

Da gestão da informação à gestão do conhecimento: uma proposta para a e-Administração em Portugal

Alexandra Lourenço^a, José Carlos Ramalho^b, Pedro Penteador^c

^a*Universidade do Nova de Lisboa, Portugal, m.alexandra.lourenco@gmail.com*

^b*Universidade do Minho, Portugal, jcr@di.uminho.pt*

^c*Universidade do Nova de Lisboa, Portugal, ppenteador@fcsb.unl.pt*

Resumo

O trabalho proposto tem como objetivo apresentar um projeto de elaboração de um sistema de informação que visa gerir o desenvolvimento de uma ontologia promovida pelo organismo de coordenação da política arquivística para a classificação e avaliação da informação na Administração Central e Local, em Portugal. Entre os produtos propostos a partir do referido sistema, conta-se um website no qual é possível consultar toda a informação contida na ontologia, bem como descarregar versões da mesma, com níveis diferentes de semântica.

A referida ontologia foi promovida pelo órgão coordenador da política arquivística portuguesa, e tem por base uma Lista consolidada (LC) de processos de processos de negócio da Administração Pública (AP). Esta lista de natureza incremental e colaborativa foi desenvolvida a partir da identificação de uma macroestrutura representativa das funções exercidas pela AP, a Macroestrutura Funcional (MEF).

Concretizado este produto e colocado à disposição da comunidade, na página oficial da Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB), importa potenciar a sua aplicabilidade. Neste sentido, encontra-se em curso um projeto que envolve os autores, em contexto universitário, orientado para o desenvolvimento de um vocabulário formal, um modelo de dados que represente este conjunto de conceitos referentes aos processos de negócio e aos relacionamentos entre eles.

Pretende-se que esta ontologia possa vir a ser disponibilizada em listas ou diretórios de ontologias com mecanismos de pesquisa (bibliotecas de ontologias) de modo a incrementar a sua utilização na web semântica, para além da sua utilização como esquema de classificação em sistemas eletrónicos de gestão de arquivos (SEGA), business intelligence systems e sistemas de gestão do conhecimento.

Com esta comunicação pretende-se dar a conhecer à comunidade de profissionais um projeto de aplicação transversal para todas as entidades públicas e para as empresas com interesse na área.

Palavras-chave: Gestão do conhecimento; Gestão da Informação; Governo eletrónico; Interoperabilidade semântica; Macroestrutura Funcional; Ontologia; Web semântica.

Keywords: Knowledge management; Information management; E-Government; Semantic interoperability; Macrostructure Functional; Ontology; Semantic Web.

Introdução

O organismo de coordenação da política arquivística nacional, atual Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB), criou o Programa "Administração Eletrónica e Interoperabilidade Semântica", de adesão voluntária, com o objetivo geral de contribuir para o fomento e implementação da interoperabilidade semântica na Administração Pública (AP) e nas entidades públicas com que esta se relaciona e executem funções de Estado.

O Programa, atualmente com mais de meia centena de organismos aderentes provenientes da Administração Central e Local, tem como objetivos específicos:

- “a) O desenvolvimento de uma Macroestrutura funcional (MEF), estrutura semântica para classificar ao 1.º e 2.º nível os documentos de arquivo das referidas entidades, bem como instrumentos de apoio à sua utilização;
- b) O desenvolvimento de documentos técnicos que aprofundem os níveis de interoperabilidade semântica alcançados na MEF;
- c) A gestão e atualização da MEF e dos documentos indicados na alínea b), bem como do esquema de metainformação que lhes é suprajacente: o MIP (Metainformação para a Interoperabilidade)”, composto por um conjunto de 17 elementos usados para descrever recursos informativos;
- “d) A promoção da utilização dos referidos instrumentos nas entidades envolvidas na Administração eletrónica” (cf. <http://arquivos.dglab.gov.pt/programas-e-projectos/modernizacao-administrativa/macroestrutura-funcional-mef/>).

A versão 2.0 da Macroestrutura Funcional viria a ser publicada em 2013, é composta por classes de 1.º e 2.º nível representativas das funções de Estado, com a sua descrição, notas de aplicação, notas de exclusão e respetivo índice.

