

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

A importância das ontologias num Museu Virtual

Alberto Simões
ambs@di.uminho.pt

Tiago Bezerra
tiago@lsd.di.uminho.pt

Pedro Henriques
prh@di.uminho.pt

Neste documento apresentamos a importância do uso de ontologias para a classificação das peças de um Museu Virtual, bem como para permitir ao utilizador uma visita contextualizada e coerente.

1. Introdução

Um Museu Virtual pretende essencialmente disponibilizar “peças de museu” através da Internet, para que os visitantes as possam ver remotamente. No entanto, o facto de ser usada a Internet para a sua disponibilização, leva a que estes museus não sejam organizados como uma sequência de “peças” a visitar, mas um conjunto de “peças” devidamente catalogadas, às quais os visitantes possam aceder directamente, de acordo com os seus interesses.

Neste documento pretendemos frisar a importância, em geral, do uso de ontologias, ou redes semânticas, para a organização de um Museu Virtual. Particularmente, discutiremos a sua importância no Museu da Pessoa (MP), um museu virtual cujas “peças de museu” são Histórias de Vida de pessoas comuns.

Uma ontologia é um conjunto de termos classificativos relacionados entre si. Estas relações, ou ligações, correspondem a diversos tipos de associações entre os termos: de equivalência (sinonímia), de especificidade (contido em) ou generalidade (contém), e outros.

Num museu virtual os visitantes devem ser capazes de navegar entre assuntos, temas, sítios, datas, eventos, sentimentos, impressões, fazendo com que a visita seja agradável e permita realizar novas descobertas, tirar conclusões e adquirir novos conhecimentos.

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

No caso concreto dos museus virtuais, o uso de ontologias permite que se definam relações entre os termos classificativos das “peças de museu”, de modo a criar entre estas ligações que permitirão o traçado dos diversos “percursos de visita”.

Com o intuito de atingir os objectivos referidos, é necessário o uso de tecnologias da informação, tanto nos procedimentos de aquisição de informação, como no armazenamento e disponibilização, tendo como desafio principal uma minuciosa análise da “meta-informação” das peças do espólio, pois é através da dita meta-informação que se faz a catalogação e se permitem as pesquisas. No caso do Museu da Pessoa, a meta-informação das Histórias de Vida será a base para suportar o relacionamento entre os depoimentos e os factos históricos e sócio-culturais.

É evidente que a organização das “peças de museu” do MP, isto é, a classificação e relacionamento das Histórias de Vida, deve ser feita por uma equipa multidisciplinar, com sociólogos, historiadores e informáticos, de modo a assegurar a criação de uma ontologia adequada ao Museu.

Para clarificar, consideremos como caso de estudo, uma ontologia para o Museu da Pessoa. Uma história de vida contém, quase por norma, uma referência ao trabalho ou ofício do entrevistado. Esta profissão é um termo classificativo que deve ser introduzido numa ontologia, juntamente com todas as outras profissões. Cada uma destas profissões irá estar relacionada com um conjunto de histórias, mas também com um conjunto de termos classificativos da ontologia como sejam profissões relacionadas (“peixeira” e “pescador”) ou ferramentas usadas (“pescador” e “rede de pesca” ou “pescador” e “arrastão”).

Consideremos ainda outro exemplo: por vezes, as histórias de vida mencionam desastres naturais ou sociais, como “cheias”, “terramoto” ou “naufrágio.” Em vez de classificar cada documento com um destes termos específicos e também como “desastre”, um museu virtual que tire partido de uma ontologia irá classificar os documentos apenas com o termo específico. Quando o utilizador fizer uma pesquisa por “desastre”, não serão mostrados apenas os documentos directamente classificados como

“desastre” mas todos aqueles que têm um classificador que seja mencionado na ontologia como específico deste.

Como se irá mostrar nesta comunicação, o uso de uma ferramenta informática para o manuseamento de ontologias permite que se definam relações adequadas às necessidades do museu, e também que se definam propriedades matemáticas sobre os relacionamentos, para que a ontologia se mantenha consistente.

2. Ontologias

Uma ontologia é uma especificação ou *formalização de uma conceptualização*[2]. Alternativamente, uma ontologia é uma teoria lógica, a qual dá uma explicação de uma conceptualização, projectada para ser partilhada por agentes com vários objectivos[3]. Uma conceptualização é descrita por um conjunto de conceitos e pelas relações entre si.

Uma ontologia também pode ser vista como uma forma de descrever um entendimento comum partilhável, sobre o tipo de objectos e de relacionamentos que estão a ser descritos, tal que a comunicação possa acontecer entre pessoas e sistemas[4]. Noutras palavras, é a terminologia de um domínio (define o universo de discurso). Como um exemplo real, pode-se considerar um thesaurus usado para a procura uniforme sobre um conjunto de web sites semelhantes, porém independentes.

