, le thesaurus BPCE compte onze domaines, ou micro-thesauri (table des matières complète en Annexe 2) :

1. Traitement de l'information³⁹
2. Sciences économiques
3. Institutions et politique économiques
4. Questions juridiques⁴⁰
5. Monnaie / crédit
6. Système financier
7. Epargne et marchés de capitaux
8. Relations internationales
9. Social
10. Secteurs d'activité
11. Entreprise

Les domaines 5, 6 et 7 constituent le cœur du thesaurus, puisqu'ils sont consacrés aux activités centrales de l'entreprise. Par conséquent, ce sont les plus développés. La structure générale, ainsi que l'ordre utilisés dans le plan de classement ont été conservés, afin de ne pas perturber les

³⁹ Ce domaine est quant à lui plutôt réservé à un usage interne au service.

⁴⁰ Ces trois domaines transversaux ne font pas l'objet d'une veille thématique, en-dehors de la veille réglementaire quotidienne concernant les secteurs de la banque et de l'assurance.

habitudes de travail des documentalistes. Ces 11 micro-thesauri, et en particulier les sous-domaines qui les composent, correspondent de près aux thématiques traitées par les produits de veilles documentaires, de la même façon que celles-ci orientent les choix des documentalistes pour la constitution du fonds documentaire.

7.1.2 Description de la structure hiérarchique

7.1.2.1 Structure utilisée pour le thesaurus spécialisé BPCE

L'un des axes d'évolution de la nouvelle norme ISO 25964 encourage une plus grande hiérarchisation de la structure théssaurale, ce qui s'applique particulièrement bien à un thesaurus spécialisé multi domaines, tel que ceux mis en œuvre dans les entreprises. Nous avons choisi d'adopter cette approche, qui devait permettre en outre de rester le plus proche possible du plan de classement qui structurait le thesaurus jusqu'ici. Dans cette optique, le parti-pris a été d'utiliser tous les niveaux de structuration proposés par la norme, à savoir :

- Domaine (micro-thesaurus)
- Champ sémantique (*subject category*, noté CS) ou sous-domaine
- Terme de tête (*top term*, noté TT) ou sous-sous-domaine
- Terme générique (noté TG)
- Terme spécifique (noté TS)

Nous avons ainsi pris le parti d'intégrer la fonction de *Top Term* (ou "terme de tête", noté TT) pour plus de structuration. « *L'affichage de la relation directe entre un concept et le concept situé en tête de la structure hiérarchique dans laquelle il se trouve peut être très utile dans le cas de thesaurus fortement hiérarchisés.* » [4, DALBIN et al., p.16] Les champs sémantiques (CS) ne sont pas des descripteurs. Le niveau de spécification sous les termes génériques a été limité à trois (TS3), en adéquation avec les besoins des utilisateurs.

Ces partis pris correspondent à une contrainte forte, posée dès la phase d'étude préliminaire à la mission avec le commanditaire. Le lien étroit avec le plan de classement, dans lequel s'insérait jusqu'ici le thesaurus, devait être respecté. Néanmoins, pour atteindre le niveau de qualité visé et la cohérence nécessaires, certains domaines ont pu être éclatés et des sous-domaines fusionnés ou ajoutés. Ainsi, le développement durable, dont on ne parlait pas il y a 20 ans, est devenu une préoccupation incontournable des organisations. Ce thème a donc tout naturellement trouvé sa place dans le nouveau thesaurus.

7.1.2.2 Hiérarchie atteinte à l'issue du projet

La hiérarchie du thesaurus se décomposait ainsi au 6 septembre 2013 :

- 39 champs sémantiques (CS)
- 63 termes de tête (ou « *top terms* » TT)
- 326 termes génériques (TG)
- 2 377 termes préférentiels
- 771 termes non-préférentiels (EP)
- 281 relations associatives, soit 562 réciproques (TA)

7.1.3 Evaluation du RAF (reste à faire)

A l'issue de la mission, l'état d'avancement du livrable était estimé à 95 %. En effet, certains domaines n'ont pu, pour des raisons de temps ou de ressources disponibles, être finalisés.

Domaines à finaliser :

- Comptabilité, à partir du thesaurus de la CNCC réceptionné en fin de mission seulement
- Termes géographiques : il est recommandé de reprendre une liste d'autorité existante, en l'adaptant au besoin

Domaines à enrichir :

- Développement durable : Il s'agit d'une création de ce nouveau thesaurus, inséré dans le plan de classement avec la côte 2600.
- Noms propres : liste d'autorité à reprendre depuis JLB (mise à jour) et à enrichir au fur et à mesure par toute l'équipe.

Le thesaurus étant un outil vivant, comme on l'a vu, on ajoutera bien sûr à cette liste tous les autres domaines à compléter et adapter à la réalité future.

7.2 Outils mis en place pour l'exploitation du thesaurus

7.2.1 Identification des besoins

Le besoin primaire était la refonte et la mise à jour du thesaurus en usage à la DG BPCE. L'essentiel du travail s'est porté sur la version électronique/papier du thesaurus. Une présentation thématique et hiérarchisée, on l'a vu, a permis d'apporter plus de lisibilité par rapport à l'ancienne présentation, thématique et alphabétique. Cet outil représentait le livrable final attendu par le commanditaire.

Le second besoin, qui n'a pu être traité lors de cette mission, est le nettoyage du lexique de mots-clés proposé actuellement dans le progiciel de gestion des ressources documentaires JLB-Net.

Les termes appartenant au thesaurus devront être retraités en fonction des modifications apportées et les notices réindexées. Les mots-clés issus de DIP devront également être intégrés à cette reprise.

7.2.2 Présentation adoptée pour le thesaurus « papier »

7.2.2.1 La présentation thématique hiérarchisée

La présentation choisie pour le thesaurus BPCE est celle qui correspond le mieux à l'usage actuel en interne du thesaurus. Elle reste donc très proche de celle d'un plan de classement enrichi d'un index, comme c'était le cas jusqu'à présent. La différence majeure apportée est que le lexique n'est plus présenté par ordre alphabétique, sur la partie gauche, complété par les relations propres à chaque terme, recensées sur la partie droite.

Désormais, les termes sont présentés de façon hiérarchisée, et l'arborescence est matérialisée par indentations successives, de la manière suivante :

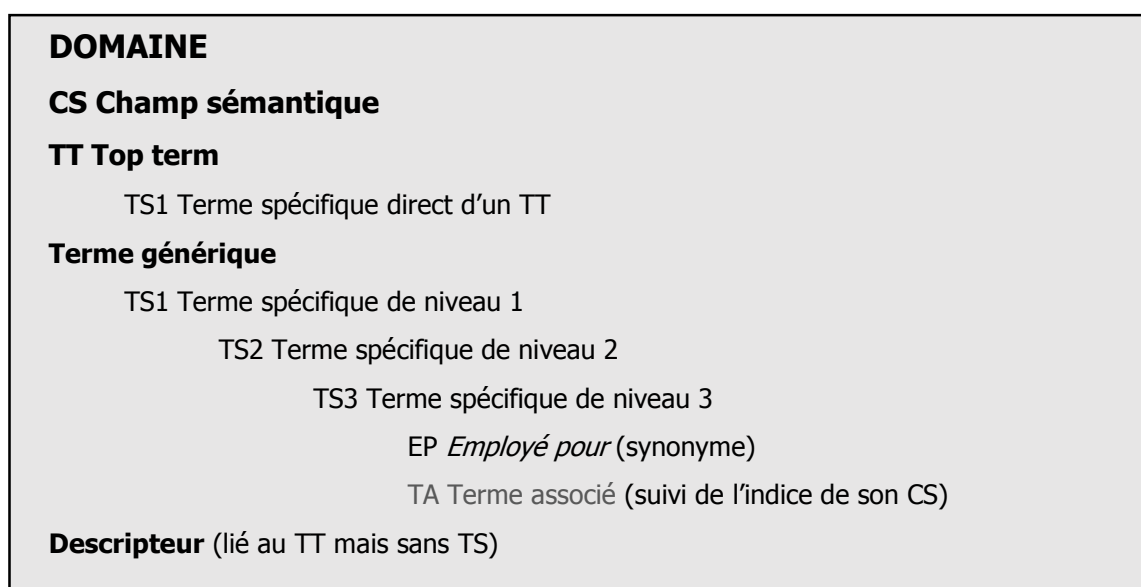


FIGURE III : PRESENTATION SCHEMATIQUE DU THESAURUS BPCE

7.2.2.2 Principes suivis pour la présentation

La présentation adoptée pour le thesaurus « papier » (disponible bien sûr en version numérique) a été conçue de façon à garantir l'accès le plus rapide et le plus efficace possible à l'information recherchée par les utilisateurs, les documentalistes. En fonction de l'utilisation du langage d'indexation à BPCE (Cf. 3.3), cette structure semblait la plus performante.

L'ordre des différents termes, selon la fonction qui leur est attribuée dans le thesaurus, a été pensé de façon à limiter tout risque de mauvaise manipulation future dans l'éditeur de texte, lors des mises à jour. Ainsi, le niveau de spécification d'un terme est toujours précisé (TS1, TS2 ou TS3), au cas où la tabulation serait perturbée. De la même façon, les synonymes (EP) et termes

associés (TA) sont toujours présentés directement après le descripteur auquel ils se rattachent. D'autre part, une édition en format PDF est conservée en permanence comme version de référence, assortie d'une politique de versioning permettant la gestion des éditions successives.

Ces principes de présentation paraissent en effet indispensables dans l'attente de l'implémentation du nouveau thesaurus dans l'outil informatique indexant les ressources du service. Seule cette mise en œuvre dans le logiciel de SIGB permettra de garantir la pérennité du thesaurus et de son usage.

7.2.3 Outils logiciels pour l'exploitation du thesaurus

7.2.3.1 Implémentation dans le module de gestion de thesaurus intégré au logiciel de SIGB JLB-Net

Le second outil est le logiciel de gestion de bibliothèque JLB-Net. Les transformations apportées au thesaurus devront être implémentées dans son module de gestion de thesaurus. Ce travail est actuellement en cours de réalisation, confié à chaque experte sur ses propres domaines. Pour les y aider, une note de procédure (Annexe 4) a été rédigée expliquant les différentes manipulations nécessaires. Bien que contraignante, cette opération apporte aussi des bénéfices essentiels, que sont une prise en main totale du thesaurus sur ses domaines, ainsi qu'une étape de vérification et de correction sur le travail effectué.

Néanmoins, cette étape cruciale pour la suite du projet et la transition effective vers le nouveau thesaurus s'avère délicate à mener, sans l'appui du chef de projet externe dont a bénéficié la première phase du projet. Des tensions internes et une motivation émue par le retour au rythme habituel du service font regretter à sa responsable de n'avoir pas mieux pris en compte la durée nécessaire à cette seconde phase dans son estimation des délais de l'intervention. Ainsi, des semaines précieuses ont été passées à mettre au point une méthodologie acceptée par toutes pour réaliser ces tâches.

7.2.3.2 En complément, préconisation d'un logiciel autonome

Les fonctionnalités proposées par le logiciel JLB-Net pour l'exploration du thesaurus s'avérant extrêmement limitées, il nous a semblé opportun d'en proposer un complément. En parallèle de la mission, nous avons donc procédé à des tests fonctionnels et à la validation d'un outil complémentaire, à mettre en place pour aider à l'exploration et à l'affichage du thesaurus : le logiciel autonome (ou *standalone*) de création et gestion de thesaurus ThManager (Cf. Annexe 5 pour le guide utilisateur).

7.2.3.2.1 *Présentation du logiciel ThManager*

ThManager 2.0 est un logiciel de création et gestion de thesaurus développé par le Laboratoire des Systèmes d'Information Avancés de l'Université de Saragosse, Espagne⁴¹. Il a été choisi essentiellement car sa licence est libre et gratuite (*Open source*). Néanmoins, la licence GNU (Licence Publique Générale Limitée) permet d'intégrer cette application à un logiciel de licence commerciale (pas de clause de *copyleft*). D'autre part, ThManager génère des thesaurus en format SKOS (Simple Knowledge Organization system), langage correspondant aux standards du W3C (World Wide Web Consortium) concernant les vocabulaires contrôlés et structurés. Cela garantit la meilleure interopérabilité avec tout autre logiciel qui nécessiterait l'utilisation du thesaurus à l'avenir.

7.2.3.2.2 *Obstacle rencontré et préconisations*

Suite à un incident survenu au cours d'une demande à l'assistance informatique interne, les données du thesaurus déjà implémentées dans ThManager ont été perdues. En conséquence, une étude de marché sur les autres logiciels de gestion de thesaurus disponibles a été menée. A l'issue de celle-ci, ThManager a été confirmé comme un outil satisfaisant et adapté au besoin. Il présentait en outre l'avantage d'être désormais disponible dans l'environnement informatique BPCE et d'avoir déjà été pris en mains. Il a donc été préconisé de le conserver, dans la mesure où la direction informatique pourrait garantir une mise en place sécurisée et stable sur les postes des futurs utilisateurs. En tout état de cause, il s'avère essentiel de procéder à un export du thesaurus à chaque étape de son implémentation. Sa conservation en format SKOS, au besoin sur différents supports, assure de toujours pouvoir le réimporter dans le logiciel en cas de perte des données accidentelle.

Cependant, si les fonctionnalités de ThManager correspondent globalement au besoin d'exploration du thesaurus BPCE, il ne permet pas de le structurer dans toute la richesse des relations prévue. L'usage déterminera donc l'opportunité d'examiner d'autres logiciels plus adaptables, mais sans doute payants (budget moyen 1 500/2 000 € la première année). D'autre part, ThManager adoptant les recommandations du W3C, il est probable qu'une prochaine version offre plus de possibilités, dans le sens de la norme ISO 25964. Il sera donc important de suivre ces évolutions.

⁴¹ <http://iaaa.cps.unizar.es/>

7.2.3.2.3 Documents supports

Outre le guide intégré à l'application (en anglais), un guide utilisateur a été rédigé au cours de la mission (Annexe 5), en s'appuyant sur les tests d'utilisation portant sur la création du domaine « Entreprise » du thesaurus. Par ailleurs, une séance individuelle de prise en main a été organisée pour la documentaliste qui a accepté de prendre la suite du projet en s'occupant de la maintenance du thesaurus.

7.3 Préconisations pour l'amélioration du thesaurus BPCE

Le travail réalisé en à peine plus de deux mois a été, grâce à l'implication et au sérieux de l'équipe, d'une ampleur remarquable. L'objectif initial de reprendre l'intégralité du thesaurus semblait en effet très largement utopique, pourtant l'essentiel du travail a pu être achevé dans le temps imparti. Néanmoins, il est certain qu'il demeurerait impossible d'aboutir à un outil parfait en si peu de temps, d'autant moins sans disposer d'une équipe entièrement dédiée au projet. Il est donc nécessaire de procéder *a posteriori* à une critique constructive de l'outil obtenu. Les quelques éléments évoqués ici, suivis de propositions de pistes d'améliorations, sont autant de préconisations pour enrichir et pérenniser le travail ambitieux qui a été entrepris.

7.3.1 Analyse critique du livrable obtenu

7.3.1.1 Une structure qui reste plus proche du plan de classement

7.3.1.1.1 Un outil conçu selon une contrainte forte

Les contraintes principales qui se sont appliquées aux travaux menés lors de la refonte de ce thesaurus étaient l'impossibilité de modifier le plan de classement, d'une part, et la nécessité exprimée par le commanditaire de rester au plus près de la structure initiale du thesaurus, c'est-à-dire celle d'un lexique inséré dans le plan de classement, d'autre part. Nous avons abordé plus haut la stratégie utilisée pour transcrire cette organisation dans le langage d'indexation ainsi révisé. Néanmoins, ce faisant, celui-ci est naturellement resté plus proche de la logique du plan de classement que de celle du thesaurus.

7.3.1.1.2 Un outil qui répond à l'usage actuel

C'est d'ailleurs cette approche qui est matérialisée dans la présentation choisie pour le livrable papier qui a été obtenu, c'est-à-dire une présentation thématique hiérarchisée, où seuls les termes génériques, leurs spécifiques et les autres simple descripteurs sont présentés en ordre strictement alphabétique. Cette présentation reste actuellement celle qui correspond le mieux à l'usage qui est fait de ce langage d'indexation par les documentalistes, à la fois de guide pour déterminer la cote d'un ouvrage et d'aide pour l'indexation dans le logiciel JLB, qui ne présente qu'un affichage alphabétique des termes.

7.3.1.2 Une démarche de type « pragmatique antifondationaliste »

Il nous semble important de tenter de définir, avec le recul, le point de vue qui a orienté nos travaux. La posture adoptée a bien sûr été impulsée par la démarche épistémologique en usage dans le contexte de mise en œuvre. Elle découle donc de tous les éléments de contexte, de positionnement et de mission au sein de l'organisation qui ont été présentés tout au long de ce mémoire, pour le service Documentation générale de BPCE.

Nous nous appuyerons pour ce faire sur l'article de Joseph Tennis paru récemment dans la revue *Etudes de Communication* [20, TENNIS], ainsi que sur son excellente intervention⁴² lors de la 9^e conférence ISKO France, en octobre 2013 (actes à paraître en 2014). De sa présentation des cinq positions théoriques idéales⁴³ pour la conception de systèmes d'organisation des connaissances (SOC), nous retiendrons celles qui nous semblent s'appliquer à nos travaux. A première vue, il semble que la conception du thesaurus – qui est une forme particulière de SOC – relève pour ce service de documentation (et ce dans sa continuité historique) du **pragmatisme fondationaliste**. Celui-ci serait caractérisé, selon Tennis, par 5 postulats :

- 1) « le langage ne façonne ce que nous savons que dans la mesure où il nous permet d'agir dans le monde ;
- 2) le langage renvoie à un groupe de personnes (petit ou grand) et à leurs actions (nous nous jugeons aptes à observer cette réalité particulière) ;
- 3) nous ne pouvons savoir qu'une fois que nous étudions ce groupe son langage et ses actions (il y a donc des réalités multiples) ;
- 4) nous pouvons représenter dans nos systèmes comment ce groupe utilise le langage et donc les connaissances ;
- 5) nous faisons confiance au langage utilisé par les concepteurs du SOC pour représenter les connaissances de ce groupe parce que nous observons leur utilisation du SOC dans leur travail. »

En effet, selon ces critères, cette approche semble la plus logique pour la conception d'un SOC s'appuyant sur un domaine spécifique, en référence à l'action d'une communauté d'utilisateurs bien identifiée et facilement observable. Cependant, il s'agit selon Tennis d'une position ontologique (au sens philosophique du terme), qui revendique une représentation « correcte », au sens de vérité absolue, des connaissances dans un domaine ou contexte particulier, tout en reconnaissant paradoxalement la nécessité d'une adaptation constante aux évolutions de ce domaine.

⁴² Intitulée " *Rhythm, Mode, and Form: The Culture, Context, and Philosophy of Time, History, and Intention in Knowledge Organization Systems*"

⁴³ Les trois autres étant le réalisme, la pensée critique dialectique et le relativisme linguistique.

Pourtant, nous pensons l'avoir démontré tout au long de ce mémoire, le critère principal ayant orienté toutes les décisions prises pour la conception de ce langage contrôlé, était celui de l'**utilité** pour les utilisateurs. Or, selon Tennis, l'utilité est une des notions clés d'une autre démarche – celle que lui-même encourage : le **pragmatisme antifondationaliste**. Là où la notion de correct impose une représentation complète, la notion d'utile permet, comme ce fut le cas ici, de sélectionner uniquement les éléments du domaine qui vont effectivement servir pour l'action envisagée. Du même coup, le changement ne représente plus une menace pour les fondements de ce SOC, qui peut réellement se développer comme un outil nous aidant à appréhender le monde. C'est cette approche, basée sur le critère d'utilité et donc sur l'intention, qui permet la création de valeur à travers les SOC en entreprise. C'est pourquoi elle est la seule qui nous semble envisageable dans un tel contexte.

Mais pour réellement ambitionner une telle représentation pragmatique d'un domaine d'activité, il est indispensable d'intégrer la question de la valeur sociale du SOC : à qui doit-il être utile ? Pour garantir une représentation du domaine fidèle à la conception de son destinataire, celui-ci ne devrait-il pas toujours être mis à contribution dans sa conception ?

7.3.1.3 Une réflexion terminologique réalisée en interne

La norme présente dans la Clause 13.1.3.2, *Human resources* [1, ISO, p.89] les différents types de ressources humaines que devrait réunir une équipe pour la conception de thesaurus⁴⁴ :

- *« la supervision d'un cadre dirigeant, et idéalement un engagement au niveau de l'organisation dans le projet de thesaurus ;*
- *l'intérêt et le soutien des utilisateurs, en particulier d'experts terminologiques d'une ou plusieurs spécialités du domaine couvert ;*
- *des professionnels de l'information possédant de solides connaissances du domaine, des aptitudes et de préférence une formation à la conception de thesaurus, pour mener à bien les tâches principales de construction et maintenance du thesaurus ;*
- *le support de professionnels des technologies de l'information pour aider à l'installation/l'adaptation du logiciel de gestion de thesaurus, toute programmation requise pour intégrer le thesaurus aux systèmes dans lesquels il sera utilisé, et la bonne marche des applications. »*

Dans le contexte de cette mission, l'ensemble de l'équipe était composée de membres du service de veille et documentation, dont la responsable avait recruté elle-même le chef de projet. Si le

⁴⁴ Traduit de la norme [1, ISO] par l'auteur.

projet n'avait pas de visibilité au niveau organisationnel, elle bénéficiait néanmoins du soutien et de l'intérêt de son supérieur direct, le directeur du pôle Immobilier et Services généraux. On a présenté plus haut les documentalistes impliquées, dont le niveau de compétences dans la conception de thesaurus était globalement satisfaisant. Enfin, le projet n'a pas eu besoin d'une assistance technique à ce stade, l'essentiel de la réflexion étant de nature conceptuelle et terminologique. Elle deviendra éventuellement nécessaire au cours de l'implémentation des modifications dans JLB-Net, mais surtout si les ambitions pour ce vocabulaire venaient à être étendues (Cf. 8.3.2.1).

La seule lacune sur ce point est donc l'absence d'implication directe des utilisateurs dans la réflexion. Bien sûr, les documentalistes ont pu s'appuyer sur leur connaissance des utilisateurs et des requêtes les plus fréquentes. Mais pour valider les choix terminologiques et pousser plus loin la construction des relations sémantiques, l'aide ponctuelle d'expert métiers, au sens strict cette fois, se serait avérée précieuse. Il est à regretter que la durée du projet comme sa réalisation limitée aux ressources internes au service ne nous aient pas permis d'y recourir.

7.3.2 Principaux axes de développement à exploiter

Le contexte de la mission était marqué, à la fois, par un délai très court et par, en corollaire, une organisation basée, comme on l'a vu, sur la délégation des différents domaines en fonction des responsabilités des veilles thématiques. Chacune des expertes a donc travaillé sur chaque partie du thesaurus qui lui était confiée de la façon la plus efficace possible, afin de balayer l'ensemble des concepts qui la composaient et de les compléter, puis d'en consolider la hiérarchie. Aujourd'hui, la prochaine étape consiste à reprendre point par point ce travail, afin de le valider ou de le corriger. Voici les principales pistes à exploiter en priorité.

