

Web Semântica e e-Learning juntos por uma boa causa

Vitor Gonçalves¹, Eurico Carrapatoso²

¹ Escola Superior de Educação, Instituto Politécnico de Bragança,
Campus de Santa Apolónia, Apartado 1101, 5301-856 Bragança, Portugal
vg@ipb.pt

² Faculdade de Engenharia da Universidade do Porto, INESC Porto,
Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal
emc@fe.up.pt

Abstract. A sociedade em que vivemos exige-nos uma contínua e rápida actualização dos conhecimentos para que possamos adaptar-nos e reagir às constantes mudanças profissionais e sociais. A interiorização do conceito de educação ao longo da vida tem vindo a impulsionar novas formas de auto-aprendizagem, tais como: a aprendizagem através de conteúdos Web dispersos localizados por motores de busca e a aprendizagem através de conteúdos disponibilizados por sistemas de e-Learning. No entanto, o crescimento da Web tornou cada vez mais problemática a descoberta e a recuperação desses objectos de aprendizagem. Neste sentido, propõe-se, por um lado, o recurso a metadados e ontologias para a pesquisa de objectos de aprendizagem dispersos e, por outro, o uso de mapas de conceitos, descritos através da tecnologia XTM (*XML Topic Maps*), para organizar e apoiar a recuperação de objectos de aprendizagem em cursos de e-Learning e, conseqüente, geração de planos de formação personalizados. Assim, este artigo sugere uma arquitectura para um sistema de pesquisa de conteúdos de aprendizagem baseado em tecnologias de e-Learning, tecnologias da Web Semântica e tecnologias de Agentes.

1 Introdução

A Internet assume-se como um ambiente repleto de recursos e serviços para formandos e formadores, mas também para todos aqueles que já interiorizaram o conceito de educação contínua. Com a Internet surgiram novas estratégias de difusão de informação, novos modelos de comunicação e novas formas de ensino/aprendizagem.

Aparentemente, a World Wide Web (WWW ou simplesmente Web) pode comparar-se a uma gigantesca mediateca de recursos ao dispor de todos aqueles que tenham acesso a um equipamento com ligação à Internet. Professores, alunos e cidadãos em geral dispõem de oportunidades de aprendizagem ou de geração de novo conhecimento que nem sempre é possível concretizar dadas as dificuldades de localização das fontes de informação úteis. Mas a sociedade actual necessita de pessoas que aprendam e se adaptem depressa, pelo que os motores de busca e as arquitecturas de ensino à distância baseadas na Web têm que ser repensados com vista a responder de forma rápida, precisa e personalizada aos requisitos dos utilizadores.

A Web foi criada com a visão de que seria um espaço onde a informação teria um significado bem definido, facilitando a cooperação e a comunicação entre as pessoas e

os agentes de software [Berners-Lee et al. 2001]. Passada mais de uma década do seu lançamento, essa visão ainda não foi alcançada. A Web continua organizada na perspectiva dos humanos (linguagem natural e ênfase na apresentação através da linguagem HTML), o que torna a busca de informação uma tarefa árdua. Há que encontrar soluções de representação, organização, integração, intercâmbio e compreensão semântica da informação tanto na óptica dos humanos, como na óptica das máquinas ou dos agentes de software. É neste contexto que surge a denominada Web Semântica (*Semantic Web*) [W3C 2001] como um cenário desejável para a Web.

Convictos de que esta evolução beneficiará as actividades de aprendizagem, mas cientes de que o processo de aprendizagem (como aprender) e os modelos de estruturação de conteúdos (como ensinar) não devem ser descurados, o presente artigo explora as vantagens que a visão da Web Semântica pode provocar nos sistemas de e-Learning, sugerindo o recurso metadados, ontologias e a mapas de conceitos (e XML *Topic Maps*), para orientar os agentes de software na localização, recuperação e reutilização de objectos de aprendizagem.