Na sequência da sua publicação desta versão, a DGLAB apresentou uma metodologia complementar para apoio ao desenvolvimento de planos de classificação conformes à MEF. Nesse sentido, desenvolveu um projeto cooperativo para a criação de uma linguagem comum a 3.º nível, denominado “Harmonização de classes de 3.º nível em planos de classificação conformes à MEF”. Estas classes são representações conceptuais de processos de negócio comuns ou específicos. Entre outros produtos, publica na Web uma Lista consolidada dos 3.ºs níveis harmonizados, ligados com os níveis superiores e visando, desta forma, a sua reutilização. Reúne as propostas das distintas entidades participantes da Administração Central e Local. A publicitação da Lista consolidada destina-se a apoiar o processo de construção de novos planos de classificação e a sistematizar a informação produzida, proporcionando a partilha de conhecimento sobre os processos desenvolvidos pela Administração.

No total, a Lista consolidada, constituída por três níveis, é composta atualmente por mais de um milhão de instâncias, que representam as funções e as subfunções da Administração Pública (MEF) e os processos de negócio em que estas se consubstanciam. A Lista consolidada permanece em crescimento, estando aberta à inclusão de processos de negócio ainda não identificados ou descritos parcialmente. Tem uma natureza incremental e visa integrar a representação da totalidade dos processos de negócio levados a cabo pela Administração.

Complementarmente encontra-se a decorrer o Projeto “Avaliação suprainstitucional da informação arquivística (ÁSIA)” que pretende, a partir de uma metodologia relacional conjugada com os dispositivos legais que determinam ou influenciam direta ou indiretamente a conservação da informação, determinar prazos e destinos para o produto informacional resultante dos processos de negócio identificados, bem como a(s) entidade(s) responsáveis pela sua preservação. O resultado final pode conduzir a algumas alterações / melhorias na Lista consolidada atualmente disponível.

A mudança do foco da criação de uma mera macroestrutura classificativa para a gestão, utilização e reutilização das instâncias da Lista consolidada evidencia a necessidade das entidades envolvidas terem um sistema de informação com alguma complexidade, para este efeito.

Este produto reforça ainda a posição do país na implementação das políticas e estratégias europeias de interoperabilidade, promoção do acesso e reutilização da informação pública, na sequência, respetivamente, da Decisão n.º 922/2009/EC e da Diretiva 2013/37/EU do Parlamento e do Conselho. Na continuidade daquela Decisão, foi implementado o Programa ISA – “Interoperability Solutions for European Public Administrations” e a criação da plataforma colaborativa Joinup, onde têm sido publicados documentos relevantes, indicações sobre os principais repositórios de metainformação e o catálogo de ativos semânticos disponíveis (cf. <https://joinup.ec.europa.eu/>).

Para que se possa integrar estas iniciativas europeias e potenciar o seu desenvolvimento nacional é necessário transformar a referida Lista de processos de negócio da Administração pública num

Da gestão da informação à gestão do conhecimento

vocabulário formal, passível de vir a integrar o catálogo de ativos semânticos disponíveis na Joinup, bem como o projeto LOD 2 - Linked Open Data, o qual visa enriquecer os dados das Administrações públicas através de tecnologias semânticas, na Web, tornando-os mais utilizáveis. Para este efeito, é fundamental, entre outros aspetos, efetuar a modelação destes dados e a sua conversão para RDF ("Resource Description Framework"), SKOS ("Simple Knowledge Organization System") e OWL ("Ontology Web Language") (Lourenço & Penteadó, 2014).

A necessidade de um novo formato

Em primeiro lugar, no âmbito deste projeto, procurou-se responder à necessidade das entidades envolvidas possuírem um sistema de informação, com alguma complexidade, que permita gerir a estrutura classificativa (Lista consolidada - LC, que já integra a MEF). Uma das primeiras decisões a tomar foi sobre qual o formato que deveria suportar esta representação conceptual.

A LC tem sido mantida num documento do MSWord e em folhas de cálculo do MExcel. Ferramentas do domínio comum no seio das entidades envolvidas. No entanto, à medida que foi crescendo, a sua gestão, manutenção e consulta têm-se tornado cada vez mais difíceis.

