

Descrição de recursos em uma estrutura de metadados pautada no modelo FRBR

Elisabete Gonçalves de Souza

Doutora; Universidade Federal Fluminense; elisabetes.souza@gmail.com

Darlene Alves Bezerra

Mestre; Universidade Federal Fluminense; darlinglene@yahoo.com.br

Wellington Freire Cunha Costa

Mestrando; Universidade Federal Fluminense; freirew@yahoo.com.br

Resumo: Analisa os impactos da web semântica nos processos de organização e acesso à informação e as perspectivas trazidas pelo modelo conceitual FRBR para o campo da catalogação descritiva. Objetiva testar a aplicação de atributos FRAD em um repositório de modo a averiguar os limites e as possibilidades de uso de modelos conceituais na estrutura descritiva deste tipo de biblioteca digital. Demonstra como o uso de modelos conceituais quando associado a outras tecnologias, como as boas práticas indicadas pelo uso de *linked data*, pode contribuir para enriquecer os processos de busca e recuperação da informação. Conclui que a possibilidade de substituir metadados textuais por identificadores universais permitirá que os agentes de *softwares* identifiquem com precisão as entidades essenciais à recuperação dos diferentes recursos, como as pessoas (criadoras das obras científicas), interligando o repositório, de forma lógica e segura, a outros dados disponíveis sobre elas na Web.

Palavras-chave: Representação descritiva. Modelo conceitual FRBR. Repositório digital. Web semântica.

1 Introdução

O desenvolvimento das novas tecnologias, como a *Web* e a internet, modificou o paradigma teórico-metodológico que fundamentava os estudos no campo da representação descritiva deslocando-os do item para o usuário, colocando este no centro dos debates acerca dos processos de organização e recuperação da informação. Nesse novo tipo de abordagem, destaca-se a aproximação da IFLA

(*International Federation of Library Associations and Institutions*) às discussões em torno do modelo Entidade-Relacionamento (E-R) e o desenvolvimento do modelo conceitual FRBR (*Functional Requirements for Bibliographic Records*) para a representação das entidades bibliográficas, seja no âmbito dos catálogos ou das bibliotecas digitais, como os repositórios institucionais, objeto de estudo deste artigo.

Segundo Café et al. (2003), os repositórios surgem na década de 1990 com um complemento do sistema tradicional de comunicação científica; um espaço de divulgação e socialização de conhecimentos através do qual as instituições e/ou comunidades científicas armazenam, preservam, gerenciam e disponibilizam o acesso à produção científica de seus pesquisadores.

O desenvolvimento de modelos conceituais, como os FRBR (*Functional Requirements for Bibliographic Records*), aponta novas possibilidades para a organização e acesso à produção documental armazenada nos repositórios digitais. A modelagem da estrutura Dublin Core (DC) de acordo com os novos fundamentos da catalogação dará às bibliotecas digitais uma maior flexibilidade, pois aproxima os campos DC das entidades FRBR, qualificando os registros e tornando-os singulares; individualizando-os a partir dos atributos específicos associados às entidades correspondentes.

O pressuposto que move a pesquisa é o uso do modelo conceitual FRBR na estruturação descritiva de metadados DC, especialmente a ampliação de novos campos que abriguem atributos relacionados às entidades de forma que se tenha uma descrição mais rica, à medida que a esses campos podem-se associar identificadores estáveis que interliguem as entidades a outros dados disponíveis no ambiente *Web*.

As questões que mobilizaram essa investigação foram: como o novo modelo conceitual pode contribuir para melhorar o desempenho dos repositórios digitais? O padrão DC suporta a aplicação do modelo conceitual FRBR? De que forma a aplicação do modelo melhora o processo de busca e recuperação da informação?

Para responder essas questões foram traçados os seguintes objetivos: a) conhecer os novos paradigmas que fundamentam a criação do modelo conceitu-

al FRBR; b) aplicar, por meio de simulações, o novo modelo conceitual na modelagem de recursos disponibilizados em repositórios.

Para atingir esses objetivos, buscamos respaldo nos teóricos da área. No que diz respeito à representação descritiva da informação e seus novos fundamentos pautados no modelo conceitual FRBR, utilizaram-se os estudos de Mey e Silveira (2009), Moreno (2006, 2007) e Tillett (2003). Para discutir a organização dos repositórios digitais, contamos com as contribuições de Weitzel (2006) e Marcondes (2012). Nos estudos sobre o Dublin Core (DC) e a criação de repositórios digitais, utilizamos as análises de Souza, Vendrusculo e Melo (2000) e Marcondes (2005).