7.3.2.1 Distinction formelle entre le plan de classement et le thesaurus

La première étape indispensable pour permettre la construction d'un thesaurus à part entière sera bien sûr d'extraire du plan de classement le lexique qui constitue actuellement le thesaurus. Le thesaurus pourra alors être reconfiguré selon sa propre logique, sans avoir à correspondre à celle du plan de classement.

7.3.2.1.1 Deux logiques distinctes

Un thesaurus ne peut être calqué sur l'organisation d'un plan de classement, car ces deux langages répondent à des objectifs spécifiques, obéissant chacun à une logique différente. En effet, le plan de classement suit une logique essentiellement pragmatique, puisque sa fonction principale est d'organiser physiquement le fonds documentaire dont dispose le centre de ressources qui l'a conçu. Il décrit donc cette réalité spécifique.

Or, un thesaurus est pensé pour représenter la réalité d'un domaine conceptuel, dans une logique toute théorique. Si le choix des domaines, l'étendue des concepts qui y sont représentés ou encore le niveau de spécificité dans leur description sont certes dictés par l'environnement d'utilisation du thesaurus, son organisation conceptuelle et les relations entre les concepts doivent, quant à eux, tendre vers la représentation la plus consensuelle possible. Sans un tel principe, aucune correspondance, aucun alignement entre thesauri ne serait envisageable, et l'idée même d'interopérabilité serait impossible à appliquer.

7.3.2.1.2 Une distinction nécessaire quand les usages évoluent

Certes, jusqu'ici, l'outil d'indexation hybride employé à la Documentation générale répondait dans les grandes lignes aux besoins et pratiques de ses utilisateurs. Mais cela correspondait à un système limité à un logiciel de SIGB, utilisé uniquement par les documentalistes internes, qui en maîtrisaient les codes. L'extension de la consultation de la base de données, et donc de l'accès au thesaurus grâce auquel elle est indexée, à l'ensemble des utilisateurs du groupe, sur le portail OPAC via l'intranet, va changer la donne. Le thesaurus devra répondre à une logique partagée par ces deux types d'utilisateurs, ce qui ne pourra se faire qu'en s'éloignant d'une logique basée sur la description d'un fonds documentaire, pour se rapprocher de la description d'une réalité partagée par l'ensemble d'une communauté : l'ensemble les collaborateurs du groupe BPCE.

7.3.2.2 Examen systématique des relations conceptuelles

Une fois le thesaurus rendu autonome vis-à-vis du plan de classement, il deviendra possible de travailler sur une structure théaurale à proprement parler. Pour ce faire, la première étape consistera à réévaluer chaque relation hiérarchique individuellement.

7.3.2.2.1 Validation des relations hiérarchiques

Chaque relation définie comme hiérarchique dans le thesaurus BPCE actuel devra être réexaminée. L'objectif est de s'assurer qu'elle appartient bien à l'une des trois catégories valides, à savoir les relations de type générique, partitif ou d'instance (Cf. 1.2.3.2). A nouveau, ce travail d'ordre conceptuel doit être fait en prenant de la distance par rapport aux ressources habituellement indexées ou consultées pour la veille. Ce type de travail nécessite un approfondissement de la compréhension des concepts abordés, et probablement l'intervention ponctuelle d'experts des domaines concernés exerçant dans l'entreprise. En cela, cette étape constituerait une avancée décisive vers un référentiel commun à l'entreprise.

7.3.2.2.2 Elargissement des associations

Les relations initialement qualifiées de hiérarchiques, qui ne satisferont pas à l'exigence évoquée ci-dessus, seront donc nécessairement redéfinies en relations associatives.

Jusqu'ici, à la fois pour satisfaire à une logique plus proche du plan de classement et pour simplifier la réflexion lors de la mission, les relations associatives ont été limitées à des concepts relevant au minimum de termes de têtes différents (*Top terms*, c'est-à-dire sous-sous-domaines). Cependant, une fois établie une structure conceptuellement plus juste, il sera possible d'étendre ce type de relation à tous les niveaux du thesaurus.

7.3.2.3 Validation et enrichissement du travail sur les descripteurs

7.3.2.3.1 S'assurer de l'unicité de la relation concept / terme préférentiel

On l'a vu, les différents domaines du thesaurus, ou micro-thesaurus, ont été retravaillés par des expertes travaillant de façon indépendante la plupart du temps. Malgré les étapes de mise en commun organisées régulièrement, les premiers travaux menés après la fin de la mission pour la finalisation du thesaurus ont montré que subsistaient çà et là certains doublons. Il s'agit plus précisément de descripteurs différents, mais qui peuvent représenter le même concept. Un travail de repérage de ces termes et concepts reste donc à réaliser, pour choisir un seul terme préférentiel parmi les deux proposés et définir le domaine auquel il appartient selon la logique propre au thesaurus BPCE. Une relation associative vers un autre concept pourra bien sûr venir remplacer le doublon supprimé dans le second domaine concerné.

7.3.2.3.2 Introduction de la post-coordination

Le thesaurus BPCE comporte un grand nombre de termes pré-coordonnés, correspondant aux nombreux concepts complexes et précis qui sont en jeu dans les domaines de la banque et la finance notamment. Néanmoins, leur usage semble excessif et l'introduction plus systématique du principe de post-coordination permettrait d'affiner la structure théssaurale. Ainsi, les termes composés de deux concepts significatifs en eux-mêmes devraient être proposés en mode post-coordonné. Dans ce cas, une note d'application vient préciser que cette notion peut être représentée en associant les deux descripteurs concernés, de la façon suivante :

Ex. : *Marketing bancaire*

NA utiliser **Marketing + Banque**

Ce type d'association ponctuelle de concepts peut se rapprocher de la notion de relations syntagmatique évoquée plus haut (Cf.1.2.3). Elle est évoquée dans la norme par la Clause 8.5 sous le titre *Representation of complex concepts by a combination of terms*⁴⁵. La Clause 7.3, *Deciding whether or not to admit a complex concept*, permet quant à elle de déterminer si un terme composé doit être intégré tel quel ou s'il convient de scinder le concept en deux

⁴⁵ « Représentation de concepts complexes par combinaison de termes ». [1, ISO, p. 49-50]

descripteurs (*splitting the concept*) [1, ISO, p. 39-43]. On se reportera à l'ensemble de la Clause 7 pour les autres aspects touchant aux concepts complexes et leur traitement [1, ISO, p.37-44].

Bien sûr, la pré-coordination reste nécessaire dans bien des cas pour représenter précisément certains concepts pointus d'un domaine (et ils sont nombreux dans celui de la banque et de la finance), ou pour des notions bien établies (Ex : « Aide au développement »). De même, si une expression se retrouve continuellement dans les documents exprimée sous une certaine forme, il est sans doute inutile de la décomposer de façon artificielle. Bref, comme l'explique M. Hudon :

« Le concepteur du thésaurus doit se fier à sa connaissance du domaine et à son propre jugement pour décider de l'opportunité de fragmenter ou non un terme composé. Il faut avant tout respecter l'usage, les habitudes et les besoins [...]. » [9, HUDON, p.121]

7.3.2.3 Repérage des synonymes non présents

Le recensement exhaustif des relations d'équivalence n'a pas, faute de temps, constitué une priorité lors du déroulement de la mission. Néanmoins, l'utilisation du thésaurus étant d'ores et déjà étendue au portail, permettant une recherche de mots-clés par l'ensemble des collaborateurs du groupe, ce travail devrait être envisagé à brève échéance. En effet, les résultats de requêtes gagneraient ainsi sensiblement en pertinence, en permettant de réduire les risques de silence qui pourraient persister sur certaines notions. En particulier, si les équivalents anglo-saxons, largement utilisés dans le domaine de la finance, ont été pris en compte dans les cas les plus courants, l'utilisation extensive de ce jargon par les utilisateurs des métiers concernés nécessiterait sans doute de rendre systématique la présence de ces termes, définis, dans le cadre d'un thésaurus monolingue, nécessairement comme synonymes. D'ailleurs, dans les cas où le terme anglophone est le plus courant, il a souvent été choisi ici comme descripteur, au détriment de son équivalent français. C'est l'usage de la communauté métier (utilisateurs du groupe, mais aussi sources utilisées pour la veille) qui a dicté ce choix, à l'encontre parfois des préconisations publiées au Journal Officiel.

7.3.2.4 Introduction de facettes pour enrichir la structure théssaurale

Nous l'avons vu, les relations associatives ont été interdites entre termes de même sous-domaine, afin d'encourager une meilleure hiérarchisation des concepts. Néanmoins, ce choix a pu créer localement des déséquilibres. A la fois pour les résoudre et pour passer à l'étape suivante de la réflexion sur la hiérarchisation des concepts, il serait profitable d'intégrer un regroupement par facette à la structure du thésaurus BPCE. Nous avons présenté plus haut les préconisations de la norme concernant l'analyse par facettes (Cf. 1.2.2.2). Voyons maintenant quelques pistes pour leur mise en place.

Le père de la théorie des facettes, l'indien S. R. Ranganathan, en a lui-même donné une définition volontairement ambiguë, dont Hjørland, entre autres, a largement critiqué la dimension idéaliste [18, MANIEZ]. Il pensait avoir abouti à une syntaxe absolue de l'énonciation des sujets, pouvant être réduite à seulement cinq éléments fondamentaux : Temps, Espace, Energie, Matière et Personnalité, dont chaque facette pouvait être considérée comme une manifestation. « *Ceux-ci contiennent sous une forme latente et non manifeste toutes les facettes possibles de tous les sujets possibles ; dans des sujets différents ils se manifestent sous des noms différents et des fonctions différentes.* »⁴⁶

Une facette n'est ni une sous-classe d'un ensemble, ni même un critère de division logique. Jacques Maniez a proposé de définir ce qu'il appelle des facettes classificatoires « catégorielles » comme un « *petit nombre de catégories universelles telles que "entités", "actions", "états", applicables à toute classification de concepts indépendamment des domaines* » [18, MANIEZ]. Il recommande de ne pas multiplier le nombre des facettes appliquées à un thesaurus, notamment lorsqu'on souhaite subdiviser des catégories jugées trop générales pour un domaine d'application, au risque de perdre en cohérence. Il cite par ailleurs les principes établis par Brian Vickery, permettant d'établir de façon expérimentale les catégories fondamentales d'un domaine étroit, par réduction progressive du spécifique au générique. On choisira ainsi des facettes de type universel ou spécifique, c'est-à-dire des facettes fonctionnelles ou structurelles d'un domaine particulier, selon les besoins et l'objectif du thesaurus.

Dans l'exemple ci-dessous (FIGURE IV), nous avons essayé de retravailler le champ sémantique « Activité bancaire » à l'aide de facettes extraites des catégories générales proposées par la norme. Il est certain qu'une réflexion plus poussée serait nécessaire afin d'identifier les facettes qui seraient le plus pertinentes pour ce thesaurus spécialisé en particulier. Néanmoins, on peut déjà imaginer l'avantage qu'une telle répartition représente, en termes de recherche d'information notamment. Les éléments présents dans le thesaurus ont ici seulement été redistribués selon ce principe. Mais on voit bien dans cet exemple que cette simple opération suffit à nourrir la réflexion, en guidant vers des concepts qui pourraient enrichir ce champ sémantique.

⁴⁶ Shiyali Ramamrita Ranganathan, *Classification, codification et appareillage de recherche*, discours devant l'UNESCO, Paris, 30 juin 1950, p. 21 (publié en français le 21 décembre 1951). http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=133325&set=4B1BA8F9_1_463&database=qed&qp=0&mode=e&lin=1&ll=f

TT Activité bancaire	TT Activité bancaire
TS1 Concurrence bancaire	(Acteurs)
TA Concurrence	Banquier
TS1 Opération de banque	Concurrence bancaire
EP <i>Produit bancaire</i>	TA Concurrence
TA Crédit	(Actions)
TA Assurance des opérations bancaires	Distribution des produits et services bancaires
TS2 Compte de dépôt	Opération de banque
TS3 Compte rémunéré	TA Crédit
TA Rémunération de l'épargne	TA Assurance des opérations bancaires
TS2 Convention de compte	TS1 Intermediation financière
TS2 Intermediation financière	EP <i>Intermediation bancaire</i>
EP <i>Intermediation bancaire</i>	TA Désintermediation
TA Désintermediation	TS1 Opération déplacée
TS2 Opération déplacée	TS1 Opération internationale
TS2 Opération internationale	(Document)
TS1 Service bancaire	Convention de compte
TS2 Affacturage	(Lieux)
TA Société d'affacturage	Agence bancaire
TS2 Coffre-fort	EP <i>Guichet bancaire</i>
TS1 Service bancaire de base	Banque à distance
TA Droit au compte	EP <i>Banque à domicile</i>
TS1 Technique bancaire	(Moyens)
Distribution des produits et services bancaires	Guichet automatique de banque
TS1 Agence bancaire	EP GAB
EP <i>Guichet bancaire</i>	EP Libre-service bancaire
TS1 Banque à distance	TA Distributeur automatique de billets
EP <i>Banque à domicile</i>	Technique bancaire
TS1 Guichet automatique de banque	(Produits)
EP GAB	Compte de dépôt
EP Libre-service bancaire	TS1 Compte rémunéré
TA Distributeur automatique de billets	TA Rémunération de l'épargne
Métier de la banque	Service bancaire
TA Métier	TS1 Affacturage
TS1 Banquier	TA Société d'affacturage
	TS1 Coffre-fort
	TS1 Service bancaire de base
	TA Droit au compte

FIGURE IV : ESSAI D'APPLICATION DE FACETTES AU THESAURUS BPCE

8 Consolider et valoriser : une organisation apprenante au cœur de l'entreprise

8.1 Retour d'expérience : quels bénéfices pour l'équipe du service Documentation générale ?

Lors de la phase de bilan du projet, un *feedback* global a été demandé à chaque membre de l'équipe, afin :

- d'évaluer la gestion et le déroulement du projet ;
- d'évaluer le livrable obtenu et son utilisation ;
- d'identifier les compétences déployées et les savoir-faire acquis au cours du projet.

Chacune s'est montrée heureuse d'avoir participé à cette expérience et fière du résultat obtenu, malgré ses imperfections – et la conscience du travail restant à accomplir ! C'est sur le dernier point, en forme d'autoévaluation, qu'il nous semble important de revenir maintenant.

8.1.1 Renforcement de l'« intelligence informationnelle »

Les chercheurs en sciences de l'information et de la communication prônent aujourd'hui de plus en plus souvent « *une différenciation entre culture de l'information de l'utilisateur et culture informationnelle des professionnels de l'information* » [25, GARDIES]. En introduction de l'ouvrage collectif *Le développement de l'intelligence informationnelle*, rédigé sous sa direction [24, BOISVERT, p.11-13], Danielle Boisvert définit cette notion, en parallèle à celle d'intelligence émotionnelle, développée en psychologie par Daniel Goleman. Nous nous appuyons sur des éléments de sa définition pour illustrer certaines compétences indispensables aux professionnels de la documentation, et comprendre comment elles peuvent être enrichies grâce à la (re)conception de thesaurus.

8.1.1.1 Formulation du besoin de l'utilisateur

« En milieu documentaire, l'utilisateur doit être capable d'exprimer ses besoins en information et nous, les professionnels, sommes là pour lui faciliter la tâche. »

En remplaçant le terme d'émotion, présent dans le concept original, par celui de besoin d'information, D. Boisvert laisse intacte la dimension hautement délicate de leur expression, qui n'a rien d'instinctif. Formuler une requête lors d'une recherche d'information est une opération qui requiert, au minimum, une familiarisation initiale avec les outils documentaires. Pour aller au bout de cette recherche, l'aide de professionnels de ces systèmes est bien souvent indispensable. Mais la compréhension profonde d'un besoin d'information est tout aussi complexe que celle des émotions que ressent l'autre. Formuler le besoin d'information d'un utilisateur demande de

l'empathie, ou son équivalent en entreprise : la compréhension profonde de son domaine d'activité et de ses besoins inhérents. Formulé ainsi, on comprend sans peine les bénéfices indéniables qu'a pu retirer de ce projet l'équipe de la DG dans ses activités quotidiennes, qu'il s'agisse des recherches ponctuelles ou de la confection des veilles documentaires.

8.1.1.2 « Auto-motivation » face au changement

« Les changements dans nos milieux sont grands et l'utilisateur ainsi que nous-mêmes sommes constamment mis en situation de nous réajuster pour demeurer performants. [...] le monde documentaire se transforme à un point tel que nos rôles traditionnels sont remis en question et pourraient peut-être ne plus exister. »

Le service de veille et documentation de BPCE s'est trouvé en effet confronté à de nombreuses transformations ces dernières années, les bouleversements organisationnels venant s'ajouter aux changements technologiques dans l'activité, sémiologiques et structurels concernant ses domaines de connaissance. L'adaptation des pratiques et des missions s'est toujours faite de la façon la plus efficace possible, en fonction des moyens alloués par l'entreprise, avec pour objectif permanent la préservation de la qualité de service. A cet égard, les documentalistes ont démontré une implication et un professionnalisme irréprochables, y compris lorsqu'elles avaient suivi une formation initiale différente.

Mais il est souvent difficile de prendre la mesure du changement lorsque l'on est tributaire de tâches quotidiennes nombreuses et répétitives. Participer à un projet, réalisé en équipe et centré sur des techniques essentielles au métier de la documentation, l'indexation et l'organisation des connaissances, est l'occasion d'en prendre conscience. Cela correspond parfaitement à la nécessité de se mettre en « *situation permanente de réapprentissage avec humilité* » évoquée par C. Volant [23, VOLANT]. Les documentalistes ont ainsi entamé une réflexion sur leurs pratiques d'indexation, en intégrant le paramètre nouveau que constitue l'ouverture de l'OPAC sur l'intranet du groupe. De telles occasions de se recentrer sur l'évolution de son contexte de travail et son propre métier – lorsqu'on est au service des métiers de ses utilisateurs – est certes un facteur important dans l'entretien de sa propre motivation.

8.1.2 Compétences développés et capitalisés

Les tâches effectuées durant le projet ont également été l'occasion d'entretenir et de développer des compétences et des savoir-faire indispensables aux activités documentaires. Les documentalistes ont ainsi pointé comme compétences générales particulièrement mobilisées :

- Rigueur
- Capacités d'analyse et de synthèse
- Travail collaboratif

- Organisation du travail
- Recherche d'information et même « curiosité »

Toutes ces compétences sont bien sûr essentielles, mais pas uniquement dans les métiers de l'Infodoc. Nous aimerions maintenant faire le point sur les savoir-faire opérationnels qui ont été développés, et qui constituent une montée en compétences profitable au service collectivement, et au management en particulier pour la réussite de ses missions.

8.1.2.1 Compétences documentaires : révision des classiques et acquisition de bonnes pratiques

Différents aspects du travail sur le thesaurus ont conduit les membres de l'équipe à s'interroger sur leurs pratiques quotidiennes, et à entamer une réflexion visant à améliorer leur qualité. Evidemment, chacune a évoqué une importante montée en compétences sur la conception de thesaurus, le travail terminologique et les relations sémantiques. La prise en compte des préconisations de la nouvelle norme ISO 25964 a permis d'approfondir les connaissances de celles ayant déjà une bonne expérience dans ce domaine, en particulier la responsable du service, à l'origine du précédent thesaurus. Deux éléments sont à noter en corollaire de ce premier point. D'une part, l'équipe a pris conscience de l'intérêt de faire vivre, d'entretenir et d'enrichir constamment ce langage, qui constitue une ressource essentielle pour leur activité (Cf. 8.2.1). D'autre part, la plupart ont été amenées à repenser leur pratiques d'indexation, avec une meilleure compréhension des enjeux et des conséquences de leurs choix pour la recherche d'information. La nécessité de mettre en place un cadre de bonnes pratiques est ainsi apparue comme une conséquence logique de la démarche de refonte du thesaurus.

8.1.2.2 Compétences métier : approfondissement des connaissances sur les domaines de veille

Le second aspect notable de cette trajectoire de montée en compétence porte sur la spécificité de l'activité de ce service, essentiellement occupé par la veille comme nous l'avons vu (Cf. 3.2.2). Dès l'origine, le projet était sous-tendu par des attentes d'ordre managérial de la part de son commanditaire. En effet, le rôle confié au chef de projet se limitait à organiser le travail et à « faire travailler les documentalistes », en leur servant de guide pour la mise en œuvre. Mais faire travailler chacune individuellement sur ses thématiques de veille habituelles n'était pas seulement une démarche logique et efficace. Le but était avant tout de leur permettre de prendre le temps d'approfondir leurs connaissances de ces domaines. De leur fournir une occasion de le faire, dont elles se sont saisies avec plaisir et empressement. Elles se sont d'ailleurs tant prises au jeu qu'elles regrettèrent *a posteriori* le peu de temps accordé pour le mener à bien.

On citera donc parmi les points de montée en compétences les plus notables pour la pratique de la veille documentaire développés au cours du projet :

- la validation des sources d'information pertinentes et l'enrichissement du répertoire des ressources sur le Web ;
- une meilleure connaissance du public, grâce à une réflexion sur les activités des utilisateurs et sur la culture d'entreprise de BPCE, afin de circonscrire les domaines à représenter dans le thesaurus et la profondeur de spécification ;
- une maîtrise accrue des concepts d'un domaine grâce au travail sur les définitions et sur leur hiérarchisation ;
- la structuration des connaissances et leur mise à jour pour certains domaines qui ont beaucoup évolué.

8.2 Quel ROI pour le management ?

On vient d'évoquer les ambitions managériales qui accompagnaient la conduite de ce projet. Deux autres éléments de « retour sur investissement » sont déjà identifiables, peu après l'issue du projet. Au niveau méthodologique, le service a naturellement entamé une réflexion sur la politique de maintenance qui doit être associée au nouveau thesaurus. Au niveau organisationnel, ce projet a développé certains aspects d'un mode de management adapté à une organisation apprenante.

8.2.1 La mise en place d'une politique de gestion du thesaurus

L'implication de l'équipe, les connaissances acquises ou renforcées sur le sujet, la réflexion mise en œuvre sont autant de bénéfices liés au projet qui devraient faciliter l'adoption d'une politique simple, efficace et connue de tous concernant la gestion du nouveau thesaurus. Celle-ci constitue un point essentiel pour permettre et garantir son maintien et son évolution.

8.2.1.1 Une réflexion commune et continue

Des rencontres périodiques sur le sujet devront être maintenues quoi qu'il arrive, selon le besoin/la charge de travail. La documentaliste en charge de la gestion du thesaurus s'est proposée de les organiser une semaine sur deux dans un premier temps, en-dehors des réunions de service, pour maintenir un focus spécifique. Il s'agit de celle qui n'avait pu réaliser la première tentative de refonte du thesaurus faute de moyen. Pour le management, sa forte implication dans le projet a permis de confirmer sa maîtrise théorique de la conception de thesaurus et de valider sa compétence pour l'entretien du nouveau langage. Elle aura en outre démontré son intérêt et sa motivation à assurer l'implémentation complète et le maintien du thesaurus selon des critères de qualité élevés. La confirmation des qualifications de cette ressource en interne est à lui seul un ROI non négligeable.

8.2.1.2 Généraliser l'utilisation des candidats descripteurs

L'utilisation des candidats descripteurs devra être généralisée, assortie d'une gestion rigoureuse et d'un dialogue au sein de l'équipe autour de cette pratique. Une note de procédure *ad hoc* a été produite pour accompagner les documentalistes dans cette nouvelle pratique.