Estavam criadas as condições para se pensar a LC para além de um conjunto de descrições textuais, como um sistema de informação com alguma complexidade, como veremos.

A nova solução para a LC deveria ser capaz de dar resposta aos seguintes requisitos principais:

- Ser fácil de consultar e de utilizar;
- Ser fácil de manter, permitindo acrescentar novas classes, alterar descrições existentes, etc;
- Estar disponível em formatos processáveis por máquinas para ser automaticamente incluída em aplicações de software, sem recurso a trabalho e utilização de mão de obra extra.

Do ponto de vista técnico, havia que escolher uma abordagem à sua representação. Atualmente há três grandes modelos para o armazenamento e persistência da informação:

- a) o modelo relacional (as famosas bases de dados relacionais);
- b) o modelo orientado ao objeto (aplicável sobretudo no suporte de modelos hierárquicos onde o encapsulamento é importante);
- c) o modelo de grafo (mais lato e que permite representar tudo o que se possa pretender).

Numa primeira análise, a LC é uma hierarquia de classes, pelo que podemos eliminar o modelo relacional e focar-nos nos outros dois. Qualquer um dos outros modelos suporta uma hierarquia de classes. No entanto, os processos de negócio na Administração Pública têm um ciclo de vida, excluem-se uns aos outros (cf. Notas de exclusão) bem como, no contexto do projeto ÁSIA, foi necessário compreender e especificar o conjunto de relações entre processos (ou entre a informação que lhes dá corpo) pertinentes para a sua valoração, nomeadamente: processos que dão sequência a outros, que precisam de informação de outros para prosseguir, que são complementares a outros, ou cuja informação se encontra sintetizada noutros processos. Ou seja, vão existir relações não hierárquicas entre estes processos (relações de sucessão, relações cruzadas, relações de complementaridade, relações de síntese e outras) que se pretende representar no modelo. Ora, a necessidade de também representar estas relações não hierárquicas entre processos, deixam-nos apenas uma opção de representação, o modelo de grafo.

O modelo de grafo é também o modelo usado para a representação de ontologias. No nosso contexto, uma ontologia é uma especificação formal de conhecimento. O termo formal quer dizer que a especificação deve ser inteligível por humanos mas também por máquinas. A MEF começou por ser uma hierarquia de classes, mas com a adição das relações não hierárquicas passou a ser uma ontologia.

Portanto, numa primeira instância foi preciso escolher um formato neutro (independente de plataformas de software e hardware) capaz de representar um modelo ontológico (grafo).

Devido à natureza hierárquica da especificação e aos requisitos de neutralidade relativamente a plataformas tecnológicas a escolha óbvia foi o XML ("eXtended Markup Language"), uma metalinguagem de anotação que nos permite, de forma descritiva, dar estrutura e semântica a um texto.

Assim, desenvolveu-se uma pequena linguagem XML com o objetivo de suportar a descrição da Lista. Nesta linguagem, além de podermos descrever cada classe também podemos relacioná-las hierarquicamente e de forma não hierárquica. O desenvolvimento desta linguagem partiu da análise dos documentos usados para armazenar a informação. Desta análise resultou uma linguagem formalizada através de um XML Schema (Lee & Chu, 2000). No exemplo que se pode observar na figura 1 encontra-se aplicado à Macroestrutura Funcional (MEF) integrada na LC, ou seja, aos 2 primeiros níveis classificativos.