Em termos metodológicos, a presente pesquisa define-se como uma pesquisa qualitativa de caráter descritivo, que, segundo Gil (2002, p. 42), “[...] tem como objetivo primordial a descrição das características de determinados fenômenos ou o estabelecimento de relações entre variáveis.”. Nossa proposta é demonstrar o quanto o acréscimo de novos atributos – metadados referentes à entidade pessoa presentes no modelo FRAD (*Functional Requirements for Authority Data*) – pode melhorar a recuperação da informação e o acesso aos documentos em repositórios digitais à medida que sua estrutura é acrescida de atributos fundamentais para identificação dos produtores dos recursos (os pesquisadores) e suas linhas de pesquisa. Por exemplo: atributos identificando os grupos de pesquisa, quando associados à forma do nome e forma autorizada do nome, aperfeiçoam as buscas pelos autores. Isso é possível porque os repositórios são desenvolvidos no formato DC e seus metadados descritos em RDF (*Resource Description Framework*), o que permite que os robôs de busca recuperem essas informações dando maior visibilidade às atividades científicas realizadas no âmbito das instituições.

O campo empírico escolhido para aplicação do modelo conceitual FRBR foi o Repositório ALICE (Acesso Livre à Informação Científica), desenvolvido pela Embrapa (Empresa Brasileira de Pesquisa Agropecuária) com tecnologias padronizadas adotadas pela comunidade científica mundial e interoperável com os demais sistemas de acesso aberto, o que o torna uma eficaz fonte de informação científica no âmbito das ciências agrárias.

2 Marco teórico

A complexidade do ambiente digital e a diversidade de suportes demandaram a criação de sofisticados recursos tecnológicos compatíveis com o grau de desenvolvimento do campo informacional. Esse cenário trouxe à tona uma série de discussões, mudanças e transformações que atingiram significativamente a área de organização da informação.

O marco inaugural foi o Seminário sobre Registros Bibliográficos de 1990, realizado na cidade de Estocolmo (Suécia), conhecido como Seminário de Estocolmo. Dentre as questões abordadas e debatidas pelos participantes do evento, uma das principais foi “[...] a importância de satisfazer as necessidades dos usuários e resolver eficazmente a ampla gama de necessidades associadas aos diferentes tipos de materiais, assim como os vários contextos em que se utilizam os registros bibliográficos [...]” (IFLA, 1998, p. 1-2, tradução nossa). Esse debate trouxe uma mudança nos paradigmas da Representação Descritiva, pois colocou o usuário e suas necessidades informacionais como a categoria central para se repensar os Princípios da Catalogação. Conforme ressalta Mey “[...] o cerne da representação desloca-se do item para o usuário, visando permitir-lhe as tarefas de encontrar, identificar, selecionar e obter uma ‘entidade’ adequada a seus propósitos.” (MEY, 2003, p. 1). Outra importante questão que veio a debate estava relacionada ao uso de novas tecnologias e de um novo arcabouço teórico e metodológico que adequasse as práticas catalográficas às novas exigências informacionais impostas pelo desenvolvimento da *Web*.

Tendo em vista estas questões, neste seminário, foram aprovadas nove resoluções e uma delas conduziu diretamente ao estudo que tornou possível o desenvolvimento de um modelo conceitual para o domínio bibliográfico e, portanto, coube à IFLA a responsabilidade de definir os requisitos funcionais para os registros bibliográficos, os quais estariam primordialmente pautados nas necessidades do usuário. Este estudo ocorreu por um período ininterrupto de oito anos e foi conduzido pela *IFLA Study Group on the Functional Requirements for Bibliographic Records*. A atividade do grupo consistiu em definir um con-

junto de elementos semânticos relacionados ao universo bibliográfico, visando empregar a metodologia de análise de entidades do Modelo E-R focando-se em “[...] definir claramente as entidades de interesse dos usuários dos registros bibliográficos, os atributos de cada uma das entidades e os tipos de relacionamentos que se operam entre as entidades.” (IFLA, 1998, p. 3). A intenção era produzir um modelo conceitual que integrasse os atributos específicos e os relacionamentos visando às distintas tarefas que os usuários realizam quando consultam os registros bibliográficos, a fim de recuperar informações pertinentes. As “tarefas dos usuários” sintetizam requisitos básicos que devem ser previstos pelo sistema com vistas a auxiliar o usuário a navegar, encontrar, identificar, selecionar e obter um recurso.