En effet, l'entretien d'un thesaurus exige un travail constant de mise à jour et d'enrichissement lexical et structurel. Ce travail passe essentiellement par l'identification de termes absents du thesaurus, mais nécessaires au travail d'indexation, ainsi que des relations sémantiques qui lui correspondent. Un candidat-descripteur est un terme proposé pour devenir descripteur au sein du thesaurus, mais qui reste en attente de validation. Ce statut permet de tester l'utilisation d'un terme à l'indexation – puisqu'il est proposé dans la liste des mots-clés, précédé d'une * dans JLB-Net – sans modifier initialement la structure du thesaurus. Lorsqu'un utilisateur du thesaurus constate qu'un terme, qui lui serait nécessaire pour indexer une ressource, est absent du thesaurus, plusieurs cas se présentent à lui :

- il est possible de décrire le concept manquant par la combinaison de plusieurs concepts existants dans le thesaurus. Dans un premier temps du moins, cette solution est à privilégier ;
- le concept à décrire ne possède aucun terme équivalent et ne peut être représenté autrement à l'aide du thesaurus. Dans ce cas, et après avoir identifié (au besoin par une recherche) le terme préférentiel à proposer, il est opportun d'ajouter un candidat-descripteur.

Une séance de travail dédiée à la validation des candidats-descripteurs devra être mise en place de façon régulière et suivie. Même si peu de candidats sont à traiter, il est préconisé de conserver ce rendez-vous afin que l'équipe reste consciente de l'importance d'entretenir le thesaurus. Chaque candidat-descripteur est présenté à l'ensemble des documentalistes, pour évaluer sa pertinence, son utilité (en regard des notices effectivement indexées avec ce terme) et s'assurer que le concept retenu n'est pas déjà représenté dans le thesaurus. Lorsqu'un candidat est validé, il s'agit ensuite d'examiner les relations qu'il convient de créer pour décrire son environnement sémantique : ses synonymes non-descripteurs (EP) et ses relations hiérarchiques, en l'associant éventuellement à un terme générique et/ou à un ou plusieurs termes spécifiques. Enfin, chaque spécialiste des différents domaines du thesaurus étudie l'opportunité de proposer un terme associé dans les domaines connexes.

8.2.1.3 Gestion des outils logiciels associés

L'implémentation des changements dans le progiciel JLB-Net Bibliothèques par les documentalistes de la DG est une problématique qui a été évoqué plus haut (Cf. 7.2.3.1). La préconisation initiale prévoyait que le nouveau thesaurus soit complet et opérationnel sous 3 mois

maximum – compte-tenu d'un volume de reprise inégal suivant les documentalistes et des tâches quotidiennes incompressibles. Au cours de cette étape, un travail parallèle doit également être mené sur les mots-clés issus de DIP, qui n'ont pu être intégrés à la mission faute de temps. Cependant, cette étape a pris du retard du fait de l'émergence de tensions interpersonnelles au sein de l'équipe, en raison de la division des responsabilités concernant le thesaurus d'une part, l'outil informatique de l'autre. Cette situation a contraint la responsable à intervenir afin de trouver un terrain d'entente et imposer une méthodologie. Ce développement a confirmé sa conviction que la gestion de tels projets transversaux, qui se superposent aux activités « normales », nécessitent d'être pilotés intégralement par une ressource externe au service, au positionnement spécifique. Dans ces conditions, la prise en main et le déploiement de l'outil complémentaire ThManager (ou l'identification d'un autre logiciel payant si nécessaire) a du être repoussée.

8.2.1.4 Adapter le plan de classement

Dernière conséquence logique pour la politique documentaire du service : mettre à jour le plan de classement. Malgré la contrainte évoquée, des arbitrages ont été indispensables pour permettre la cohérence du thesaurus. Il s'agira donc d'intégrer ces modifications, ainsi que les changements appliqués dans les termes adoptés pour les champs sémantiques, notamment. Un sous-domaine est également à ajouter, celui concernant le développement durable. Cette mise à jour pourra être l'occasion de procéder à une évaluation du fonds documentaire, d'en analyser la composition, la répartition des ressources, même si un désherbage a été réalisé récemment, suite à la fusion.

8.2.2 Manager pour développer une organisation apprenante

Dans son article intitulé *Réseau documentaire... organisation apprenante*, paru en 1998 [23, VOLANT], Christiane Volant développe une vision des modes de management adaptés à l'évolution de ces systèmes complexes qui n'a pas pris une ride. Sans aucun doute, un service de veille et documentation en entreprise fait partie de ces systèmes dynamiques qui ne peuvent « être managé[s] qu'avec des principes facilitant la créativité, l'émergence de savoirs nouveaux, le développement de comportements d'échanges permanents de points de vue. »

8.2.2.1 Encourager l'apprentissage collectif

S'adapter à un contexte en mutation constante, faire face aux exigences croissantes des systèmes d'information et des utilisateurs exige du système de s'adapter collectivement en s'appuyant sur les apprentissages individuels de chacun de ses membres. Etre en situation d'apprentissage permanent rend l'individu de plus en plus apte à apprendre. Pour encourager un apprentissage collectif efficace, C. Volant a déterminé cinq pivots sur lesquels peut s'appuyer le management d'une organisation apprenante :

- une démarche active, permanente et progressive ;
- l'instauration d'un climat de confiance ;
- la fluidité de la circulation de l'information ;
- l'organisation en réseau ;
- l'appui sur les ressources internes, par le repérage des personnes-ressources.

Ces principes décrivent tous une organisation dynamique, favorisant la synergie de ses membres grâce à leur autonomisation et la valorisation des compétences individuelles. Autant de principes que nous avons tenté de mettre en œuvre à travers la gestion de ce projet.

8.2.2.2 Des méthodes de management adaptées à ces systèmes

Justement, le premier mode de management préconisé par C. Volant est le pilotage par projet plutôt que par structure. Selon elle :

« Les structures sont bien sûr indispensables au fonctionnement, mais elles prennent fréquemment une place prioritaire et intangible, devenant alors une fin en soi alors qu'elles ne sont que des moyens d'organisation au service de missions, de projets, etc. »

Manager par projet, c'est instaurer des responsabilités réellement partagées et constamment redistribuées, donner au système la capacité de se transformer, encourager les médiations humaines et les relations d'échange, d'entraide. Ce sont ces situations sans cesse renouvelées qui, en favorisant les capacités d'adaptation, encouragent l'autonomisation des acteurs du système. C'est le second mode de management que C. Volant encourage :

« Ainsi, chaque unité constituante peut avoir une grande autonomie lui permettant de faire émerger les solutions adaptées à ses problèmes et aux conditions changeantes de fonctionnement, pour peu que la cohérence du système soit garantie par un ensemble de règles stimulant la solidarité entre les unités. »

Dernier élément, pour réellement adopter cette méthode constructiviste de conception et d'action, il est nécessaire de parvenir d'abord à modifier les comportements. En effet, la confrontation des représentations individuelles engage des qualités d'écoute, de confiance, d'évolution personnelle, de mobilisation des savoirs. Tous ces savoir-être doivent être développés au service de l'apprentissage collectif. Enfin, une organisation apprenante exige du management une capacité à adapter en permanence sa stratégie pour obtenir une implication sincère de ses membres.

Il nous semble que le projet de refonte du thesaurus a permis de créer pour la DG de BPCE, avec sa composition et ses préoccupations actuelles, un précédent allant dans le sens de ces

préconisations. Une dynamique de base a été insufflée, encouragée par une intervention externe – et par conséquent neutre, que le management pourra mettre à profit et encourager grâce au suivi des opérations d'implémentation, de gestion et de maintenance du thesaurus.

8.3 Impacts organisationnels : effets sur l'image de marque du service DG au sein de l'organisation BPCE

8.3.1 Gains immédiats

8.3.1.1.1 Valorisation des compétences documentaires auprès de la direction

Ce projet a été l'occasion de communiquer auprès du directeur de la DISG, qui chapeaute la DG, sur la réalité des pratiques du service et sur les compétences documentaires à l'œuvre. Dans ce but, la note de cadrage lui a été remise en début de projet, et il a été invité à assister à la présentation du bilan de projet. Ce fut l'occasion d'insister sur ce qu'est un thesaurus et son utilité pour les différentes missions du service, sur les savoir-faire mis en œuvre par ces travaux et leur consolidation, enfin sur la connaissance irremplaçable des utilisateurs et de l'organisation que possèdent les veilleur-documentalistes.

Ce type d'action de valorisation très ciblée auprès d'une personne décisionnaire peut se révéler bénéfique à très court terme. Au moment de négocier le budget de la Documentation, il est toujours préférable pour le responsable du service de s'adresser à un interlocuteur déjà sensibilisé à ce métier, en particulier lorsqu'il s'agit d'une entreprise où ce dernier n'est que peu pris en considération. Ce point est d'autant plus déterminant pour un centre de ressources dépendant d'un organe central, dont la vocation est de fédérer les moyens mis à disposition de l'ensemble du groupe. Bien sûr, à long terme, un tel allié peut se révéler indispensable pour la pérennité même du service, dans un contexte où la réduction des budgets et effectifs alloués aux services de documentation traditionnels se généralise, y compris dans les grands groupes.

8.3.1.1.2 Valorisation du fonds documentaire proposé aux utilisateurs

Le déploiement du nouveau thesaurus dans l'OPAC, dès son implémentation dans JLB-Net, permettra de proposer un vocabulaire de référence, pertinent et cohérent, couvrant l'ensemble des activités du groupe, aux utilisateurs de l'application via l'intranet. Ce nouveau service, dès que les tests fonctionnels l'auront validé et qu'il aura été communiqué sur son existence, ne manquera pas d'attirer un regard nouveau sur les ressources offertes aux collaborateurs de BPCE. Or, nous avons vu l'importance d'un langage documentaire bien conçu et à jour pour la qualité et l'efficacité du repérage de l'information.

8.3.2 Capitalisation à plus long terme

Le travail effectué pour mettre à jour ce thesaurus, et encore plus une fois qu'il aura pu être amélioré en fonction des préconisations avancées (Cf. 7.3.2), constitue selon nous une base, qui pourrait être mieux exploitée au profit de l'entreprise, en l'étendant à d'autres services essentiels pour la circulation de l'information au sein de l'entreprise.

8.3.2.1 Pistes de réflexion pour l'utilisation future du thesaurus au sein du réseau documentaire du Groupe BPCE

8.3.2.1.1 *Déclinaison du thesaurus en taxonomie pour enrichir le portail documentaire*

Le travail sur le thesaurus peut être étendu pour aller plus loin dans la démarche de revalorisation des ressources du service Documentation générale et en permettre un accès toujours plus facile. Ainsi, extraire une taxonomie, basée à la fois sur les domaines et champs sémantiques du thesaurus et sur les statistiques d'indexation, permettrait de proposer une navigation par facettes, encourageant l'exploration du catalogue. Ces facettes pourraient dans un second temps être pondérées en fonctions des requêtes de l'utilisateur, afin de lui permettre d'affiner sa recherche selon des critères pertinents et accéder plus facilement aux ressources adéquates.

Un autre service qui pourrait être proposé grâce au portail est la mise en avant des derniers ouvrages acquis, éventuellement sur la base de profils plus ou moins personnalisés. La mise en place d'une DSI⁴⁷ ou de flux RSS⁴⁸, concernant l'actualité du service et des ressources proposées, serait également facilitée par l'application des facettes à la création de ces profils. Ces deux vecteurs de communications, s'ils étaient généralisés auraient un impact important sur la présence et l'image du service et de la fonction documentaire au sein de BPCE⁴⁹.

8.3.2.1.2 *Adoption par les autres services de documentation du groupe pour indexer des ressources unifiées*

Le nouveau portail JLB-Net est déployé sur l'intranet du groupe et consultable par l'ensemble des collaborateurs. L'objectif principal était d'ouvrir ce service aux 90 % du groupe qui ne bénéficient pas de services documentaires en interne. Pour les quelques services documentaires qui subsistent, l'adoption du thesaurus BPCE permettrait d'unifier l'indexation de l'ensemble des fonds documentaires et de proposer des recherches unifiées sur les différentes bases de données. En pratique, plusieurs solutions sont possibles en fonction des langages documentaires utilisés selon les bases : alignement entre les vocabulaires, en fonction de l'interopérabilité, réindexation automatique du fonds... Cela nécessiterait, là aussi, que le réseau sémantique et la structure aient

⁴⁷ Diffusion sélective de l'information

⁴⁸ *Really Simple Syndication* : format de données utilisé pour la syndication de contenus Web.

⁴⁹ Toutes ces préconisations sont bien sûr soumises à la faisabilité technique et aux fonctionnalités proposées par JLB.

été en amont consolidées pour le thesaurus BPCE, et probablement l'intervention d'une assistance technique. La plupart de ces services étant de taille réduite, cela ne devrait pas être trop lourd.

Ainsi, les utilisateurs accéderaient à des résultats de requêtes étendus, permettant de réduire le silence sur certains sujets, qui ne manque pas de concerner les services disposant de très peu de ressources. Le catalogue leur indiquerait alors les ressources présentes sur place et celles disponibles au siège de BPCE SA. Des pratiques de prêt entre services pourraient se développer, mais la solution la plus efficace serait le développement de l'acquisition de droit d'accès en ligne aux bases de ressources numériques (comme c'est déjà le cas pour les études économiques DAFSA et Xerfi). Dans cette optique, il pourrait être envisagé de fédérer les budgets des différents services de documentation intéressés, qui délègueraient la négociation et la gestion de ces ressources au service Documentation générale de l'organe central.

8.3.2.1.3 Impact organisationnel pour la valorisation des métiers de la documentation

Au niveau organisationnel, une telle initiative serait l'occasion de mettre le focus sur la fonction documentaire et les différentes formes qu'elle prend en fonction des structures où elle s'exerce. Elle permettrait, en reprenant les ambitions du Club métier évoqué plus haut (Cf. 3.1.2.2), de mieux fédérer les professionnels de la documentation dispersés dans le groupe autour de bonnes pratiques communes. Ce mouvement serait en outre un moyen de mettre en valeur des fonctions souvent mal comprises et peu valorisées au sein de l'entreprise, en communiquant sur les produits et services rendus, les savoir-faire spécifiques, et de pérenniser ainsi des emplois trop souvent mis en péril.

Cette démarche gratifiante devrait en outre limiter l'inévitable résistance au changement de la part de petites structures, souvent composées d'une seule personne, attachées à leurs usages. Un travail commun pour l'intégration au thesaurus des aspects spécifiques à chaque organe, enfin, serait la dernière brique pour engager un travail réellement collaboratif, dont l'ensemble des collaborateurs du groupe tireraient grand profit.

8.3.2.2 Quelques recommandations pour étendre sa valorisation à l'organisation

8.3.2.2.1 Application aux ressources de l'intranet du groupe pour optimiser le moteur de recherche d'entreprise

L'intranet du Groupe BPCE est un outil réellement tentaculaire, sur lequel chaque métier, chaque service dispose de son, voire de ses propres espaces. Il constitue une jungle, dans laquelle il est bien difficile de se retrouver. De ce fait, il est tentant de se contenter d'utiliser les niches que l'on connaît et maîtrise, et de passer à côté d'une mine de renseignements. Le moteur de recherche d'entreprise utilisé au sein de cet intranet s'appuie sur une recherche en texte intégral.

Il génère de fait un bruit⁵⁰ considérable et le taux de précision⁵¹ est très faible. La mise en place d'un langage contrôlé unique sur lequel pourrait s'appuyer ce moteur, commun à toutes les applications, constituerait donc un gain de temps non négligeable.

De la même façon, ce vocabulaire pourrait être employé pour générer les métadonnées de tous les documents produits dans les différents outils collaboratifs et applicatifs métiers utilisés dans le groupe. Toutes ces ressources seraient ainsi indexées automatiquement avec le même référentiel, permettant au moteur de recherche d'entreprise de puiser dans ces métadonnées sémantiques à forte valeur ajoutée. La quantité exponentielle de données non structurées émanant de l'entreprise représente aujourd'hui l'une des problématiques essentielles rencontrées par les directions des systèmes informatiques, à laquelle les langages documentaires traditionnels apportent une solution pertinente. Et bien sûr, cela représente un ROI considérable pour l'organisation, considérant ce volume d'informations potentiellement noyées dans l'« infobésité » croissante, et pourtant stratégiquement essentielles.⁵²

8.3.2.2 *Exprimer la culture de l'entreprise : le thesaurus BPCE comme base de travail pour concevoir un référentiel commun*

Le groupe BPCE est né en 2009 de la fusion de deux grands groupes bancaires fortement marqués historiquement et aux valeurs affirmées. Cette construction n'a pu se faire sans heurts et des effets résiduels de la forte résistance au changement qui s'est exprimée subsistent aujourd'hui encore. Ainsi, la culture d'entreprise BPCE (violette) peine à s'imposer face aux cultures implantées de longue date de la Caisse d'Épargne (rouge) et de la Banque Populaire (bleue). On m'a d'ailleurs prévenue dès la première semaine que je risquais de constater « *un déficit de culture violette pour fédérer les cultures rouge et bleue* ». Cela ne manque pas d'amplifier les tensions, chacun pouvant se réclamer de sa « culture d'origine » au moindre différent. Malgré le discours officiel et la volonté d'unité affichée, cet état de fait s'exprime d'ailleurs clairement dès la page d'accueil de l'Intranet. En effet, chacun disposait encore récemment, outre les espaces communs à tous « Mon groupe » et « Mon entreprise », d'un espace CE ou BP, selon son organisme d'origine⁵³. Outre ces deux branches essentielles, le groupe compte de nombreuses filiales, à commencer par la plus grande d'entre elles, la banque d'affaires Natixis. Cette complexité organisationnelle, la longue histoire et les cultures spécifiques de ses actionnaires

⁵⁰ « Ensemble de documents non pertinents trouvés en réponse à une question, lors d'une recherche d'information. » ADBS, Vocabulaire de la documentation

⁵¹ « Mesure de l'efficacité d'un système d'indexation et de recherche établie à partir du ratio entre le nombre de documents pertinents trouvés lors d'une recherche documentaire et le nombre total de documents trouvés en réponse à la question. C'est un indicateur de mesure du bruit. » Idem.

⁵² Pour en savoir plus à ce sujet, les enjeux et les solutions techniques, on pourra consulter le très complet : VUILLEQUEZ Jean-Yves. *Le moteur de recherche d'entreprise : quels enjeux organisationnels et technologiques ?*. Mémoire professionnel INTD, Titre I, Ingénierie documentaire, 2013.

⁵³ Ces deux espaces ne sont devenus accessibles à tous, quelle que soit leur origine dans le groupe, qu'en février 2013.

principaux font donc de la construction d'une culture d'entreprise forte et unitaire, propre à BPCE, un enjeu capital aujourd'hui encore, plus de quatre ans après la fusion.

Or, on a évoqué l'outil puissant que pouvait représenter un référentiel commun, tel qu'un thesaurus d'entreprise, dans l'incarnation de l'identité et des valeurs d'une organisation (Cf. 2.3.2). Nous avons vu plus haut que le SOC conçu par la DG de BPCE, notamment à travers son langage d'indexation, répondait à la logique du pragmatisme antifondationaliste (Cf. 7.3.1.2). Un des points faibles de cette démarche découle justement de son absence de fondement éthique pour justifier sa représentation [20, TENNIS]. Dans un tel contexte, cet aspect peut au contraire représenter un atout majeur, puisqu'à un positionnement éthique de type moral peut tout à fait se substituer une interprétation du réel dictée par la culture d'entreprise que l'on souhaite promouvoir. La valeur du SOC, selon cette approche, pourrait dès lors se développer selon deux axes : l'utilité informationnelle pour les usagers du SI, l'utilité politique au niveau organisationnel.

Partant de ce constat, il ne saurait être que profitable à BPCE d'entamer une réflexion sur l'opportunité de mener une démarche de ce type, dont l'équipe du service Documentation générale serait le pivot. En effet, leurs compétences documentaires, comme la connaissance des différents métiers qui composent son public, font de ses membres les relais et intermédiaires privilégiés des nombreux experts nécessairement impliqués dans un tel projet. Inclure les professionnels de l'Infodoc dans les projets à dimension organisationnelle, portant sur l'usage des langages contrôlés dans les nouvelles technologies, est justement l'un des objets de la seconde partie de la norme ISO 25964, parue en mars 2013.

9 Les référentiels terminologiques aujourd'hui : évolution et perspectives pour l'entreprise

A partir des préconisations réunies dans la seconde partie de la norme ISO 25964, dédiée à l'interopérabilité entre les vocabulaires [2, ISO], nous aimerions pour conclure ce mémoire, faire un point sur les technologies liées à ce qu'on nomme le « Web sémantique » et leur utilité en entreprise.

9.1 ISO 25964 : une nouvelle norme pour encourager la transition vers le Web sémantique

Après une période où les avancées de la recherche en texte intégral sur le Web et de l'indexation collectives semblaient avoir enterré les référentiels traditionnels, ils reviennent aujourd'hui sur le devant de la scène, s'affirmant plus nécessaires que jamais.

« Le succès du Web sémantique [...] dépendra de la capacité des ordinateurs à agir de façon coordonnée afin que des agents intelligents puissent extraire et manipuler l'information de multiples sources en réseau. » [5, DEXTRE CLARKE]

9.1.1 Une norme adaptée à l'évolution des pratiques

Un nombre croissant d'environnements, qu'il s'agisse des bibliothèques, des centres de documentation, ou plus largement au sein des entreprises, sont confrontés à la nécessité de partager leurs référentiels terminologiques et d'unifier leurs ressources. Des fonds indexés avec différents vocabulaires contrôlés, à différents niveaux de structuration, doivent être accessibles simultanément à la recherche via une interface unique. Devant l'impossibilité de réindexer de tels volumes de documents à l'aide d'un même vocabulaire – opération qui devrait de surcroît être recommencée pour chaque nouvelle source d'information – la question de l'interopérabilité entre les vocabulaires en place devenait cruciale. Or, la solution étant souvent une partie du problème, c'est bien l'explosion des techniques liées à l'Internet qui a encouragé la standardisation des langages informatiques, avec le soutien du W3C, qui permettent aujourd'hui de mettre largement en œuvre cette interopérabilité. Sans rentrer dans les détails techniques la norme présente les principaux types de référentiels que l'on trouve aujourd'hui sur le Web, et propose un cadre commun pour les aligner avec les thesaurus.

Car s'il est un constat que ce mouvement permet, c'est bien celui de l'omniprésence des thesauri, souvent cités par les acteurs du secteur comme le plus représentatif des référentiels utiles pour

le Web de données. Un sondage intitulé *Do controlled vocabularies matter ?*⁵⁴, mené au printemps 2011 par Semantic Web Company (éditeur du système de gestion de thesaurus PoolParty) auprès de 158 professionnels du secteur de l'Infodoc en témoigne. Si la qualité du sondage peut être critiquée (méthode, panel), ses conclusions n'en sont pas moins intéressantes. Elles prouvent le succès des vocabulaires contrôlés, utilisés par plus de 85 % des personnes interrogées, dont près de la moitié depuis au moins cinq ans, bien plus longtemps dans le secteur public. Les langages les plus utilisés sont (plusieurs réponses possibles) les taxonomies (72,6%), les ontologies (63%) et les thesauri (59,3%) – les deux premiers étant surtout en cours d'élaboration. Un autre constat notable, qui montre bien l'intérêt pour ces questions, est l'importance accordée au format SKOS : plus des ¾ le jugent pertinent, voire très important.