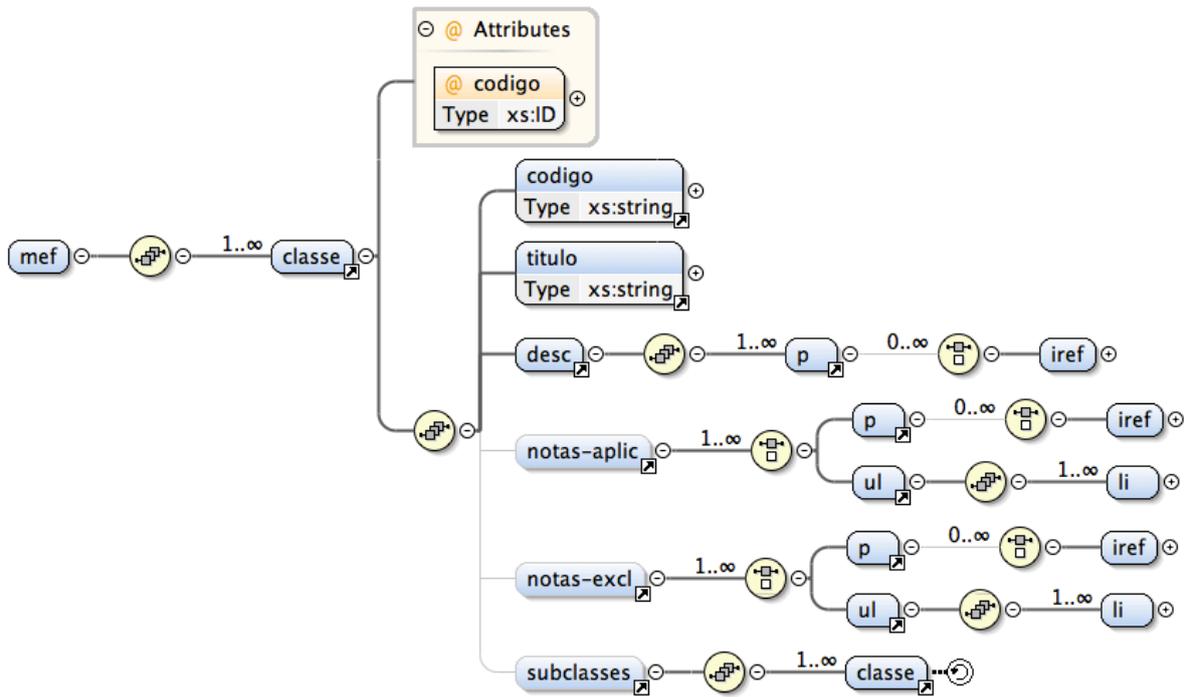


FIGURA 1: XML SCHEMA DA MEF

Como se pode ver, a MEF é constituída, em primeiro lugar, por uma lista de classes de 1.º nível. Cada classe é constituída por um conjunto de elementos de informação: código, título, descrição, notas de aplicação, notas de exclusão e uma lista de classes de 2.º nível ou subclasses. Por sua vez, cada subclasse obedece à mesma estrutura da classe. O código e o título são elementos textuais simples. A descrição e as notas são elementos semiestruturados, essencialmente constituídos por texto que pode conter elementos estruturais que devem ser anotados: parágrafos, listas de itens e referências internas a outras classes. Estes elementos semiestruturados podem, no futuro, vir a ser enriquecidos com mais elementos estruturais, caso se sinta essa necessidade.