2 Tecnologias para o e-Learning

Um sistema de e-Learning representa um ambiente de aprendizagem, no qual a distribuição de conteúdos multimédia, a interacção social e o apoio na aprendizagem são suportados pela Internet ou por uma Intranet ou Extranet. Genericamente, o e-Learning preocupa-se, por um lado, com a comunicação entre o professor e o aluno (interacção social e intra-pessoal) e, por outro, com os conteúdos do curso (interacção com os recursos). É esta última componente que interessa no âmbito deste artigo.

Nesta última década têm vindo a surgir no mercado plataformas adequadas à criação de ambientes de aprendizagem hipermédia ou sistemas de e-Learning que integram tecnologias de comunicação que promovem a interacção síncrona e assíncrona, tecnologias Web que permitem a distribuição dos conteúdos e a implementação de ambientes colaborativos e ferramentas de autor para a criação de aplicações multimédia/hipermédia. Actualmente, as plataformas que mais se destacam são: *WebCT*; *Lotus*; *Virtual-U*; *TopClass*; *Blackboard*; *Luvit*; *Atutor* e *Moodle*, entre outras plataformas de gestão de conteúdos na Web, tais como: *Zope*, *Mambo* ou *Joomla!*.

Não menosprezando as vantagens de cada uma destas plataformas, a verdade é que muitas das experiências de implementação de sistemas de e-Learning não têm obtido resultados satisfatórios devido, essencialmente, aos conteúdos disponibilizados. A transposição simples e directa dos conteúdos do ensino presencial para um formato digital (por exemplo, PDF) não é a solução adequada para criar cursos de e-Learning com qualidade e susceptíveis de proporcionar resultados benéficos.

As tecnologias, por si só, não são suficientes para obter sistemas de e-Learning adequados às necessidades dos professores e alunos, pelo que o modelo pedagógico inerente à concepção, organização e estruturação dos conteúdos não deve ser descurado. O modelo educacional desejável para a Sociedade da Informação e do Conhecimento assenta na abordagem construtivista, segundo a qual a aprendizagem é um processo de índole social, e não apenas cognitivo e individual, através do qual o conhecimento é construído pelo aluno aquando da interacção da sua base de conheci-

mentos com as novas experiências de aprendizagem e com outros intervenientes do processo educativo. Muitas vezes, a aprendizagem deriva do acesso a informações bem estruturadas, o que corresponde a uma forma inovadora de aprendizagem que requer um papel mais pró-activo por parte do aluno.

Neste sentido, acreditamos que o uso de Mapas de Conceitos para representar o conhecimento nos sistemas de e-Learning (e nos sistemas de gestão de conteúdos) favorecerá não só a organização, localização, recuperação e reutilização de recursos educativos (objectos de aprendizagem), mas também o processo de aprendizagem (quer na planificação de cursos, quer na organização mental dos conceitos).

Os mapas de conceitos baseiam-se na teoria construtivista da aprendizagem significativa formulada por Ausubel e constituem uma representação visual que relaciona conceitos hierarquicamente. Constituem simultaneamente um instrumento para apoiar processos de auto-aprendizagem, uma estratégia de organização de materiais educativos e um modelo para localizar e elucidar o significado dos recursos [Paz et al. 2005].