9.1.2 Encadrer et faciliter l'interopérabilité entre les vocabulaires

Le principal objet de cette extension de la norme est d'apporter des préconisations de schémas d'alignement ou de correspondance (*crosswalk*) entre les thesauri et d'autres vocabulaires. S'appuyant cette fois sur les compétences informatiques nécessaires pour la mise en œuvre des langages contrôlés, elle utilise le terme de « vocabulaire », défini dans la première partie (Cf. 1.1.2).

9.1.2.1 Objectif de la norme

La norme l'exprime ainsi : « *Pour la recherche d'information, le but principal de l'interopérabilité entre vocabulaires est de permettre à une expression formulée à l'aide d'un vocabulaire d'être convertie en (ou complétée par) une expression correspondante d'un ou plusieurs autres vocabulaires.* » [2, ISO, Clause 5.1, p. 16] Il peut s'agir d'une requête ou des métadonnées attachées à un document, et les vocabulaires en question peuvent être exprimés dans des langues naturelles différentes. L'interopérabilité permet également d'augmenter un vocabulaire ou même de le fusionner avec d'autres. Dans tous les cas, la première étape consiste à cartographier les concepts afin de pouvoir les faire correspondre un à un et les rendre interchangeables. Cette opération (*mapping*), qui présente des difficultés à la fois théoriques et techniques, est cruciale pour la réussite de l'alignement et la qualité des résultats obtenus. C'est pourquoi cette norme apporte des éléments pour concevoir des modèles structurels permettant d'établir ces correspondances entre les différents types de vocabulaires.

9.1.2.2 Modèles structurels proposés par la norme

La Clause 6 propose trois modèles génériques sur lesquels baser son schéma d'alignement :

⁵⁴ « Les vocabulaires contrôlés comptent-ils ? » Accessible en ligne : http://issuu.com/andreas_blumauer/docs/survey_do_controlled_vocabularies_matter_2011_june

- **Unité structurelle** (*structural unity*) : tous les vocabulaires participants partagent la même structure de relations hiérarchiques et associatives entre les concepts, quels que soient la langue, notation ou code.
- **Lien direct** (*direct-linked*): ce modèle concerne des vocabulaires appartenant à différents types de langages et de structures. Ils nécessitent d'être mis en relation un à un (avec chaque fois deux schémas symétriques), pour former un réseau permettant de rechercher dans l'ensemble des documents indexés à partir de n'importe lequel de ces vocabulaires mis en relation. Ce modèle très efficace étant assez lourd à gérer, il est préconisé lorsque peu de vocabulaires sont en jeu.

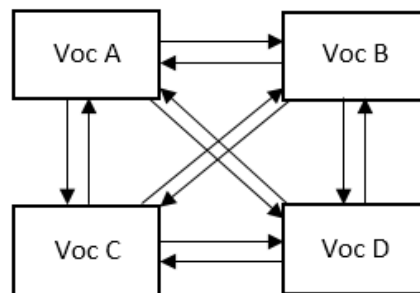


FIGURE V : EXEMPLE D'APPLICATION DU MODELE "LIEN DIRECT" A 4 VOCABULAIRES

- **Structure pivot** (*hub structure*) : plutôt que d'établir un grand nombre de schémas d'alignement, il peut être plus pratique de désigner l'un des vocabulaires comme « pivot », avec lequel chacun des autres vocabulaires sera aligné. Néanmoins, les résultats peuvent être de moins bonne qualité lorsque la conversion nécessite deux étapes (Voc C à Voc B par exemple). De plus, si aucun des vocabulaires présents ne permet de couvrir tous les autres, il est nécessaire soit d'étendre le plus complet d'entre eux pour couvrir l'étendue et la profondeur des autres, soit d'en concevoir un nouveau qui corresponde au besoin.

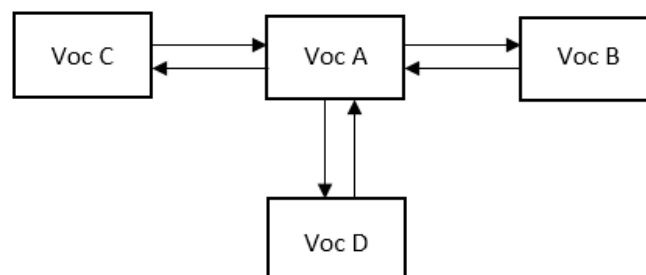


FIGURE VI : EXEMPLE D'APPLICATION DU MODELE "STRUCTURE PIVOT" A 4 VOCABULAIRES

Bien sûr, ces solutions de type « idéal » doivent être adaptées en fonction du besoin, dicté par l'environnement applicatif dans lequel le réseau sémantique sera mis en œuvre et l'utilisation réelle de ces langages. Ainsi, un langage pivot mis en œuvre sur un portail permettrait d'effectuer des recherches à travers tous les sites et applications fédérés, sans avoir besoin d'un alignement

symétrique vers les langages d'indexation utilisés dans chacun d'eux. D'autre part, l'étendue du schéma d'alignement peut être adaptée au besoin, qui n'est pas forcément de couvrir l'ensemble des vocabulaires, mais seulement les domaines concernés par l'application en question. La norme parle alors d'alignement sélectif (*selective mapping*). Enfin, ces trois modèles devront parfois être combinés afin d'obtenir la solution la plus efficace selon l'objectif du projet, tout en prenant en compte les besoins de maintenance nécessités par chaque modèle et les ressources effectives.

9.1.2.3 Des correspondances analogues à la structure théaurale

Les relations établies entre les vocabulaires (*mapping*⁵⁵) correspondent à trois types d'alignements, découlant de la même logique que le thesaurus (Clauses 7 à 10):

- **Alignement d'équivalence** : cette relation doit être établie entre tous les concepts correspondants dans les différents vocabulaires. Contrairement au thesaurus, aucune distinction de statut n'est établie entre eux. Il sera souvent nécessaire de préciser le niveau d'équivalence cependant. La norme prévoit l'équivalence simple (identique) et l'équivalence composée. Dans ce second cas, un concept complexe retenu dans un vocabulaire doit être associé à deux concepts simples dans un second, plusieurs concepts sont recouverts par un seul (*cumulative compound*), ou encore un concept se trouve à l'intersection de deux autres (*intersecting compound*).
- **Alignement hiérarchique** : relation établie lorsqu'un concept est clairement le générique d'un second, selon les mêmes règles que pour le thesaurus [1, ISO, Clause 10.2]
- **Alignement associatif** : lorsque la relation n'appartient à aucune des deux premiers types, mais qu'il existe un lien sémantique entre les deux concepts, leur association apportera un complément intéressant pour la recherche d'information.

L'équivalence étant le type d'alignement le plus fréquent, la Clause 11 décline plus en détail les différents cas : équivalence exacte, inexact ou partielle. Il est proposé, afin de parfaire le schéma de correspondances, de préciser ce niveau d'équivalence lorsque celle-ci est simple.

Les Clauses 12 à 14 fournissent des préconisations plus précises sur les moyens nécessaires et les méthodes à appliquer selon les cas, en fonction des vocabulaires impliqués et de problématiques plus spécifiques nécessitant des traitements plus complexes. La Clause 13 concerne en particulier la pré-coordination. Enfin, la Clause 15 traite de la gestion des données, et la Clause 16 présente des propositions de présentation pour les tables de correspondances.

⁵⁵ Ce terme est employé dans la norme pour parler à la fois des relations entre les concepts issus de vocabulaires différents et de l'opération consistant à établir ces relations. N'ayant pu trouver de traduction bien établie en français, nous avons choisi de parler ici de schéma d'alignement ou de correspondance.

9.1.3 Panorama des principaux types de vocabulaires, caractéristiques et interopérabilité

Les autres types de vocabulaires pouvant interopérer avec les thesauri n'ayant généralement pas fait l'objet d'une normalisation spécifique, la norme ISO 25964 propose un rappel de leurs principales caractéristiques comparées au thesaurus et des recommandations pour l'alignement (Clauses 17 à 24). Afin d'en présenter une vision synthétique, nous avons résumé ces informations dans le tableau qui suit. Nous rappelons au préalable que des tables d'alignement ne concernent que les concepts, les relations hiérarchiques propres à chaque vocabulaire n'y étant pas représentées. Le vocabulaire source est celui de départ, le vocabulaire cible (*target*) est celui au sein duquel sont recherchés les concepts correspondants.

VOCABULAIRE	CARACTERISTIQUES	STRUCTURE SEMANTIQUE	RECOMMANDATION D'ALIGNEMENT
Plan de classification	Organisation des ressources Enumérative, analytico-synthétique ou à facettes Concept = classe Notation : code/indice et label unique par classe	Mono-hiérarchie souple Pré-coordination élevée Index décrivant le champ du concept Associations : Voir aussi	Classe représentée par sa notation codée Equivalences composées fréquentes Index non inclus (aide)
Plan de classification en entreprise	Records management, portail Activités et environnement de l'entreprise (logique) Non adapté à la recherche Identifiant alphanumérique à chaque niveau	Mono-hiérarchie souple Agrégation de termes Même label utilisé dans différents domaines Métadonnées liées à la politique documentaire	Idem ci-dessus Choisir un mode de représentation unique Résultats à la recherche moins efficaces qu'en indexant directement avec le thesaurus
Taxonomie	Environnements en réseau Navigation et recherche Concept = catégorie Notation : label et parfois identifiant unique Multilinguisme fréquent	+/- proche d'un thesaurus Mono ou poly-hiérarchie Catégories associées Relations d'équivalence Label unique non systématique Pré-coordination fréquente	Adapter en fonction de la structure effective et des notations disponibles Catégories non-sujets problématiques (comparables aux facettes)
Répertoire de vedettes matières	Représentation synthétique des sujets traités dans les doc. Utilisation proche du thesaurus Concept = vedette Vocabulaire contrôlé	Hiérarchie : vedettes, subdivisions Relations associatives Relations d'équivalence Pré-coordination surtout pour les subdivisions	Pas de solution idéale : concepts simples seuls, concepts pré-coordonnés ou uniquement les vedettes utilisées. Si thesaurus = source, aligner tous les concepts

Ontologie	"Spécification formelle et explicite d'une conceptualisation partagée" ⁵⁶ d'un domaine de connaissance Langage : RDFS ou OWL (W3C) Concept = individu ou objet Notation : identifiant unique (URI) pour chaque élément But : concevoir des affirmations (<i>assertions</i>) sur les instances	Composition : - Classes ou type d'objets - Propriétés ou attributs - Axiomes ou règles - Individus ou instances Hiérarchie entre les classes et transmission des propriétés et axiomes Pas de synonymes Pas de désambiguïsation	2 logiques totalement différentes : ontologie orientée machine Les relations présentes dans un thesaurus n'ont aucun sens en ontologie Objectif : remplacer ou compléter un thesaurus Coût élevé et complexité : besoin et faisabilité ?
Terminologie	Glossaire, base terminologique... Traduction, communication... Concept = sa dénomination Différents degrés de contrôle ISO 1087-1:2000 et 704:2009 "Travaux terminologiques" Standard TBX (ISO 30042)	Concepts : - Génériques = termes - Individuels = appellation Définitions (= NA) Taux d'acceptabilité/terme Synonymes : préférentiel, admis ou désapprouvé Relations hiérarchiques génériques ou partitives Relations associatives Equivalences multilingues	Unification des ressources terminologiques d'une entreprise Enrichissement/extension d'un thesaurus Construction d'un "thesaurus de recherche" Très proche de l'alignement entre thesauri Choix le moins ambiguë, identifiant unique si dispo.
Liste d'autorité	Liste de noms et entités Représentation unique et contrôlée Concept = nom Projet VIAF	Parfois structurée (hiérarchie partitive) Equivalences : différentes appellations, transcriptions Relations associatives Relations chronologiques Informations supplém. (homonymie, entité) Relations de composition	Objectifs et contenus très différents Correspondances sélectives Plus appropriée : équivalence exacte Désignation la moins ambiguë : nom préférentiel ou identifiant si existant
Cercles de synonymes (<i>synonym ring</i>)	Énumération de synonymes et autres formes représentant un même concept Objectif : exhaustivité, aide à la recherche en texte intégral Notation : terme et parfois un identifiant unique	Aucune hiérarchie Tous les termes sont équivalents (implicite) Pas de désambiguïsation Occurrences multiples	Interopérabilité rarement utilisée mais faisable, surtout si un identifiant unique est disponible Construction/enrichissement de thesaurus (EP) Recherche sur ressources non ou sous-indexées

TABLEAU 3 : RECOMMANDATIONS POUR L'INTEROPERABILITE AVEC D'AUTRES LANGAGES

Dans tous les cas, la structure précise d'un vocabulaire devant être aligné avec un thesaurus devra être étudiée attentivement pour établir la méthode la plus adaptée et garantir les résultats à la recherche les plus efficaces.

⁵⁶ Définition, tirée de Gruber et retenue dans la norme, de Studer et al. *Knowledge engineering: principles and methods*, in Data & knowledge engineering, 1998, vol. 25, n°1-2, p.161-197 [traduction de l'auteur]

Le cas des ontologies est un peu particulier, n'étant que très rarement utilisées initialement pour la recherche d'informations. De ce fait, les projets d'alignement entre un thesaurus et une ontologie prennent des formes un peu différentes, s'inscrivant toujours dans le contexte du Web sémantique. Le premier reste un accès *user-friendly* au langage d'indexation au niveau de l'interface utilisateur, tandis que la seconde enrichit les possibilités offertes par l'application pour le traitement sémantique des requêtes et l'enrichissement de résultats. La norme reste très prudente sur les développements futurs de ces aspects, se contentant de remarquer la diversité des projets en cours d'élaboration.

9.2 Les technologies du Web sémantique au service de l'entreprise

De la même manière, le Web sémantique se définit au fur et à mesure des projets menés, à travers les réalisations qu'il fait vivre. C'est un jeu empirique de rêves et de concrétisations successifs qui permet de délimiter peu à peu les prérogatives et les possibilités du Web sémantique, dont les limites sont chaque fois repoussées un peu plus loin par les équipes qui le construisent au fil de l'eau. Tentons malgré tout de présenter simplement son fonctionnement et ses enjeux.

9.2.1 Qu'entend-on par Web « sémantique » ?

9.2.1.1 Du Web de données au Web 3.0 ?

Parfois désigné comme le « Web 3.0 », le Web sémantique représente une évolution majeure des technologies liées à l'Internet. Son objectif est de simplifier l'échange de données entre ses acteurs, en mettant en place une structure garantissant l'interopérabilité entre les données, quels que soient leurs sources, systèmes ou applications d'origine. L'ambition est d'utiliser la puissance du Web pour favoriser l'émergence de nouvelles connaissances à partir des données liées (*linked data*). L'explosion des volumes de données en ligne, parfois appelée « infobésité » rend en effet de plus en plus urgente la nécessité de rendre les machines qui les exploitent plus intelligentes.

L'expression a été lancée dès 2001 par Tim Berners-Lee, l'inventeur du World Wide Web et directeur du W3C, l'organisme de standardisation des formats informatiques utilisés sur l'Internet. Les travaux se poursuivent à l'heure actuelle pour déployer le Web sémantique dans toute son ampleur. De nombreux projets sont menés, initiés par les Etats ou des organismes de recherche principalement, visant aussi bien à encourager la découverte scientifique qu'à permettre le partage de données entre entreprises. Ces nouvelles technologies représentent de fait des opportunités intéressantes pour les acteurs commerciaux désireux d'optimiser leur impact sur la toile.

9.2.1.2 L'architecture du Web sémantique

Le « *layer cake* » (Cf. FIGURE I) proposé par T. Berners-Lee a fixé la superposition des briques nécessaires pour construire l'environnement où pourra se déployer le Web sémantique. Ce schéma permet de comprendre comment les données y sont indexées, à l'usage des agents informatiques. Le Web repose à l'origine sur le langage HTML, qui définit l'emplacement des objets sur une page. Pour répondre aux ambitions du Web sémantique, il a été nécessaire en complément de développer de nouveaux langages spécifiques pour la description des données (métadonnées) :

- Le XML (*eXtensible Markup Language*) permet de représenter une information structurée
- Le format RDF (*Resource Description Framework*) est à la base de toute indexation sur le Web ;
- SPARQL est un langage de requête et un protocole s'appliquant aux données RDF pour les bases de données sur le Web ;
- SKOS (*Simple Knowledge Organization System*) permet de représenter des vocabulaires contrôlés et structurés.

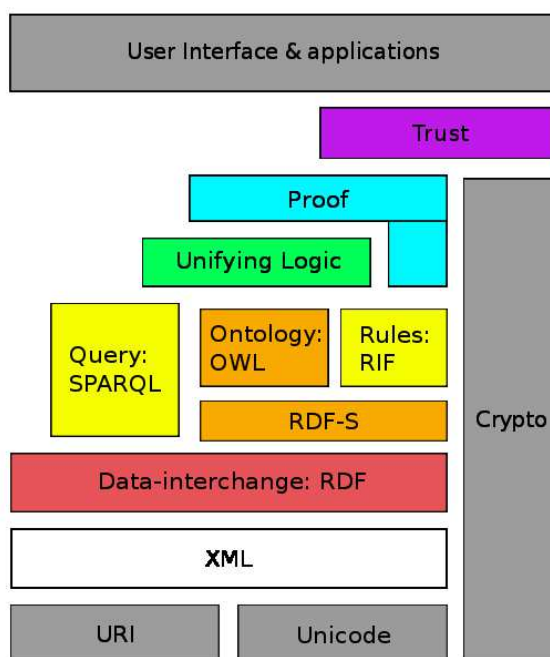


FIGURE VII : LE « LAYER CAKE » DU WEB SEMANTIQUE – D'APRES TIM BERNERS-LEE⁵⁷

Ces vocabulaires permettent aux machines d'appréhender le sens des concepts en les renseignant sur leurs définitions, leurs usages et les relations sémantiques qui les lient entre eux. Le tri des données disponibles sur le Web, rendu plus efficace, garantit ainsi un accès facilité et rapide à

⁵⁷ Source : article Wikipédia sur Tim Berners-Lee : http://fr.wikipedia.org/wiki/Tim_Berners-Lee (publié par Mhermans sous licence GNU, 2007).

une information pertinente pour l'utilisateur. Le Web de données est aujourd'hui installé de façon pérenne, avec ses langages et processus propres. La prochaine étape consiste désormais à mettre en place la logique d'unification, ainsi que les mécanismes de preuve et d'intégrité, qui constituent le sommet de l'architecture du Web sémantique.

9.2.1.3 Le triplet RDF, à la base des technologies sémantiques

Parler de sémantique en informatique renvoie à trois types de métadonnées indispensables pour qualifier un objet, décrites par le triplet du modèle RDF :

- le type de ressource ;
- son agent, c'est-à-dire le programme qui permet de lire cette ressource ;
- le concept, ou l'information qu'elle contient et qui permet de la catégoriser.

Cette utilisation systématique, structurée et normalisée des métadonnées permet de construire des données interprétables par des machines, qui procèdent par relation d'inférence. Plus la collection de termes ainsi décrits prend un aspect formel et complexe dans la description des concepts, plus on s'approchera d'une « ontologie de domaine ». Parmi tous les types d'organisations des connaissances qui ont été développées jusqu'ici pour permettre à l'humain de communiquer aux machines le sens des mots, les ontologies informatiques sont les plus puissantes. Basées sur des langages spécifiques (RDF Schema, puis OWL *Web Ontology Language* développé par le W3C), elles permettent de décrire l'environnement sémantique d'un concept de la façon la plus détaillée et formalisée à ce jour (Cf. TABLEAU 3).

9.2.2 Quelles perspectives pour l'entreprise

Si le Web sémantique reste une nébuleuse difficilement déchiffrable pour les non-initiés, l'application croissante de ses technologies au monde de l'entreprise rend ses possibilités de plus en plus reconnaissables. Elles s'inscrivent ainsi aujourd'hui comme des solutions de premier plan pour les organisations, confrontées chaque jour aux problématiques d'une information non structurée exponentielle. On entend souvent que celle-ci représenterait 80% de l'information disponible, mais Alain Garnier, président de Jamespot, évalue quant à lui cette proportion à 99%, sans même mentionner le *Big data* ! Nous nous appuyerons pour présenter l'actualité de ce secteur sur le livre blanc que lui a récemment consacré une grande société de SSII française [30, ECONOCOM-OSIATIS].

9.2.2.1 Technologies sémantiques et TAL, deux approches complémentaires

Les liens entre linguistique et informatique sont bien connus. Dès les années 1950 les recherches, autour de l'intelligence artificielle notamment, ont rapproché les deux disciplines, donnant naissance à une branche importante pour la convergence entre informatique et SIC, le traitement

automatique des langues, communément appelé TAL. Ces deux objets principaux sont, jusqu'à aujourd'hui, l'analyse et le traitement de textes écrits d'une part, faciliter le dialogue homme/machine d'autre part. Si le second objectif reste largement à réaliser, le premier a déjà livré au cours des ans de nombreux outils : la traduction et le résumé automatiques, l'annotation de corpus, etc.

Les travaux liés à la vision d'un Web dit « sémantique », développée par T. Berners-Lee, sont venus compléter cette approche il y a une dizaine d'années, à la faveur de l'émergence du Web de données, comme on l'a vu plus haut. L'ambition ici est proche du second objectif du TAL, puisqu'il s'agit de permettre aux agents informatiques de comprendre le sens des mots employés, pour répondre de façon plus pertinente à une requête et l'enrichir en explorant le Web des données. Les techniques proposées par le TAL (désambiguïsation de termes, extraction d'entités, classification ou indexation automatiques) sont désormais complétées et soutenues par les standards conçus pour le Web sémantique et encadrés par le W3C, sans lesquels l'interopérabilité resterait une gageure. Cette convergence s'opère alors même que les phénomènes de l'*Open data* (consistant à rendre les données publiques) et du *Linked data* (interconnecter les données de sources différentes) viennent multiplier les opportunités d'exploitation des données disponibles sur les réseaux Internet. L'un des pivots de ce mouvement général, qui en symbolise bien ses possibilités, est DBpedia. Ce projet communautaire a pour but d'exploiter l'encyclopédie collaborative Wikipedia, pour en extraire une base de données normalisée au format du Web sémantique, permettant d'aligner les données présentes sur le Web [26, BIZER].

9.2.2.2 Vers des réseaux d'entreprise sémantiques

Si l'on a beaucoup entendu parler des projets menés au niveau institutionnel, en particulier autour des questions récentes sur l'*Open data*, les entreprises se sont également peu à peu approprié ces technologies, pour les mettre au service de leurs propres objectifs. Les exigences en termes de gestion de l'information se font toujours plus pressantes et exigeantes, à mesure que le volume s'accroît et que ce capital immatériel de l'entreprise acquiert une valeur primordiale. Or, les technologies sémantiques permettent de donner à l'organisation des connaissances une ampleur totalement différente, dans leur diffusion, mais surtout dans leur conception. Un système construit autour d'une ontologie, par exemple, révolutionne le modèle conceptuel appliqué aux données. Le fonctionnement à partir d'un modèle relationnel, qui doit être redéfini à chaque adaptation, y est remplacé par le modèle d'instanciation encadré par le triplet RDF sujet-prédicat-objet (Cf. 9.2.1.3), facilitant les modifications.