As ontologias podem ser usadas para:

- Criar um vocabulário central estruturado, para ser usado e validado por um conjunto de actores numa comunidade;
- Definir e usar relacionamentos lógicos e regras entre os conceitos, permitindo um uso eficiente de agentes inteligentes;
- Desenvolver, manter e publicar conhecimento (que pode sofrer mudanças rapidamente) sobre uma organização (toda ou parte dela), provendo facilmente visões distintas sobre a mesma.

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

Num museu, as ontologias servem de base para a catalogação do seu espólio. Assim como numa biblioteca se usa uma taxonomia para a classificação dos livros, também é importante a existência de uma estrutura clara para a classificação das “peças de museu”.

Para um museu virtual, as ontologias têm uma importância redobrada. São cruciais não só para a classificação das suas obras, como também para permitir o seu inter-relacionamento, facilitando ao visitante diferentes percursos de exploração.

No Museu da Pessoa definimos um thesaurus para a classificação e estruturação das Histórias de Vida recolhidas. Esta ontologia enquadra informação de diferente tipo, como informação geográfica (locais, cidades, países), histórica (monumentos, grandes guerras, revoluções) e social (profissões, festas).

Segue-se um pequeno extracto do thesaurus do Museu da Pessoa, e respectiva explicação do formato usado:

pesca

BY pescador

BT actividade

NT pesca do bacalhau

RT barco, naufrágio

barco

RT naufrágio, pescador

profissão

NT violeiro, jornaleiro, tanoeiro

NT leiteira, carpinteiro, artesão

NT empregada de servir, agricultor

NT comerciante, pedreiro, amolador

NT relojoeiro, metalúrgico, padeira

NT torneiro, taberneiro, marinheiro

NT pescador, peixeira, lavrador

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

Cada bloco deste texto corresponde à definição de um termo ou conceito. A primeira linha de cada um destes blocos contém o termo representativo do conceito. Neste extracto, definiu-se três conceitos: pesca, barco e profissão. Segue-se um conjunto de linhas que definem relações entre este conceito e outros conceitos. No início de cada linha, é definido o relacionamento do conceito actual com os outros conceitos: RT (Related Term) indica conceitos relacionados; BT (Broader Term) refere-se a conceitos mais amplos; NT (Narrow Term) refere-se a conceitos mais específicos; BY indica quem realiza o conceito; etc.

Por exemplo, no extracto diz-se que a "pesca" é uma tarefa relacionada com barcos, e naufrágios. Também, é um caso particular de uma actividade (conceito mais genérico). Por sua vez, a "pesca do bacalhau" é um conceito mais específico, já que se refere apenas ao bacalhau.

3. Anotação de Peças de Museu

A anotação, ou catalogação, das peças de museu é o processo mais importante para o bom funcionamento de um museu virtual. Sem uma catalogação consistente não será possível aceder às peças de museu e inter-relacioná-las.

Este processo deve envolver uma equipa multidisciplinar, incluindo informáticos que ajudem na organização da base de dados, historiadores que contextualizem historicamente as "peças de museu", sociólogos para relacionar as peças com a vivência actual da sociedade, museólogos e outros técnicos.

Para exemplificar a importância da anotação, e como esta pode ser obtida, apresentamos de seguida um resumo deste processo no Museu da Pessoa[1].

No Museu da Pessoa, as "peças de museu" são, como já se disse, Histórias de Vida das pessoas comuns. Estas histórias são recolhidas numa entrevista, gravada em áudio ou vídeo. Posteriormente, são transcritas, anotadas e editadas. A anotação é o processo que aqui queremos realçar. Esta é feita no próprio documento, depois de transcrito,

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

adicionando meta-informação através da adição ao texto de etiquetas ou marcas adequadas, como se pode ver no exemplo seguinte:

```
<h tipo="Episódio" tit="O Dione que ficou sem leme">
  E numa dessas viagens aos <ref>Açores</ref>, no
  navio Dione, à vinda para <ref>Lisboa</ref> o
  <ref>barco</ref> largou o leme. A gente vinha a governar
  em cima e o leme andava sempre à roda em baixo. Tentávamos
  guiar, mas não conseguíamos governar o barco porque
  não tinha o leme debaixo. Andámos oito dias à deriva em
  alto-mar a apanhar vagas muito altas. Nenhum salva-vidas
  nos foi buscar, foi um arrastão de pesca que nos deu
  reboque até aos estaleiros. Tivemos o
<ref>cozinheiro</ref>
  que ficou louco, começou a ver coisas e queria matar-me.
Houve
  uma altura, em que pensámos matar o cão para comer, porque
  não tínhamos nada para comer. O cão, infelizmente, também
  caiu pela borda fora e foi para o mar. Há coisas que
naquele
  tempo passavam-se com maior dificuldade.
</h>
```

