

Fronteiras incertas: da integração dos sistemas à expansão dos serviços

MARIA INÊS CORDEIRO | JOAQUIM DE CARVALHO

PALAVRAS-CHAVE | TECNOLOGIAS *WEB* | SISTEMAS DISTRIBUIDOS | INTEROPERABILIDADE
SERVIÇOS *WEB* | SISTEMAS DE INFORMAÇÃO DE BIBLIOTECA

R E S U M O

A evolução da tecnologia *web* é uma realidade que extravasa o campo tecnológico e que reflecte alterações mais profundas, de vária ordem, que tomam várias formas à superfície dos sistemas de informação. A compreensão da tecnologia é, por sua vez, um factor fundamental para analisar e muitas vezes reorientar, essas transformações.

Neste artigo é analisada a tecnologia dos serviços *web* nessa perspectiva, realçando o seu alcance e significado. O artigo conjuga uma explanação dos componentes técnicos e conceptuais subjacentes aos serviços *web* com perspectivas sobre as suas potencialidades aplicadas à gestão de serviços de informação, especialmente do ponto de vista das bibliotecas.

A B S T R A C T

The nature of the web technology evolution goes far beyond the technological realm and reflects transformations that are, in many respects, more profound. Yet, the understanding of technology as such is a fundamental step to perceive, and often to re-orientate, such transformations.

In this paper the web services technology is approached with that perspective in mind, underlining their reach and meaning.

The paper combines an explanation of the technical and conceptual constituents of web services with prospects about their potentials applied to the management of information services, especially taken from the libraries' point of view.

INTRODUÇÃO

As drásticas mudanças por que tem passado o ambiente informacional em que vivemos nos últimos anos têm sido geradoras de potencialidades e expectativas que modificam a relação prática que existe entre informação, documentação e conhecimento e entre os agentes que operacionalizam estes conceitos. A discussão sobre a importância ou sobre a prevalência, entre a agência da tecnologia e agência humana ou social, nessas transformações, tem alimentado uma dicotomia cuja razão de ser e utilidade é, as mais das vezes, irrelevante.

A evolução da WWW e da sua tecnologia oferece diversos casos que são exemplo da forma recursiva e simbiótica como factores humanos, organizacionais e sociais interagem com aspectos técnicos (RENNIE & MASON 2003) para gerar avanços que são depois, tantas vezes, tidos como puramente tecnológicos. Na perspectiva de que os dois tipos de realidade se influenciam mutuamente, cada vez mais e de forma mais acelerada, este artigo¹ propõe uma reflexão sobre o modelo subjacente a uma nova tecnologia – os serviços *web* – acreditando que ele pode contribuir para se compreender o que está a mudar, em termos conceptuais e estratégicos, na natureza das interacções em rede. Para o âmbito específico que nos interessa, essa compreensão carece, desde logo, de uma apreensão em traços gerais, daquilo que têm sido, para as bibliotecas, esses modelos de interacção.

Nos sistemas de gestão de biblioteca, há mais de uma década que os serviços de interoperabilidade são uma realidade. Neste âmbito, as soluções mais comuns que se podem encontrar cobrem dois aspectos principais: por um lado, a pesquisa e recuperação de dados bibliográficos, através de implementações do protocolo Z39.50²; por outro, a gestão de transacções envolvendo vários sistemas, como é o caso do protocolo ILL³ para gestão de empréstimos entre diferentes instituições ou o das soluções EDI⁴ para integração dos processos de aquisições das bibliotecas com os sistemas de livreiros fornecedores. Todos estes serviços de integração têm sido baseados no modelo cliente-servidor clássico, isto é, implicando conhecimento mútuo e uma estreita interdependência entre os sistemas em causa.

Especialmente desde o desenvolvimento da Internet e da generalização da WWW, que as atenções se têm focado essencialmente nas questões da pesquisa e recuperação

de informação simultânea em vários sistemas. Neste aspecto o Z39.50 tem sido um tema fundamental, por ser a tecnologia mais usada para expor e explorar catálogos bibliográficos disponíveis em rede. No entanto, a expansão da WWW e os avanços da tecnologia *web* têm vindo a diversificar quer os requisitos dos serviços quer as soluções possíveis, ao mesmo tempo que emergem novos modelos funcionais para sistemas federados de descoberta e acesso a recursos de informação, tais como *gateways* e catálogos colectivos virtuais⁵, portais⁶ e sistemas dinâmicos de ligações de referência⁷.

Estes novos modelos não são exclusivos entre si, isto é, os seus objectivos e características não são exactamente coincidentes e podem ser combinados. Em comum têm o facto de consubstanciarem as mesmas tendências que apontam para: I) a necessidade de fazer face a um ambiente informacional mais alargado e diversificado, o que para as bibliotecas tende a ultrapassar a geografia, as funções e as funcionalidades do domínio biblioteconómico tradicional; e II) explorar ao máximo as potencialidades da natureza distribuída e navegacional da própria WWW.

Os novos modelos e soluções acima referidos podem ser vistos como serviços *web* num sentido genérico, porque fornecem ao utilizador final uma nova camada de serviços baseados em tecnologia *web*, nomeadamente usando o protocolo HTTP e muitas vezes o XML. Mas o termo tem também uma acepção diferente e mais específica, que envolve não apenas o que se vê nessa camada superior do serviço final, mas uma concepção funcional diversa, com uma realização tecnológica de características inovadoras e particulares. O que é, então, a tecnologia dos serviços *web* e que diferença e valor vem introduzir no actual panorama? Para responder a esta questão torna-se necessário sintetizar a natureza e vantagens dos serviços *web*, por um lado, e perceber o que as tendências da tecnologia *web* têm até agora produzido em relação com as bibliotecas.

SERVIÇOS WEB: MAIS DO QUE UMA NOVA TECNOLOGIA, UM NOVO CONCEITO

Os serviços *web* são um conjunto de tecnologias e *standards* de interoperabilidade vocacionados especialmente para as características e potencial do ambiente WWW e emergiram para facilitar a integração de aplicações, através da partilha de serviços entre sistemas. A noção de serviço é, pois, fundamental neste contexto:

um serviço *web* refere-se a um componente de *software* disponível em rede que pode ser invocado e fornecer uma determinada informação/funcionalidade ao sistema ou aplicação que o invoca. Nesta acepção, um serviço *web* distingue-se claramente da acepção genérica de serviços *web*, em que os serviços são principalmente accionados por humanos num *browser*, como explica CERAMI (2002:5-6).