Tendo como base a MEF, o passo seguinte foi migrar toda a informação recolhida até ao momento, guardada em documentos do MSOffice, para este formato neutro. Com aplicação e combinação de várias tecnologias conseguiu-se realizar esta migração. Primeiro constatou-se que não era viável migrar diretamente do MSWord. Testaram-se vários dos formatos de exportação da ferramenta, inclusive o XML, mas a quantidade de “lixo” gerado junto com a informação era tão grande que invalidou a sua migração desta forma. O caminho seguido foi a exportação em PDF e depois a transformação do PDF num formato XML intermédio correspondente à lista sequencial das classes. Depois programou-se um migrador que a partir deste formato gerou uma descrição de acordo com o formato apresentado na figura 1.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <mef xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
2 xsi:noNamespaceSchemaLocation="new-mef.xsd">
3   <classe codigo="c100">
4     <codigo>100</codigo>
5     <titulo>ORDENAMENTO JURÍDICO E NORMATIVO</titulo>
6     <desc>
7       <p>Relativo à construção e interpretação das Normas, no sentido lato do termo
7 (com e sem carácter coercivo): elaboração, aprovação e publicitação dos atos de carácter
7 dispositivo e de carácter orientador que regulam as ações e relações entre os diversos
7 atores sociais - da legislação aos regulamentos, regras internas de funcionamento,
7 instruções procedimentais e normas técnicas.</p>
8       <p>Inclui os processos de vinculação do Estado português a convenções
8 internacionais, bem como os avisos relativos à vinculação dos restantes Estados-parte
8 naquelas convenções.</p>
9     </desc>
10    <notas-aplic>
11      <p>Os atos de carácter dispositivo podem incidir sobre a sociedade portuguesa

```

FIGURA 2: DESCRIÇÃO MEF NO FORMATO XML

A partir desta base em XML, cuja metodologia é extensível da MEF para toda a Lista consolidada, todo um número de possibilidades se abre, estando atualmente em curso:

- Geração automática em PDF de uma monografia descritiva da referida Lista para os utilizadores poderem ter um guia em papel/digital semelhante ao que existe atualmente. O gerador será programado em XSLT e criará uma versão XSLFO e, posteriormente, PDF;
- Geração automática de um website, no qual é possível consultar toda a informação contida na LC e navegar de forma simples nas relações entre classes. A primeira versão já está implementada. Neste caso, o gerador feito em XSLT gera um conjunto de páginas HTML relacionadas;
- Geração automática de uma ontologia baseada na LC com a qual é possível alimentar automaticamente vários sistemas de software e trabalhar a Lista ao nível semântico, desenvolvendo interrogações semânticas, inferindo automaticamente novas relações e novo conhecimento, etc. Neste caso, o gerador de SKOS já está feito em XSLT e já foi testado numa situação real, que se descreve mais à frente;
- Geração automática de planos de classificação a partir da fonte XML ou da ontologia (um plano de classificação não é mais do que uma subárvore da LC, a árvore classificativa original, i. e.: são extraídas da LC as classes que representam os processos comuns e os processos específicos que corporizaram o plano de determinada entidade ou conjunto de entidades, possibilitando deste modo a comunhão de códigos, base para a interoperabilidade); será realizado brevemente e será implementado através de uma stylesheet XSLT parametrizada com um ficheiro XML onde se declaram os ramos da árvore da LC que se pretendem ver incluídos em determinado plano de classificação.

A ontologia gerada a partir dos dados recolhidos corresponderá à materialização de uma parte significativa do referido sistema de informação.

MACROESTRUTURA FUNCIONAL (MEF)

VERSÃO 2.0

Título: Macroestrutura funcional (MEF)
Tipo Id: SIARQ/Documentum
Id.: DSAN_13_I/000868
Produtor: Direção Geral do Livro, dos Arquivos e das Bibliotecas
Autor: Alexandra Lourenço, Cecília Henriques e Pedro Penteadó (coordenador)
Descritores: classificação, interoperabilidade, governo eletrónico, administração pública, classes, índice.
Data/Hora: 2013-03-28

Índice de Classes

- [100 - ORDENAMENTO JURÍDICO E NORMATIVO](#)
 - [100.10 - ELABORAÇÃO DE DIPLOMAS JURÍDICO-NORMATIVOS E DE NORMAS TÉCNICAS](#)

Índice Remissivo de Termos

- [ADSE \(participações\) : 650.10](#)
