

Linguagens documentárias e vocabulários semânticos para a *web*

Fernanda Maria Melo Alves

Departamento de Biblioteconomía y Documentación

Universidad Carlos III de Madrid, España

RECENSÃO CRÍTICA

A Web atual continua a oferecer deficiências quando à recuperação de informação. Assim o demonstra a presença persistente de sobrecarga informativa, a ineficácia das palavras-chave quando se pretende fazer recuperações precisas, a desordem terminológica consequente da falta de autoridade literária ou a ausência crónica de sistemas com PLN. Com ideia de superar estas limitações, a última década teve um crescimento e uma adaptação intensa dos vocabulários e das linguagens que organizam o conhecimento. A sua estrutura e a sua complexidade ficam determinadas pelo uso de léxico alargado, contando inclusive com estruturas verbais, pela ampliação das relações existentes entre os seus componentes, pela superação da referência a objetos do pensamento e objetos do mundo, enfim, pelo emprego de estruturas lógicas e de mapas conceituais para uma organização dos documentos cada vez mais complexa.

Por todas estas razões, procura-se conseguir um tratamento mais apurado das informações da web. Para organizar, representar e recuperar o conhecimento digital gerou-se um conjunto de tecnologias que tentam dar sentido e acrescentar semântica interpretável pelas máquinas ao conteúdo dos documentos. De tal forma que atuam como ferramentas de representação do conhecimento.

Os sistemas de organização e de representação, que empregamos na atualidade, surgiram por impulso de uma web que precisa de novas estruturas para responder a novas necessidades, facto que forçou a sua adaptação às linguagens documentais preexistentes. O conjunto das linguagens documentais geradas no passado e dos novos sistemas de

representação agrupa-se ao redor do nome Vocabulários para a web. Nesse contexto, este livro caminha mais perto das transformações que tiveram os elementos componentes e as estruturas de relacionamento dos vocabulários, para se adaptar a um meio tão exigente, do que dos sucessivos modelos de dados de XML necessários para expressar o conhecimento em estruturas simples predefinidas, como é o caso do emprego de SKOS para codificar tesouros, ou qualquer linguagem documental, em formato RDF.

Sem desatender totalmente esses modelos, prima aqui uma perspectiva de análise de conteúdo documental, em que se diferenciam os elementos linguísticos que compõem os vocabulários e as linguagens, de acordo com o emprego de conceitos inferidos, termos unívocos ou palavras livres. Assim como os relacionamentos semânticos que se possam estabelecer entre os referidos elementos, segundo a responsabilidade e a intervenção das pessoas na construção das linguagens e dos vocabulários, e inclusive de acordo com a aplicação que se estabeleça com a representação de objetos e o uso aberto ou restringido dos vocabulários. É, sem dúvida, uma visão parcial, mas necessária, para entender a origem, a transformação e a projeção dos vocabulários e das linguagens documentais que se empregam para organizar o conhecimento na web, bem como o respetivo enquadramento teórico.

A partir das necessidades de resposta à trajetória seguida pela representação do conhecimento temático na web, foi obrigatório considerar duas magnitudes diferentes, a Web Semântica e a Web Social (W2.0), que estabeleceram a adequação dos sistemas de representação às características e às necessidades da cada uma delas. O texto dá uma atenção preferente à diferenciação semântica na representação do conhecimento, seguindo o eixo fixado pela distinção entre o emprego de palavras, termos ou conceitos na constituição dos vocabulários. Este termo veio ocupar o local do seu precedente das linguagens documentarias, incluindo-as. E está a converter-se, gradualmente, numa denominação genérica para falar de estruturas ou de elementos padrão que representam e organizam a informação em Internet. A sua polissemia pode limitar-se de acordo com o contexto utilizado. De forma que na Web semântica o conceito de “ontologia” é sinónimo de vocabulário, enquanto na Web Social é o de “folksonomía”. Para os estandares de metadados, vocabulário é o conjunto de elementos descritivos e de propriedades. Mas no âmbito dos KOS (Knowledge Organization Systems), os vocabulários são as estruturas

tradicionais de representação como os tesouros ou as classificações de assunto. Finalmente, para as corporações e as instituições prefere-se o termo "taxonomia".

Deste modo, a obra teve que analisar e descrever estes diferentes vocabulários, bem como percorrer os principais sistemas de representação. Então, se a web impulsionou as estruturas de organização do conhecimento, mediante a intervenção das Ontologias e das Taxonomias, também teve que se adaptar às linguagens preexistentes, como é o caso dos Tesouros, da representação livre mediante palavras-chave, agora folksonomias, e dos mapas conceituais para desenhar os Topic maps. Como a maioria dos sistemas de organização e de representação surgiu antes que existisse o ciberespaço, há menos de uma geração, ainda estamos submersos numa forma cultural que aparece inconclusa, de maneira que os documentalistas do século XXI enfrentam o problema de criar uma nova geração de sistemas simbólicos ou de melhorar os precedentes. O conteúdo desta monografia mostra essas novidades e variações, cuja apresentação esquemática e emprego de uma linguagem direta e cuidada servem para atingir uma boa exposição conceitual. Por isso, devemos estar gratos à Editora da Universidade Federal da Bahia, pela oportunidade de divulgar outra obra deste autor, com o intuito de levar à comunidade lusófona um dos enquadramentos que está a atravessar maiores mudanças e exigências na área da Informação-Documentação.

Referência Bibliográfica

Moreiro González, José Antonio - *Linguagens documentárias e vocabulários semânticos para a web. Elementos conceituais*. Salvador de Bahia: EDUFBA, 2011. ISBN 978-85-232-0824-0