Si le Web sémantique tel que le rêve le W3C n'est pas pour tout de suite, les intranets sémantiques sont eux à portée de main, nous promettent les éditeurs de logiciels exploitant ses technologies (Mondeca, Antidot, Temis, Exalead, pour citer les plus connus). Elles offrent en outre une solution

avantageuse pour pallier l'une des grandes problématiques des DSI : la multiplication et la superposition des briques logicielles et des silos de données qui ont constitué le système d'information au cours de son développement. En appliquant une couche logique commune à l'ensemble du système, elles permettent d'unifier les applications et les ressources qui y sont stockées, permettant un usage réellement partagé et leur valorisation.

9.2.2.3 « Emancipation des référentiels » : quelles opportunités pour les professionnels de l'Infodoc ?

On l'a compris, le Web sémantique s'appuie d'abord et avant tout sur des experts terminologiques, sans qui il serait impossible de concevoir les outils, au premier rang desquels les ontologies, nécessaires pour rendre la machine « intelligente » et les données « intelligibles » par elle. En entreprise, au-delà de l'interopérabilité technique, la question de l'alignement des différents référentiels, avec leurs terminologies et leurs structures propres, est à la racine de tout projet mettant en œuvre les technologies du Web sémantique. Un travail préliminaire de cartographie de l'existant est indispensable pour obtenir une vision claire des vocabulaires qui devront être interconnectés, à associer à une typologie de l'information disponible et nécessaire. Dans son livre blanc, la société de conseil Econocom-Osiatis le préconise clairement :

« Dans l'idéal, il faudrait une personne dédiée à cette tâche : un architecte de l'information dont le rôle serait de définir la typologie et la classification des contenus ainsi que d'optimiser la navigation et la recherche d'information en fonction des objectifs du public visé. » [30, ECONOCOM-OSIATIS]

Connaissance des publics, maîtrise des référentiels terminologiques et de l'analyse documentaire, capacité à mettre en place des systèmes d'organisation des connaissances... autant de compétences intrinsèquement liées aux professions de l'information-documentation. Des compétences et des savoir-faire métier que l'entreprise a tout intérêt à identifier et mettre à profit, si elle souhaite évoluer vers une gestion optimisée de ses connaissances et de l'information qu'elle produit ou disponible.

Conclusion

La conduite d'un projet de réingénierie de thesaurus a donc bien des effets bénéfiques multiples sur le management d'un centre de ressources documentaires en entreprise, et sur les membres de son équipe. Culture informationnelle, compétences documentaires, travail collaboratif et bien d'autres aptitudes indispensables aux organisations apprenantes ont ainsi été encouragées et approfondies. Le positionnement d'un chef de projet externe pour mener à bien un projet de ce type, en appui du management du service, apparaît effectivement comme un atout méthodologique, mais aussi sur le plan humain. Les nouvelles bases qui auront été posées à cette occasion seront déterminantes pour la pérennité et l'évolution de cet outil et son impact sur la gestion documentaire.

Comme dans de nombreuses organisations, ce thesaurus documentaire, mais aussi les autres référentiels de l'entreprise, restent des ressources internes, qui n'ont pas vocation à être publiées sur le Web ni mis à la disposition de tiers, comme c'est le cas pour des acteurs institutionnels. Des ambitions plus larges pour ces vocabulaires conçus en interne impliquent donc nécessairement un positionnement à un plus haut niveau structurel, qui seul permet de mettre en place une politique d'envergure, apte à synchroniser des objectifs communs. Nous avons tenté de montrer les avantages que représenterait pour l'organisation le fait de disposer d'un référentiel commun, regroupant le vocabulaire décrivant l'ensemble des métiers et activités de l'entreprise, son environnement concurrentiel, légal, politique... dans l'articulation et l'optimisation de l'ensemble de son système d'information.

Proposer, organiser et encadrer la réflexion pour la constitution de ce langage commun, construit conjointement avec des experts issus de tous les départements du groupe, dans ses différentes filiales, constitue une mission à la hauteur des compétences observées au sein du service Documentation générale de BPCE⁵⁸. Surtout, un tel projet, s'il pouvait être compris et porté par la direction du groupe, permettrait de remettre enfin les métiers de la documentation à la place qui devrait être la leur dans toute organisation aujourd'hui, quel que soit son domaine d'exercice, ce qui n'est malheureusement pas le cas à l'heure actuelle au sein de BPCE SA. Plus que jamais, les compétences documentaires doivent accéder à une plus grande visibilité au sein des organisations, et s'imposer comme les plus capables de développer une gouvernance de l'information qui mette pleinement à profit les langages contrôlés développés pour transcrire efficacement les préoccupations de l'entreprise.

Cette mission est donc loin de constituer un aboutissement, et ce même si l'outil livré semble correspondre aux usages pour lesquels il avait été initialement nécessaire. Sa réussite résidera plutôt dans sa capacité à engendrer une réflexion sur les suites envisageables à donner à ce

⁵⁸ A l'heure où nous achevons ce mémoire, le service a enfin obtenu d'être renommé « Veille et documentation », afin de mieux refléter dans l'organisation la réalité de ses missions et activités actuelles.

travail, dans le contexte organisationnel et stratégique particulier de BPCE SA aujourd'hui. Valoriser les compétences documentaires au sein du groupe, porter la fonction documentaire au service de l'entreprise dans sa globalité, et non plus seulement dans ses quelques services de veille et de documentation : telles sont les ambitions légitimes de l'équipe avec laquelle j'ai eu le plaisir de mener cette mission, et pour laquelle le travail entamé ici sur le thesaurus peut constituer un préalable déterminant.

Bibliographie

La bibliographie qui suit a été arrêtée au 18 novembre 2013. Tous les liens proposés concernant l'accès aux documents référencés étaient valides à cette date. Seuls les liens vers les ressources accessibles gratuitement en ligne ont été retenus.

Il s'agit d'une bibliographie analytique, organisée autour des grands thèmes traités dans ce mémoire :

- Le thesaurus
- Les systèmes documentaires et les SOC
- Le management des compétences documentaires
- Les référentiels pour le Web
- La gestion de projet

Ces thématiques principales sont elles-mêmes divisées en sous-thématiques, afin de permettre au lecteur un accès rapide aux ressources les plus pertinentes selon sa problématique.

Les références comportent une numérotation correspondant à celle reprise dans le corps du mémoire sous la forme [n°, NOM de l'auteur]. Dans chaque sous-thématique, les documents sont classés par ordre alphabétique des noms d'auteur, à l'exception de la norme ISO qui a volontairement été placée en tête des références.

La rédaction des références bibliographiques est conforme aux normes :

- Z44-005. décembre 1987. Documentation. Références bibliographiques : contenu, forme et structure et à la norme
- NF ISO 690-2 Février 1998 Information et documentation. Références bibliographiques Documents électroniques, documents complets et parties de documents

L'encyclopédie en ligne Wikipédia, qui n'est pas citée dans cette bibliographie, a également été utilisée régulièrement, comme première approche pour cerner certaines notions, ou pour obtenir des informations factuelles (méthodes Agiles, Scrum, termes techniques utilisés en français, etc.).

1. Thesaurus

Norme

[1] ISO 25964-1:2011. Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. 1^e édition. Genève, International Organization for Standards, Août 2011. 152 p. [en anglais]

Le report à la norme s'avère indispensable pour mener un travail en profondeur sur un thesaurus. Celle-ci présente l'intérêt de proposer de nombreux exemples, bien plus que dans les éditions précédentes. La Clause 2 regroupe en outre 64 définitions normalisées, qui en constituent l'une des nouveautés remarquables. Aucune traduction en français de la norme n'est prévue à l'heure actuelle, mais plusieurs exemples en français y sont proposés.

Voir aussi : ISO 25964-1 Schema and Data Model. [en ligne]

<www.niso.org/schemas/iso25964/#schema>

Cette page Web propose un résumé des informations relatives à la norme, ainsi que de nombreuses ressources complémentaires, utiles pour en savoir plus.

[2] ISO 25964-2:2013. Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies. 1^e édition. Suisse, ISO, Mars 2013. 99 p. [Anglais]

La seconde partie de la norme est parue près d'un an et demi plus tard. Elle est consacrée aux autres langages documentaires et leur interopérabilité avec les thesauri, notamment dans le contexte du Web sémantique.

[3] ADBS, Groupe « Langages documentaires ». Les normes de conception, gestion et maintenance de thésaurus : évolutions récentes et perspectives. Documentaliste-Sciences de l'information, ADBS, 2007, Vol. 44, n°1, p. 66-74. ISSN 0012-4508*

Un article intéressant pour retracer les étapes successives qui ont mené à la production d'une nouvelle norme internationale sur la conception de thesaurus.

[4] DALBIN Sylvie, YAKOVLEFF Nathalie, ZYSMAN Hélène, et al. Livre blanc sur la norme ISO 25964-1 « Thésaurus pour la recherche documentaire » [PDF en ligne]. AFNOR, BiVi, mis en ligne le 29 janvier 2013, [consulté le 22 octobre 2013]. <<http://www.bivi.fonctions->

* Les articles suivis d'un astérisque sont issus du numéro spécial « Langages documentaires et outils linguistiques » de la revue Documentaliste-Sciences de l'information, paru en février 2007. Seules les contributions les plus utiles à la rédaction de ce mémoire ont été retenues ici.

documentaires.afnor.org/livres-blancs/livre-blanc-sur-la-norme-iso-25964-1-thesaurus-pour-la-recherche-documentaire-parue-en-version-anglaise-en-aout-2011>

Cette ressource est précieuse, car c'est la seule en français sur la norme, qui n'a pas été traduite par l'AFNOR. Elle constitue de plus un excellent résumé de ses points essentiels.

[5] DEXTRE CLARKE, Stella G., LEI ZENG, Marcia. Standard Spotlight: From ISO 2788 to ISO 25964: The Evolution of Thesaurus Standards towards Interoperability and Data Modeling. Information Standards Quarterly, 2012, Winter 24(1):20-26
<http://www.niso.org/apps/group_public/download.php/8289/SP_clarke_zeng_isqv24no1.pdf>

Cet article est intéressant pour compléter les points de vue des professionnelles interrogées pour ce mémoire (Cf. 1.1.3). Il est cosigné par la responsable de l'équipe internationale chargée de la rédaction de la nouvelle norme, S. Dextre Clarke.

Conception et réingénierie de thesaurus

[6] CYROT, Catherine. Réingénierie de thésaurus : une étude de cas. Christain Preuss, collab. Documentaliste-Sciences de l'information, ADBS, 2009, Vol. 46, n°3, p. 4-13. ISSN 0012-4508

Cet article constitue un retour d'expérience précieux, car il aborde principalement les aspects techniques, indissociables aujourd'hui de la réingénierie de thesauri, liés à leur utilisation dans les systèmes d'informatique documentaire.

[7] DEGEZ, Danièle. Construire un thesaurus. Archimag, SERDA, 2009, n°222, p.44-45. ISSN 0769-0975

Cet article très pédagogique reprend point par point les principales étapes de la construction d'un thesaurus, du cahier des charges à la maintenance. Elle est co-auteure du Thésauro-glossaire des langages documentaires, paru à l'ADBS en 2001, qui n'a pas été utilisé pour ce mémoire.

[8] DUPUICH, Jonas, CARR, Gabe. Case Study: Developing the PLOS Thesaurus. Bulletin of the American Society for Information Science and Technology, déc/jan 2013, Vol. 39, n°2, p. 22-25. ISSN 1532-2890

Cet article retrace les principaux enjeux de la réingénierie d'un thesaurus utilisé pour une application au sein d'un site Web, dans un contexte collaboratif. Il est intéressant de comparer la méthode utilisée par le prestataire engagé pour ce projet. On trouvera en outre quelques pistes de réflexion pour le couplage avec un outil d'indexation automatique dans le cas du texte intégral, ainsi que sur les règles à définir pour garantir la pertinence de ses propositions.

[9] HUDON, Michèle. Guide pratique pour l'élaboration d'un thesaurus documentaire. Danièle Dégez, Dominique Ménillet, collab. 2^e édition [1994]. Montréal, Les Editions ASTED (diffusion France : ADBS), 2009. 274 p. ISBN 978-2-923563-17-6

A l'heure actuelle, cet ouvrage est la ressource la plus récente et complète pour accompagner des professionnels confrontés à la conception de thesaurus. Un incontournable, même s'il s'appuie sur la précédente norme, qui aborde tous les aspects de ce type de projets. M. Hudon enseigne à l'EBSI de l'Université de Montréal.

[10] TÊTE, Caroline, et al. Le thesaurus, outil d'accès à l'information. Documentaliste-Sciences de l'information, ADBS, 2013, Vol. 50, n°1, p. 12-13. ISSN 0012-4508

Un autre exemple de mission de refonte de thesaurus, portant cette fois sur la fin de vie et les soins palliatifs. L'objectif ici était de lui donner une ampleur bien plus importante, pour l'imposer comme une ressource incontournable du domaine au niveau international.

Outils informatiques et mise en œuvre de thesaurus

[11] DALBIN, Sylvie. Thesaurus et informatique documentaires : partenaires de toujours ? Documentaliste-Sciences de l'information, ADBS, 2007, Vol. 44, n°1, p. 42-55. ISSN 0012-4508*

[12] DALBIN, Sylvie. Thesaurus et informatique documentaires : les Noces d'Or. Documentaliste-Sciences de l'information, ADBS, 2007, Vol. 44, n°1, p. 76-80. ISSN 0012-4508*

Ces deux articles d'une des principales spécialistes du thesaurus se penchent sur l'évolution des outils informatiques dédiés à la gestion de thesaurus. Le premier présente plus spécifiquement les fonctionnalités nécessaires et les types de logiciels. Le second se penche sur l'historique de leur diffusion et les perspectives de l'époque.

[13] DALBIN, Sylvie. Descripteurs : site dédié aux thesaurus et autres vocabulaires contrôlés pour l'accès à l'information, [blog en ligne]. <<http://dossierdoc.typepad.com/descripteurs/>>

Du même auteur, un blog régulièrement actualisé où sont réunis des informations indispensables sur l'actualité des thesauri. L'auteur y publie également certains des articles qu'elle écrit pour différentes publications.

[14] LACASTA, Javier et al. ThManager : An Open Source Tool for Creating and Visualizing SKOS. Information Technology and Libraries, 2007, vol. 26, n°3, p. 39-51. [en ligne] <<http://ejournals.bc.edu/ojs/index.php/ital/article/view/3274>>

Cet article permet de mieux comprendre les principes de fonctionnement de l'outil de création et gestion de thesaurus Open source que nous avons mis en œuvre au cours du projet. A lire en complément du guide utilisateur écrit pour l'équipe, fourni en annexe.

2. Systèmes documentaires et SOC

Langages documentaires et indexation

[15] CHAUMIER, Jacques. Les techniques documentaires. 9^e édition. Paris, PUF, Que sais-je ?, 2002. Chap. IV, Les langages documentaires, p. 43-65. ISBN 978-2-13-052424-3

Ce chapitre en particulier, au sein d'un ouvrage quelque peu daté mais utile pour faire le tour des techniques documentaires, offre un panorama intéressant pour s'y retrouver au sein de la famille des langages documentaires. Il rappelle notamment les principaux avantages et inconvénients de chacun. La partie dédiée au thesaurus est la plus étendue.

[16] COURBIERES, Caroline. Les langages documentaires : éléments d'organisation des connaissances. **In** GARDIES Cécile, dir. Approche de l'information-documentation. Concepts fondateurs. Toulouse, Cépaduès Editions, 2011. p. 131-150. ISBN 978-2-85428-982-4

Un complément plus récent au précédent article, qui présente également les principes et différents types de langages documentaires.

[17] COURBIERES, Caroline. L'analyse documentaire : éléments d'organisation des connaissances. **In** GARDIES Cécile, dir. Approche de l'information-documentation. Concepts fondateurs. Toulouse, Cépaduès Editions, 2011. p. 151-164. ISBN 978-2-85428-982-4

Cet article permet de faire le point sur les différents aspects de l'analyse documentaire, qui précède la représentation de l'information, formée de l'indexation et du résumé. L'auteur y affirme la nécessité de placer le besoin du document avant celui de l'utilisateur dans ce processus d'interprétation.

Organisation des connaissances

[18] MANIEZ, Jacques. Des classifications aux thésaurus : Du bon usage des facettes. Documentaliste-Sciences de l'information, ADBS, 1999, Vol. 36, n°4-5, p. 249-262. ISSN 0012-4508. [en ligne] <<http://www.adbs.fr/des-classifications-aux-thesaurus-du-bon-usage-des-facettes-13338.htm>>

Bien qu'un peu daté, ce long article reste une ressource précieuse pour qui veut y voir plus clair dans les différentes acceptions de la notion de facette et leur évolution. Il pose des bases conceptuelles qui restent très actuelles, malgré l'évolution des outils.

[19] POLITY, Yolla, HENNERON, Gérard, PALERMITI, Rosalba (éd.). L'organisation des connaissances : approches conceptuelles. Paris, L'Harmattan, 2005. [Actes du 4^e congrès ISKO-France, Grenoble, 3-4 juillet 2003]. Chap. 4, Ontologie et Thesauri, p.177-222. ISBN 2-7475-8274-4

Ce chapitre aborde la question des ontologies conçues non comme un outil du Web sémantique, mais comme une forme de SOC. La dernière contribution du chapitre, signée M. Hudon, se penche quant à elle sur les « Macrothésauri et systèmes d'information gouvernementaux accessibles au grand public » (p. 213-222).

[20] TENNIS, Joseph T. Le poids du langage et de l'action dans l'organisation des connaissances : position épistémologique, action méthodologique et perspective théorique. **In** HUDON, Michèle, MUSTAFA EL HADI, Widad (coord.). Organisation des connaissances : épistémologie, approches théoriques et méthodologiques. Etudes de Communication, Université Charles-de-Gaulle – Lille 3, n° 39, 2012/2. p. 15-40. ISBN 978-2-917562-08-6 [traduit de l'anglais]

Cet article, seule ressource disponible en français à l'heure actuelle de cet auteur américain (University of Washington, Seattle), nous a été très précieux pour replacer notre travail sur le thesaurus BPCE dans le contexte épistémologique de la recherche sur les SOC et leur conception.

3. Management des compétences documentaires

[21] VOLANT, Christiane. Le management de l'information dans l'entreprise : vers une vision systémique. Paris, ADBS Editions, 2002. 106 p. Collection Sciences de l'information, série Etudes et techniques. ISBN 2-84365-063-1

Cet ouvrage reste un incontournable pour qui souhaite aborder les notions de systémique, de complexité ou la théorie des organisations aux services documentaires en entreprise.

[22] REGIMBEAU, Gérard. Médiation. **In** GARDIES Cécile, dir. Approche de l'information-documentation. Concepts fondateurs. Toulouse, Cépaduès Editions, 2011. p. 75-114. ISBN 978-2-85428-982-4

Le documentaliste est avant tout un médiateur de l'information. Cet article interroge ce concept, en revenant sur ses définitions et approches théoriques, pour mieux cerner en quoi consiste la médiation documentaire.

Organisation apprenante

[23] VOLANT, Christiane. Réseau documentaire... organisation apprenante. Documentaliste-Sciences de l'information, ADBS, 1998, Vol. 35, n°3, p. 144-146. ISSN 0012-4508

Bien qu'un peu daté, cet article reste à nos yeux une ressource indispensable pour une première approche du concept d'organisation apprenante appliqué aux services documentaires et des modes de management adaptés pour les faire évoluer.

Culture informationnelle

[24] BOISVERT, Danielle, dir. Le développement de l'intelligence informationnelle : les acteurs, les défis et la quête de sens. Montréal, Editions ASTED, 2010. 219 p. ISBN 978-2-923563-18-3

Cet ouvrage n'est pas en lien direct avec notre sujet, puisqu'il traite de l'enseignement des compétences informationnelles en milieux scolaires. Il nous a été utile cependant pour définir le concept d'intelligence informationnelle appliqué aux professionnels de la documentation.

[25] GARDIES, Cécile. Culture de l'information, culture informationnelle. In GARDIES Cécile, dir. Approche de l'information-documentation. Concepts fondateurs. Toulouse, Cepaduès Editions, 2011. p. 187-208. ISBN 978-2-85428-982-4

Cet article offre un retour intéressant sur l'évolution de la notion de culture de l'information, jusqu'au concept aujourd'hui émergent de culture informationnelle, en passant par les différentes approches de l'éducation à l'information.

4. Référentiels pour le Web

[26] BIZER, Christian et al. DBpedia - A crystallization point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 7, Issue 3, septembre 2009, p. 154-165. ISSN 1570-8268 [*preprint* en ligne] <http://jens-lehmann.org/files/2009/dbpedia_jws.pdf>

Cet article, écrit par des chercheurs des universités dont émane le projet, présente de façon relativement simple les principes de DBpedia ainsi que des pistes d'applications concrètes. Bien sûr, celles-ci pourraient être complétées, la publication remontant à 2009.

[27] Collectif. Web sémantique, web de données... Quelle nouvelle donne ? Documentaliste-Sciences de l'information, ADBS, 2011, Vol. 48, n°4, 80 p. ISSN 0012-4508

Ce numéro spécial de DocSI apporte des clés de compréhension pour l'application des technologies du Web sémantique à la documentation. L'article de Sylvie Dalbin est consultable sur son blog, les autres accessibles aux adhérents sur le site de l'ADBS. <<http://dossierdoc.typepad.com/descripteurs/2012/02/le-web-s%C3%A9mantique-en-entreprise-quelques-cas-dusage.html>>

Web sémantique et entreprise

[28] BARBAUX, Aurélie. L'industrie s'empare du Web sémantique. Usinenouvelle.com, Technos et Innovations, 06/04/2010, [en ligne] <<http://www.usinenouvelle.com/article/l-industrie-s-empare-du-web-semantique.N128850>>

Cet article de presse offre un autre discours sur les applications du Web sémantique, cette fois au monde de l'industrie. Dans un second temps, il offre un tour d'horizon des projets existant en 2010, qui sont toujours d'actualité.

[29] Collectif. Blog Antidot, [société commerciale, en ligne] <<http://blog.antidot.net/>>

Ce blog très actif est tenu par la société française Antidot, éditeur de solutions d'accès à l'information basées sur les technologies du Web sémantique. Les billets toujours très précis mais accessibles permettent de mieux comprendre l'actualité de ce secteur et ses applications concrètes au monde de l'entreprise et des collectivités. On lira en particulier l'article « Réinventer les référentiels », sur les apports pour les moteurs de recherche. <<http://blog.antidot.net/2010/03/25/reinventer-les-referentiels/>>

[30] ECONOCOM-OSIATIS. Les technologies sémantiques : quel avenir pour l'entreprise ? Etat des lieux et apports fonctionnels. 1^e édition. Livre blanc, octobre 2013, 60 p. [en ligne] <http://www.osiatis.com/medias/upload/files/fr_FR/livre-blanc-tech-semantiques-vfweb_1381992554.pdf>

Ce récent livre blanc nous a été précieux pour mieux cerner les apports des technologies issues du Web sémantique pour les SI en entreprise. Après avoir rappelé l'historique et les grands principes de ces technologies, il présente les possibilités d'applications concrètes et l'offre des principaux éditeurs de solutions du marché. Paru en octobre 2013, il vient saluer le rapprochement en septembre de ces deux importantes SSII françaises.