- [ADSE \(inscrições\) : 400.10](#)
- [Abaixo assinado : 950.20](#)

FIGURA 3: PÁGINA PRINCIPAL DO WEBSITE GERADO

Na figura 3, pode-se observar parte da página principal do site gerado especificamente para a MEF com alguma metainformação e os dois índices de navegação principais: o índice hierárquico de classes e o índice remissivo de termos. De realçar que o tipo de navegação permitido por estes índices torna muito mais simples e fácil o acesso e a consulta da MEF, algo que começa a ser complicado de fazer na versão PDF/papel. A solução proposta é, obviamente, replicável para o caso da LC.

Na próxima secção, vamos focar-nos na geração dos vários formatos da ontologia e tentar responder às seguintes questões: Qual a importância do modelo ontológico neste contexto? Porque se vão gerar vários formatos para a ontologia?

Ontologia

As ontologias e a sua especificação sempre foram uma área trabalhada em disciplinas como a Filosofia e áreas como a Arquivística e a Biblioteconomia sempre tiveram necessidade de utilizar vocabulários controlados, taxonomias classificativas e ontologias.

Em termos tecnológicos, a grande evolução dá-se quando começa a haver necessidade de aplicar estas metodologias à Web, numa iniciativa que ficou conhecida como Web Semântica (Berners-Lee, Hendler & Lassila, 2001). Esta não é uma nova Web mas sim uma extensão da existente. Acrescentam-se atributos semânticos aos dados de modo a associar-lhes significado de uma forma que as máquinas conseguem interpretá-los, tarefa essa que até agora só os humanos conseguiam fazer. A utilização desta extensão está relacionada com o crescimento exponencial da informação na Web e com a necessidade de conseguir reduzir as listas de resultados nas pesquisas.

Há toda uma família de normas que se desenvolveram no seio desta iniciativa, relevantes para o projeto que se quer desenvolver: RDF, SKOS e OWL. Possuem algumas características comuns e todas servem para representar, com mais ou menos detalhe, um domínio de conhecimento, tendo no XML o seu formato de representação.

Como não podia deixar de ser, existe uma relação entre os três formatos, que clarificamos de seguida, na sua descrição.

RDF

O RDF é uma norma W3C (“World Wide Web Consortium”) para a descrição de recursos. O que levanta a questão: o que é um recurso? Esta é uma questão que continua a suscitar amplas discussões. No nosso contexto, vamos assumir que um recurso é algo que conseguimos identificar. O leitor é um recurso, a sua página pessoal é um recurso, esta comunicação, o número um e o leão no filme “Lion King” da Walt Disney também o são (por sua vez, o filme e a Walt Disney também poderiam ser considerados recursos).

O RDF descreve um recurso através de asserções que têm a forma de triplos:

(sujeito, predicado, objeto)

Um sujeito relaciona-se através de um predicado com um objeto. Vamos representar, por exemplo, a classe 100 da MEF ilustrada na figura 2. Da sua descrição podemos extrair algumas relações:

(c100, temCódigo, “100”)

(c100, temTítulo, “ORDENAMENTO JURÍDICO E NORMATIVO”)

(c100, temDesc, “<p>Relativo à construção...”)

(c100, temSubclasse, c100.10)

(c100, temSubclasse, c100.20)

Um sujeito é sempre um recurso e normalmente é representado por um identificador. Um predicado é o identificador de uma relação e o objeto pode ser uma de duas coisas: ou é um identificador de um recurso (permitindo estabelecer relações entre recursos), ou é um valor literal (permitindo definir o valor de um atributo do sujeito).

Podemos concluir que o RDF é uma linguagem muito simples mas, em simultâneo, que nos permite representar qualquer domínio de conhecimento pois tudo se pode reduzir a uma lista de relações.

SKOS

Quando trabalhamos num contexto global, a interoperabilidade é uma característica chave e importante. O RDF não nos garante nenhum tipo de interoperabilidade semântica pois é completamente livre; ou seja, qualquer utilizador pode utilizar os predicados que bem entender para formalizar o seu domínio de conhecimento. Foi desta preocupação com a interoperabilidade que surgiu o SKOS (“Simple Knowledge Organization System”), para ajudar a normalizar os descritores usados no domínio das taxonomias, dos esquemas de classificação e dos thesauri. Ao fixar um conjunto de predicados RDF, esta iniciativa veio permitir o intercâmbio e a troca de especificações.