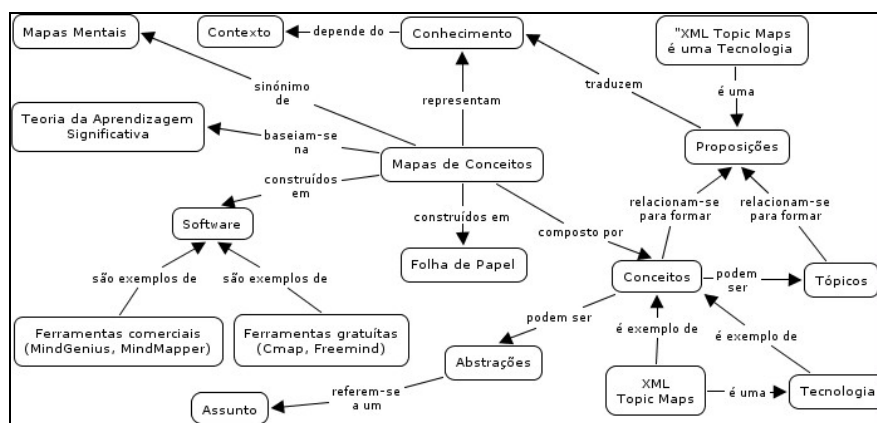


Fig. 1. Um Mapa de Conceitos sobre “Mapa de Conceitos” editado com CmapTools

Mais recentemente, surgiu uma abordagem análoga, designada por Mapas de Tópicos (*Topic Maps*) [Park e Hunting 2003]. Com vista à descrição, manipulação e intercâmbio de mapas de tópicos no âmbito da Internet, foi desenvolvida uma linguagem baseada em XML (*eXtensible Markup Language*) designada XTM (*XML Topic Maps*) [Pepper e Moore 2001]. Embora com finalidades diferentes, os Mapas de Conceitos assemelham-se aos Mapas de Tópicos, sendo possível a tradução de uns nos outros. Existem algumas ferramentas para a criação e navegação em Mapas de Conceitos, das quais destacamos a CmapTools que permite a importação/exportação do mapa de conceitos no formato normalizado para mapas de tópicos XTM, bem como no formato proprietário e específico XCM (*XML Concept Maps*), não esquecendo a possibilidade de exportar para o formato HTML que permite a navegação nos tópicos *browsers* Web [Cañas 2004].

Um conteúdo bem estruturado e de fácil utilização favorece a ocorrência de aprendizagem. Nesta perspectiva, os objectos de aprendizagem (LOs – *Learning Objects*) assumem-se como agentes catalizadores na forma de planear e produzir conteúdos para e-Learning, devido às suas potencialidades de reutilização e interoperabilidade.

Embora o LOMWG (*Learning Objects Metadata Working Group*) do IEEE-LTSC apresente uma definição mais ampla, vamos assumir que um LO é um recurso digital (texto, imagem, som, vídeo, *applet* Java, filme flash, programa de simulação, etc.) que pode ser reutilizado para apoiar a aprendizagem [LOMWG 2002].

A granularidade de um LO pode variar da simples imagem até ao currículo completo de uma lição (por exemplo, site ou conjunto de páginas Web). Não obstante, quanto maior for a dimensão do LO, menor será a sua possibilidade de reutilização.

Com vista a normalizar a produção de conteúdos de aprendizagem, a iniciativa ADL desenvolveu um modelo de referência designado por SCORM (*Sharable Content Object Resource Model*). Este modelo é um conjunto de normas, especificações e orientações técnicas para o desenvolvimento de conteúdos de aprendizagem, de forma a garantir a reutilização, interoperabilidade, durabilidade e acessibilidade [ADL 2004].

O Modelo SCORM não é mais do que um manual de boas práticas composto de 4 manuais técnicos. O recurso à norma de metadados LOM e à norma IMS *Content Packing* (pacote em formato zip contendo todos os conteúdos - SCO (*Shareable Content Packing*) - e um ficheiro xml (*imsmanifest.xml*) que inclui os metadados, a sequência de navegação e todos os recursos associados) é crucial para o desenvolvimento e a reutilização e intercâmbio de objectos de aprendizagem [IMS 2004]. Para auxiliar este processo existem *Metadata Generator Tools* (ferramentas de geração de metadados) e ferramentas gráficas baseadas no modelo SCORM: *Reload SCORM Editor* e *Reload SCORM Player*, *SCORMxt* ou *eXe - E-learning XHTML Editor*.

3 Tecnologias para a Web Semântica

A primeira geração da Web preocupou-se com a implementação da infra-estrutura tecnológica do serviço Web (o conteúdo era disponibilizado através de páginas estáticas), a segunda incidiu na construção de aplicações Web (o conteúdo passou também a ser disponibilizado através de páginas dinâmicas geradas a partir de base de dados) e, actualmente, a terceira avança rumo a uma Web mais inteligente – a Web Semântica. O primeiro passo para desencadear esta evolução será organizar e estruturar as informações e o segundo será adicionar semântica às informações da Web, de tal forma que os agentes de software possam compreendê-las [Goñi et al. 2002], tal como ilustra a infra-estrutura genérica da Web Semântica (figura 2).

