

NOUVELLES TECHNOLOGIES ET SYSTÈMES D'INFORMATION : EST-IL POSSIBLE « D'ÉCOUTER » D'AUTRES DISCOURS ? ¹⁰

Marilda Lopes Ginez de Lara

Introduction

Dans le cadre du thème « Réseaux, documents et mémoires », dans lequel l'organisation de la Journée International Mussi 2012 met en valeur, entre autres, les actions de valorisation du patrimoine documentaire et celles visant à favoriser la participation sociale je cherche à explorer la question des langages documentaires, ou systèmes d'organisation des connaissances, dans le cadre du Web sémantique, en considérant les possibilités qu'il ouvre pour la diffusion et la réutilisation des connaissances construites. Je cherche également à introduire une discussion sur la réutilisation de ces vocabulaires d'un point de vue instrumental, en la contrastant en contre point avec l'approche qui cherche à incorporer, dans cet environnement, des manières non conventionnelles d'organiser les vocabulaires.

Actuellement, la plupart des systèmes d'information sont déjà sur Internet, ce qui permet de guider la recherche de documents par des systèmes de classification, des thésaurus ou d'autres vocabulaires en ligne. Cependant, en règle générale,

¹⁰ Originalmente publicado em: LARA, M. L. G. Novas tecnologias nos sistemas de organização do conhecimento: possibilidades de 'escutar' outros discursos ? *In*: II Jornada Científica Internacional da Rede Mussi. **Anais** [...] Rio de Janeiro: IBICT-UFRJ , ICICT-FIOCRUZ, 2012.

Tradução de Gustavo Saldanha.

Revisão de Viviane Couzinet.

chaque système fonctionne selon sa propre sémantique, ce qui nécessite d'effectuer, sur une thématique donnée, des recherches différentes dans chaque système d'information.

En termes génériques, cela peut être comparé à la plupart de nos recherches sur Internet. Leurs résultats font apparaître d'innombrables *sites* reliés entre eux par des liens morphologiques. Ces résultats peuvent être plus précis en fonction des combinaisons syntaxiques utilisées dans l'équation de recherche. Cependant, comme dans le cas précédent, ils proviennent d'institutions différentes et ne se connectent que par les mots. Pour cette raison, les réponses aux recherches varient d'un utilisateur à l'autre en fonction de sa compétence, de sa connaissance des outils et du filtrage de l'information. A partir des ressources de liens et hyperliens du langage HTML, la navigation traditionnelle s'effectue site par site et la récupération est basée sur des coïncidences au niveau morphologique, mais pas au niveau sémantique. Individuellement, chaque utilisateur accepte ou rejette un résultat car il fait des choix basés sur des critères de sens (sémantiques), mais ceux-ci proviennent de ses connaissances particulières et non des ressources offertes par les programmes.

La récupération sémantique est recherchée depuis Paul Otlet et Vannevar Bush, mais, dans le contexte actuel, Tim Berners-Lee est à l'origine de changements majeurs. En 1989, il promeut et défend l'idée d'une puissante structure de connaissances permettant de relier données, documents et informations. Cet auteur a d'abord appelé cette structure le Web (1989), puis le Web sémantique (1998) et enfin les données liées (2006). Comme il le prédisait dans *Weaving the Web* (1999), le langage HTML a permis le réseau hypertextuel de documents, tandis que le RDF - *Resource Description Framework* - et les technologies sémantiques permettront un Web de données interconnectées (Méndez, Greenberg, 2012). Sous cette terminologie étrange pour beaucoup d'entre nous, il y a une gamme de possibilités « Loin d'être une simple proposition purement technologique, la proposition du Web sémantique est

assez large et ambitieuse » (Marcondes, 2012, p.176). Je reviendrais sur ces expressions et concepts tout au long de ce texte.

La différence entre le Web « syntaxique », ou Web hypertexte, et le Web sémantique est que le premier a ses recherches basées sur la coïncidence de caractères (Berners-Lee, 2006) et, pour cette raison, il lie des pages ou des documents en HTML (publie des hypertextes, mais les unités restent isolées d'un point de vue sémantique) ; Le Web sémantique relie des données structurées (Peset, Ferrer-Sapena, Subirats-Coll, 2011, p. 166), reliant les idées.

