



**ANÁLISE COMPARATIVA
DA APLICAÇÃO DO
MODELO RELACIONAL
E DO FORMALISMO
RDF À MODELAÇÃO
DE DADOS LEGISLATIVOS**

A obra pretende contribuir para uma desejada aproximação entre a ciência jurídica e as linguagens de modelação oriundas das ciências da computação. Nos últimos anos, a acelerada evolução e conseqüente enriquecimento das técnicas e linguagens de modelação conceptuais tornaram possível a sua manipulação por profissionais de todas as áreas do saber. Se há alguns anos atrás era improvável que, por exemplo, um profissional da área das ciências sociais participasse activamente na concepção da arquitectura e definição das funcionalidades de uma aplicação a ele destinada, actualmente, para além de participar, poderá ser um dos principais responsáveis pelas opções a adoptar. A motivação da autora para realizar o trabalho que está subjacente a esta obra é uma consequência deste envolvimento do denominado “utilizador final” na definição da solução computacional. Dada a proliferação de linguagens de modelação actualmente disponíveis no mercado, não sendo a autora uma especialista na área, sentiu a necessidade de optar pela linguagem que mais bem se adequava ao domínio em causa – informação jurídica – e que, dado que seria ela a utilizá-la, mais bem se adaptava aos seus próprios conhecimentos.

O objectivo da obra consiste na análise comparativa de duas linguagens de modelação largamente divulgadas, o Resource Description Framework (RDF) e o Modelo Relacional, à especificação de modelos de dados legislativos. Refere a autora que a proliferação de instrumentos normativos dificulta

o conhecimento de todas as normas a acatar e dificulta também a visão de conjunto das mesmas, nomeadamente devido ao complexo relacionamento dos actos normativos. O trabalho desenvolve-se no âmbito da Informática Jurídico-Documetal, enquanto ciência que se ocupa dos problemas relativos ao armazenamento, recuperação e transmissão de dados jurídicos, visando garantir a certeza do Direito.

Ao longo do texto, a autora, com muito rigor, mas sempre recorrendo a exemplos retirados da realidade, confronta o potencial das duas linguagens, utilizando para tal métricas variadas, tais como facilidade de utilização, grau de completude, nível de ambiguidade, etc. O resultado final é antecedido por várias etapas, a autora começa por analisar jurídico-documentalmente os diplomas legais, identificando os seus elementos descritivos e as suas relações com outros actos normativos e outros documentos jurídicos não normativos. De seguida, apresenta os conceitos fundamentais da modelação de dados, distinguindo conceptualmente as noções de esquema, ontologia e metadados e, por último, introduzindo os principais mecanismos de abstracção. Depois de abordar as principais teorias relativas à qualidade dos modelos de dados e proceder à respectiva adaptação para a avaliação dos formalismos em estudo, a autora constrói uma grelha de avaliação quantitativa dos metamodelos, definindo a ponderação de cada critério em função das dimensões semântica e sintáctica da análise.

Apesar de a motivação inicial da autora ter sido a de, com o maior rigor possível, seleccionar a linguagem que mais bem se adequava às suas necessidades, o principal interesse da obra não reside na

conclusão a que a autora chegou, mas na forma como descreve as opções que tomou para suportar a decisão. A fundamentação utilizada, naturalmente suportada por trabalhos de outros autores, e a forma minuciosa e rigorosa como ela é apresentada são extremamente úteis para qualquer leitor que se depare com a necessidade de analisar a adequação de linguagens de modelação a problemas concretos.

Pedro Ramos

PATRÍCIO, Helena Margarida Mendes Ferrão Simões
– *Análise comparativa da aplicação do modelo relacional e do formalismo RDF à modelação de dados legislativos.*

Lisboa: Vice-Presidência dos Mestrados do Departamento de Ciências e Tecnologias da Informação do ISCTE, 2007. 170 p. ISBN 978-972-97251-2-8



ORGANIZAÇÃO DO CONHECIMENTO E REPRESENTAÇÃO DE ASSUNTOS: CAMINHOS PARA UMA EFECTIVA RECUPERAÇÃO DA INFORMAÇÃO EM AMBIENTES DE REDES

Publicação da dissertação de mestrado em Ciências Documentais apresentada à Universidade de Évora por Maria Luísa Santos, actual Directora

de Serviços Bibliográficos Nacionais, esta obra constitui um dos mais recentes títulos das Publicações Técnicas editadas pela Biblioteca Nacional de Portugal. A autora sustenta que a introdução das tecnologias da informação e comunicação (TIC) se traduziu em alterações pouco significativas nos processos de recolha, processamento e difusão de informação no seio dos sistemas documentais. Os novos desafios decorrem, assim, da necessidade crescente de um trabalho de normalização imposto pela exigência de satisfazer emergentes vontades da sociedade, conjecturando-se a criação de modelos inovadores.