A natureza dos serviços *web* é, assim, orientada para aplicações, isto é para servir processos entre aplicações independentes (W3C 2001a, 2002, 2002a), sendo que cada sistema ou aplicação pode actuar como fornecedor ou consumidor desses serviços. Neste sentido, os serviços *web* facilitam mecanismos de automatização da própria WWW, o que é um aspecto fundamental da chamada *web* semântica (BERNERS-LEE, HENDLER & LASSILA 2001; W3C 2001). Um serviço *web* pode, assim, definir-se como um serviço orientado para aplicações que é disponibilizado na WWW usando um sistema de mensagens standard não dependente de nenhum sistema operativo ou linguagem de programação em particular já que usa o XML e o HTTP.

Numa breve síntese das principais características dos serviços *web* destaca-se em primeiro lugar a universalidade que provém dum ambiente tecnológico base como a Internet, através do uso dum conjunto de *standards* absolutamente neutros e independentes como o XML, SOAP¹⁰, WSDL¹¹ e HTTP. Esta neutralidade faz a diferença comparativamente a anteriores plataformas de interoperabilidade como CORBA e DCOM (BLOOMBERG 2001; OGBUJI 2002; STAL 2002). Este aspecto não só reforça a ideia da Internet como uma plataforma de computação em si própria, mais do que um mero mecanismo de transporte de dados (CARPENTER 2003) como também aumenta a ubiquidade que os serviços podem ter: os serviços *web* podem estar em qualquer lugar onde chegue a Internet.

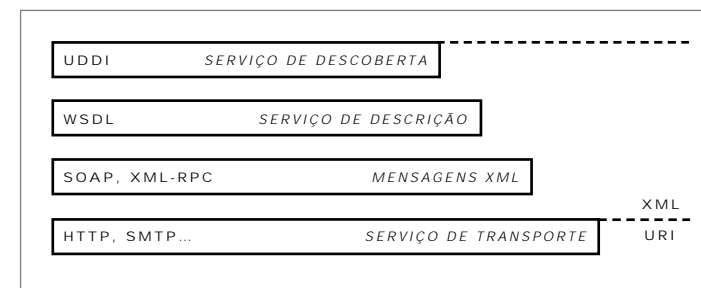


FIGURA 1
PRINCIPAIS CAMADAS/STANDARDS DOS SERVIÇOS WEB

Na vertente funcional, destacam-se os dois pontos fundamentais e inovadores dos serviços *web*: o facto de se auto-descreverem, através de uma linguagem XML convencional para o efeito (WSDL) e de poderem ser descobertos e usados duma forma dinâmica, já que são publicados em directórios especialmente vocacionados para tal (ex. UDDI¹²) e disponibilizados na rede. Estes aspectos constituem o cerne do modelo subjacente aos serviços *web* que pode ser assim esquematizado¹³:

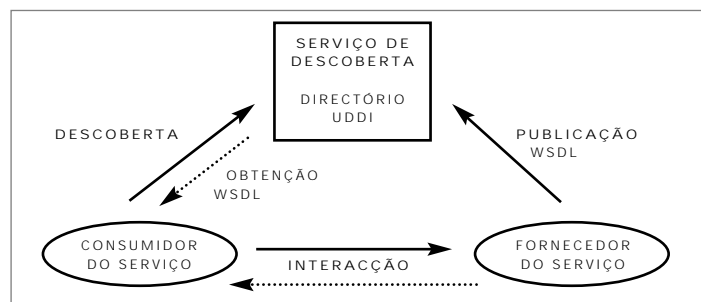


FIGURA 2
MODELO DOS SERVIÇOS WEB

O que as características e modelo atrás apontados consubstanciam é uma nova geração de soluções de interoperabilidade, que abre caminho para uma muito maior flexibilidade e facilidade de implementação de soluções de integração (SAMTANI & SADWANI 2001). As vantagens derivam das três "qualidades" mais frequentemente reconhecidas nos serviços *web*: o facto de serem **encapsulados**, isto é, em que os pormenores de implementação são "escondidos" do sistema consumidor e podem por eles ser ignorados; de não requererem uma estreita interdependência¹⁴ entre sistemas, o que permite que "fornecedor" e "consumidor" sejam geridos e se alterem de forma completamente independente; e **contratados**, isto é, fornecidos com tudo o que é necessário para serem usados por qualquer "consumidor" que lhes aceda, dentro de condições que não deixam de ser controladas pelo "fornecedor".

Esta nova geração de soluções não só modifica o tradicional modelo cliente-servidor das interações em rede (CARPENTER 2003), como também se traduz em modelos de interoperabilidade mais descentralizados, dinâmicos e orientados pela procura (LAYMAN 2001; SEDHUKIN 2001). Ambos os aspectos sugerem novos modos de pensar a constituição de aplicações distribuídas, cuja arquitectura

tende a crescer em modularidade e a recorrer cada vez mais frequentemente a um maior número de componentes externos, reutilizáveis e focalizados em serviços cada vez mais específicos (BURBECK 2000; STEVENS 2002).

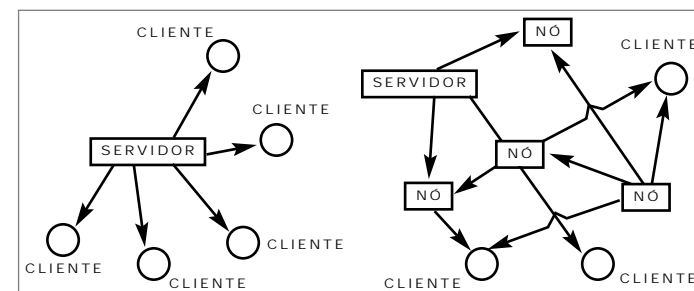


FIGURA 3
EVOLUÇÃO DO MODELO DE INTERACÇÃO

Estas características, vantagens e tendências têm implicações não apenas na gestão da tecnologia em si, mas reflectem-se igualmente numa maior capacidade de exploração e utilização de recursos de informação contidos em cada sistema. No seu todo, podem significar uma maior agilidade dos sistemas de informação em se readaptarem ao contexto e em explorar novas oportunidades (DOVEY 2001; LORENZO 2002; JACOBSON 2002; POWELL & LYON 2002; MCDONALD 2003), quer acrescentando serviços quer revalorizando conteúdos pelo simples facto de os mesmos ficarem mais facilmente reutilizáveis noutros domínios. Este é um ponto particularmente relevante para as bibliotecas, dada a riqueza de informação e funcionalidade que as suas bases de dados encerram e que, pela sua especificidade técnica, não tem sido fácil rentabilizar noutras áreas.