Langages informatiques

[31] BAKER, Thomas et al. Key choices in the design of Simple Knowledge Organization System (SKOS). Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 20, mai 2013, p. 35-49. ISSN 1570-8268, [en ligne] <<http://www.sciencedirect.com/science/article/pii/S1570826813000176>>

Cet article de fonds permet de comprendre en détail la structure et le fonctionnement du langage SKOS, initialement conçu pour représenter les thesauri.

[32] CARACCILO, Caterina, STELLATO, Armando et al. Thesaurus Maintenance, Alignment and Publication as Linked Data: The AGROVOC Use Case. International Journal of Metadata, Semantics and Ontologies vol. 7, n°1, 2012, p. 65-75. ISSN [en ligne] 1744-263X <<http://www.inderscience.com/storage/f611712853124910.pdf>>

Cette étude de cas présente la gestion et la maintenance du thesaurus AGROVOC de la FAO. Cet outil est formulé en SKOS à l'aide de l'outil dédié VocBench, ce qui permet de le relier à d'autres ressources dans le cadre du Linked data. Cet article illustre bien, au niveau institutionnel, les possibilités de partage de thesaurus grâce au Web de données.

Ontologies

[33] DJAMBIAN, Caroline. Le métier : son savoir, son parler. **In** Terminologie & Ontologie : Théories et applications, Actes de la conférence TOTh, Annecy, 26-27 mai 2011. Annecy, Institut Porphyre, 2011, p. 75-91. ISBN 978-2-9536168-4-2 [en ligne]
<<http://www.porphyre.org/toth/files/actes/TOTh-2011-actes.pdf>>

Ce premier exemple de mise en œuvre d'ontologie est particulièrement intéressant, car il développe une approche de conception basée sur l'analyse de la terminologie métier, particulièrement applicable en entreprise.

[34] GRESSER, Jean-Yves. Ontologies des risques financiers – Continuité d'activité, gestion de crise, protection des *infrastructures critiques* financières. **In** Terminologie & Ontologie : Théories et applications, Actes de la conférence TOTh, Annecy, 3-4 juin 2010. Annecy, Institut Porphyre, 2010, p.155-176. ISBN 978-2-9536168-1-1 [en ligne]
<http://www.porphyre.org/toth/files/actes/TOTh-2010_08_Gresser.pdf>

Cet article illustre bien la conception et le fonctionnement d'une ontologie de domaine, tout en offrant un exemple de projet appliqué au domaine bancaire.

5. Gestion de projet

Ces deux ouvrages utilitaires ont été utilisés pour la conception de la gestion de projet et la prévention des risques en cours de route.

[35] BODET, Guillaume. Scrum en action. Montreuil, Pearson France, 2012. IX, 168 p. ISBN 978-2-7440-2564-8

Cet ouvrage se présente sous la forme d'une fiction, dont le personnage principal est un chef de projet logiciel dans l'impasse avec un gros client. Sa rencontre avec un consultant spécialiste de Scrum va bouleverser sa vision de la gestion de projet. C'est une façon originale et agréable de découvrir cette méthode Agile, à partir de problématiques concrètes.

[36] CORBEL, Jean-Claude. L'essentiel du management de projet. Paris, Groupe Eyrolles, 2013. 132 p. Les Essentiels. ISBN 978-2-212-55648-3

La démarche de cet ouvrage est originale, puisqu'il s'agit d'un « anti-guide » de la gestion de projet. A partir de son expérience en ingénierie automobile, l'auteur se focalise sur ce qu'il ne faut pas faire, les écueils à éviter pour garantir le succès du projet. En 10 étapes clés d'un projet, présentant chacune une poignée de risques majeurs, il nous livre l'essentiel d'une gestion de projet efficace, très utile en situation.

Mémoires INTD

Pour terminer, j'aimerais citer ces quelques mémoires, issu de la même formation, dont les thèmes peuvent apporter un complément au présent mémoire.

OLIVIER-DE CONINCK, Florence. Réingénierie du thésaurus d'une base de données en ligne et mutualisée dans le secteur culturel : le cas de RéseauDocs. 2011. 145 p. Mémoire pour obtenir le titre professionnel de chef de projet en ingénierie documentaire, niveau I. Conservatoire National des Arts et Métiers - Institut National des techniques de la Documentation, INTD. 2011

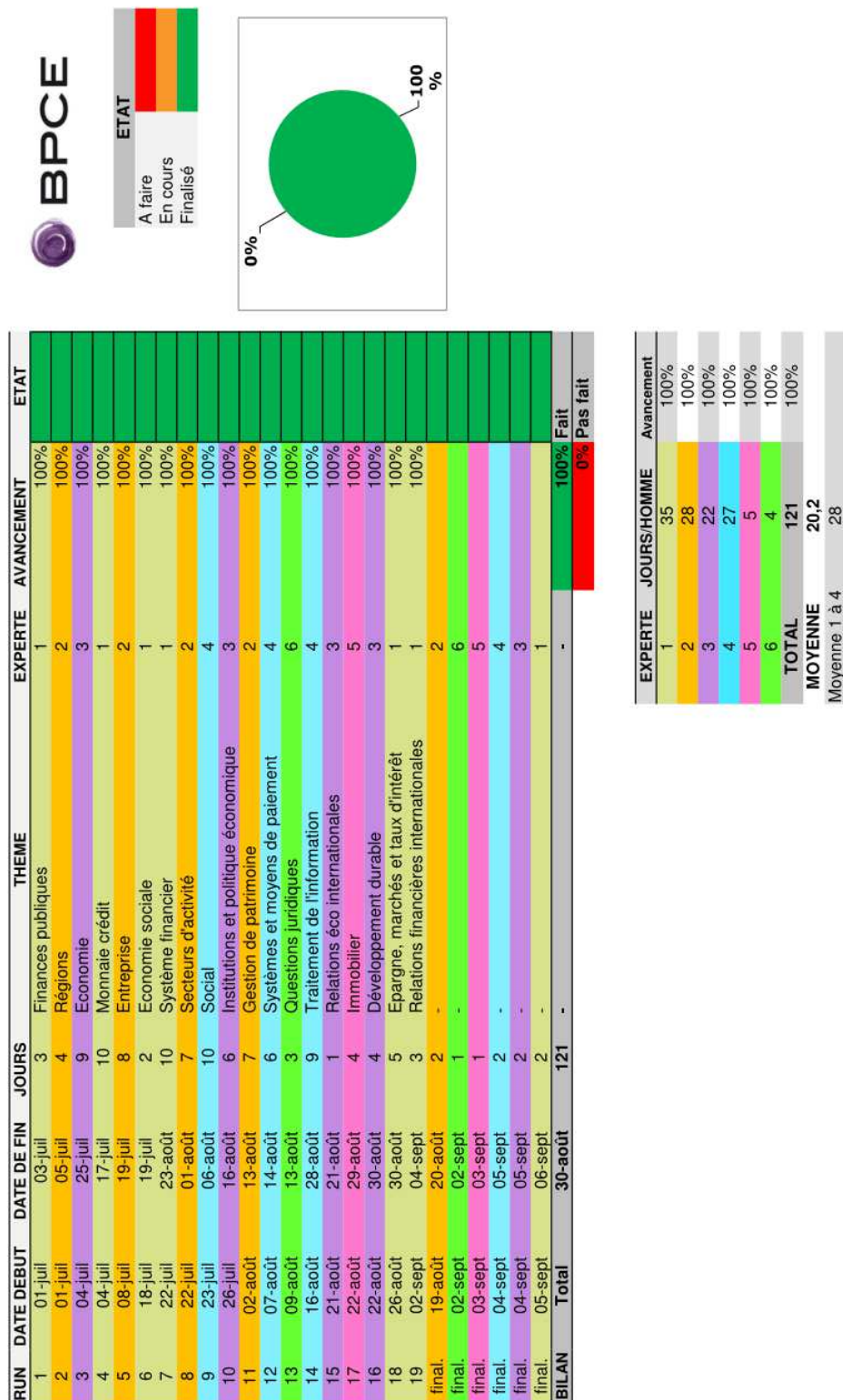
PIERRE, Béatrice. L'avenir des langages documentaires dans le cadre du Web sémantique : conception d'un thésaurus iconographique pour le Petit Palais. 2010. 115 p. Mémoire pour obtenir le titre professionnel de chef de projet en ingénierie documentaire, niveau I. Conservatoire National des Arts et Métiers - Institut National des techniques de la Documentation, INTD. 2010. [en ligne] <<http://memsic.ccsd.cnrs.fr/docs/00/57/50/53/PDF/PIERRE.pdf>>

RAVET, Brice. De l'usage des outils du Web de données pour une recherche efficace sur des ressources disséminées et hétérogènes : la mise en place d'un portail de recherche fédérée pour le Musée National du Sport. 2011. 187 p. Mémoire pour obtenir le titre professionnel de chef de projet en ingénierie documentaire, niveau I. Conservatoire National des Arts et Métiers - Institut National des techniques de la Documentation, INTD. 2010. [en ligne] <<http://memsic.ccsd.cnrs.fr/docs/00/67/98/84/PDF/RAVET.pdf>>

RUIZ LEPORES, Domingos. Des grandes classifications au Web de données et l'émergence de l'indexation sémantique : le cas du tagging sémantique dans le portail Histoire des Arts. 2011. 118 p. Mémoire pour obtenir le titre professionnel de chef de projet en ingénierie documentaire, niveau I. Conservatoire National des Arts et Métiers - Institut National des techniques de la Documentation, INTD. 2011 [en ligne] <http://memsic.ccsd.cnrs.fr/docs/00/67/99/06/PDF/RUIZ_LEPORES.pdf>

Annexes

Annexe 1 : Planning global et répartition des *runs*



Annexe 2 : Composition du thesaurus BPCE

Table des matières

Table des matières	1
Domaine : Traitement de l'information (11)	4
TT Informatique (111)	4
TT Science de l'information	5
TT Technologie de l'information et de la communication (112)	6
Domaine : Sciences économiques (1)	7
CS Analyse économique (12)	7
TT Méthodologie (121)	7
TT Pensée économique (13)	8
CS Structure économique (142)	9
CS Economie nationale	10
TT Comptabilité nationale (122)	10
TT Activité économique	11
Domaine : Institutions et politique économiques (2)	12
CS Vie politique (201)	12
TT Administration publique	12
TT Institution politique	12
CS Régions (22)	13
TT Institution régionale et locale (221)	13
TT Aménagement du territoire (222)	14
TT Innovation (23)	15
CS Politique économique (24)	15
TT Politique économique	15
CS Finances publiques (25)	16
TT Finances publiques	16
TT Fiscalité (252)	17
TT Trésor Public (253)	19
CS Développement durable (26)	19
TT Développement durable	19
TT Environnement	20
TT Responsabilité sociale des entreprises	21
Domaine : Questions juridiques (21)	22
CS Droit public (211)	22
CS Droit privé (212)	23
TT Droit civil (214)	23
TT Droit des affaires (213)	23
TT Droit pénal (2121)	25
TT Procédure	26
TT Profession juridique	26
CS Droit international – droit comparé (215)	26
Domaine : Monnaie crédit (31)	26
CS Moyens de paiement	26
TT Monnaie	27
TT Politique monétaire	28
TT Système de paiement (313)	29
CS Crédit (314)	29
TT Crédit	29
Domaine : Système financier (32)	32
CS Structure et réglementation du système financier (321)	32

CS Gestion bancaire (323).....	34
CS Activité bancaire.....	35
TT Activité bancaire.....	35
TT Gestion de patrimoine.....	35
CS Etablissements du secteur bancaire (324).....	36
TT Etablissement de crédit.....	36
CS Assurance (325).....	38
TT Assurance.....	38
CS Autres organismes du secteur financier (327).....	40
Domaine : Epargne et marchés de capitaux (34).....	41
CS Politique de l'épargne (341).....	41
CS Produits d'épargne (342).....	42
CS Instruments financiers (343).....	43
TT Titre financier.....	43
TT Contrat financier.....	44
TT Placement collectif.....	45
CS Marché des capitaux (344).....	46
TT Marché financier.....	47
CS Gestion d'actifs (345).....	49
CS Taux d'intérêt (35).....	49
Domaine : Relations internationales (4).....	50
CS Echanges économiques et commerciaux (40).....	50
TT Echanges internationaux.....	50
CS Relations économiques internationales (41).....	51
TT Coopération internationale.....	51
CS Relations financières internationales (42).....	51
TT Relations financières internationales.....	51
CS Union européenne (43).....	53
TT Union européenne.....	53
Domaine : Social (5).....	55
TT Ethique (50).....	55
CS Société (51/52).....	55
TT Population (51).....	55
TT Education.....	57
TT Sciences sociales.....	57
CS Travail (53/54).....	58
TT Organisation du travail.....	58
TT Emploi (54).....	58
TT Relations sociales.....	59
TT Ressources humaines.....	60
CS Protection sociale (55/56).....	61
TT Santé (550).....	61
TT Sécurité sociale (551).....	62
TT Retraite (56).....	62
Domaine : Secteurs d'activité (6).....	63
CS Secteurs économiques.....	63
TT Secteur primaire (611-612).....	63
TT Secteur secondaire (613-614).....	64
TT Secteur tertiaire (615).....	65
CS Immobilier (616).....	66
TT Immobilier.....	66
TT Financement de l'immobilier.....	67
TT Investissement immobilier.....	68
TT Logement.....	68
Domaine : Entreprise (7).....	70
CS Structure et gestion des entreprises (712).....	70
TT Entreprise.....	70
TT Gestion de l'entreprise.....	72
CS Gestion commerciale et marketing (7121).....	73
TT Marketing.....	73
TT Vente.....	74
CS Finance d'entreprise (714).....	75
TT Finance d'entreprise.....	75
CS Economie sociale et solidaire (716).....	76

Annexe 3 : Un domaine issu du thesaurus BPCE

Domaine : Entreprise (7)

CS Structure et gestion des entreprises (712)

TT Entreprise

- EP *Firme*
- TS1 Entreprise en difficulté
 - TA Procédure collective (213)
 - TS2 Défaillance d'entreprise
- TS1 Entreprise étrangère
- TS1 Entreprise innovante
 - EP *Start-up*
- TS1 Entreprise multinationale
 - EP *Multinationale*
- TS1 Entreprise privée
 - TA Secteur privé (1410)
 - TA Privatisation (2410)
- TS1 Entreprise publique
 - EP *Entreprise nationalisée*
 - TA Secteur public (1410)

Catégorie d'entreprise

- TS1 ETI
 - EP *Entreprise de taille intermédiaire*
- TS1 Grande entreprise
- TS1 Microentreprise
- TS1 PME - PMI
 - EP *Petite et moyenne entreprise*
 - EP *Petite et moyenne industrie*
 - TA Clientèle commerciale (3232)
- TS1 TPE
 - EP *Très petite entreprise*

Concentration des entreprises

- EP *Cartel*
- EP *Pool*
- TS1 Cession d'entreprise
- TS1 Fusion d'entreprises
 - EP *Acquisition d'entreprise*
- TS1 Scission d'entreprise

Création d'entreprise

- TA Business plan (7140)
- TS1 Siège social
- TS1 Société en formation
- TS1 Statuts de l'entreprise

Dissolution de société

Forme juridique de l'entreprise

- EP *Forme sociale de l'entreprise*
- TA Droit des sociétés (213)
- TA Groupement foncier agricole (611)
- TA Société civile immobilière (6160)
- TS1 Entreprise individuelle
 - TA Travailleur indépendant (51)
 - TS2 Autoentreprise
 - EP *Autoentrepreneur*
- TS2 Entreprise individuelle à responsabilité limitée
 - EP *EURL*
- TS2 Entreprise unipersonnelle à responsabilité limitée
 - EP *EURL*
 - EP *S4SU*
- TS2 Société par actions simplifiée unipersonnelle
 - EP *GIE*
- TS1 Groupement d'intérêt économique
 - EP *GEIE*
- TS1 Groupement européen d'intérêt économique
 - EP *GEIE*
- TS1 Société de capitaux
 - TA Société commerciale (213)
 - TS2 Société anonyme
 - EP *SA*
- TS2 Société à responsabilité limitée
 - EP *SARL*
- TS2 Société en commandite par actions
 - EP *SCA*
- TS1 Société de personnes
 - TA Société civile (213)
 - TS2 Société en nom collectif
 - EP *SWC*
- TS1 Société par actions simplifiée
 - EP *SAS*

<p>Groupe de sociétés</p> <ul style="list-style-type: none"> TS1 Coentreprise <ul style="list-style-type: none"> EP <i>Joint venture</i> TS1 Filiale TS1 Groupe industriel TS1 Holding TS1 Société en participation TS1 Société mère <p>Transmission d'entreprise</p> <ul style="list-style-type: none"> TS1 Rachat d'actions par l'entreprise TS1 Rachat d'entreprise <ul style="list-style-type: none"> TS2 Leveraged buy-out <ul style="list-style-type: none"> EP <i>LBO</i> TA Gestion d'actifs (345) TS1 Reprise d'entreprise par les salariés <ul style="list-style-type: none"> EP <i>RES</i> <p>TT Gestion de l'entreprise</p> <ul style="list-style-type: none"> TA Culture d'entreprise (5300) TA Information financière (1120) TS1 Gestion de crise <ul style="list-style-type: none"> EP <i>Risk management</i> TA Risque pays (41) TA Value at risk TS2 Division des risques <ul style="list-style-type: none"> TA Coassurance (325) TA Réassurance (325) TS2 Plan de continuité d'activité <ul style="list-style-type: none"> EP <i>PCA</i> TS2 Risque opérationnel <ul style="list-style-type: none"> TA Risque bancaire (3230) TA Risque financier (344) TS1 Gestion prévisionnelle TS1 Qualité <ul style="list-style-type: none"> TA Cercle de qualité (5300) TA Normalisation (23) TS2 Gestion de la qualité <ul style="list-style-type: none"> EP <i>Management de la qualité</i> TS1 Sous-traitance TS1 Vie sociale de l'entreprise 	<ul style="list-style-type: none"> TS2 Assemblée générale TS2 Conseil d'administration TS2 Conseil de surveillance <ul style="list-style-type: none"> TA Conseil d'orientation et de surveillance (324) TS2 Directoire <p>Contrôle de gestion</p> <ul style="list-style-type: none"> EP <i>Contrôle budgétaire</i> TA Rationalisation des choix budgétaires (2510) TS1 Audit interne <ul style="list-style-type: none"> TA Gestion bancaire (323) TS1 Contrôle interne <ul style="list-style-type: none"> EP <i>Conformité</i> TA Gestion bancaire (323) TA Normalisation (23) TS1 Système d'information de gestion <ul style="list-style-type: none"> TA Système d'information (11) TS2 Tableau de bord <p>Gouvernement d'entreprise</p> <ul style="list-style-type: none"> EP <i>Gouvernance d'entreprise</i> EP <i>Corporate governance</i> TS1 Gouvernance de l'information <ul style="list-style-type: none"> TA Gestion de l'information (111) TA Gouvernance informatique (111) <p>Stratégie de l'entreprise</p> <ul style="list-style-type: none"> TA Stratégie bancaire (3230) TS1 Modèle d'entreprise <ul style="list-style-type: none"> EP <i>Business model</i> <p>Structure de l'entreprise</p> <ul style="list-style-type: none"> EP <i>Organigramme</i> <p>CS Gestion commerciale et marketing (7121)</p> <p>Communication externe</p> <ul style="list-style-type: none"> TA Marketing mix TS1 Identité visuelle <ul style="list-style-type: none"> EP <i>Logo</i> TA Commerce (615) TS2 Enseigne
---	---

- TS1 Publicité
 - TS2 Concours publicitaire
- TS1 Relations publiques
 - TS2 Mécénat
 - TS2 Sponsoring
 - EP Parrainage
- TS1 Salon professionnel

TT Marketing

- TS1 Marketing bancaire
 - TA Gestion bancaire (323)
- TS1 Marketing des services
- TS1 Marketing direct
 - TS2 Fidélisation
 - TS2 Prospection commerciale
 - EP Démarchage
- TS1 Marketing industriel
 - EP Business to business
 - EP B to B
- TS1 Marketing relationnel
- TS1 Marketing sportif

Etude de marché

- TS1 Benchmarking
 - EP Référenciation
 - TA Gestion de la qualité (7120)
- TS1 Comportement du consommateur
 - EP Panel/ de consommateurs
 - TA Psychologie (51)
- TS2 Prescripteur
- TS1 Part de marché
 - TA Concurrence (141)
- TS1 Segmentation
 - TA Sociologie (51)
- TS1 SWOT
 - EP Strength weakness opportunity threat

Innovation commerciale

- TA Innovation (23)

Relation client

- TA Endettement des ménages (3410)

- TS1 Gestion de la relation client
 - EP Customer Relationship management
 - EP CRM
- TS1 Médiateur bancaire
- TS1 Protection du consommateur
 - EP Médiation

Stratégie marketing

- EP Marketing opérationnel/
- TS1 Marketing mix
 - EP Mix marketing
 - TA Communication externe
 - TA Politique de produit
 - TS2 Politique de prix
 - TS2 Stratégie de distribution
 - EP Politique d'implantation
 - EP Réseau de distribution
- TS1 Plan marketing
 - EP Plan de développement commercial
- TS1 Politique commerciale

TT Vente

- TS1 Point de vente
 - TA Commerce (615)
 - TS2 Animation commerciale
 - TS2 Distribution des produits
 - TS2 Merchandising
- TS1 Promotion commerciale
 - TS2 Publicité sur le lieu de vente
 - EP PLV
- TS1 Technique de vente
 - TS2 Démarchage téléphonique
 - EP Call center
 - EP Phoning
 - TS2 Publipostage
 - EP Mailing
 - EP E-mailing
- TS1 Vente par correspondance
 - EP VPC
 - TA Banque à distance (3233)

Canal de distribution

- TS1 Commerce électronique
 - TA Paiement en ligne (3131)
- TS1 Multicanal
- TS1 Vente directe

Performance commerciale

- EP Force de vente

Politique de produit

- EP Gamme de produits
- EP Prix de vente
- EP Produit nouveau
- TA Marketing mix
- TA Produit bancaire (3233)
- TS1 Cycle de vie du produit
- TS1 Vente liée
 - EP Package

CS Finance d'entreprise (714)

TT Finance d'entreprise

Analyse financière

- TS1 Business plan
 - TA Création d'entreprise (712)
- TS1 Equilibre financier
- TS1 Indicateur financier
 - EP Ratio financier
- TS1 Rentabilité
 - EP Profit
 - TA Seuil de rentabilité (713)

Financement de l'entreprise

- TA Augmentation de capital (344)
- TA Emission de valeur mobilière (3442)
- TA Risque client (314)
- TS1 Autofinancement
 - TS2 Capacité d'autofinancement
- TS1 Capital investissement
 - EP Private equity

- TA Gestion d'actifs (345)
- TS2 Capital développement
- TS2 Capital risque
 - EP Venture capital/
- TS1 Compte courant d'associé
- TS1 Coût du capital
- TS1 Endettement de l'entreprise
 - TA Crédit aux entreprises (314)
 - TS2 Crédit interentreprises
 - TS2 Effet de levier
- TS1 Escompte
- TS1 Fonds propres
- TS2 Sous-capitalisation
- TS1 Loi Dailly
 - TA Cession de créance (213)
 - TA Crédit aux entreprises (3140)

Gestion financière

- TS1 Budget de l'entreprise
- TS1 Ingénierie financière
 - TS2 Défaillance
 - EP Defeasance
 - EP Désendettement de fait
 - TS2 Fusion-acquisition
 - TA Fusion d'entreprises (712)
- TS1 Planification financière
 - TS2 Plan de financement
- TS1 Structure financière de l'entreprise

Trésorerie

- TS1 Flux de trésorerie
 - EP Cash flow
 - TS2 Tableau de financement
 - EP Tableau de flux de trésorerie
- TS1 Gestion de trésorerie
 - EP Cash management

Valeur de l'entreprise

- TS1 Goodwill
 - EP Ecart d'acquisition
 - EP Survaleur
 - TA Gestion actif-passif (323)

CS Economie sociale et solidaire (716)

Economie sociale et solidaire

- EP ESS
 - TA Banque mutualiste ou coopérative (324)
 - TS1 Association
 - EP Association loi 1901
 - TS2 Association culturelle
 - TS2 Association d'Alsace-Lorraine
 - TS2 Association diocésaine
 - TS2 Association sportive
 - TA Sport (52)
 - TS2 OGECE
 - EP Organisme de gestion des établissements
- catholiques*
- TA Enseignement privé (51)
 - TS2 Secte
 - TS1 Congrégation
 - TS1 Coopérative
 - TS2 Coopérative de consommation
 - TS2 Coopérative de production
 - TS1 Fondation
 - TS1 Fonds de dotation
 - TS1 Insertion par l'activité économique
 - EP Chantier d'insertion
 - EP Entreprise d'insertion
 - EP Entreprise intermédiaire
 - TS1 Mutuelle
 - TA Mutuelle 45 (325)
 - TA Mutuelle d'assurance (325)
 - TS1 ONG
 - EP Organisation non gouvernementale
 - TS1 Organisme sans but lucratif
 - EP OSBL
 - TA Comité d'entreprise (54)
 - TA Monnaie complémentaire (31)
 - TA Organisation syndicale (54)
 - TA Parti politique (201)

Annexe 4 : Note de procédure pour JLB-Net

NP : Mise à jour du thesaurus dans JLB-Net – Rédacteur : Loraine KELLER



Note de procédures : Mise à jour du thesaurus dans JLB-Net

Service : Documentation générale BPCE SA

Date de diffusion : 05 septembre 2013

Rédacteur : Loraine Keller

Contexte :

La refonte du thesaurus, menée par l'équipe des documentalistes durant les mois de juillet et août 2013, devra dans les semaines qui suivront sa validation donner lieu à une série de modifications sous le logiciel de SIGB JLB-Net, afin de faire concorder les deux langages d'indexation.