Notre présentation a pour but de discuter certaines des initiatives liées au thème du Web sémantique, des données interconnectées et de l'interopérabilité, en concentrant et en problématisant la discussion particulièrement en relation avec les vocabulaires sémantiques. Ce travail est le résultat partiel d'un projet de recherche développé lors d'un stage senior à l'Université Carlos III de Madrid avec le soutien du Conseil National de Développement Scientifique et Technologique (CNPq), Brésil.

Le Web sémantique et l'initiative *Linked Open Data*

Le Web sémantique constitue une extension du Web dont nous disposons actuellement. Contrairement à un Web syntaxique basé sur des modèles de formes, le Web sémantique est ancré dans une structure de contenu. Il utilise un modèle qui modélise sémantiquement les données pour permettre leur identification et les relier les unes aux autres selon des propriétés et des classes de propriétés, permettant ainsi la construction de déclarations.

Pour assurer une telle mise en relation, le projet Web sémantique travaille à partir de données ouvertes structurées. L'ouverture les rend disponibles et la structuration conduit à la possibilité de leur relation. Dans la même lignée que le projet Open Access, appliqué à promouvoir une large diffusion de la documentation scientifique, l'initiative *Open Data* (OD), et plus particulièrement *Linked Open Data* (LOD), *Interconnected Open Data*,

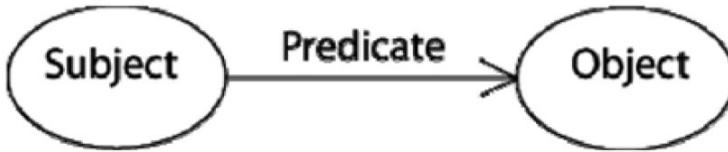
visé l'ouverture des données visant à « préparer l'affichage de contenus utilisables et réutilisables pour le Web sémantique » (Peset, Ferrer-Sapena, Subirats-Coll, 2011, p.166, notre traduction). Les vocabulaires en libre accès mentionnés sont disponibles dans le catalogue du projet *Linked Open Vocabularies* (LOV): <http://labs.mondeca.com/dataset/lov/> (1).

L'initiative découle des propositions de Berners-Lee (2006) et est liée aux recommandations et normes du W3Consortium, une communauté internationale qui développe des normes ouvertes pour assurer la croissance à long terme du Web. Parmi les principes du W3C figurent ceux du partage des connaissances, en les rendant accessibles à tous, « quels que soient leurs hardware, software, infrastructure de réseau, langue, culture, situation géographique ou capacité physique ou mentale » (W3C) (2).

Les données doivent être disponibles avec une licence ouverte ; elles n'ont pas de format propriétaire, c'est-à-dire protégé par un brevet ou un droit d'auteur ; elles doivent nécessairement être structurés (par exemple, une feuille Excell, au lieu d'une image numérisée d'un tableau). Sa base est le standard RDF défini par le W3C (3), un langage de métadonnées doté d'une sémantique qui normalise les descriptions et permet aux agents logiciels d'identifier les concepts.

RDF est un langage adapté à la représentation sur Internet. Avec RDF, on dispose de sources de données, ou métadonnées, c'est-à-dire des données qui font référence à des données. La structure d'une expression en RDF est basée sur des triplets dont chacun est constitué d'une expression qui met en relation un sujet, un objet, un prédicat ou une propriété, identifiés individuellement par des URI – *Uniform Resource Identifier* (W3C).