A linguagem SKOS é composta por um conjunto de classes (sujeitos) e de propriedades (predicados) RDF com os quais é possível especificar taxonomias.

Mais uma vez, exemplificando com a MEF:

Term: c100.10 - ELABORAÇÃO DE DIPLOMAS JURÍDICO-NORMATIVOS ...

Used for:

Compreende os processos de elaboração/alteração de legislação, de regulamentos e de diretivas políticas ou operacionais portuguesas...

Broader terms:

100: ORDENAMENTO JURÍDICO E NORMATIVO

Narrower terms:

100.10.001: Produção e comunicação de atos legislativos
100.10.002: Aprovação e comunicação de convenções internacionais
100.10.003: Comunicação de decisões dos tribunais ...
100.10.200: Produção e comunicação de atos regulamentares gerais
100.10.400: Produção e comunicação de atos regulamentares locais
100.10.600: Produção e comunicação de regras institucionais
100.10.800: Produção e comunicação de normas técnicas

Related terms:

150.10
150.20
300.30
450.20

Scope Note:

No que respeita à elaboração de diplomas jurídico-normativos, deve ser sempre considerada nesta área a iniciativa (sob a forma de projeto ou de proposta) e os procedimentos orientados à fixação da respetiva redação final...

Estas classes e propriedades correspondem aos nossos blocos de construção que podemos usar para construir taxonomias: skos:concept, skos:broader, skos:narrower, skos:scopeNote, etc.

OWL

Até agora, já analisamos um formato, o RDF e uma linguagem baseada neste formato, o SKOS. No entanto, em muitos cenários, o SKOS/RDF não é suficiente. Como já se referiu no início, numa segunda fase, pretende-se guardar um registo de contextualização das classes da LC. Neste registo aparecem relações como sucessão e complementaridade que possuem propriedades matemáticas que é importante captar na especificação. Pegando nos dois exemplos dados, temos que a relação de complementaridade é simétrica: se A é complementar de B, então B é complementar de A; a relação de sucessão é assimétrica, por isso terá uma relação inversa: normalmente, nestes casos, usa-se um par de relações: sucede/antecede; e pode ser transitiva: se B sucede a A e C sucede a B, então C sucede a A. A inclusão destas propriedades matemáticas na especificação da ontologia permite que se criem uns programas especiais designados por “reasonners” que conseguem inferir conhecimento novo a partir do que já está especificado de forma declarativa.

Foi para estes cenários que surgiu a OWL (“Ontology Web Language”), uma linguagem baseada em RDF, RDFS, SKOS, OIL e DAML+OIL que permite especificar ontologias com o grau de complexidade que se pretender e que hoje é a norma do W3C nesta área.

Não iremos detalhar aqui as características da OWL. Apenas a mencionamos e descrevemos a razão da sua existência para ilustrar a sua pertinência neste projeto.

Aplicação

Como prova de conceito e dada a proximidade de um dos autores a uma equipa de desenvolvimento de uma aplicação de descrição arquivística, decidiu-se implementar o gerador de SKOS, suportado pela aplicação em causa. A ideia era verificar qual o esforço necessário, depois de termos a LC num formato neutro, para a incorporar numa aplicação de software real, passível de utilizar sobretudo na descrição de documentação capturada e classificada em sistemas de arquivo, seja em sistemas eletrónicos de gestão de arquivo (SEGA), com informação ativa, seja em aplicações de descrição de património arquivístico, por exemplo, num momento seguinte¹.

Os resultados ficaram acima das expectativas. A stylesheet XSLT que implementa o gerador de SKOS desenvolveu-se em cerca de uma hora e a ingestão da ontologia SKOS pela aplicação foi automática.

Foi possível abrir janelas da aplicação de descrição arquivística onde o utilizador está a navegar na MEF (ou na LC completa) para selecionar a classe que pretende associar ao registo que está a editar. Deste

¹ A LC, que integra a MEF, foi concebida para representar os atuais processos de negócio das entidades da Administração, Constitui-se como base para a construção de planos de classificação institucionais ou de comunidades de interesse. Quando integrar todos os processos desenvolvidos por estas entidades será o plano de classificação da Administração. Encontra-se orientada ao presente e ao futuro. Todavia, tal propósito, não é impeditivo para uma visão diacrónica dos processos, que contemple a sua relação evolutiva, passível de ser representada em sistemas integrados de descrição da informação.