La présente note de procédure a pour but de faire le point sur les différents cas qui peuvent être rencontrés et sur les étapes à suivre pour chacun d'eux. Elle constituera par la suite un élément de référence pour l'équipe, et un support de formation à destination des nouveaux arrivants dans le service.

Contenu

1. Supprimer un terme	2
2. Ajouter un terme.....	4
3. Modifier des relations.....	4
4. Réattribuer les notices	5

1. Supprimer un terme

Une campagne de suppression des doublons de l'index des mots-clés issus de DIP a déjà permis d'éliminer près de 240 termes, auxquels s'ajoutent 250 descripteurs obsolètes issus de Techlib, se rapportant aux produits et services de la Caisse d'Epargne.

Dans **Thesaurus > Lexique**, sélectionner le lexique *Mots clés*, puis suivant les cas :

- Le terme à supprimer n'indexe aucune notice et ne participe à aucune relation, il suffit de double-cliquer dessus pour le supprimer de la liste des mots-clés. Quelle que soit l'opération effectuée dans JLB, une boîte de dialogue apparaît systématiquement, demandant de confirmer l'action demandée en cliquant sur OK.

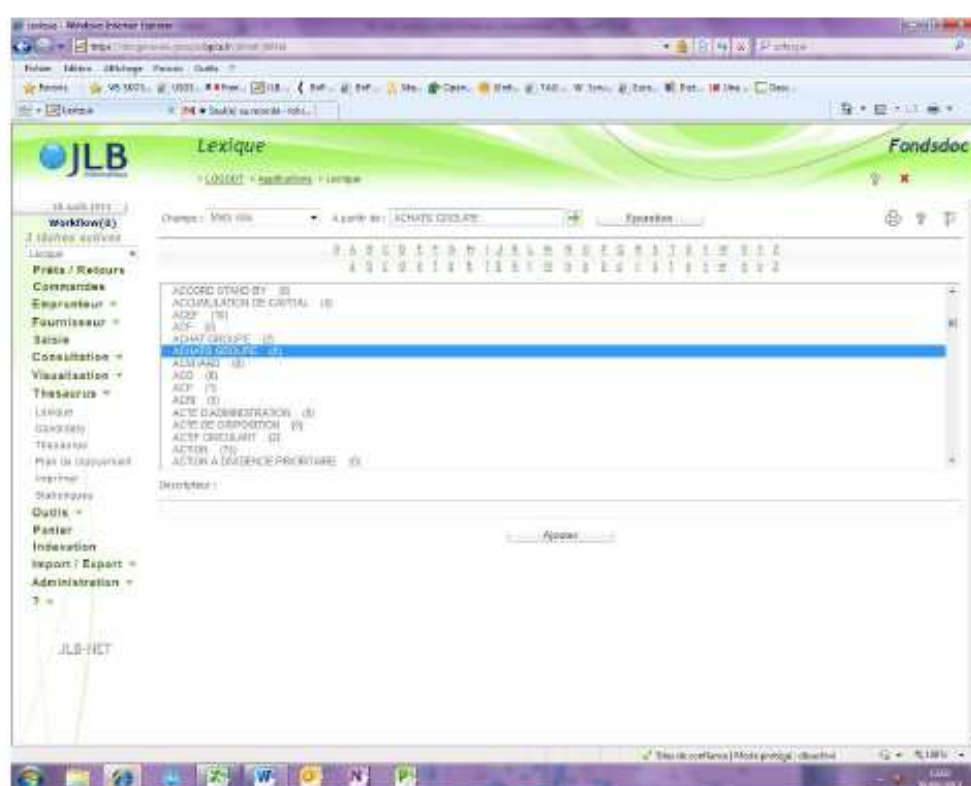


Figure 1

- Le terme n'indexe aucune notice mais possède des relations : celles-ci doivent d'abord être supprimées dans l'onglet **Thesaurus**, rubrique **Thesaurus** (toujours sélectionner l'index des *Mots clés*). De la même façon, on supprime une relation en double-cliquant dessus. Cliquer ensuite sur le bouton *Valider les relations* pour confirmer l'opération. Puis revenir à la rubrique Lexique et procéder comme expliqué ci-dessus¹.

¹ Il arrive qu'un bug empêche de supprimer un terme alors qu'aucune relation n'apparaît dans le module de visualisation du thesaurus. Dans ce cas, supprimer ces relations fictives en double-cliquant de la même façon en haut de chaque case (vide). Cette action semble régler le problème et il est ensuite possible de supprimer le terme.

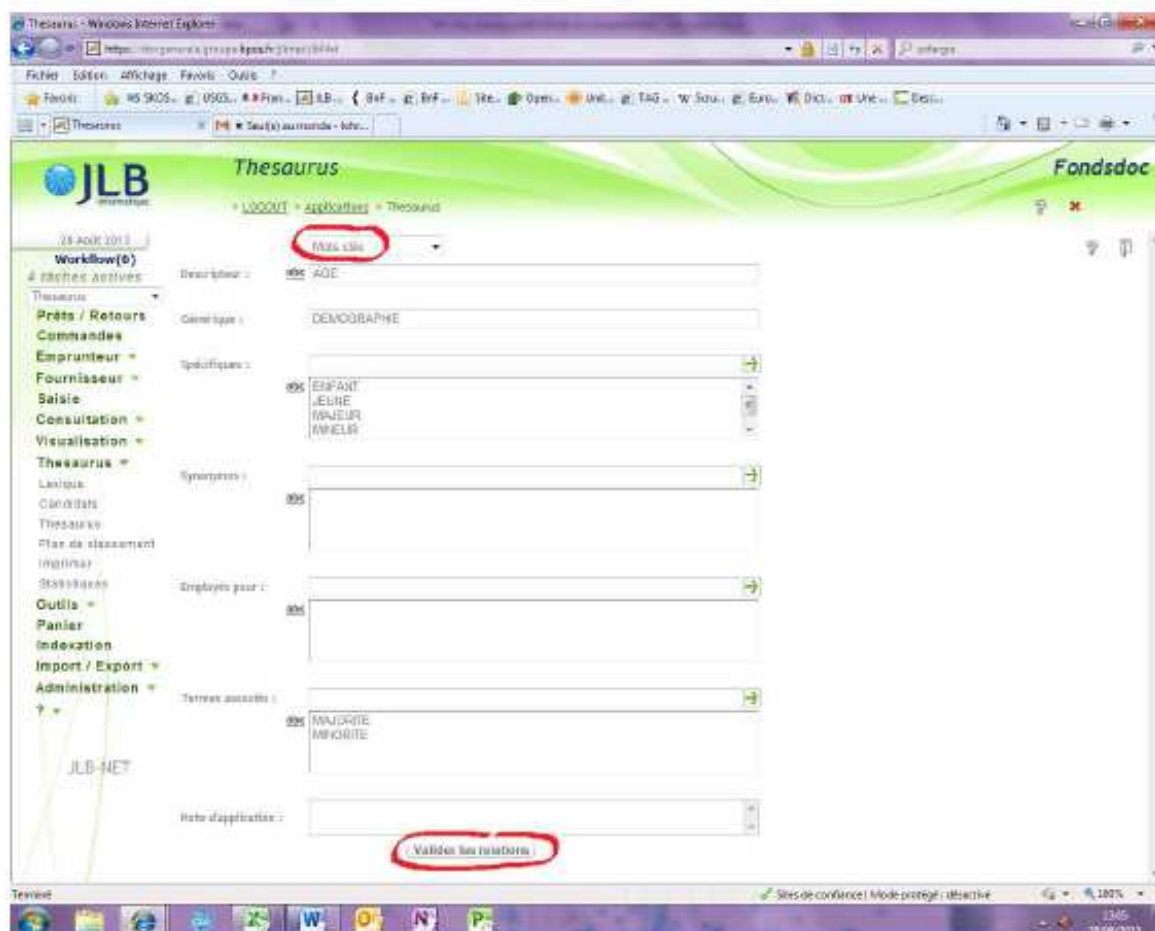


Figure 2

- Le terme à supprimer indexe des notices : si un autre terme a été choisi pour le remplacer, il faut lui réattribuer ces notices (Cf. point 4 Réattribuer les notices). Si aucun autre descripteur ne le remplace, supprimer uniquement ce terme dans l'indexat des notices concernées (en principe très peu). Une fois l'une ou l'autre de ces étapes réalisée, supprimer le terme comme expliqué précédemment.

Attention : Avant de supprimer un terme de l'ancien thesaurus qui n'a pas été conservé dans un domaine, toujours s'assurer qu'il ne l'a pas été dans un autre domaine (utiliser éventuellement la fonction "recherche" de Word dans le thesaurus : Ctrl + F). S'il a été réaffecté, modifier les relations si nécessaire.

2. Ajouter un terme

Pour ajouter un terme au thesaurus sous JLB, la démarche comporte deux étapes :

a) Création du mot-clé

Dans l'onglet **Thesaurus**, rubrique **Lexique**, choisir la liste *Mots-clés*. Puis, dans le champ de saisie libre, sous l'encadré d'affichage du lexique, taper le terme à ajouter. Cliquer sur "Ajouter" pour créer le nouveau terme dans le lexique. (Cf. Figure 1Figure 1)

b) Intégration du nouveau mot clé au thesaurus

Pour attribuer un rôle et des relations au mot clé ainsi créé, se rendre ensuite dans la rubrique **Thesaurus**. Pour créer la notice du mot-clé, choisir la liste *Mots clés* et sélectionner le terme à modifier. Il ne reste plus qu'à sélectionner, selon les cas, un terme générique (TG ou TT), un ou des termes spécifiques (TS), les synonymes (EP) et termes associés (TA) éventuels. Bien sûr, tous les termes avec lesquels vous voudrez mettre en place une relation devront préexister, ou avoir été créés au préalable dans le lexique des mots-clés, comme indiqué ci-dessus. (Cf. Figure 2)

c) Adaptation de la structure du thesaurus BPCE à JLB

JLB-Net ne gère pas, au moment où nous rédigeons, autant de niveaux de hiérarchie que le prévoit le thesaurus BPCE. La fonction de "terme de tête", en particulier, n'est pas proposée. Les Top Terms (TT) seront donc traités comme n'importe quel autre terme générique. Les seules relations qui lui seront attachées sont celles qui le concernent directement : employés pour (EP), termes associés (TA) et termes spécifiques directs (TS1).

Si nécessaire, on pourra éventuellement traduire la relation plus lâche qui lie un top term à un autre terme générique par une relation associative, afin de la faire apparaître dans le thesaurus sous JLB (mais pas systématiquement).

3. Modifier des relations

Pour modifier une relation, quelle qu'elle soit, il est nécessaire de procéder en deux étapes dans JLB :

a) Supprimer une relation obsolète

Dans **Thesaurus** > **Thesaurus**, sélectionner la liste *Mots clés*, choisir le terme générique de la relation et supprimer celle-ci, puis cliquer sur *Valider les relations*.

b) Configurer une nouvelle relation

Dans un second temps, si une autre relation doit remplacer la précédente, ouvrir à nouveau la fiche du générique de la relation, puis paramétrer la nouvelle relation.

Exemple :

Le terme "défaillance d'entreprise" était un non-descripteur, employé pour "entreprise en difficulté". Suite à la refonte, il a été sélectionné comme descripteur et devient un terme spécifique de "entreprise en difficulté".

- La première étape consiste, dans la notice du mot-clé "entreprise en difficulté" à supprimer l'employé pour "défaillance d'entreprise". Cliquer sur le bouton *Valider les relations* pour enregistrer la suppression.
- Ensuite, dans la même notice, sélectionner comme spécifique le mot clé "défaillance d'entreprise" (sous lequel l'ancienne mention EM n'apparaît plus). Cliquer à nouveau sur *Valider les relations*.
- On peut désormais vérifier dans la notice du mot-clé "défaillance d'entreprise" que le terme "entreprise en difficulté" est apparu dans le champ *Générique*, confirmant que la nouvelle relation a bien été prise en compte.

NB S'il s'agit d'une relation associative (termes associées : TA), l'opération doit être effectuée des deux côtés, car la suppression n'est pas automatique. De la même façon, pour créer une relation associative, il est nécessaire de la paramétrer dans chaque notice de descripteur concernée, car la réciproque n'est pas automatique.

4. Réattribuer les notices

Quand un terme indexant déjà des notices a été remplacé, il est nécessaire, une fois le nouveau descripteur créé dans le thesaurus, d'opérer la réaffectation des notices. Pour cela, on utilise la fonction "traitement par lot" de JLB-Net.

- 1) Dans **Consultation > Formulaire simplifié**, sélectionner le Mot-clé à supprimer pour isoler les notices concernées.
- 2) Dans la liste des notices obtenue, sélectionner toutes les notices et cliquer sur le bouton "Modifier en série", au-dessus du tableau au milieu de l'écran.



Figure 3

- 3) Dans le formulaire **Modification en série**, sélectionner d'abord dans la liste l'index *Mots clés*. Dans le champ *Ancienne valeur*, sélectionner le terme à supprimer. Puis, dans le champ *Nouvelle valeur*, sélectionner le descripteur par lequel vous souhaitez le remplacer.
- 4) Cliquer que le bouton *Lancer la modification en série*, en bas du formulaire. Suivant le nombre de notices à traiter, l'opération peut prendre de quelques secondes à plusieurs minutes.

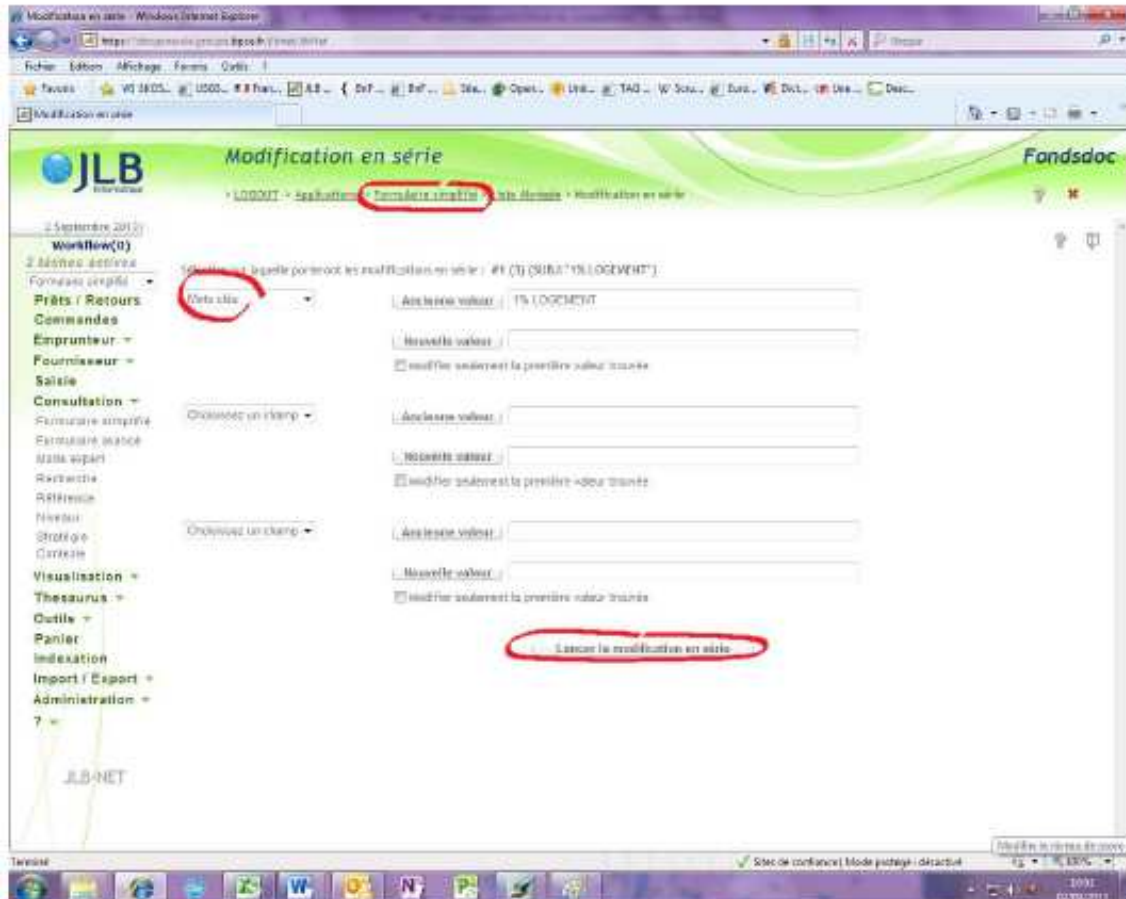


Figure 4

- 5) Revenir au **Formulaire simplifié** en cliquant sur le lien présent dans le fil d'Ariane, en haut de l'écran. Relancer la recherche sur le mot clé à supprimer pour s'assurer qu'il n'indexe plus aucune notice.
- 6) Revenir ensuite dans l'onglet **Thesaurus > Lexique** pour supprimer le terme, qui doit désormais être suivi d'un zéro (il peut être nécessaire de rafraîchir la page en relançant la recherche sur ce mot clé).

Annexe 5 : Guide utilisateur pour ThManager

Guide utilisateur du logiciel ThManager – Rédacteur Loraine Keller – 6 Septembre 2013



Guide d'utilisateur ThManager : logiciel de gestion de thesaurus



Sommaire

Présentation du logiciel	2
1. Fonctionnalités : pourquoi utiliser un logiciel ad hoc ?	2
Prise en main du logiciel	2
1. Ouvrir le logiciel et accéder au thesaurus BPCE	2
2. Visualisation du thesaurus.....	4
Enrichissement du thesaurus	4
1. Créer un concept.....	4
2. Attribution des termes exprimant le concept : <i>Attributes</i>	6
3. Configuration des relations entre les concepts : <i>Relations</i>	7
Mise en place de la structure du thesaurus	8
1. Adaptation de la structure du Thesaurus BPCE à la logique de ThManager	8
Opérations de suivi	9

Ressources :

- ✓ Plus d'informations sur le site de l'éditeur : <http://thmanager.sourceforge.net/>
- ✓ Un guide utilisateur (en anglais) est disponible dans le logiciel à partir de l'onglet **Help** en haut à gauche, et également téléchargeable sur le site.
- ✓ Informations en français sur le logiciel à cette adresse : <http://georezo.net/wiki/main/logiciels/thmanager>

Présentation du logiciel

ThManager 2.0 est un logiciel de création et gestion de thesaurus développé par le Laboratoire des Systèmes d'Information Avancés de l'Université de Saragosse, Espagne (<http://iaaa.cps.unizar.es/>). Il a été choisi essentiellement car la licence est libre et gratuite (Open source) Néanmoins, la licence GNU Licence Publique Générale Limitée permet d'intégrer cette application à un logiciel de licence commerciale (pas de clause de *copyleft*).

D'autre part, ThManager génère des thesaurus en format SKOS (Simple Knowledge Organization system), langage correspondant aux standards du W3C (World Wide Web Consortium) concernant les vocabulaires contrôlés et structurés. Cela garantit la meilleure interopérabilité avec tout autre logiciel qui nécessiterait l'utilisation du thesaurus à l'avenir.

1. Fonctionnalités : pourquoi utiliser un logiciel ad hoc ?

Dans le cadre de la mission de refonte du thesaurus de BPCE, menée à l'été 2013, il est apparu nécessaire de compléter les outils alors disponibles pour l'exploitation du thesaurus.

En effet la version électronique sous Word (présentation thématique hiérarchisée) et son implémentation sous JLB (présentation alphabétique hiérarchisée) n'offrent pas des possibilités suffisantes pour l'affichage, l'exploration et l'export éventuel du thesaurus.

De plus, ThManager permet de mettre en place une navigation par arborescence plus proche des niveaux de hiérarchie effectivement prévus par cette refonte, inscrite dans le cadre de la nouvelle norme ISO 25964:1 Information et documentation – Thésaurus et interopérabilité avec d'autres vocabulaires (août 2011).