Figure 1. Graphique d'un triplet



Source: Uniform Resource Identifier (W3C). RDF 1.1. Concepts and Abstract Syntax. <http://www.w3.org/TR/rdf-concepts/>

A travers ces triplets, il est possible de représenter un énoncé reliant chacun des nœuds. Le sujet peut être une personne, une entité à laquelle on se réfère; le prédicat est la propriété que l'on souhaite attribuer au sujet, et l'objet, la valeur de la propriété ou autre ressource avec laquelle la relation est établie (Peset; Ferrer-Sapena; Subirats-Coll, 2011, p.166). On peut représenter en triplet une affirmation comme : « 2^{ème} journées scientifiques du Réseau Mussi sont formées par des chercheurs français et des chercheurs brésiliens ». « La signification est exprimée en RDF » (Berners-Lee, Hendler, Lassila, 2001) et à partir de cette norme il est possible de promouvoir l'interopérabilité.

Dans les bibliothèques numériques, l'interopérabilité fait référence à la possibilité d'effectuer des recherches dans des systèmes d'information hétérogènes (Marcondes, Sayão, 2001, p. 27), mais la rendre opérationnelle se heurte, entre autres, aux normes et formats utilisés par chaque système. « En accédant à ces catalogues sur le Web, l'utilisateur devient prisonnier de ces contextes systémiques et institutionnels spécifiques, avec pratiquement aucune possibilité de naviguer de l'extérieur vers l'intérieur ou de l'intérieur vers l'extérieur de ceux-ci » (Marcondes, 2012, p. 182). Le problème est qu'à la base de l'interopérabilité se trouve l'exigence de partager des normes aux niveaux syntaxiques (formats, protocoles), syntaxique-sémantique (métadonnées) et

organisationnel (règles d'accès, maintenance des collections, etc.), ce qui justifie l'utilisation de RDF.

Pour évaluer l'importance de l'interopérabilité, il est intéressant de rappeler que, bien que le terme ait été mis à l'ordre du jour par la technologie, une finalité similaire, sous un autre vocabulaire, existait déjà depuis les débuts de la Documentation, tant avec Paul Otlet (dans la Classification Décimale Universel), qu'avec Melvil Dewey (dans la Classification Décimale de Dewey) basée sur la recommandation d'attribuer une notation unique ou équivalente aux documents pour représenter leurs sujets dans toutes les unités d'information. Il est également possible de relier le principe d'interopérabilité à des expériences encore plus anciennes, comme les efforts des moines franciscains qui, au XIII^e siècle, se rendaient dans différents monastères avec un catalogue, cherchant à s'assurer qu'ils étaient tous informés de la collection disponible dans chacun des monastères (Moreira, Lara, 2012).

Normes et recommandations pour promouvoir l'interopérabilité entre les thésaurus et autres vocabulaires

Dans le vaste univers des questions sur le Web sémantique, nous soulignons, pour nos besoins, celles liées à l'interopérabilité des systèmes d'organisation des connaissances, tels que les thésaurus et les vocabulaires contrôlés.

La norme ISO 25964-1/2 :2011/2012, *Thesauri and interoperability with other vocabularies*, récemment publiée, donne la priorité aux concepts plutôt qu'aux mots, ce qui est lié au besoin de formalisation pour promouvoir l'interopérabilité. La version de cette norme succède, entre autres, à la BS 8723-5: 2008, *Structured vocabularies for information retrieval: guide: interoperability between vocabularies and other components of information storage and retrieval systems* et à ANSI/NISO Z39.19 : 2005, *Guidelines for the construction, format and management of monolingual controlled vocabularies*, en complément des normes précédent liées.

L'accent mis sur les concepts, malgré les discussions dans le domaine de la Terminologie contemporaine sur le fait que les différences entre les concepts et les mots dépendent fortement des types de discours impliqués (le discours des sciences sociales n'a pas la même forme que celui des sciences dites exactes) et peuvent être interprétées comme visant à éviter les équivalences entre les formes, en veillant à considérer d'abord la charge sémantique exprimée par les termes. La motivation sous-jacente est le besoin de formalisation, afin de fournir les conditions d'interopérabilité, c'est-à-dire que les machines « comprennent » une expression qui relie un signifiant à un signifié.