exercício se constata que a disponibilização da MEF ou de toda a LC num formato neutro e fácil de processar é uma vantagem para todos e pode ajudar a disseminar este esquema de classificação.

Conclusões

Este projeto, desenvolvido em contexto universitário, procura responder à necessidade do organismo de coordenação da política arquivística, bem como de outras entidades envolvidas na aplicação da ontologia (Lista consolidada ou LC) criada para a classificação da informação pública, possuírem um sistema de informação que a permita gerir de forma eficiente.

Foram ponderados vários formatos mas devido à natureza hierárquica da especificação, aos requisitos de neutralidade relativamente a plataformas tecnológicas, entre outros fatores, optou-se por desenvolver uma linguagem formalizada em XML, com o objetivo de suportar a descrição da Lista.

A principal conclusão que se pode tirar deste trabalho é que o formato XML criado para a LC veio facilitar a sua utilização em vários contextos e aplicações.

No entanto, não se pode estar a gerar este formato sempre que é lançada uma nova versão destes instrumentos, principalmente porque o processo de migração não é totalmente automático. Por isso, está em curso um novo projeto para tornar o site que se gerou para a LC num sistema colaborativo para a edição e manutenção da norma. Neste sistema, o formato de armazenamento é o formato XML definido neste trabalho.

Será também necessário pensar de que forma é que se irão disponibilizar estes novos recursos associados à LC: site, ontologias nos vários formatos, etc. Os autores esperam que, após a sua formulação no âmbito deste projeto, estes recursos possam ser aceites e geridos pela DGLAB, disponibilizados na Web e utilizados, pelo menos, pela comunidade.

A disponibilização desta ontologia em listas ou diretórios de produtos similares com mecanismos de pesquisa (bibliotecas de ontologias), poderá potenciar a combinação com outros vocabulários normalizados.

Pretende-se ainda que esta ontologia possa ser aplicada em múltiplos contextos - de produção, utilização e reutilização da informação pública, quer em sistemas de informação organizacionais, (como os SEGA, os sistemas de informação da área de negócio, os business intelligence systems); quer em sistemas que potenciem a Web semântica.

Este projeto pode ainda potenciar um serviço central de classificação, responsável pelas atualizações e alterações do sistema de classificação, que disseminaria de imediato a informação para SEGA (records system) desenvolvidos de acordo com os requisitos MoReq 2010, com serviço modular de classificação partilhado.

Neste sentido, esperamos que esta comunicação possa constituir mais um importante contributo para a promoção de soluções favoráveis ao desenvolvimento da sociedade e da economia do conhecimento.

Bibliografia

ALLEMANG, D. & HENDLER, J. (2011) - *Semantic Web for the working ontologist. Effective modeling in RDFS and OWL*. 2 ed. Burlington, MA: Morgan Kaufmann.

Da gestão da informação à gestão do conhecimento

BERNERS-LEE, T.; HENDLER, J. & LASSILA, Ora (2001) - The semantic web. *Scientific American*. (Maio), pp. 28–37.

DLM Forum (2011) - *MoReq2010®: Modular Requirements for Records Systems — Volume 1: Core Services & Plug-in Modules. Version 1.1*. [Em linha]. [Consult. 13 Jul. 2015]. Disponível na Internet <URL: <http://moreq2010.eu/>>.

LEE, D. & CHU, W. W. (2000) - Comparative Analysis of Six XML Schema Languages. *ACM SIGMOD Record* 29, 3, (September). [Em linha]. [Consult. 13 Jul. 2015]. Disponível na Internet <URL: <http://www.cobase.cs.ucla.edu/tech-docs/dongwon/sigmod-record-00.html>>.

LOURENÇO, A. & PENTEADO, P. (2014) – Una estrategia para mejorar el acceso y la reutilización de la información pública en Portugal: el papel de la interoperabilidad semántica. *Girona 2014: Archivos e Industrias Culturales*. Girona: ICA. [Em linha]. [Consult. 13 Jul. 2015]. Disponível na Internet <URL: <http://www.girona.cat/web/ica2014/ponents/textos/id200.pdf>>.

RAMALHO, J. C. (2014) – *Arquivos e Web Semântica*. Lisboa: BAD, 2014. [Em linha]. Acedido Jul. 13, 2015, em <http://hdl.handle.net/1822/29735> (Apresentação efetuada em "Arquivos e Web semântica: ciclo de reflexões e debates", Coimbra, 2014).