Esses estudos consolidaram-se com a publicação, ainda em 1998, do relatório final sobre os FRBR contendo as novas diretrizes para a descrição de recursos em catálogos, repositórios e outras ferramentas bibliográficas. Após a publicação desse documento, a IFLA decidiu formar novos grupos reunindo especialistas em catalogação de diversos países para a conclusão de uma versão atualizada dos princípios internacionais de catalogação. Em 2003, no Encontro de Peritos realizado em Frankfurt, foi aprovado um documento preliminar para a criação de um novo código de catalogação internacional, e aprovada a Nova Declaração de Princípios da Catalogação, abarcando não só as obras textuais, mas todo tipo de material, incluindo os recursos digitais.

O relatório final dos FRBR apresenta dois objetivos básicos: a) fornecer um quadro estruturado, claramente definido, para relacionar dados indicados em registros bibliográficos às necessidades dos usuários destes registros; e b) recomendar um nível básico de funcionalidade para registros criados por agências bibliográficas nacionais (IFLA, 1998, p. 7, tradução nossa). Esses objetivos visam enfatizar a relação entre usuário e suas necessidades, mostrando que as informações registradas em bases de dados em linha devem apresentar relações bibliográficas mais claras, servindo de forma mais útil aos usuários; e estabelecer um nível básico de funcionalidade entre as entidades, relacionamentos e atributos descritos no modelo a elementos de dados específicos que devem ser incluídos no registro (MORENO; BRASCHER, 2007).

Segundo Tillett (2003), antes dos FRBR as regras de catalogação apresentavam pouca clareza no que tange ao uso das palavras “obra”, “edição”, ou “item”. Por exemplo, quando se diz “livro” para se descrever um objeto físico que tem páginas de papel e uma encadernação, os FRBR chamam esse objeto de um “item”. Quando se diz “livro” no sentido de identificar algum tipo de “publicação”, os FRBR chamam essa instância de “manifestação”. Quando se diz “livro”, no contexto de “quem o traduziu”, os FRBR chamam isso de “expressão”.

Nesse aspecto, podemos dizer que a reestruturação da descrição bibliográfica significou uma nova maneira de refletir a estrutura conceitual de buscas de informação, levando em conta não só a diversidade de usuários, mas também dos materiais (textuais, musicais, cartográficos, audiovisuais, gráficos e tridimensionais); dos suportes físicos (papel, filme, registros sonoros, iconográficos, meios ópticos de armazenagem, etc.); dos formatos (livros, folhas, discos, cassetes, fotografias, cartuchos, etc.) e das diferentes formas de expressão de uma obra.

2.1 As entidades da família FRBR

A família FRBR modelo E-R tem 10 entidades divididas em três grupos. As entidades do grupo 1 são as responsáveis pelos produtos do trabalho intelectual ou artístico que correspondem aos interesses do usuário: obra, expressão, manifestação e item. A entidade obra é abstrata e refere-se a uma criação intelectual ou artística distinta. De acordo com Mey e Silveira (2009, p. 19), é “[...] o conteúdo intelectual em si, independentemente de seu suporte ou de sua forma [...]”. Pode ser assunto de outra obra nas relações do FRBR e é reconhecida como entidade por meio de suas diversas expressões e manifestações.

A expressão também é abstrata e refere-se à realização do conteúdo intelectual ou artístico de uma obra, ou seja, é a maneira de expressar o conteúdo intelectual, que pode ser através de uma tradução, de uma adaptação, de uma notação alfanumérica, musical ou coreográfica, entre outras possíveis formas. Se o conteúdo intelectual ou artístico sofrer mudanças, ocorrerá o mesmo com a expressão.

A manifestação é a materialização de uma expressão de uma obra, ou seja, a concretização em termos físicos, dentre os quais podem ser livros, monografias, periódicos, filmes, entre outros. Ao publicar dois textos iguais – sendo um como livro impresso e outro como gravação sonora – teremos a mesma obra, a mesma expressão e duas diferentes manifestações, pois não se trata do mesmo suporte físico.

O item é uma entidade concreta que corresponde a um exemplar individual de uma manifestação. Há exceções quando um objeto é dividido em volumes, como uma monografia publicada em três volumes. O item representa tanto um objeto físico disponível numa biblioteca tradicional quanto um objeto em rede visualizado em uma biblioteca digital.