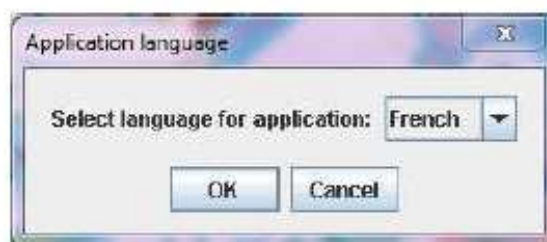
Prise en main du logiciel

1. Ouvrir le logiciel et accéder au thesaurus BPCE

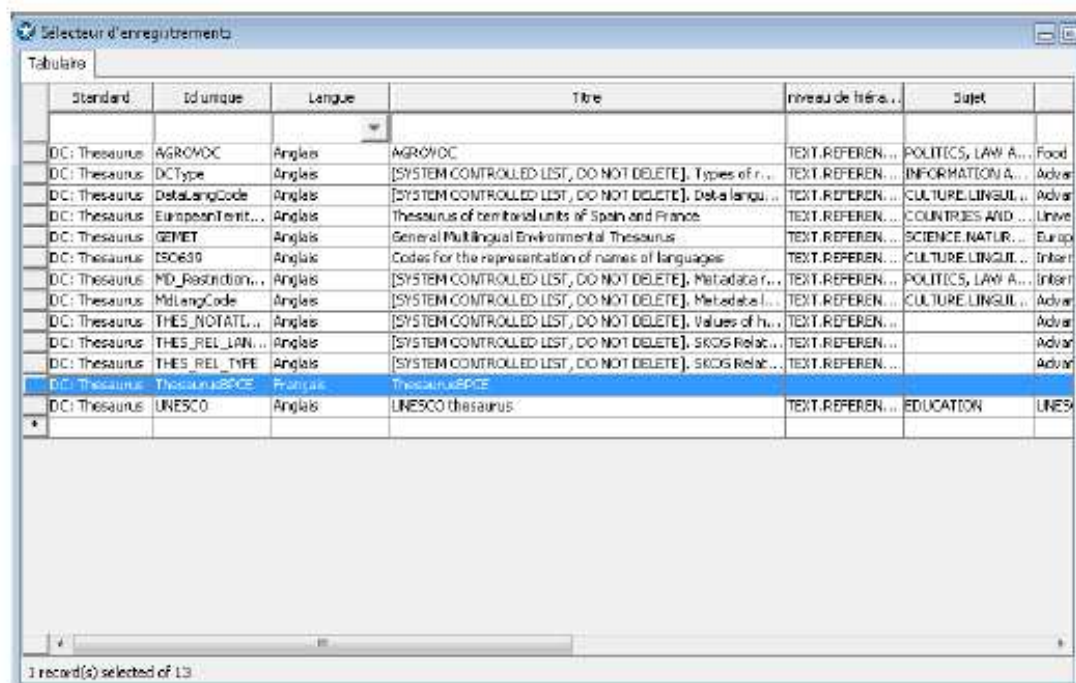


Pour ouvrir le logiciel, cliquer sur l'icône  dans le menu **Démarrer > Tous les programmes > Applications**.

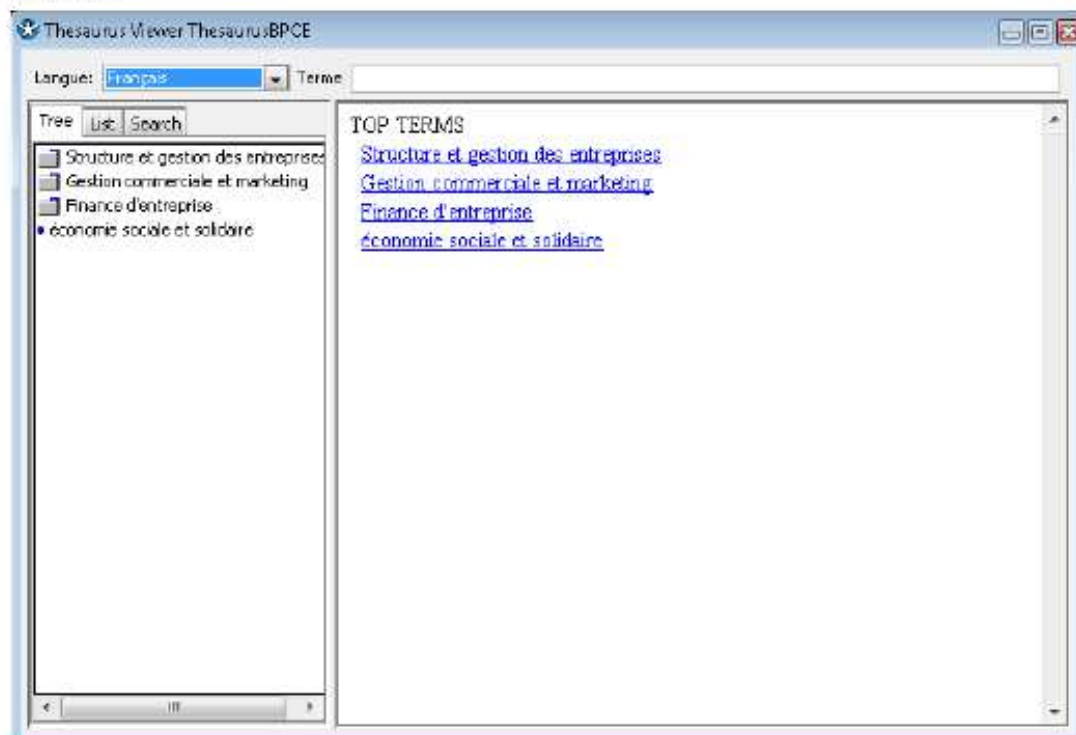
Dans la fenêtre **Application language**, sélectionner la langue *French* puis cliquer sur *OK* :



L'application **Sélecteur d'enregistrement** apparaît à l'ouverture du logiciel, permettant d'accéder à différents thesaurus préprogrammés. Pour accéder au thesaurus de BPCE, double-cliquer sur la ligne concernée.



L'application **Thesaurus Viewer** s'ouvre alors (il peut être nécessaire d'agrandir la fenêtre).



Les différentes applications sont accessibles de plusieurs manières :

- Depuis l'onglet **Dossier** en haut à gauche de la fenêtre ;
- Grâce aux icônes de chaque application, dans la barre d'accès rapide placée en-dessous du menu principal ;
- L'onglet **Fenêtre** permet de basculer rapidement entre les différentes applications déjà ouvertes.

2. Visualisation du thesaurus



L'application **Thesaurus Viewer** permet de naviguer dans le thesaurus, grâce à un réseau de liens hypertexte. Pour commencer la navigation, la liste des *Top Terms* est proposée (correspondant dans le thesaurus BPCE aux champs sémantiques). En cliquant sur l'un d'eux, on accède aux termes de têtes (TT) qui le composent (définis ici comme des spécifiques des Top concepts), et ainsi de suite.

Au fur et à mesure de la navigation au sein du thesaurus, l'arborescence se déploie dans la colonne de gauche, permettant d'accéder directement à différents emplacements. Une icône représentant un fichier indique une arborescence descendante. Pour déployer les spécifiques d'un concepts, il suffit de cliquer sur celui-ci, puis sur le signe + qui apparaît alors à gauche du dossier, afin de déployer son arborescence spécifique.

La colonne de droite permet la visualisation de la notice complète d'un concept, ses attributs et ses relations. Chacun des termes apparaissant en bleu sont des liens hypertextes permettant d'accéder aux notices respectives des concepts en relation avec le concept analysé (terme générique, terme(s) spécifique(s) et terme(s) associé(s)).

Enrichissement du thesaurus



L'application **Thesaurus editor** permet d'enrichir le thesaurus en créant ou modifiant des notices de concept.

1. Créer un concept

Créer un concept signifie en premier lieu lui attribuer un identifiant unique, ou URI (*Uniform Resource Identifier*), qui permettra à l'ordinateur de toujours le reconnaître. Cela permet, par exemple, de pouvoir très facilement changer le terme préférentiel (ou descripteur) traduisant ce concept, sans bouleverser les relations déjà en place.

Cela signifie également qu'il n'existe pas de fiche spécifique pour les non-descripteurs (EP), puisque ceux-ci n'existent qu'en tant que termes alternatifs pour représenter un concept. On ne peut de ce fait, avec ce logiciel, effectuer de recherche sur les termes non préférentiels.

2. Attribution des termes exprimant le concept : *Attributes*

Une fois la notice d'un concept créée, la première étape est de lui assigner des termes le représentant :

- Un descripteur unique ou terme préférentiel (*Preferred label*), noté TRA_ ;
- Eventuellement, un ou plusieurs synonymes non préférentiels (*Alternative label*), notés SYN ;
- Il est également possible, dans cette même partie de la notice, d'ajouter une définition (*Definition*) et/ou une note d'application (*Scope note*), notés DEF et SN.

URI: Mfo://ThesaurusPCE/Concept#91

ATTRIBUTES

Top concept

Type	Language	Value
TRA_	Français	SWOT
SYN	Français	Strength weakness opportunity threat
DEF	Français	Force faiblesse opportunité menace
*		

Type: Language:

Value:

Étapes :

- 1) Dans le premier tableau, cliquer dans la case blanche en dessous de l'étiquette **Type**. Une nouvelle ligne apparaît.
- 2) En-dessous du tableau, dans la liste déroulante **Type**, sélectionner *Preferred label*.
- 3) Dans la liste déroulante **Language**, à droite, sélectionner *Français*¹. Il est indispensable de choisir la langue pour pouvoir enregistrer le terme.
- 4) Dans le cadre de saisie libre **Value**, en-dessous, saisir le libellé du terme préférentiel choisi pour ce concept.

Attention : Si l'initiale du terme à enregistrer comporte un accent, il est important de ne pas utiliser la majuscule pour permettre de le retrouver lors d'une recherche. Ainsi, la recherche sur une chaîne de caractères permet de faire remonter celle-ci quelle que soit sa position dans le terme descripteur (index alphabétique et index permuté).

Ex. : Concept#91 « étude de marché » (et non « Etude de marché »)

- 5) Procéder de la même façon pour ajouter un ou plusieurs synonymes, en sélectionnant cette fois *Alternative label*. Un enregistrement spécifique doit être créé pour chaque terme.
- 6) Vous pouvez enfin définir une définition, en sélectionnant *Definition* dans la liste **Type**, ou encore une note d'application, en sélectionnant *Scope note*.

¹ Par commodité, le français est la seule langue qui a été paramétrée pour ce thesaurus. Elle sera donc sélectionnée y compris dans les cas où l'équivalent anglais a été retenu, que ce soit en terme préférentiel ou non.

3. Configuration des relations entre les concepts : *Relations*

Les relations entre les concepts se créent de façon très simple, en déclarant le type de relation qui relie un concept à un autre, puis en sélectionnant le second concept dans une liste. Il est donc nécessaire d'avoir au préalable créé les notices de chacun des concepts à relier.


3 types de relations sont proposés ici :

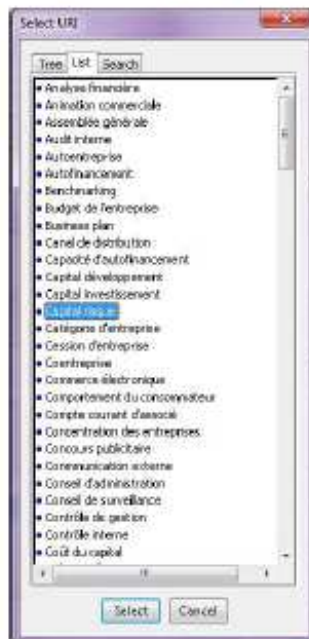
- Du concept à son générique (TG) : *Broader*, noté BT
- Du concept à son spécifique (TS) : *Narrower*, note NT
- Du concept à son terme associé (TA) : *Related*, note RT

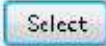
Type	URI
BT	http://thesaurusPCE/Concept#103
NT	http://thesaurusPCE/Concept#105
NT	http://thesaurusPCE/Concept#106
RT	http://thesaurusPCE/Concept#74
RT	http://thesaurusPCE/Concept#124
*	

Type LRI 


Etapes :

- 1) Dans le tableau du bas, cliquer sur la case blanche en dessous de **Type** pour créer la première relation.
- 2) Dans la liste défilante **Type**, sélectionner le type de relation que vous souhaitez définir, comme expliqué plus haut.
- 3) Pour sélectionner le concept à associer, cliquer sur le bouton  pour accéder à la fenêtre *Select URI* présentant la liste des termes préférentiels déjà créés.




- 4) Choisir l'onglet **List**, sélectionner le terme recherché et cliquer sur . L'URI du concept sélectionné apparaît dans le tableau des relations, en face du sigle représentant le type de relation défini.

NB Il est toujours possible de modifier la relation, en cliquant sur la ligne à modifier, puis en sélectionnant une autre valeur dans la liste déroulante **Type**, ou en sélectionnant un autre concept.

- 5) Continuer ainsi de suite, pour créer l'ensemble des relations nécessaires.
- 6) Le logiciel ne crée pas automatiquement la réciproque de la relation. Il ne faut donc pas oublier, pour chaque relation configurée, d'aller créer la relation inverse (pour *Broader* et *Narrower*) ou équivalente (pour *Related*) dans la notice du second concept.
- 7) Ne pas oublier d'enregistrer la notice du concept en cliquant sur l'icône **Save**  en haut de la colonne de gauche. Le terme préférentiel est alors généré dans la liste alphabétique présentée dans cette colonne.

NB Il arrive qu'un bug se produise au moment de l'enregistrement (un rectangle gris apparaît alors souvent sur l'écran). Dans ce cas, le fait de passer d'une fenêtre d'affichage à l'autre permet de rafraîchir le formulaire de l'éditeur.

- 8) Vérifier que la relation a bien été créée, soit directement dans le **Thesaurus Viewer**, soit en cliquant sur l'icône **Preview hierachy**  (à droite de **Save**), qui permet d'ouvrir une fenêtre de prévisualisation.

NB La relation n'apparaîtra pas si elle n'a pas au préalable été créée des deux côtés. Cette opération demande beaucoup de rigueur, car ce logiciel ne dispose pas d'un système d'alerte prévenant l'opérateur qu'une relation n'a pas été paramétrée de façon symétrique.

Seule la relation associative apparaît dès qu'elle a été enregistrée dans la fiche de l'un des deux concepts ainsi reliés. Néanmoins, il ne faut pas oublier de la reproduire dans la seconde fiche.

Mise en place de la structure du thesaurus

1. Adaptation de la structure du Thesaurus BPCE à la logique de ThManager

Pour correspondre à la structure adoptée au cours de la refonte du thesaurus BPCE et permettre une navigation simple et la plus pertinente, il a été décidé de ne pas faire apparaître les domaines (correspondant à l'origine aux sections du plan de classement).

Il en résulte que :

- La navigation débute au niveau des champs sémantiques, enregistrés comme concepts portant un URI de type *http://ThesaurusBPCE/SemanticField#xx*, auxquels on a attribué la fonction de top concept. Pour cela, il suffit de cocher la case **Top concept**, située en haut de l'encadré *Attributes* dans la fiche du concept.

L'application génère automatiquement une liste de top terms issus de ces top concepts. Ces top terms s'affichent dès l'ouverture du module de visualisation du thesaurus.

Ils sont accompagnés d'une note d'application (*Scope note*) précisant à quel sous-domaine du plan de classement ils correspondent, ainsi que la mention " Ne pas utiliser comme descripteur".

TERM http://ThesaurusBPCE/SemanticField#1 Structure et gestion des entreprises * SCOPE NOTE: Correspond au champ sémantique 712. NE PAS UTILISER COMME DESCRIPTEUR
BROADER TERMS TOP TERMS
NARROWER TERMS Entreprise Gestion de l'entreprise

- Les Top Terms (ou termes de tête, notés TT) du thesaurus sont définis comme termes spécifiques (*Narrower*) des Top concepts : il s'agit du 2nd niveau de navigation.

NB Lorsqu'un même terme cumule dans le thesaurus papier les fonctions de champ sémantique (CS) et de top term (TT), il doit être enregistré une seule fois (une URI par concept) et défini comme un Top concept, afin d'apparaître dès le 1^{er} niveau, en Top term. Dans ce cas, sa note d'application précisera qu'il peut être utilisé en tant que top term.

- L'ensemble des descripteurs placés sous un Top Term (qu'ils soient ses spécifiques directs ou des descripteurs dépendants du même sous-domaine) sont définis comme ses termes spécifiques ou *Narrower* : il s'agit du 3^{ème} niveau de navigation. Le logiciel ThManager permet ici d'enrichir la présentation du thesaurus sous JLB, en faisant remonter les Top Terms définis pour affiner la structure du thesaurus.²
- Le 4^{ème} niveau de navigation présente de façon traditionnelle un terme générique (TG) et la liste de ses spécifiques (TS).
- Les derniers niveaux de navigation concernent les termes spécifiques, généralement un voire deux niveaux de spécification supplémentaire au maximum.

Opérations de suivi

La création et l'enrichissement du thesaurus nécessitent des opérations de suivi strictes, indispensable pour garantir la cohérence la fiabilité du thesaurus géré sous cette application.

1. Suivi des identifiants de concepts (URI)


Pour faciliter l'enrichissement du thesaurus, surtout si plusieurs utilisateurs sont amenés à le faire, il est important d'assurer un suivi strict des URI déjà enregistrées. Un tableau Excel

² JLB-Net ne permet pas de définir le rôle de Top Terms. Il a donc été choisi d'ignorer ce niveau de hiérarchisation dans JLB, pour ne pas rendre trop complexe l'affichage du thesaurus en multipliant les termes spécifiques. Au contraire, cela ne pose pas de problème avec ThManager.

a été créé à cet effet, nommé "URI_ThManager", et placé dans le dossier :
 O:\DISG\DISG_DOC\InfoDoc\PROJETS 2013\THESAURUS\Outils projet\ThMANAGER

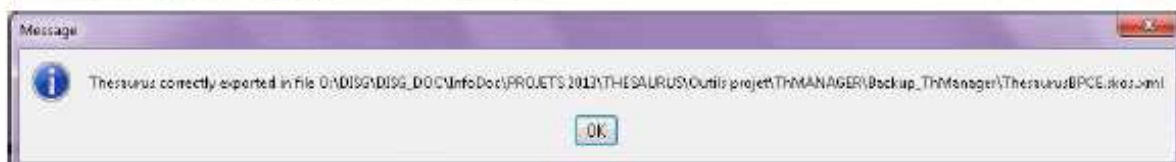
2. Sauvegarde du thesaurus

Ce dossier contient également l'export du thesaurus. Il est indispensable de procéder à une sauvegarde complète du thesaurus (sous format SKOS) après chaque enrichissement dans l'application ThManager.

Pour cela, avant de fermer l'application, cliquer sur l'icône  ou dans l'onglet **Dossier**, sur la fonction *Export thesaurus*. Dans la fenêtre qui s'ouvre alors, sélectionner (si ce n'est pas déjà proposé) le dossier dans lequel vous souhaitez enregistrer le fichier. Enfin, cliquer sur le bouton *Export thesaurus* afin de valider l'action.




Vous devez voir apparaître ce message :



Dans le dossier où est sauvegardé le thesaurus, vous devez trouver deux fichiers sous la forme :

Nom	Modifié le	Type	Taille
 ThesaurusBPCE.MD.DC	06/09/2013 11:17	Document XML	2 Ko
 ThesaurusBPCE.skos	06/09/2013 11:17	Document XML	81 Ko

Pour plus de sûreté, il est préférable de faire une seconde sauvegarde régulièrement sur un autre emplacement ou sur une clé USB.

Si l'on souhaite importer le thesaurus dans un logiciel ThManager nouvellement installé sur un poste, il suffit à l'inverse de cliquer sur l'icône , ou dans l'onglet **Dossier**, sur la fonction *Import thesaurus*. Dans la fenêtre qui s'ouvre alors, sélectionner le fichier SKOS contenant le "ThesaurusBPCE", puis cliquer sur le bouton *Import thesaurus*.

Annexe 6 : Répertoire de thesauri en ligne

THESAURUS GENERALISTES

- ECLAS - Catalogue des bibliothèques de la Commission Européenne :
http://ec.europa.eu/eclas/F/IMGBMTHTS6DXCS2ULV2LJQ1VRGMC6GDUIXM33XTCB3MH3BJ61V-72806?func=file&file_name=thesfre.htm
- MOTBIS - Ministère de l'éducation nationale : <http://www.cndp.fr/thesaurus-motbis/site/>
- EuroVoc - Thesaurus multilingue de l'Union Européenne :
<http://eurovoc.europa.eu/drupal/?q=fr>
- UNBIS - Système d'information bibliographique de l'ONU :
[http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf/\\$\\$searchf?OpenForm](http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf/$$searchf?OpenForm)
- UNESCO : <http://databases.unesco.org/thesfr/>

SECTEUR ECONOMIQUE, FINANCE...

- DELPHES - Chambre de Commerce et d'Industrie de Paris :
http://static.canalblog.com/storagev1/brich59.canalblog.com/docs/Thes_Delp_Matiere_Hiera2004.pdf
- Terminologie multilingue du FMI : <http://www.imf.org/external/np/term/fra/index.htm>

POLITIQUE, INSTITUTIONS

- CERES : Council of Europe REtrival System - Liste bilingue de mots-clés sujets :
http://www.coe.int/t/dgal/dit/ilcd/tools/ceres_FR.asp
- TAG - Thesaurus de l'activité gouvernementale (Québec) :
<http://www.thesaurus.gouv.qc.ca/tag/accueil.do>
- Thésaurus d'Information publique - La Documentation française, Collection des discours publics : <http://www.vie-publique.fr/discours/thesaurus-collection-discours-publics.html>

POPULATION, TRAVAIL, SECURITE SOCIALE...

- AISS - Association Internationale de la Sécurité Sociale (thesaurus multilingue 2006) :
<http://www.issa.int/fre/Media/Files/Advanced-Search/Thesaurus/ISSA-Social-Security-Thesaurus>
- CISTHES - Thesaurus de la sécurité et de la santé au travail :
http://www.ilo.org/dyn/cisdoc/cisthes_search.ilostart?p_lang=F
- POPIN - Thesaurus multilingue de population, 1993
<http://www.cicred.org/Eng/Publications/Books/ThesaurusF/thesaurusF.htm>

Navigation dans le thesaurus en ligne (3ème version adaptée, 2011) sur le site de l'INED : <https://bibliotheque.web.ined.fr/popin/>

- TEF - Thesaurus Européen de la Formation du CEDEFOP :
Téléchargement : <http://libserver.cedefop.europa.eu/ett/fr/download.htm>
Consultation : <http://www.vocabularyserver.com/ettfr/>
- TESS - Travail, Emploi, Santé, Solidarité : <http://194.199.119.234/TESS/>
- Thésaurus du BIT - Organisation Internationale du Travail :
<http://www.ilo.org/public/french/support/lib/tools/aboutthes.htm>
- Thesaurus de la formation du Centre Inffo - Centre de documentation sur la formation et le travail (CNAM) : <http://docser.cnam.fr:8080/ListRecord.htm?list=table&table=5>

ENVIRONNEMENT

- GEMET - Thesaurus de l'Agence Européenne pour l'Environnement :
<http://www.eionet.europa.eu/gemet/>
- Institut des risques majeurs - Irma Rhône-Alpes : http://www.irma-grenoble.com/05documentation/00bibliotheque_recherchethesaurus.php
- Thésaurus Ecoplanète - Portail d'information du ministère de l'Ecologie et du DD :
http://www.side.developpement-durable.gouv.fr/cda/portal.aspx?INSTANCE=exploitation&PORTAL_ID=medd_P0_R_RechercheFocus_Ecoplanete.xml

URBANISME, LOGEMENT

- doC@UE - Centres d'Architecture, Urbanisme et Environnement :
<http://www.docaue.archi.fr/thesaurus.php?idMenu=outils&sousMenu=thesaurus-aue&thesaurus=aue>
- URBAMET (version 2001 et mises à jour fréquentes) :
<http://portail.documentation.developpement-durable.gouv.fr/notx/Urbanisme/thesaurus/navigation.xhtml>

THESAURUS GEOGRAPHIQUE

- GeoEthno : <http://www.mae.u-paris10.fr/dbtw-wpd/bibliotheque/GEO-DOC-v2.HTM>

Annexe 7 : Schéma de l'activité de production d'une veille documentaire