TECNOLOGIAS WEB E BIBLIOTECAS

Os desenvolvimentos que a tecnologia *web* tem suscitado com interesse mais directo para a área das bibliotecas revelam tendências que têm muito de comum com as questões a que os serviços *web* pretendem responder: como "expor" serviços e dados de sistemas heterogéneos, nomeadamente como fazer face à enorme variedade de metadados, e conseguir métodos para a sua descoberta, transformação e reutilização por uma diversidade de sistemas. O resultado desses desenvolvimentos tem sido uma atenção especial às técnicas

de *harvesting*, transformação e agregação de dados usando o XML, o que tem originado novos *standards* para representação e transmissão de metadados (GARDNER, J. R. 2001; CARVALHO & CORDEIRO 2002) tais como Dublin Core, OAI MHP, XMLMARC, METS e OpenURL.¹⁵

Fazendo uso do XML como peça chave da tecnologia *web*, estes novos *standards* vêm facilitar a "autonomização" dos dados relativamente aos respectivos sistemas de origem e facilitar o seu fluxo num ambiente mais alargado. No entanto, até ao momento, tais *standards* têm sido usados essencialmente para construir arquitecturas *a priori*, isto é, ou para montar novos sistemas intermediários ou para acrescentar funcionalidades a sistemas existentes, mas na forma tradicional. Ou seja, trata-se de implementações que requerem um conhecimento prévio – no momento em que é desenhado/montado o sistema/serviço – da variedade de sistemas implicados, tendo muitas vezes ainda que se entrar em linha de conta com as particularidades dos respectivos ambientes tecnológicos. Na sua maior parte, essas arquitecturas visam soluções do tipo B2C (*business-to-consumer*), isto é, são desenhadas para fornecer serviços ao utilizador final e não tanto para explorar a função dos sistemas como consumidores de informação doutros sistemas, no sentido das aplicações B2B (*business-to-business*).

A tecnologia dos serviços *web* vem acrescentar, às vantagens de flexibilidade e escalabilidade do XML aplicado a metadados, um conjunto de metodologias com enorme potencial para qualquer tipo de função: pesquisa, visualização, reutilização ou ligações de referência entre conteúdos de sistemas distintos. Muitos dos serviços federados de descoberta e acesso a recursos de informação, acima mencionados, podem ser mais facilmente viabilizados, e de forma mais flexível, através da montagem de conjuntos de serviços *web* fornecidos directamente pelos sistemas "fonte" da informação, e não através de sistemas terceiros, intermediários.

Actuando como fornecedor de serviços *web*, um sistema pode expor e disponibilizar os seus dados e funcionalidades de modo independente, e até fazê-lo de várias formas, para corresponder a diferentes finalidades/tipos de utilização. Os sistemas consumidores desses serviços podem integrá-los nos seus próprios serviços já existentes, ou simplesmente agregá-los para criar novas funcionalidades e aplicações. Os serviços podem ser descobertos e invocados de forma dinâmica,

e não somente planeados *a priori*. Podem ser usados exactamente na forma como são publicados ou podem fazer parte de um agregado mais complexo de serviços que use diferentes fontes de informação e que reprocessse os dados para finalidades e audiências diversas das originais.

É nestes aspectos que a filosofia dos serviços *web* difere substancialmente da que tem presidido às formas mais comuns de federação de sistemas de biblioteca e afins. De certa forma, esta nova filosofia aproxima-se da que está subjacente à norma OpenURL, que normaliza a forma como os metadados são codificados num endereço *web*. O OpenURL permite que sejam gerados *links* entre dados de sistemas diferentes, numa forma que é entendível por, mas independente de, cada um dos sistemas envolvidos. Este processo de "separar" os dados das ligações dos sistemas de origem e o destino é o que confere a flexibilidade, facilidade de gestão e escalabilidade do chamado *smart linking*: ele expande-se de forma automática e dinâmica à medida que aumenta o número de OpenURL disponíveis para uma comunidade de sistemas.

Embora o OpenURL possa ser visto como uma espécie de "contrato" de alguma forma semelhante aos serviços *web*, existe uma diferença fundamental: o OpenURL apenas normaliza o "caminho" de acesso aos conteúdos, mas não fornece a "inteligência" necessária sobre o formato desses conteúdos para permitir o seu reprocessamento noutros sistemas e aplicações. É neste ponto que cabe falar do Z39.50, precisamente porque é um protocolo estruturado para permitir o reprocessamento de dados por outros sistemas, e perceber a sua relação com os novos modelos de interoperabilidade proporcionados pela tecnologia *web*, nomeadamente no que respeita a serviços *web*.

O Z39.50 E A EVOLUÇÃO PARA SERVIÇOS WEB

Até recentemente, o protocolo Z39.50 tem sido o principal foco das atenções e expectativas quando se fala de interoperabilidade nos sistemas de biblioteca, à volta daquilo que é a parte fundamental dos dados e funcionalidade desses sistemas: os registos bibliográficos e a pesquisa e recuperação de informação bibliográfica. Sendo uma norma que se desenvolveu e evoluiu quase exclusivamente no seio da comunidade bibliográfica, o Z39.50 está particularmente conotado com este domínio e com as suas especificidades

técnicas e normativas. Com excepção de mais alguns domínios, que também acontece serem maioritariamente sectores de serviço público, tais como os de informação governamental, museus e arquivos, o protocolo nunca angariou um interesse muito substancial da parte dos fornecedores comerciais, para além dos que desenvolvem produtos e serviços para bibliotecas ou relacionados com utilizadores de biblioteca, como acontece com o *software* pessoal de gestão bibliográfica. Apesar de a informação contida em sistemas de informação de bibliotecas ser altamente importante para muitos outros sectores, a aplicação do Z39.50 tem sido limitada quase exclusivamente ao domínio das bibliotecas – entre bibliotecas e para bibliotecas (NEEDLEMAN 2000).

Vários factores têm contribuído para esta situação, a começar pelo desenvolvimento um tanto tardio da norma face ao ambiente tecnológico para que foi originalmente pensada (LYNCH 1997). De facto, só após quase duas décadas de desenvolvimento e experimentação, o Z39.50 viu surgir uma massa crítica de implementações, a partir de meados dos anos 90. Nessa altura, com o desenvolvimento e expansão da Internet e WWW, já o contexto em que se enquadravam os modelos de exploração do Z39.50 estava a mudar (MOEN 2000), assim como começavam a ser sentidas algumas limitações e inconvenientes da tecnologia subjacente ao protocolo.