As camadas mais baixas desta estrutura são cada vez mais usadas nos sistemas de informação para Web e constituem a camada esquema (*Schema Layer*). São as camadas responsáveis por estruturar os dados e definir o seu significado para que se possa elaborar um raciocínio lógico. Nesta camada, para além da definição de que todos os documentos, objectos e eventos são recursos identificados por URIs (*Uniform Resource Identifiers*) e que a sua identificação universal é garantida pela norma Unicode (norma internacional para representar ou codificar caracteres como inteiros, ou seja, norma que associa um número a cada caracter), a XML e a RDF(S) constituem os principais elementos chave na estruturação e atribuição de significado aos dados. Por sua vez, a camada ontológica (*Ontology Layer*) fornece o significado dos termos

e das teorias do domínio compartilhadas por todos os participantes do domínio. Os modelos de raciocínio com dados semânticos são executados na camada lógica (*Logic Layer*), visto que as camadas superiores da estrutura pretendem fornecer a troca das provas para reconhecer a verdade das inferências obtidas.

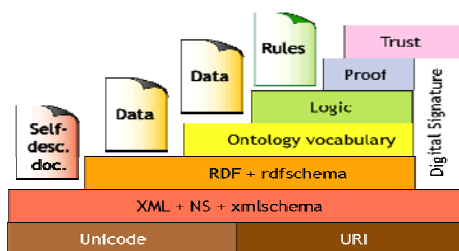


Fig. 2. Infra-estrutura da Web Semântica

Enquanto as páginas possuem apenas informação léxica, os agentes de pesquisa, mesmo os mais avançados, encontrarão um ambiente hostil para a realização das suas tarefas, porque tanto o conteúdo das páginas como o relacionamento entre elas é difícil de ser compreendido pelas máquinas, uma vez que a linguagem HTML apenas descreve como uma página deve ser exibida e não oferece qualquer descrição acerca dos dados. O primeiro passo para solucionar este problema passa pela utilização da tecnologia XML para descrever a estrutura do documento e não apenas a sua apresentação ou aparência (separação entre os rótulos de descrição e a apresentação do conteúdo). Mas a XML nada diz acerca do significado dessa estrutura, pelo que essa tarefa será deixada para as tecnologias de metadados ou metalinguagens (RDF, DCMES, LOM, MPEG7) que permitem descrever a informação de forma não ambígua para depois ser processada pelas máquinas [Thompson 2004].

No contexto particular da Web Semântica a linguagem RDF assume-se como a norma recomendada. Esta linguagem é composta por três tipos de objectos – *resource* (recurso), *property* (propriedade) e *statement* (declaração, sentença, expressão ou afirmação), através dos quais é possível descrever o conteúdo da informação disseminada na rede, definindo recursos (um Website, uma página Web simples, parte de uma página Web ou objectos acerca de um conceito bem definido, devidamente identificados por URIs - *Uniform Resource Identifiers* - que possuem propriedades e valores).

RSS (*RDF Site Summary*), FOAF (*Friend Of A Friend*) e XMP (*eXtensible Metadata Platform*) são algumas das aplicações mais relevantes da linguagem RDF. Podemos afirmar que a RDF está para os computadores assim como a HTML está para as pessoas. Contudo, os metadados, por si só, não permitem estabelecer relações entre os recursos. Se um agente de software, durante o processo de busca, encontrar dois documentos que usam URIs diferentes para o mesmo conceito, desorientar-se-á. Para resolver este problema recorreremos a ontologias (RDFS, OWL). Genericamente e independentemente da ferramenta usada na sua criação (por exemplo: Protégé, SemanticWorks, OilEd, OntoEdit ou SWOOP, entre outras), as ontologias mostram as relações entre os conceitos e fornecem o vocabulário e estruturas de dados compartilhadas, indispensáveis para a comunicação entre os agentes de software [Thompson 2004]. Não obstante, para que os agentes possam raciocinar sobre as

estruturas de dados, tirando partido das relações entre os objectos dessas estruturas esclarecidas através de ontologias, devem previamente ser definidas regras e mecanismos de inferência [Daconta et al. 2003].