O primeiro capítulo justifica uma das grandes vantagens da informatização pela facilidade de criação de bases de dados bibliográficas de tipo cooperativo, possibilitando a criação de catálogos colectivos, responsáveis pela antecipação do conceito de rede em bibliotecas. Esta parece ser a razão segundo a qual a obra dedica particular atenção à estrutura da informação em bases de dados cooperativas ou catálogos colectivos, aprofundando a temática do acesso por assuntos à informação.

A automatização do processamento documental e a necessidade de proceder à permuta de informação contida nas bases de dados bibliográficas conduziram, inevitavelmente, ao desenvolvimento de redes e protocolos de comunicação, sendo os da Internet os melhores exemplos, face à transformação radical operada relativamente aos sistemas fechados. A autora defende a norma ISO 2709 e os formatos Marc como os responsáveis pela génese das bases de dados bibliográficas e enumera um amplo conjunto de sistemas de suporte às referidas bases de que fazem parte redes, protocolos de comunicação, formatos e linguagens de escrita

e armazenamento. Apresenta, de seguida, uma síntese de alguns dos mais emblemáticos projectos de cooperação e normalização da descrição bibliográfica e do controlo de autoridade e corrobora as teorias de Michael Gorman ao afirmar não ser possível o controlo bibliográfico sem um controlo de autoridade e que este é importante para se conseguir um determinado nível de normalização (p. 28). Enumera, ainda, as características gerais das bases de dados cooperativas, apontando diferentes modelos de catálogos colectivos.

No segundo capítulo, apresenta uma breve caracterização da Base Nacional de Dados Bibliográficos (PORBASE) enquanto catálogo colectivo, dissertando sobre as actuais possibilidades em aceder por assuntos aos conteúdos informativos. Apresenta as principais fases do seu desenvolvimento enquanto projecto cooperativo e caracteriza o conteúdo da base de dados e o universo das bibliotecas cooperantes. Finalmente, expõe a arquitectura sobre a qual assenta a PORBASE.

A autora aponta esta via de catálogo virtual como o “que melhor se adapta às novas exigências da difusão da informação e da partilha de recursos” (p. 48), o que não significa que não haja necessidade de se vir a recorrer a novas soluções tecnológicas, alteração de procedimentos técnicos e diferentes processos de disponibilização e acesso à informação. O acesso por assuntos à informação é o tema do terceiro capítulo. Nele são apresentados os modelos tradicionais de representação do conhecimento e de recuperação por assuntos, de que fazem parte as classificações universais (Decimal e Dewey), ainda hoje relevantes

instrumentos de organização e de representação de conteúdos temáticos. Seguem-se os sistemas de indexação integrados em tesouros (línguas pós-coordenadas), enquanto instrumentos utilizadores de uma linguagem controlada mais próxima da linguagem natural e, por essa razão, mais capazes de satisfazer um crescente conjunto de necessidades informativas, impedindo a perda de relações semânticas relativamente ao conceito representado. As listas de cabeçalhos de assuntos (línguas pré-coordenadas) apresentam-se com uma vantagem de maior precisão na descrição de assuntos, contra uma menor rapidez de processamento face à utilização das línguas controladas.

A autora relata ainda a existência de outros instrumentos de apoio à análise e representação temáticas. Aborda, também, as dificuldades relacionadas com a representação de assuntos e conseqüente recuperação em catálogos colectivos ou bases de dados cooperativas, sintetizadas na diversidade de critérios subjacentes à aplicação das línguas documentais, diferentes níveis de análise e de representação, dependentes de políticas de indexação implementadas e irregularidades de processamento.

O aumento exponencial da informação e a emergência do ambiente Web seriam os grandes responsáveis por uma séria reflexão acerca das características da informação disponibilizada por tais línguas, abrindo vias para novas formas de recuperação assentes em pesquisa em texto livre e pelo testar de “novos instrumentos para permitir a organização e representação dos conteúdos temáticos

dos objectos digitais e do ambiente web, concebidos tendo como ponto de partida a complementaridade entre línguas controladas e linguagem natural” (p. 61).

Os portais temáticos são na obra apresentados como “resultado de uma reflexão aberta e partilhada sobre a selecção da informação a circular na rede e reflectem o trabalho conjunto de diferentes entidades, concretizado em diversos projectos desenvolvidos com o objectivo de criar formas unívocas de aceder, de um modo consistente, à informação existente na rede sobre determinados assuntos” (p. 62).

Como novas formas de representação de determinados domínios do conhecimento, a autora refere ainda as ontologias, as quais apresentam como grande vantagem de utilização uma grande flexibilidade em termos de partilha e reutilização, podendo conter termos descritivos variáveis e “expressar toda a estrutura semântica e todas as relações numa linguagem compreendida pelo computador” (p. 65).