Algumas das características do Z39.50 tornaram-se desfasadas do ambiente tecnológico corrente e esse tem sido um aspecto especialmente referido como pouco atractivo para novos implementadores. Entre essas características salienta-se o facto de ser um protocolo binário, fortemente orientado à conexão, *stateful*, e prescrevendo sintaxes de uso não generalizado como é o caso de ASN1/BER (GATENBY 2000, 2002; NEEDLEMAN 2000). Por outro lado, há a complexidade da própria norma – que cresceu em resultado das muitas componentes opcionais acrescentadas ao longo das várias versões – e o conhecimento especializado que é necessário, sobre as funcionalidades e estruturas dos dados dos sistemas bibliográficos, para a implementar na prática. Quando esse conhecimento não existe as implementações são difíceis e os resultados podem ser também, muitas vezes, decepcionantes (LYNCH 1997; MOEN 2001).

A tecnologia dos serviços *web* abre expectativas para se ultrapassarem alguns dos aspectos acima focados, especialmente os de ordem tecnológica. Várias iniciativas surgiram desde 2001 no sentido de transportar o Z39.50

para realizações tecnológicas mais consentâneas com a tecnologia *web*, com destaque para a actividade liderada pelo *ZING – Z39.50 International: Next Generation*⁶. Como resultado, e fazendo uso dos desenvolvimentos já existentes no campo dos formatos XML para dados bibliográficos, um novo *standard* surgiu no final de 2002, designado SRW – *Search/Retrieve web Service*⁷.

O SRW combina serviços *web* com o Z39.50 clássico, proporcionando assim uma compatibilidade retrospectiva. Em termos tecnológicos destaca-se o uso de XML/SOAP/RPC para a troca de mensagens usando o protocolo HTTP, o recurso a uma linguagem para formulação de *queries* (*CQL – Common Query Language*) baseada em CCL (*Common Command Language*) e expressa em XML; e o uso de XML para a sintaxe dos registos, considerando *standards* como o Dublin Core, ONIX¹⁸, MODS e MARCXML. Em termos de maior facilidade de utilização face à especificidade das normas biblioteconómicas subjacentes, é de referir o estabelecimento de pontos de acesso "*flat*", isto é combinações estaticamente definidas de atributos Z39.50, em vez do uso directo desses próprios atributos.

Estas características visam, à partida, reduzir algumas das barreiras que têm obviado a uma utilização mais alargada do protocolo. Sendo essencialmente uma versão *light* das principais funções do Z39.50, o SRW não cobre todos os serviços e funcionalidades da versão clássica do protocolo, mas pode facilitar a sua exploração em termos básicos e assim aumentar as possibilidades de um uso mais alargado dos repositórios de dados bibliográficos, especialmente por entidades/sistemas fora da área das bibliotecas. Diversos projectos e sistemas planeiam, ou experimentam já, implementações do SRW¹⁹.

Explicadas genericamente as vantagens, é importante reforçar a clarificação entre Z39.50 e SRW, sublinhando que cada um deles aborda níveis de questões diferentes, e que o SRW é, face ao Z39.50, essencialmente uma estratégia de realinhamento tecnológico mais do que uma substituição. Enquanto o SRW representa uma nova forma de realização dos serviços do Z39.50, o protocolo clássico permanece necessário, assim como toda a funcionalidade especializada que nele é modelada, nos seus perfis e arquitecturas de atributos, e que continua válida e essencial para a interoperabilidade interna ao domínio das bibliotecas.

A este respeito, é de notar que continuam relevantes todos os esforços que têm vindo a ser desenvolvidos desde 1999 no sentido de melhor aprofundar,

e normalizar, os perfis de aplicação e os conjuntos de atributos do Z39.50 usados na área das bibliotecas (NEEDLEMAN 2000; MOEN 2000, 2001)²⁰. A prática com o Z39.50 tem demonstrado resultados aquém das expectativas devido a níveis deficientes de acerto de parâmetros comuns, semânticos e funcionais, entre os diferentes sistemas. Tal tem sido documentado em diversos estudos publicados desde 1998 que mostram que os reais impedimentos para uma cabal obtenção dos resultados esperados com o Z39.50 estão para lá da tecnologia do protocolo e dizem sobretudo respeito a correspondências semânticas e funcionais (como as estruturas de dados, índices, etc. disponíveis em cada sistema) que devem existir, para que o protocolo funcione razoavelmente numa comunidade de sistemas (MOEN 2001, 2002). O SRW, só por si, não resolve esse nível de problemas.

PARA LÁ DO Z39.50, UM CAMPO ABERTO DE POSSIBILIDADES

Como tantas vezes acontece, as primeiras aplicações de uma nova tecnologia focam normalmente novas soluções para problemas já anteriores. Também a aplicação mais óbvia de serviços *web* em bibliotecas é a pesquisa e recuperação de informação, na sua forma clássica, e disso é exemplo o SRW. Não é, no entanto, a única forma e outras aplicações mais criativas estão ainda por explorar. Espera-se, por exemplo, que na área da produtividade profissional, os serviços *web* sejam explorados no sentido de vir a viabilizar, para colaborações organizadas, várias das funcionalidades que, para além da pesquisa e recuperação de informação, eram contempladas na última versão do protocolo. Igualmente, é de prever que as soluções de biblioteca que agora se baseiam em EDI, para áreas como as aquisições, empréstimo interbibliotecas e fornecimento de documentos, venham a fazer uso de serviços *web*, alargando assim as possibilidades de interacção com um número muito mais alargado de fornecedores e sistemas.

Um aspecto importante a destacar é o de que a tecnologia dos serviços *web* pode facilitar grandemente a integração de aplicações de diferentes fornecedores, praticamente em modo "*plug & play*". Este é um aspecto promissor no sentido de ultrapassar muitos dos constrangimentos dos tradicionais sistemas automatizados de biblioteca que têm tido uma natureza demasiado monolítica, e também de reduzir as dependências de uma abordagem de gestão que tem sido, até recentemente, baseada num modelo dominante de "fornecedor único".

Este aspecto dos serviços *web* pode aumentar decisivamente a competitividade no mercado das soluções disponíveis e facilitar uma gestão mais flexível das infra-estruturas tecnológicas das bibliotecas. Os sistemas de informação das bibliotecas podem tornar-se bastante mais ágeis em termos da sua adaptação a novas tecnologias e de actualização das suas funcionalidades.