La norme dispose également de recommandations importantes pour la cartographie des thésaurus et autres vocabulaires similaires, condition essentielle pour leur comparaison et leur réutilisation dans le souci d'économie, d'optimisation des efforts, et de partage. On suppose que la création de vocabulaires doit, ou peut, faire appel à des connaissances déjà existantes. En poursuivant l'interopérabilité et la réutilisation, la norme cherche à fournir les conditions d'interconnexion entre les données sémantiques, comme le suggère Berners-Lee pour le Web sémantique.

Parallèlement à l'ISO et à des entités similaires, W3Consortium a publié, en 2009, sur la base de recherches développées depuis 2002, une recommandation spécifique pour les vocabulaires dans le contexte du Web sémantique et des données ouvertes interconnectées. La recommandation de format SKOS – *Simple Knowledge Organisation System* (4) vise à décrire, partager et relier entre eux les systèmes d'organisation des connaissances sur le web. Il s'agit d'un modèle de données spécifique pour l'encodage de vocabulaires contrôlés et de langages d'indexation basé sur la reconnaissance que « de nombreux systèmes d'organisation de connaissances, tels que les thésaurus, les taxonomies, les listes de classification et de vedettes-matières partagent une structure similaire et sont utilisées dans des applications similaires [...] SKOS capture bon nombre de ces similitudes et les rend explicites pour

permettre le partage de données et de technologies entre applications » (Miles, Bechhofer, 2009). Le modèle à faible coût permet de migrer les systèmes d'organisation des connaissances existants vers des environnements du Web sémantique, ainsi que de développer de nouveaux systèmes basés sur leur partage.

L'acronyme SKOS – *Simple Knowledge Organisation System* –, a été créé à partir d'un autre – KOS – Knowledge Organization Systems, terme utilisé pour englober les différents types de schémas d'organisation des connaissances (Hodge, 2000, cité par Hjørland, 2008 ; Soergel, 2001). Le format SKOS utilise également RDF pour fournir un standard permettant de représenter les systèmes d'organisation des connaissances, de façon lisible par machine.

Suivant le standard *Linked Open Data*, chaque concept est représenté par des URI - *Uniform Resource Identifier* - et, suivant la relation entre les triplets, lié aux autres via des hiérarchies et des associations. Son support est constitué de concepts et non de mots, tout comme la norme actuelle, cherchant à faciliter le partage et l'interopérabilité sémantique, c'est-à-dire la capacité d'échanger des données avec d'autres systèmes. En plus de permettre la réutilisation de vocabulaires déjà créés, le format SKOS peut être utilisé seul ou en combinaison avec des langages de représentation formelle tels que l'*Ontology Web Language* (OWL).

Entre ISO 25964-1/2 :2011/2012 et SKOS, il existe des différences d'objectif. La norme concerne des recommandations sur la structuration et la cartographie des thésaurus et autres vocabulaires. Le format SKOS est cité par la norme, tout comme d'autres formats similaires, mais ils ne fournissent qu'un langage de formalisation pour la publication de vocabulaires sur le Web. Pour clarifier la distinction, Codina et Pedraza-Jiménez (2011, p. 562) déclarent que

[...] un thésaurus représenté dans SKOS reste un thésaurus, mais le fait de disposer d'un langage avec un niveau formel élevé et des exigences logiques comme RDF peut aider non seulement à

développer de meilleurs langages documentaires, mais aussi à faciliter l'interconnexion entre les langages développés et spécifiés à l'aide de SKOS.

L'utilisation du standard et du format SKOS permet de réutiliser des vocabulaires existants.

Problèmes conceptuels dans la formalisation des vocabulaires

La transformation d'un thésaurus ou autre vocabulaire existant au format SKOS est une opération qui ne peut pas être réalisée en privilégiant uniquement l'aspect instrumental de formalisation. Au contraire, la démarche requiert une vision critique, puisque c'est à ce moment là que les problèmes structurants originaux se précisent.