O quarto capítulo apresenta um conjunto de abordagens às questões relativas à análise, representação e recuperação de assuntos, mediante o exame extenso de três projectos, correspondentes a diferentes abordagens: o RENARDUS é pela autora apresentado como “um exemplo de organização do conhecimento e de interoperabilidade entre diferentes sistemas de classificação e um modelo comum” (p. 73), sendo que o derradeiro propósito do projecto consistiria em conduzir os proprietários dos sistemas de classificação envolvidos nessa empresa a manterem um mapeamento com a CDD (Classificação decimal Dewey), impondo-se como necessária ainda

uma normalização para vocabulários envolvidos, as suas relações semânticas e diferentes tipos de mapeamento; o MACS (Multilingual Access to Subjects) surgiu da necessidade de corresponder à intenção expressa pela CELN (Conference of European National Librarians) em ver resolvida a dificuldade do acesso por assuntos nas bases de dados bibliográficas europeias. No projecto, foram envolvidas bibliotecas utilizando línguas de indexação normalizadas monolíngues com o objectivo de desenvolver instrumentos que possibilitassem o acesso a bases de dados multilíngues, mediante a utilização de pontos de acesso (cabeçalhos) equivalentes, oriundos de catálogos de autoridade de assuntos. Tal empresa assumiu relevância pela possibilidade de desenvolvimento de cooperação internacional; o projecto HILT (High Level Thesaurus) consistiu numa experiência de interoperabilidade entre línguas de indexação em diferentes sistemas documentais (bibliotecas, arquivos e museus), em diferentes línguas. O objectivo final seria a elaboração de um roteiro para terminologias e procurar estabelecer requisitos de um serviço de terminologias por assuntos, considerando a sua compatibilidade internacional.

Tais projectos são pela autora defendidos com o objectivo de “apontar caminhos que permitam desenvolver, numa nova perspectiva, os catálogos colectivos ou bases de dados cooperativas” (p. 16).

O capítulo final sugere uma proposta de soluções a encetar com o objectivo de melhorar o acesso à informação

através da PORBASE, sugerindo uma reflexão da mesma. Para o efeito, analisa os vários aspectos da sua gestão, apresentando o diagnóstico de um conjunto de dificuldades e de situações geradoras de problemas, que dificultam a sua actualização permanente. Maria Luísa Santos evidencia ainda algumas estratégias de desenvolvimento anteriormente anunciadas, as quais poderão passar por “uma redefinição mais precisa do [seu] conteúdo” (p. 97), reforçando a síntese dos resultados dos projectos anteriormente referidos para apontar caminho no domínio do acesso por assuntos à informação em rede, tendentes a conseguir uma interoperabilidade efectiva a nível dos assuntos, antevendo-se como constrangimentos: “as limitações de ordem financeira, a escassez de recursos humanos especializados, pressões de carácter institucional e factores decorrentes de uma cultura organizacional que é lenta na mudança, que dificultam a implementação de determinadas propostas e obrigam à sua adaptação em função da realidade e do ambiente que envolve as organizações (p. 99-100).

Consequentemente, a autora sublinha a vantagem de se usarem, como ponto de partida, “formas de organização e representação do conhecimento já estabelecidas, como as classificações, os tesouros, as listas de cabeçalho de assunto e outros instrumentos terminológicos usados para indexar documentos” (p. 107).

As experiências partilhadas ao longo da investigação permitiram ainda

à autora demonstrar a adaptabilidade de semelhantes linguagens à classificação automática de documentos disponíveis em texto integral, a qual tem vindo a revelar resultados satisfatórios, impondo-se, naturalmente, certas adaptações e desenvolvimentos.

O estado actual do acesso por assuntos em ambiente de rede é finalmente sintetizado pela autora nas “palavras-chave”: “interoperabilidade, normalização, cooperação, investigação, integração” (p. 110).

Maria Luísa Santos abona a apresentação dos projectos mais importantes, e que foram objecto da sua investigação, com um copioso conjunto de anexos. De referir, também, um amplo número de referências bibliográficas actualizadas, remetendo, muitas delas, para recursos electrónicos.

Pelas razões enunciadas, trata-se de uma obra que contribui para o aprofundar dos estudos no domínio da ciência da informação em Portugal, com particular destaque para as experiências nacionais e europeias no domínio da representação de assuntos. Constitui, por essa razão, um instrumento de trabalho com particular interesse para o diálogo entre profissionais de informação e documentação.

Maria Carla Proença

SANTOS, Maria Luísa F. N. dos
– *Organização do conhecimento
e representação de assuntos: caminhos
para uma efectiva recuperação
da informação em ambientes de redes.*
Lisboa: Biblioteca Nacional de Portugal,
2007. 180 p. ISBN 978-972-565-412-5.