Com efeito, os serviços *web* oferecem um vasto campo de possibilidades para melhorar e expandir a funcionalidade dos sistemas de fornecimento de informação. Podem ser usados para construir funcionalidades especializadas e serviços personalizados ou vocacionados para comunidades específicas, com base em serviços de informação obtidos de vários sistemas terceiros, de natureza variada. Ideias exploratórias dessas possibilidades incluem a construção de *listeners*, isto é, serviços em que a função de pesquisa corresponde a um modelo *pull* de recuperação de informação. À medida que se expandem as redes de informação torna-se cada vez mais difícil para cada "nó" da rede, enquanto consumidor, pesquisar fornecedores com nova informação. Por essa razão, é provável a emergência de um modelo *push* baseado em serviços *web*, em que as aplicações registem *listeners* para novas informações e recebam notificações automatizadas numa base regular.

Outra área em perspectiva é a dos serviços de *harvesting*, *data mining* e de reconstrução de redes semânticas. A informação obtida de sistemas de biblioteca através de serviços *web* pode ser misturada com informação provida de outras fontes tais como a Amazon ou o Google, para fornecer informações e recursos relacionados. Características e funcionalidades de associação, tais como "quem comprou este livro também comprou...", e a co-ocorrência de referências bibliográficas em páginas *web*, podem fornecer dados preciosos para uma exploração mais alargada e inteligente dos recursos bibliográficos. Por exemplo, podem ser desenhados sistemas automatizados que partem da informação sobre determinado livro num catálogo de biblioteca e usem um conjunto de serviços *web* para encontrar, noutras fontes, referências para recursos semelhantes ou relacionados.

SERVIÇOS WEB E BIBLIOTECAS: O EXEMPLO DE UMA APLICAÇÃO PRÁTICA

Damos em seguida um exemplo prático deste tipo de aplicação, totalmente construída usando serviços *web* disponibilizados por várias fontes. Esta aplicação

foi desenvolvida pela BookMARC como demonstração simples das possibilidades criadas pelos serviços *web*. Os serviços utilizados foram: o acesso à PORBASE via os serviços *web* do Sirius, a Amazon e o Google²¹. A aplicação foi demonstrada na conferência Internet Librarian International, em Março de 2003 (CARVALHO & CORDEIRO 2003)

A aplicação funciona como uma "meta-interface" para a informação disponibilizada pelos serviços referidos, mas o seu objectivo vai mais além do que uma simples pesquisa distribuída. A informação de cada fonte é processada e interligada com o intuito de fornecer ao utilizador final um agregado inteligente de itens relacionados entre si. O desenvolvimento da aplicação, desde o início da concepção até ao seu fim, demorou uma semana de esforço a um único programador, o que atesta a facilidade, para o desenvolvedor, destes novos paradigmas tecnológicos.

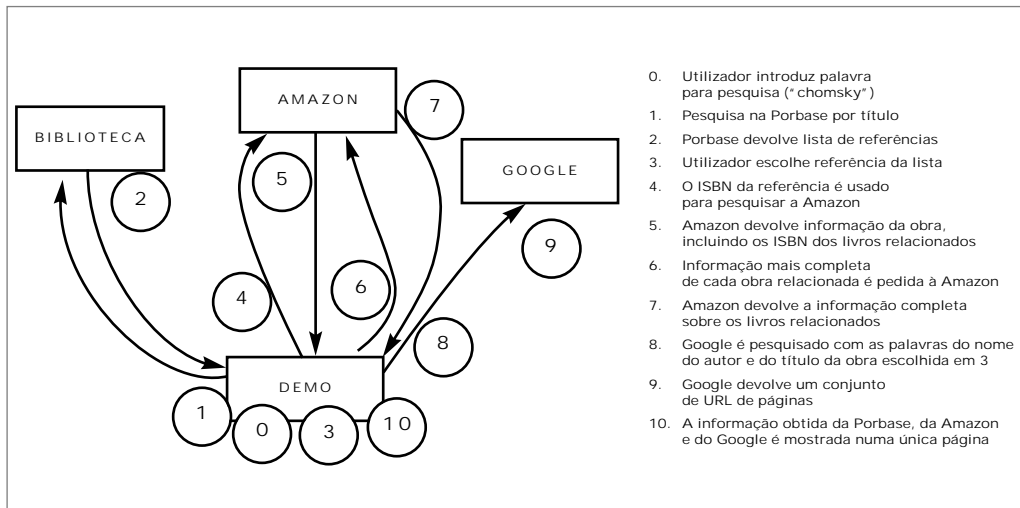


FIGURA 4
FLUXO DE INFORMAÇÃO ENVOLVENDO PORBASE, AMAZON E GOOGLE ATRAVÉS DE SERVIÇOS WEB

O cenário de utilização da aplicação de demonstração é o seguinte:

- A aplicação de demonstração apresenta ao utilizador um formulário simples de pesquisa onde podem ser introduzidas algumas palavras.
- Com essa informação a aplicação pesquisa a PORBASE via serviços *web* e apresenta uma lista de resultados ao utilizador. A informação devolvida

pela PORBASE contém os registos UNIMARC completos, pelo que é possível extrair informação de campos específicos.

- O utilizador escolhe da lista uma referência que lhe interessa particularmente. Num sistema normal esta escolha desencadearia a visualização do registo em questão. Neste caso a aplicação vai extrair informação do registo seleccionado para procurar mais informação relevante na base de dados da Amazon e no motor de pesquisa Google, apresentando uma página agregada com informação relacionada com a escolha feita.