No grupo 2 encontramos os Requisitos Funcionais de Dados de Autoridade (*Functional Requirements for Authority Data – FRAD*), um modelo direcionado ao controle de dados de autoridade. As entidades FRAD para os registros bibliográficos são: pessoa e entidade. A entidade pessoa refere-se a um indivíduo, relacionado à criação ou realização de uma obra ou de uma expressão, podendo ser: autores, editores, tradutores, artistas, compositores, entre outros, ou também relacionado como assunto de uma obra, como no caso das biografias. A entidade coletiva refere-se a uma organização ou grupo de indivíduos que incluem também grupos temporários (encontros, conferências, reuniões, festivais, etc.) e autoridades territoriais como uma federação, um estado, uma região, uma municipalidade (MORENO, 2006).

No grupo 3 estão os Requisitos Funcionais de Dados de Autoridade de Assunto (*Functional Requirements for subject Authority Data – FRSAD*). As entidades do grupo 1 e 2 também estão incluídas no grupo 3 pelo fato de poderem ser assuntos de obras. De acordo com Silva e Santos (2012), as quatro entidades do grupo 3 são definidas da seguinte forma: Conceito – noção abstrata ou ideia que pode ser o assunto de uma obra; Objeto – coisas materiais que podem ser o assunto de uma obra, podendo ser animadas ou inanimadas, fixas ou móveis, e objetos feitos pelo homem; Evento – ações e ocorrências que podem ser o assunto de uma obra, como épocas, períodos do tempo, eventos históricos; Lu-

gar – entidade que se refere a localizações terrestres, geográficas, jurisdições políticas, etc.

Como a representação descritiva das obras é o meio utilizado para aproximar o usuário das informações diante de várias manifestações e expressões das mesmas, existia uma tradição de colocar o foco sempre no item, deixando o usuário em segundo plano. Mas atualmente, com a implementação de repositórios digitais em padrão DC, há uma inversão dessa questão: procura-se dar mais ênfase ao usuário e atender melhor suas necessidades, poupando assim o seu tempo ao fazer as buscas.

Como vimos, no modelo FRBR as tarefas do usuário ganham destaque e direcionam a construção do registro bibliográfico, sendo também especificadas no modelo FRAD¹, em que a contextualização de uma entidade é ressaltada como uma tarefa primordial de modo a esclarecer, por exemplo, qual a forma mais conhecida do nome de uma entidade, além de permitir o acréscimo de outros atributos que dão singularidade ou que complementam o conhecimento sobre a ação (criativa, produtiva, etc.) de uma entidade. No âmbito dos repositórios, a aplicação do modelo conceitual FRAD permite-nos conhecer o trabalho intelectual de um pesquisador, cujas obras, em suas diferentes expressões e manifestações, povoam os repositórios institucionais.

2.2 Os repositórios no contexto da *web* semântica

Os repositórios digitais surgiram em consequência da evolução da internet, o que possibilitou o acesso rápido e fácil às informações em nível mundial. Nesse contexto, grandes instituições de pesquisa e universidades se viram diante da necessidade de dar publicidade à produção de seus pesquisadores dando maior visibilidade a eles e às suas instituições.

Como não existia uma padronização para divulgação e acesso à informação científica, foi desenvolvida uma iniciativa que fosse capaz de promover interoperabilidade entre os repositórios digitais, a *Open Archives Initiative* (OAI), responsável pela criação do protocolo *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH), com o objetivo de proporcionar a dissemina-

ção de conteúdos presentes nos repositórios digitais e assegurar a interoperabilidade entre as bibliotecas e os repositórios digitais ampliando as possibilidades de buscas e reduzindo o tempo de resposta aos usuários. A discussão que gira em torno deste protocolo diz respeito ao fato de contribuir apenas na abertura dos metadados e não exclusivamente dos textos completos dispostos nas bases de dados, o que leva à necessidade de serem desenvolvidas outras formas que sejam mais eficientes para o desenvolvimento da ciência e para o acesso livre aos pesquisadores. Segundo Kuramoto, “[...] quanto maior o acesso público à informação, maior será a possibilidade de ampliarmos a comunidade de usuários, até mesmo de levar esta informação para as comunidades que não têm este acesso”. (KURAMOTO, 2005, p. 3).