De nombreux thésaurus et vocabulaires contrôlés n'appliquent pas facilement les principes d'organisation logico-sémantique. Ces déficiences sont souvent surmontées par la médiation des personnes, contrairement à ce qui se passe avec les machines, puisque grâce à notre capacité de généralisation, de simplification et d'inférence, nous n'avons aucune difficulté à reconnaître, ou à accepter, les hiérarchies représentées par TG/TS – Terme Générique/ Terme Spécifique –, différents types de relation : relation genre/espèce, relation tout/partie, relation d'exemple. En fait, une grande partie des thésaurus utilisent la hiérarchie pour représenter à la fois les relations entre les concepts qui font référence à des types d'objets, d'entités, de phénomènes (musée / musée historique, musée des arts, musée des sciences), et pour parler de parties d'un objet (voiture / volant, sièges, volant) ou encore de cas ou instances concrets dans lesquels le concept peut être appliqué (université / USP, UFRJ, UnB). De même, les relations associatives, représentées dans les thésaurus par TR – *Related Term*, terme relié –, sont acceptées sans conteste, même si elles représentent des relations entre des concepts soutenus par des principes différents : par exemple, les relations entre processus et

conséquence (inflammation / fièvre), entre processus et produit (tissage / tissu), entre activité et lieu (exploration des sols / exploitation minière), etc., tous représentés par le même indicateur TR.

Cependant, les machines ne font pas d'inférences à moins qu'elles soient programmées pour le faire. Chaque concept doit être identifié par ses attributs, condition pour qu'il soit localisé et que ses relations avec les autres concepts soient définies. Bien que certains thésaurus fournissent ce détail, ce qui n'est pas courant, il est essentiel pour que les machines effectuent l'identification et fassent les inférences qui en résultent.

Ces problèmes sont plus petits et plus faciles à résoudre que ceux liés à une mauvaise organisation. Il n'est pas rare de constater que de nombreux vocabulaires contrôlés et thésaurus sont mal structurés, soit parce qu'ils organisent les concepts de manière aléatoire sans rendre explicites leurs points de départ ou leurs définitions, soit parce qu'ils établissent des équivalences entre termes sur la base de la coïncidence formelle des mots ou encore amalgament, en hiérarchies, des concepts qui ne sont pas liés logiquement ou sémantiquement.

Comme nous venons de le dire, nous voulons souligner que, même si la formalisation des vocabulaires peut être réalisée selon le format SKOS, si l'opération est réalisée sans esprit critique, elle peut entraîner des problèmes de déficiences conceptuelles d'origine. La simple transformation mécanique du KOS en SKOS ne suffit pas à résoudre les problèmes d'interopérabilité, car les problèmes conceptuels d'un langage documentaire doivent être résolus « ailleurs » (Sánchez-Jiménez, Gil-Urdician, 2007, p. 552), c'est-à-dire dans le domaine des principes de sa construction. Ce qui est en jeu, c'est le fait que nous travaillons avec des concepts représentés par des mots, et non avec de simples étiquettes formelles. D'autres problèmes à résoudre sont les différents degrés de spécificité ou de granularité des langages documentaires à comparer dans la cartographie des termes.

Considérations finales

Linked Open Data, standards documentaires pour la construction de systèmes d'organisation des connaissances, format SKOS pour la publication sur Internet corroborent la promotion du Web sémantique et énumère tous ses bénéfices : diffusion, partage de données, réutilisation des connaissances construites, récupération plus efficace.

Il est nécessaire de réaffirmer, cependant, que l'utilisation sans réserve de ces recommandations et technologies peut induire la reproduction de modèles idéologiques douteux qui sous-tendent de nombreux systèmes d'organisation des connaissances. Aucun système organisationnel n'est neutre et aucun d'entre eux ne répond universellement à toutes les questions et à tous les besoins. Un exemple bien connu qui peut illustrer cette situation est le cas de la Classification Décimale de Dewey, qui réserve la plupart des notations de la classe 800 aux langues américaine, anglaise, allemande, française, italienne, espagnole et portugaise, latine et grecque, et suggère d'accueillir en 890 toute la littérature dans d'autres langues (par exemple, les langues non officielles, comme c'est le cas avec la littérature africaine, entre autres). D'innombrables autres exemples pourraient être donnés concernant ce même système de classification.