Holdings

COOK, William James
 Chomsky's universal grammar : an introduction
 (Oxford : Basil Blackwell, 1988)
 ISBN: 0-631-15167-8
 Portuguese libraries where you can find this title: [UNICEL](#), [UMSA](#)

You can buy this title from Amazon

Chomsky's Universal Grammar: An Introduction
 William Cook (Blackwell Publishers) (Paperback - September, 1988)
 Out of Print Limited Availability
 Amazon price: \$27.95

Similar titles from Amazon

[Frontiers of the Theory of Syntax](#)
 Noam Chomsky (MIT Press) (Paperback - 10 March, 1965)
 Usually ships within 24 hours
 Amazon price: \$21.00

[Contemporary Linguistics: An Introduction](#)
 William O'Grady, John Archibald Mack, Russell Jones, Roger Miller, St. Martin's Press (Blackwell Publishers) (Paperback - January, 2001)
 Usually ships within 24 hours
 Amazon price: \$28.50

These pages contain related information

[Citations: Chomsky's Universal Grammar: An Introduction - Cook ...](#)
[... William Cook and Mark Newson, Chomsky's Universal Grammar: an introduction ... | Random House](#)
[Notes and Features of Language - James H. Greenberg ... | JSTOR](#)
[Introduction and Universal Grammar research by William Cook](#)
[William Cook, ... the regular and irregular New Zealand in Austroriparian \(J. Green and James, 1988\), ... | JSTOR](#)
[\[N-G\] Natural Language Processing \(NLP\)](#)
[... http://pau.uconn.edu/pub/james@uconn.edu/ftp ... me a proper citation for this?](#)
["Linguistics" William J. Cook, "Chomsky's Universal Grammar: An Introduction ...](#)
[Book: "High Logic"](#)
[... Item's Price: \\$29.95 Price subject to change: Chomsky's Universal Grammar: An Introduction by: Mark Newson, William Cook March, 1988. ...](#)
[\[uk\] Cook - Author](#)
[... Journals of Captain Cook \(Penguin Classics\) by James Cook, Philip Edwards ... Chomsky's Universal Grammar, Chomsky's Universal Grammar by William Cook, Mark Newson ...](#)

FIGURA 5
RESULTADO DA ESCOLHA DE UMA REFERÊNCIA A PARTIR DE UMA PESQUISA SOBRE CHOMSKY

- A partir do ISBN extraído do registo seleccionado a aplicação procura o livro na Amazon. Se o livro existe a aplicação recebe da Amazon a lista dos livros "relacionados", isto é, os livros que foram comprados pelas mesmas pessoas que compraram a referência escolhida. Uma segunda pesquisa na Amazon devolve os detalhes destes livros.
- Paralelamente, a aplicação utiliza os serviços *web* do Google para obter um conjunto de páginas *web* que referem o livro escolhido, pesquisando pelas palavras que compõem o nome do autor e o título.
- A aplicação compõe então uma página *web* onde consta a informação sobre a referência obtida na PORBASE, a informação comercial da Amazon, a lista de livros relacionados e as páginas da Internet que referem o livro.
- Do ponto de vista do utilizador só aconteceram dois passos: introduziu as palavras a pesquisar e, sobre o resultado da pesquisa, escolheu uma das referências. O resultado é uma página sobre a referência escolhida com informação adicional agregada de várias origens.

Esta aplicação de demonstração apenas aflora as possibilidades que as interfaces actuais das fontes de informação escolhidas permitem. As pesquisas no Google poderiam ser expandidas para incluir a informação dos livros relacionados sugeridos pela Amazon. Um desenvolvimento mais complexo poderia mesmo extrair referências bibliográficas adicionais presentes nas páginas da Internet recuperadas. A aplicação pode ser experimentada em <http://ptolemy.bookmarc.pt:8080/birmigham/>.

CONCLUSÃO

O que distingue uma aplicação, como a que foi acima exemplificada, de aplicações aparentemente semelhantes, como as que usam mecanismos de chamado *smart linking*, é o facto de possibilitar ir mais além do que a simples ligação e permitirem o reprocessamento e reutilização da informação de formas completamente independentes dos serviços fornecidos pelos sistemas de origem. Mas mais importante que a diversidade de aplicações práticas que se possam imaginar para os serviços *web*, é o próprio conceito e modelo que lhes está subjacente e que pode suscitar importantes alterações na forma como são tradicionalmente entendidos e perspectivados os sistemas de informação de biblioteca, tal como sugere HICKEY (2002).

Nesta perspectiva, a adopção de tecnologias como a dos serviços *web* coloca algumas questões de estratégia e gestão entre as quais as da pertinência de se reverem os padrões e finalidades de utilização da informação que cada sistema possui e disponibiliza. Significa isto analisar se estão alterados, e como, os âmbitos onde é relevante ser usada a informação contida num dado sistema; se fora do sistema de origem essa informação precisa ou não de ser reprocessada por outros sistemas e de que forma; e se esses sistemas são muito variáveis em termos de dimensão, características, suportes tecnológicos e *know-how*, ou seja, se tem que se fazer face, e como, a ambientes de partilha cada vez mais alargados e cada vez menos controlados.

O que estas questões suscitam, para lá da adequação de determinada tecnologia ou funcionalidade a um dado caso ou momento, é, antes de mais, o reconhecimento de que os serviços *web* representam um ponto de viragem naquilo que têm sido as soluções de interoperabilidade, uma área desde há muito considerada fundamental nos sistemas de informação de biblioteca. Essa viragem não significa apenas alterações em termos de metodologias técnicas, mas também afecta um conjunto de conceitos que pode alterar os paradigmas, os modelos e as funções dos vários tipos de elementos que concorrem para o planeamento e gestão das arquitecturas de informação que as bibliotecas congregam, transformam, produzem e distribuem.

NOTAS

¹ Parte das ideias e informações expostas neste artigo foram apresentadas pelos autores em CORDEIRO & CARVALHO (2002) e CARVALHO & CORDEIRO (2003).

² Z39.50 é usado para referir tanto a norma ANSI/NISO Z39.50 (http://www.niso.org/standards/resources/Z3950_Resources.html) como a sua correspondente ISO 23950 (<http://www.iso.org/>). A Library of Congress funciona como agência internacional de gestão do Z39.50 e disponibiliza a principal fonte de informação na matéria em <http://www.loc.gov/z3950/agency/>.

³ ILL – Inter-Library Loan refere-se a um protocolo de rede especificado no conjunto de normas ISO 10160:1997 e ISO 10161:1997. A agência de manutenção destas normas é a National Library of Canada (<http://www.nlc-bnc.ca/iso/ill/>).

⁴ EDI – Electronic Data Interchange, designa um vasto conjunto de normas para

protocolos relativos a transacções electrónicas (ver <http://www.diffuse.org/edi.html#Others>), dos quais se destacam os grupos de normas EDIFACT (ISO 9735:2002) da responsabilidade da UN/EDIFACT, e X12 da responsabilidade do ANSI Accredited Standards Committee X12 (<http://www.x12.org/>).

⁵ Como exemplo de *gateway* ver *Library of Congress Gateway to Library Catalogues* em <http://lcweb.loc.gov/z3950/gateway.html#other>. Exemplos de catálogos colectivos virtuais: *vCuc – virtual Canadian union catalogue* (<http://www.nlc-bnc.ca/8/6/index-e.html>); *M25 – Consortium of Higher Education Libraries* (<http://www.m25lib.ac.uk/>).