As linguagens normalizadas que mais se destacam no âmbito da produção de mecanismos de regras de inferência são: RuleML (*Rule Markup Language*), SWRL (*Semantic Web Rule Language*), N3 (*Notation 3*) e DAML-L (*DAML-Logic*).

4 Tecnologias de Agentes

A Web Semântica assume-se como uma rede de informação que tem como particularidade compreender os conteúdos e interrelacioná-los através de agentes inteligentes.

Com vista ao desenvolvimento de aplicações baseadas em agentes têm vindo a surgir diversas ferramentas para a construção de agentes inteligentes e infra-estruturas ou plataformas de agentes móveis, tanto fruto de projectos ou investigações científicas como de iniciativas de I&D comerciais: AgentBuilder, AgentSpace, AgentTcl ou D'Agents, Aglets ou ASDK, Ajanta, ARA, BEE-GENT, Concordia, Enago Mobile, JACK ou JDE, JADE, Mole, Plangent, TACOMA, Voyager, etc.

Os agentes devem ser capazes de analisar o ambiente em que se inserem e reagir ou responder às alterações ocorridas neste, segundo uma base de conhecimento, isto é, seguindo um conjunto de regras previamente definidas ou entretanto apreendidas, com o intuito de comunicar com o utilizador ou interagir com outros agentes. As ontologias e as regras de inferência podem constituir a base de conhecimento indispensável à actuação dos agentes de software. Ora, se os agentes de tarefas possuírem mobilidade suficiente para migrarem de um ambiente de aprendizagem para outro, então poderemos utilizá-los na construção de um mecanismo de busca baseado em metadados e ontologias, com vista a facilitar o acesso, localização e reutilização do conteúdo educativo, para além de reduzir o tráfego de dados na rede e aumentar o *throughput*.

5 Arquitectura de um Sistema para a Web Semântica

Esta secção tem por objectivo apresentar uma arquitectura baseada num sistema multiagente de agentes inteligentes e móveis e nas tecnologias da Web Semântica onde, através de ontologias será possível localizar e recuperar conteúdos educativos a partir de servidores remotos por parte dos agentes. Para tal, a informação desses servidores deverá estar organizada na perspectiva de objectos de aprendizagem (IMS/SCORM). Para facilitar a pesquisa desses objectos de aprendizagem, os mesmos devem previamente ser descritos através de metadados (RDF/LOM/MPEG7) e o seu significado representado formalmente em ontologias.

Podemos resumir os passos para a realização da pesquisa de objectos de aprendizagem da seguinte forma:

- 1) Formador e formando interagem com o sistema através de uma interface gráfica denominada Agente de Interface, através da qual submetem uma pesquisa sobre conteúdo educativo que desejam localizar.
- 2) O Agente de Interface informa o Servidor de Agentes da necessidade de realizar uma tarefa: procurar informação de carácter educativo (objectos de aprendizagem);
- 3) De acordo com a tarefa de pesquisa requisitada, o Servidor de Agentes cria e envia para a rede um agente orientado à realização dessa tarefa (Agente de Pesquisa);
- 4) O Agente de Pesquisa migra para o servidor de um sistema de e-Learning, de um servidor Web ou de outro qualquer repositório de objectos de aprendizagem;
- 5) Não sendo especialista na comunicação com as bases de dados, o Agente de Pesquisa móvel requisita essa funcionalidade ao Agente de Informação remoto;
- 6) O Agente de Informação efectua a pesquisa a fim de localizar e recuperar os objectos de aprendizagem solicitados pelo Agente de Pesquisa, recorrendo aos metadados (XML/RDF/LOM/MPEG7), às páginas Web anotadas e às ontologias disponíveis de acordo com as regras de inferência previamente estabelecidas;
- 7) Enquanto o Agente de Informação realiza a sua tarefa, o Agente de Pesquisa segue o seu itinerário, migrando para outro servidor Web;
- 8) Caso existam objectos de aprendizagem, o Agente de Informação devolve directamente o resultado ao Servidor de Agentes (lista de recursos e respectivos links);
- 9) Após cumprido o itinerário, o Agente de Pesquisa informa o Servidor de Agentes do seu regresso, para que este proceda à sua destruição.
- 10) Caso os objectos de aprendizagem se enquadrem em pelo menos um dos mapas de conceitos ou mapas de tópicos existentes no sistema, os mesmos são incluídos como links dos conceitos ou tópicos descritos em XTM.