Dans une critique du livre de Bowker et Star (1999), *Sorting Things Out : Classification and its Consequences*, Hope Olson (2002) montre, en accord avec Foucault, qu'il existe plusieurs manières de « ordonner ». Faisant référence aux origines des classifications, elle souligne que ce qu'il y a de plus significatif dans le syllogisme aristotélicien, qui est à la base de la plupart des systèmes d'organisation des connaissances, c'est que son objectif téléologique est apparemment universel. La recherche en information bibliographique a rendu conventionnels, à partir de cette émergence, les principes d'exclusivité mutuelle, de téléologie et de hiérarchie. Cependant, lors de la définition un univers et ses limites sont simultanément définies. Aussi bien la CIM –

Classification internationale des maladies -, objet d'analyse par les auteurs de l'ouvrage examiné, que la Classification Décimale Universelle devraient, selon ce principe, inclure respectivement toutes les causes possibles de décès et tous les problèmes possibles connus. Mais ces classifications ne sont évidemment pas exhaustives en ce sens « La classification est un système et les systèmes ont des limites qui excluent, car ils sont comme des territoires finis » (Olson, 2002, p. 382). Ceci explique, entre autres, le cas de la lente reconnaissance du SIDA comme maladie et comme cause potentielle de décès dans la CIM.

En analysant les problèmes imposés par ces limites, Hope commente la proposition de Bowker et Star sur les « objets frontières », qui peuvent être décrits à travers une image de fils tissés qui traversent les frontières, reliant différents brins de culture. Selon Hope, cette image s'applique également aux systèmes de classification. Mais il va au-delà de la proposition des auteurs lorsqu'il affirme que, potentiellement, nous pouvons transgresser les frontières des classifications, surtout si nous comprenons « autorité » (autorité du système de classification) à partir de l'étymologie latine du mot (*augere*, augmenter), sans se limiter à son sens de contrôle. « Si nous considérons la classification comme quelque chose qui augmente potentiellement le sens, nous pouvons surmonter les limites de l'exclusivité mutuelle en utilisant les chevauchements aux points frontaliers » (Olson, 2002, p. 389), c'est-à-dire, en considérant la classification comme faisant partie de notre environnement construit et en tenant compte des contextes d'utilisation.

Nous pouvons sortir de l'autoroute téléologique de l'évolution (partout où cela nous mène) en empruntant les chemins secondaires menant à une découverte accidentelle. Nous pouvons éviter les limites de la hiérarchie en mettant en lumière des connexions autres que celles des catégories plus petites au sein de catégories plus grandes (Olson, 2002, p. 389, notre traduction).

Hope récupère trois conseils des auteurs : reconnaître l'équilibre dans l'acte de classification, rendre les voix récupérables et être sensible aux exclusions.

À partir de Hope, notre intention était de trouver des exemples pour confirmer l'existence de problèmes idéologiques dans tout système d'organisation des connaissances, puisque l'auteur a une production significative sur les préjugés et les omissions dans ces systèmes. La recherche a cependant apporté plus: à la fin de la révision, lorsque Hope affirme que l'on peut aller plus loin et suggérer d'autres réponses basées sur la superposition des systèmes et l'identification d'objets frontières qui servent mieux les différentes communautés de pratique, il évoque brièvement l'idée, sur laquelle on avance progressivement dans le pays, sur le changement de langage et les métathésaurus. Selon lui, à partir de différents types de vocabulaire, il est possible de relier plusieurs signifiants en considérant des significations communes, sans privilégier un vocabulaire ou un vocabulaire central au détriment des autres, en cherchant à établir des relations quasi hiérarchiques.