Para as tendências actuais nesta matéria ver também o Projecto TEL – The European Library (<http://www.europeanlibrary.org/>).

⁶ Ver, p. e., Resource Discovery Network Subject Portals Project (<http://www.portal.ac.uk/spp/>), e COX & YEATES (2002)

para uma análise das soluções tecnológicas para portais de biblioteca.

⁷ Sobre este assunto ver VAN DE SOMPEL & BEIT-ARIE (2001), STERN (2001), PENTZ (2001) e uma bibliografia sobre a matéria disponível em http://www.viva.lib.va.us/viva/tech/cat/link/working_bibliography.html

⁸ A expressão na forma serviços web é doravante, no texto, usado na acepção mais específica, correspondente à expressão *web services*.

⁹ O termo *standard* é usado, no texto, para distinguir do termo normas no seu sentido estrito, para referir especificações técnicas aprovadas ou adoptadas em contextos que não são os dos tradicionais organismos oficiais de normalização, como a ISO ou congêneres.

¹⁰ SOAP – *Simple Object Access Protocol*, é um formato que especifica as mensagens para requerer serviços de aplicações remotas usando XML. É uma especificação do Consórcio 3W mantida pelo W3C XML Protocol Group. Mais informação em <http://www.w3.org/2002/ws/>.

¹¹ WSDL – *Web Services Description Language* (<http://www.w3.org/TR/wsdl>) é uma especificação XML para descrever serviços web destinada a ferramentas de software de forma a que estas possam ler uma descrição WSDL e automaticamente produzir o essencial de código necessário para uma aplicação cliente aceder ao serviço descrito.

¹² UDDI – *Universal Description, Discovery and Integration* (<http://www.uddi.org/about.html>) é uma especificação para a criação de directórios de serviços web. Inclui um modelo de dados em XML Schema e uma API para pesquisa e publicação dos dados. UDDI designa também um directório existente, lançado em 2001 (*UDDI Cloud Services*, <http://www.uddi.org/register.html>). Para mais informação ver também DOVEY (2001).

¹³ Para uma introdução geral mais detalhada, compreendendo também os produtos de tecnologia de serviços web disponíveis assim como perspectivas de aplicação, ver MCDONALD (2003).

¹⁴ Esse nível de interdependência refere-se ao que é normalmente designado por *tightly-coupled versus loosely-coupled*. No modelo cliente-servidor tradicional, uma solução de integração é normalmente *tightly-coupled*,

o que significa maior interdependência, requerendo conhecimento prévio e acertos directos entre os sistemas envolvidos. A este respeito, no modelo dos serviços web, a integração é designada *loosely-coupled*.

¹⁵ DCMES – *Dublin Core Metadata Element Set* (<http://dublincore.org/>)

MARCXML – *MARC 21 XML Schema*

(<http://www.loc.gov/standards/marcxml/>)

METS – *Metadata Encoding and Transmission Standard* (<http://www.loc.gov/standards/mets/>)

MODS – *Metadata Object Description Schema*

(<http://www.loc.gov/standards/mods/>)

OAI PMH – *Open Archives Initiative Protocol for Metadata Harvesting* (<http://www.openarchives.org/>)

DOI – *Digital Object Identifier* (<http://www.doi.org/>)

OpenURL (NISO) http://www.niso.org/committees/committee_ax.html

¹⁶ ZING – *Z39.50 International: Next Generation* (<http://www.loc.gov/z3950/agency/zing/zing-home.html>)

¹⁷ SRW – *Search/Retrieve web Service* (<http://www.loc.gov/z3950/agency/zing/srw/background.html>).

¹⁸ ONIX é um conjunto de especificações para a representação e transmissão de dados sobre livros e periódicos visando o comércio electrónico, da principal responsabilidade de EDITEUR – *The International Group For Electronic Commerce In The Book And Serials Sectors* (<http://www.editeur.org/>)

¹⁹ Sobre implementações existentes e projectos em curso ver <http://www.loc.gov/z3950/agency/zing/srw/implementors.html>.

²⁰ Para informação sobre os perfis Z39.50 conhecidos e internacionalmente registados ver <http://lcweb.loc.gov/z3950/agency/profiles/profiles.html>. Os desenvolvimentos recentes na arquitectura de atributos Z39.50 podem consultar-se em <http://lcweb.loc.gov/z3950/agency/attrarch/attrarch.html>.

²¹ Informação sobre os vários serviços web utilizados pode ser obtida em: <http://ptolemy.bookmarc.pt:8001/kb/kb000002.html> (PORBASE), <http://www.amazon.com/gp/aws/landing.html> (Amazon) e <http://www.google.com/apis/> (Google).

BIBLIOGRAFIA

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. – "The semantic web". *Scientific American*, May 2001. Disponível em: <http://www.sciam.com/>

BLOOMBERG, J. – "Web services: a new paradigm for distributed computing". *The Rational Edge: e-zine for the rational community*, Sep. 2001. Disponível em: http://www.therationaledge.com/content/sep_01/f_webServices_jb.html.

BLOOMBERG, J. – *Web services management: successfully architecting the future of your business*. Zapthink white paper. 2002. Disponível em: http://www.actional.com/downloads/info_kits/web_services_mgmt.pdf.

BURBECK, S. – *The evolution of web applications into service-oriented components with web services*. 2000. Disponível em: <http://www-106.ibm.com/developerworks/library/ws-tao/index.html?dwzone=components>.

CARPENTER, B. – "Future applications and middleware, and their impact on the infrastructure." *Future Generation Computer Systems*, vol. 19 (2003) p. 191-197.

CARVALHO, J.; CORDEIRO, M. I. – "XML and bibliographic data: the TVS (Transport, Validation and Services) model". 68th IFLA council and general conference. 2002. Disponível em: <http://www.ifla.org/IV/ifla68/papers/075-095e.pdf>.

CARVALHO, J.; CORDEIRO, M. – "Web services: next generation interoperability for Libraries". In: Nixon, C., comp. *Internet librarian international conference, 25-25 March 2003, Birmingham UK. Collected presentations*. New Jersey: Information Today Inc., p. 40-49.

CERAMI, E. – *Web services essentials*. Beijing, etc.: O'Reilly, 2002. ISBN 0-596-00224-6.

CORDEIRO, M.; CARVALHO, J. – "Web Services: what they are and its importance for libraries." *VINE*, 2002, vol. 32, n.º 4, Issue 129, p. 46-62.