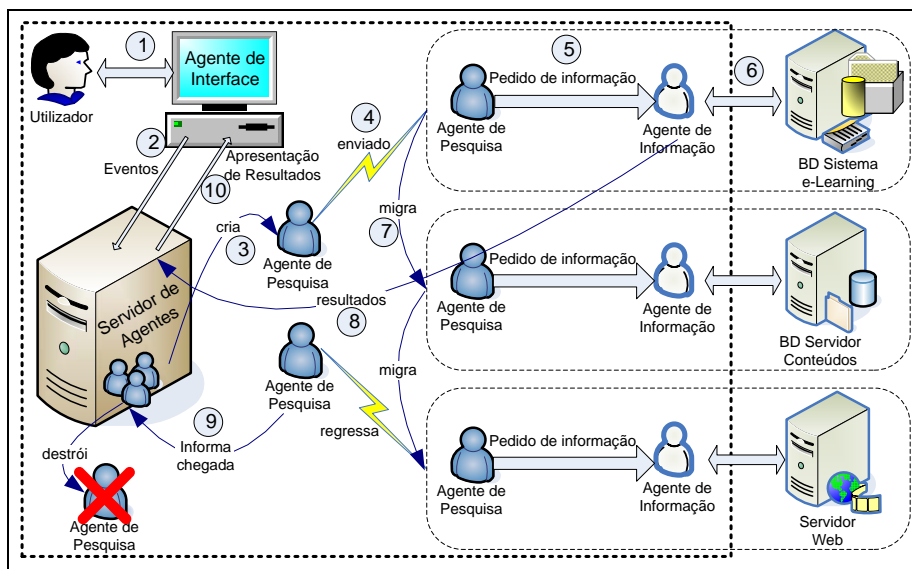


Fig. 3. Arquitectura Genérica de um Sistema de e-Learning para a Web Semântica

A figura 3 ilustra sucintamente uma arquitectura para um sistema que permita a localização de conteúdos educativos e eventual reutilização na geração automática de planos de formação sob a perspectiva de mapas de conceitos.

Em suma, este sistema (ainda em construção) pode comportar-se como um “sistema de e-Learning” suportado essencialmente por agentes que disponibilizam informação ou geram planos de formação à medida do utilizador, auto-alimentando-se com conteúdos de qualidade dispersos na Web ou distribuídos em servidores de e-Learning que permitem a descoberta e recuperação por parte do sistema proposto.

Quanto à construção do sistema multiagente, a plataforma mais indicada para a implementação da tecnologia de agentes inteligentes móveis inerentes a esta situação concreta é a plataforma Java Voyager, devido essencialmente a características, tais como: flexibilidade da arquitectura em camadas, portabilidade, interoperabilidade, segurança, confiabilidade e escalabilidade.

Genericamente, a arquitectura do servidor de agentes é composta por uma base de dados, quatro objectos que implementam as funcionalidades aos restantes componentes do sistema e um objecto que implementa as funcionalidades de gestão do sistema, desenvolvidos na linguagem Java com recurso ao NetBeans. A interface do formador e do gestor permitem a importação de mapas de conceitos ou mapas de tópicos (gerados por exemplo com a ferramenta CmapTools) para que, posteriormente, os objectivos de aprendizagem solicitados pelo formando e devolvidos pelos agentes possam ser integrados e apresentados num plano de formação descrito em XTM.