Revenant au thème central de notre présentation, nous pouvons dire que la proposition de vocabulaires sémantiques interconnectés, utilisant des données ouvertes liées et SKOS ou des formats similaires, nous permet de mettre en œuvre et d'enrichir l'idée mise en avant par Hope. Il s'agit d'explorer les ressources existantes de manière critique en vue de leur réutilisation, et non seulement de rechercher des concepts adaptés aux besoins des différentes communautés d'utilisateurs, ainsi que de concentrer les efforts pour formaliser les systèmes d'organisation des connaissances qui ne sont pas pris en charge par les grandes institutions. Parmi eux, nous retrouverons certainement des systèmes d'organisation des connaissances qui expriment des cultures locales, des points de vue originaux, qui répondent à des besoins particuliers et qui sont traditionnellement exclus et considérés comme marginaux. Dans la cartographie comparative du vocabulaire, dans une perspective critique, il est possible d'identifier les fondements qui sous-tendent chaque

système, d'analyser les conséquences de son utilisation, d'examiner les alternatives liées aux objets frontaliers et, enfin, de proposer une réutilisation sensible à la diversité culturelle et aux différents publics de l'information.

Notes

(1) Pour comprendre l'initiative, consultez Méndez & Greenberg, 2012.

(2) Source disponible sur:
<http://www.w3.org/Consortium/mission.html>.

(3) Source disponible sur : <http://www.w3.org/RDF/>.

(4) Source disponible sur : <http://www.w3.org/2004/02/skos/>.

Références

ANSI/NISO. ANSI/NISO Z39.19:2005. **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda: NISO, 2005.

BERNERS-LEE, T. Linked data. **Design Issues**, 2006.

BERNERS-LEE, T. **Weaving the Web**. London: Orion Books, 1999.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, maio 2001.

BOWKER, G. C.; STAR, S. L. **Sorting things out**: classification and its consequence. Cambridge, MA: MIT Press, 1999.

BS. **BS 8723-4:2007**: structured vocabularies for information retrieval: guide: interoperability between vocabularies. London: British Standards Institution, 2007.

BS. **BS 8723-5:2008**: structured vocabularies for information retrieval: guide: interoperation between vocabularies and other components of information storage and retrieval systems. London: British Standards Institution, 2008.

CODINA, L.; PEDRAZA-JIMÉNEZ, R. Tesauros y ontologías en sistemas de información documental. **El Profesional de la Información**, Madrid, v. 20, n. 5, p. 555-563, set./out. 2011.

HJORLAND, B. Knowledge Organization Systems (KOS). *In: Lifeboat of knowledge organization*, 2008.

ISO. **ISO 25964-1:2011**. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval. Geneva: International Standard Organization, 2011.

ISO. **ISO 25964-2:2012**. Thesauri and interoperability with other vocabularies. Part 1: Interoperability with other vocabularies. Geneva: International Standard Organization, 2012.

MARCONDES, C. H. “Linked data” – dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 34, p. 171–192, 2012.

MARCONDES, C. H.; SAYÃO, L. F. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. **Ciência da Informação**, v. 30, n. 3, p. 24–33, set./dez., 2001.

MÉNDEZ, E.; GREENBERG, J. Linked data for open vocabularies and HIVE’s global framework. **El profesional de la información**, Madrid, v. 21, n. 3, p. 236-244, maio/jun. 2012.

MILES, A.; BECHHOFER, S. (ed.). **SKOS Simple Knowledge Organization System – reference**: W3C Recommendation 18 August 2009.

MOREIRA, W.; LARA, M. L. G. Ontologias, categorias e interoperabilidade semântica. **DataGramZero: revista de Ciência da Informação**, v.13, n.4, ago. 2012.

OLSON, H. Review article: Classification and universality: application and construct. **Semiotica**, v.139, n.1/4, p.377-391, 2002.

PESET, F.; FERRER-SAPENA, A.; SUBIRATS-COLL, I. Open data y linked data: su impacto en el área de bibliotecas y documentación. **El profesional de la información**, Madrid, v. 20, n. 2, p. 165 – 173, mar-abril 2011.

SÁNCHEZ-JIMÉNEZ, R.; GIL-URDICIAN, B. Lenguajes documentales y ontologías. **El Profesional de la Información**, Madrid, v.16, n.6, p.551-560, nov.-dic. 2007.

SOERGEL, D. The representation of Knowledge Organization Structure (KOS) data: a multiplicity of standards. *In: JCDL NKOS Workshop*, VA. 2001.