Weitzel (2006) define os repositórios de acesso aberto como bibliografias especializadas, cujos serviços de indexação e resumo são constituídos pelas próprias comunidades científicas, permitindo o acesso organizado e livre às publicações, disseminando-as sem restrição, de forma pública e gratuita. Para Marcondes (2012), a questão central que envolve as bibliotecas digitais é a interoperabilidade, sintática e semântica, contribuindo para que a recuperação da informação seja satisfatória às necessidades dos usuários. A interoperabilidade é uma das principais características de um repositório, pois permite que diferentes publicações oriundas de softwares diversos possam ser encontradas no mesmo local e reunidas sem conflitos no ambiente digital.

Além de promover as instituições, os repositórios dão visibilidade às produções de seus pesquisadores permitindo a troca de informações e o compartilhamento de suas coleções com outras instituições sem que haja altos custos, sendo o uso e o reuso destas informações um dos dados mais tangíveis para medirmos a qualidade de uma instituição, pois demonstram sua relevância científica, aumentando sua credibilidade e o seu reconhecimento junto ao público (CROW, 2002).

O repositório institucional destaca-se de outras ferramentas bibliográficas pelo fato de ser feito voltado para o usuário, que como autor pode armazenar, utilizar e disseminar as informações de sua produção científica no meio digital de forma simples e fácil.

A disponibilização e a disseminação de qualquer recurso no ambiente digital envolvem não somente a inclusão da publicação na *Web*, mas também questões tecnológicas que garantam a organização, descrição, preservação e acessibilidade aos documentos publicados. Dentre essas tecnologias encontra-se o FRBR, modelo conceitual que possui uma estrutura flexível e simples que muito se assemelha aos princípios que deram origem ao formato Dublin Core usado na estruturação dos metadados descritivos em repositórios institucionais.

Sua estrutura envolve um conjunto de quinze elementos de metadados planejados para facilitar a descrição de recursos eletrônicos, de modo que os autores e *Websites* que não têm conhecimentos em catalogação possam descrever os recursos eletrônicos, “[...] tornando suas produções mais visíveis aos engenhos de busca e sistemas de recuperação.” (SOUZA; VENDRUSCULO; MELO, 2000, p. 93).

Segundo Marcondes (2005), o DC foi pensado para ser simples o suficiente e autoexplicativo de modo que o próprio autor do documento possa descrevê-lo ao publicá-lo eletronicamente, o que não seria possível em outro padrão como o *Machine Readable Cataloging* (MARC), que exige o aporte de profissionais de informação treinados, por ser mais complexo que o DC.

A escolha pelo DC no desenvolvimento de repositórios ocorreu por ser um padrão menos estruturado e mais flexível que permite a inclusão de elementos adicionais, o que contempla as peculiaridades de uma instituição, indo na mesma direção do modelo conceitual FRBR, cujo princípio visa à inclusão de elementos (entidades, atributos e relacionamentos) que atendam às necessidades dos usuários.

Conforme Rosetto (2003), os padrões de metadados determinam regras que elucidam os atributos dos recursos informacionais, a fim de gerar coerência interna entre os elementos por meio de semântica e sintaxe. Essa ação é fundamental para garantir o acesso aos recursos e permitir a interoperabilidade dos sistemas de informação. O principal objetivo de padrões de metadados é:

[...] estabelecer um grau de normalização, padronização de representação descritiva automatizada dos dados informacionais, favorecendo a compatibilidade e a transferência de dados legíveis por computador, possibilitando o compartilhamento e a cooperação de recursos e serviços entre diferentes unidades de informação. (FUSCO, 2010, p. 67).

Nessa direção, o DC destaca-se por ter as seguintes características: a) Flexibilidade – os elementos que o compõem são opcionais, podendo ser exibidos aleatoriamente e se houver necessidade podem ser repetidos ou modificados por qualificadores; b) Simplicidade – é de fácil manuseio, sendo autoexplicativo, permitindo que o próprio autor do documento faça a descrição sem a necessidade de intensos treinamentos; c) Extensibilidade – por ser um modelo simples e flexível, permite que diversas comunidades em áreas distintas utilizem o padrão DC trocando informações e obtendo o acesso a elas; d) Interoperabilidade – um modelo comum de descrição utilizado por diversas áreas proporciona aos usuários maior habilidade nas pesquisas.