Para testar a funcionalidade do sistema proposto, foram entretanto construídos dois sistemas de e-Learning (Atutor LCMS e Moodle LMS) não só por se tratar de *open source* o que os torna interessantes para instituições de ensino/formação que não pretendem adoptar uma filosofia comercial na disponibilização de conteúdos educativos, mas também porque nos pareceram as mais acessíveis para aplicar as tecnologias da Web Semântica aos conteúdos a disponibilizar. Pelas mesmas razões, foi construído um Website ou Portal de conteúdos educativos baseado na plataforma Mambo com vista a suportar a publicação dos conteúdos educativos de pelo menos um professor. Paralelamente, foi também desenvolvido um Website pessoal baseado nas tecnologias para a Web Semântica (VGportal e sistemas de e-Learning disponíveis em <http://www.vgportal.ipb.pt>).

6 Considerações Finais

À medida que a Web cresce, a descoberta e a recuperação de materiais educacionais úteis torna-se cada vez mais problemática. Com vista a resolver este problema, é necessário criar uma estrutura de metadados e ontologias para e-Learning baseada nas tecnologias da Web Semântica que satisfaça as necessidades dos utilizadores aquando da busca de materiais educativos úteis na Web. Nesta perspectiva, sugere-se, ao nível dos recursos, o uso de metadados e ontologias OWL para a pesquisa de objectos de aprendizagem dispersos e, ao nível da representação do domínio de conhecimento, o uso de mapas de conceitos, descritos através da tecnologia XTM (XML Topic Maps), para apoiar a recuperação de objectos de aprendizagem organizados em cursos de e-Learning e, conseqüente, geração de planos de formação personalizados.

Embora existam alguns projectos (dos quais se destaca o Metamorphosis da Universidade do Minho) que tentam resolver os mesmos problemas ou usam algumas das tecnologias aqui abordadas, não nos pareceu que nenhum deles tivesse a abrangência e aplicação do que aqui se propõe, uma vez que alguns desses sistemas ou propostas se limitam a Intranets ou a uma tecnologia ou camada específica da infra-estrutura da Web Semântica. Embora não dispúnhamos de informação suficiente para efectuar uma comparação válida, parece-nos que o sistema proposto poderá situar-se ao nível de uma Extranet, integrando os recursos educativos de alguns desses sistemas através dos agentes inteligentes e móveis.

Sabendo que nem tudo que luz é ouro e que a educação não é um “produto” fácil de embalar, armazenar e distribuir, acreditamos nesta visão integradora da Web Semântica e do e-Learning juntos por uma boa causa – a auto-aprendizagem.

Referências

- ADL: Sharable Content Object Reference Model. In ADL <http://www.adlnet.org> (2004)
- Berners-Lee, T. et al.: The semantic web. In Scientific American (2001)
- Cañas, A. et al.: CmapTools, Web Pages & Website : Technical Report IHMC (2004)
- Daconta, M. et al.: The Semantic Web - A guide to the future of XML, Web Services and Knowledge Management. Wiley Publishing, Inc, Indianapolis, USA (2003)
- DCMI. Dublin Core Metadata Initiative. In DCMI <http://www.dublincore.org> (2002)
- Goñi, J. et al.: E-learning e a Web Semântica. VI Congresso da RIBIE, Pontifícia Universidade Católica do Rio de Janeiro, Brasil (2002)
- IMS: Instructional Management Systems Project Learning Resource Metadata Specification. In IMS Global Learning Consortium, <http://www.imsglobal.org> (2004)
- LOMWG: Standard for Learning Object Metadata. In IEEE-LTSC Committee, <http://ltsc.ieee.org/wg12/index.html> (2002)
- Park, J e Hunting, S.: XML Topic Maps. Addison Wesley (2003)
- Paz F. et al.: XML Topic Maps e Mapas de Conceitos. In actas XATA2005 – XML: Aplicações e Tecnologias Associadas, 10 e 11 Fevereiro de 2005, Braga (2005).
- Pepper S. e Moore G.: XML Topic Maps (XTM) 1.0. TopicMaps.Org Specification, August, 2001. <http://www.topicmaps.org/xtm/1.0/> (2001)
- Thompson, J. Features: Semantic Web. In PC Pro Computing in the Real World (2004)
- W3C. Semantic Web. In WWW Consortium, <http://www.w3.org/2001/sw/> (2001)