Neste texto foram adicionadas etiquetas XML. Estas etiquetas, que em muitos casos podem levar atributos além do identificador, delimitam determinado texto, ou palavras, e associam-lhes meta-informação. A primeira etiqueta, h, define que o seu conteúdo é um episódio, e associa-lhe um título (“O Dione que ficou sem leme”). Por sua vez, as etiquetas ref associam um tipo a determinadas palavras relacionando-as (ou fazendo referência) com a ontologia existente. Neste caso, definimos que as palavras “Açores”, “Lisboa”, “barco” e “cozinheiro” serão interligadas com a ontologia, para as classificar e dar significado.

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

Um visitante do museu que procure “pescador” e navegue pelos conceitos relacionados, irá encontrar conceitos como “barco”. Visitando o termo “barco”, o sistema irá apresentar este episódio. Ao ler este episódio, o termo “cozinheiro” irá conter uma hiper-ligação que permite aceder a mais histórias relacionadas com cozinheiros.

4. Tecnologias de Suporte e Exploração

O uso de uma ontologia para dar vida a um Museu Virtual obriga à utilização de um conjunto de ferramentas de suporte (catalogação e indexação) e de exploração (navegação sobre os documentos).

Para o processo de catalogação, os termos a associar a cada documento devem pertencer a um conjunto fechado, para que duas pessoas não associem termos diferentes para o mesmo conceito. A forma mais simples de limitar o léxico é, como se tem vindo a defender, a criação de uma ontologia, e extrair os termos usados.

Uma vez que cada catalogador não conseguirá memorizar todos os termos passíveis de serem utilizados, é importante o uso de uma ferramenta para consultar a ontologia, e adicionar termos não preferenciais à ontologia. Estes termos não preferenciais funcionam como possíveis pontos de entrada no thesaurus, que os faz corresponder a termos preferenciais.

A gestão de toda a ontologia pode ser feita automaticamente. Por exemplo, o módulo Perl Biblio::Thesaurus que usamos no Museu da Pessoa permite a definição de regras matemáticas sobre os relacionamentos da ontologia para a manter coerente.

Como foi referido na secção anterior, é importante adicionar meta-informação às peças de museu. Esta meta-informação pode ser adicionada num catálogo ou base de dados, mas também pode ser adicionada a cada um dos documentos.

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

No caso do Museu da Pessoa, a meta-informação é extraída dos documentos XML, e é criado um catálogo. Este catálogo, é manuseado por um módulo Perl denominado `Biblio::Catalog`.

Neste último caso, a construção de um catálogo que liste toda a meta-informação das “peças de museu”, é um processo importante para realizar o relacionamento entre a ontologia e os documentos. Deste catálogo irão ser extraídos todos os termos usados para catalogar as peças e criados índices remissivos. Estes índices são usados para, a partir de um conjunto de termos de pesquisa, aceder directamente às respectivas obras.

A navegação na árvore da ontologia deverá permitir o acesso constante à lista de peças de museu catalogadas com o termo actualmente em visita.

Também para o processo de criação de índices remissivos, e de todo o processo de navegação sobre a ontologia e catálogo, foi desenvolvido um outro módulo Perl denominado `Biblio::WebPortal` que trata de todo o processo envolvido.

5. Conclusões

Embora não tenhamos entrado em detalhes técnicos da implementação de um Museu Virtual baseado numa ontologia, quisemos especialmente demonstrar a importância da devida catalogação das “peças de museu”, bem como da construção de uma estrutura que permita o relacionamento entre termos.

Algumas das técnicas aqui apresentadas estão a ser utilizadas no Museu da Pessoa (<http://www.museu-da-pessoa.net>) e têm vindo a demonstrar grande capacidade para a definição de visitas contextualizadas.

Bibliografia

- [1] J. João Almeida, J. Gustavo Rocha, P. Rangel Henriques, Sónia Moreira, and Alberto Simões. Museu da pessoa — arquitectura. In *ABAD - Associação Bibliotecários e Arquivistas*, 2000.

II Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural

Centro Cultural de Paredes de Coura
28 a 30 de Outubro de 2004

- [2] Thomas R. Gruber. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. In N. Guarino and R. Poli, editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Deventer, The Netherlands, 1993. Kluwer Academic Publishers.
- [3] Nicola Guarino and P. Giaretta. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In N. Mars, editor, *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, pages 25–32. Ed. Amsterdam: ISO Press, 1995.
- [4] Ann Wrightson. Topic Maps and Knowledge Representation. *Ontopia*, February, 2001. <http://www.ontopia.net/topicmaps/materials/kr-tm.html>.