COX, A.; YEATES, R. – "Library orientated portals solutions". *JISC Technology and Standards Watch Report TSW 02-03*. 2002. Disponível em: http://www.jisc.ac.uk/index.cfm?name=techwatch_report_0203.

DOVEY, M. – "UDDI (Universal Description, Discovery and Integration) standard initiative".

JISC Technology and Standards Watch Report TSW01-01:UDDI. 2001. Disponível em:

http://www.jisc.ac.uk/index.cfm?name=techwatch_report_0101

FLETCHER, P.; WATERHOUSE, M. ed. – *Web services business strategies and architectures*. Acoccks Green: Expert Press, 2002. ISBN 1904284132.

GARDNER, J. R. – Information architecture planning with XML. *Library Hi Tech*, v.19, n.º 3 (2001): 231-241.

GARDNER, T. – "An introduction to web services". *Ariadne*, 29 (Oct. 2001). Disponível em: <http://www.ariadne.ac.uk/issue29/gardner/intro.html>

GATENBY, J. – *Internet, interoperability and standards: filling the gaps*. 2000. Disponível em: <http://www.niso.org/press/whitepapers/Gatenby.html>

GATENBY, J. – "Aiming at quality and coverage combined: blending physical and virtual union catalogues". *Online Information Review*, vol. 26, n.º 5 (2002), p. 326-334.

HICKEY, T. (2002) – "Deconstructing cataloguing: a web services approach to bibliographic control." Access 2002 conference. 2002. Disponível em: <http://zeus.uwindsor.ca/library/leddy/access2002/thom.ppt>

JACOBSON, C. – "Web services: stitching together the institutional fabric". *Educause Review*, March-April, 2002, p. 50-51. Disponível em: <http://www.educause.edu/ir/library/pdf/erm0228.pdf>

LAYMAN, A. – "Web Services framework". W3C web services workshop. 2001. Disponível em: <http://www.w3.org/2001/03/WSWS-popa/paper51>

LORENZO, G. – Web services enabling technology for application integration and assembly. HEKATE (Higher Education Knowledge & Technology Exchange). 2002. Disponível em: <http://www.hekate.org/HEKATEwebserviceswhitepaper.pdf>

LYNCH, C. – "The Z39.50 Information retrieval standard. Part I: A strategic view of its past, present and future". *D-Lib Magazine*, April 1997. Disponível em: <http://www.dlib.org/dlib/april97/04lynch.html>

- MCDONALD, D. – Web Services technology. TehWatch Report 0304, May 2003. Disponível em: http://www.jisc.ac.uk/index.cfm?name=techwatch_report_0304
- MOEN, W. – "Resource discovery using Z39.50: promise and reality". Library of Congress bicentennial conf. On bibliographic control for the new millemium. 2000. Disponível em: <http://lcweb.loc.gov/catdir/bibcontrol/moen.html>.
- MOEN, W. – "Accessing interoperability in the network environment: standards, evaluation and testbeds in the context of Z39.50". In McClure, C. & Bertot, J., ed. *Evaluating networked information services: techniques, policy and issues*. Medford, NJ: Information Today, 2001, p. 85-109.
- MOEN, W. – "Z39.50 for finding it all". Access 2002 conference. 2002. Disponível em: http://zeus.uwindsor.ca/library/leddy/access2002/Z_FindingItAll.ppt .
- NEEDLEMAN, M. – "Z39.50 - a review, analysis and some thoughts on the future". *Library Hi Tech*, v. 18, n.º 2, 2000, p.156-165.
- OGBUJI, U. The past, present and future of web Services. 2002. Disponível em:
Part 1: <http://www.webservices.org/index.php/article/articleview/663/1/24/>
Part 2: <http://www.webservices.org/index.php/article/articleview/679/1/24/> .
- PENTZ, E. – "CrossRef: a collaborative linking network". *Issues in Science and Technology Librarianship*, Winter 2001. Disponível em: <http://www.library.ucsb.edu/istl/01-winter/article1.html> .
- POWELL, A.; LYON, L. – "The JISC Information Environment and web services". *Ariadne*, Issue 31 (March-April 2002). Disponível em: <http://www.ariadne.ac.uk/issue31/information-environments/> .
- RENNIE, F.; MASON, R. – "The ecology of the connection". *First Monday*, vol. 8, n.º 8 (August, 4, 2003) . Disponível em: http://www.firstmonday.dk/issues/issue8_8/rennie/index.html .
- SAMTANI, G.; SADWANI, D. – "EAI and web services: easier Enterprise Application Integration?". *web Services Architect*, Oct. 17, 2001. Disponível em: <http://www.webservicesarchitect.com/content/articles/samtani01.asp>
- SEDUKHIN, I. – "Platform for assembly, orchestration and management of web services". W3C web services workshop, 2001. Disponível em: <http://www.w3.org/2001/03/WSWS-popa/paper10>
- STAL, M. – "Web services: beyond component-based computing". *Communications of the ACM*, vol. 45, n.º 10 (Oct. 2002) p. 71-76.
- STERN, D. – "Automating enhanced discovery and delivery: the OpenURL possibilities". *ONLINE*, March 2001. Disponível em: http://www.infotoday.com/online/OL2001/ster_n3_01.html .
- STEVENS, M. – "Service-oriented architecture: introduction".
Part 1. *Developer.com*, 16 April 2002 . Disponível em: <http://www.developer.com/tech/article.php/1010451>
Part 2 *Developer.com*, 23 April 2002 . Disponível em: <http://www.developer.com/tech/article.php/1014371> .
- W3 CONSORTIUM – *Semantic web activity statement*. (2001). Disponível em: <http://www.w3.org/2001/sw/Activity> .
- W3 CONSORTIUM – *Web services activity statement*. (2001a). Disponível em: <http://www.w3.org/2002/ws/Activity> .
- W3 CONSORTIUM – *Web services architecture*. (2002). Disponível em: <http://www.w3.org/TR/ws-arch/> .
- W3 CONSORTIUM– *Web services architecture working group charter*. (2002a). Disponível em: <http://www.w3.org/2002/01/ws-arch-charter> .
- VAN DE SOMPEL, H.; BEIT-ARIE, O. – "Generalizing the OpenURL framework beyond references to scholarly works: the Bison-Futé model". *D-Lib Magazine*, vol. 7, n.º 7-8 (Jul.-Aug. 2001) . Disponível em: <http://www.dlib.org/dlib/july01/vandesompel/07vandesompel.html>