Em termos institucionais, o DC foi desenvolvido e é mantido pelo *Dublin Core Metadata Initiative* (DCMI), uma organização de caráter aberto responsável pela promoção de padrões e pelo desenvolvimento de vocabulário especializado. Os membros do DCMI estão distribuídos pelo mundo, realizando workshops e listas de discussão para troca de experiências e informações. Essa organização tem como objetivo a interoperabilidade das descrições feitas no DC, adaptando suas aplicações à linguagem *Extensible Markup Language* (XML) e à semântica da *Resource Description Framework* (RDF). De acordo com Marcondes (2005, p. 108), “[...] enquanto a XML é uma linguagem genérica para a descrição da estruturação de documentos eletrônicos, RDF é específica para criar metadados com a finalidade de localizar e identificar recursos.” São essas especificidades que tornam o DC um padrão de grande importância para o desenvolvimento da descrição dos recursos em rede, sendo o mais adequado para o cenário dos repositórios institucionais.

O *Linked Data*, termo usado por Bizer, Heath e Berners-Lee (2009, p. 1) para referir-se “[...] a um conjunto de melhores práticas para publicar e conectar dados estruturados na Web [...]” corrobora com os novos processos de organiza-

ção e recuperação da informação em meio digital. Essas melhores práticas têm caminhado rumo à configuração de uma *Web* de dados (*Web of Data*), onde *links* semânticos são desenvolvidos para serem conectados, a fim de unir diferentes fontes, ou seja, recursos informacionais diversificados na *Web*.

Tim Berners-Lee (2006) foi quem introduziu essas práticas, as quais se resumem nos seguintes princípios básicos: 1) usar URIs como nomes para recursos; 2) usar URIs HTTP para que as pessoas possam encontrar esses nomes; 3) quando alguém procura por uma URI, fornecer informações úteis, usando os padrões (RDF, SPARQL) e 4) incluir *links* para outras URIs, de forma que outros recursos possam ser descobertos (BIZER; HEATH; BERNERS-LEE, 2009).

Assim, esses quatro princípios se tornaram conhecidos como *Linked Data Principles*, devido a “[...] fornecer uma receita básica para publicar e conectar dados usando a infraestrutura da *Web* aderindo-os à sua arquitetura e padrões;” (BIZER; HEATH; BERNERS-LEE, 2009, p. 2). Por isso, Saorín (2012, p. 290) ressalta que “[...] a arquitetura de dados ligados permite descobrir fontes de dados explorando as relações, o que enriquece a recuperação de informações”. Essas relações permitem à tecnologia *Linked Open Data* oferecer maiores possibilidades para que diversificados recursos disponíveis na *Web* sejam agregados uns aos outros, por meio de *links* semânticos e identificadores universais, como URIs e padrões como o RDF.

Marcondes (2012) explica que as tecnologias de dados interligados

[...] não são propriamente tecnologias de recuperação de informações, mas sim, uma vez recuperado o recurso informacional, tecnologias para interligá-lo com outros que lhe agreguem valor semântico, cultural, cognitivo, econômico ou científico. (MARCONDES, 2012, p. 174).

Isto é, as tecnologias *Linked Data* proporcionam uma navegação pelo amplo universo *Web* incorporando significado e precisão a esse processo, de modo que a cada exploração sejam feitas novas descobertas através da navegação entre os recursos *Web*.

Os URIs apresentam *links* permanentes e estáveis, além de terem um significado preciso (MARCONDES, 2012), o que garante relações estáveis e seguras entre os diversos recursos que podem ser entrelaçados na *Web*. Outra

característica importante presente no *Linked Data* é que os links que unem as páginas têm significado, ou seja, são *links* semânticos, que expressam a relação entre os recursos (MARCONDES, 2012). Pois com o RDF é possível descrever semanticamente um recurso, ou seja, expressar o significado contido nele. Isso habilita os agentes inteligentes de software a explorar os dados descritos formalmente em RDF automaticamente, muitas vezes, agregando, interpretando ou mesclando dados. Conforme Bizer, Cyganiak e Heath (2007, introdução), “[...] a cola que une o documento *Web* tradicional são os links hipertexto entre as páginas HTML. A cola de dados web são os links RDF.”.

O que atesta o valor do padrão RDF no contexto da proposta *Linked Open Data* é a sua compatibilidade com a extensão da *Web* atual,² (BERNERS-LEE; HENDLER; LASSILA, 2001), sendo essa linguagem uma das tecnologias indispensáveis no desenvolvimento da *Web* Semântica, o que evidencia explicitamente a compatibilidade do formalismo apresentado pelo padrão RDF com a *Web* em expansão. No âmbito dos repositórios, o uso dessas práticas permite que dados relacionados aos atributos de uma entidade possam ser buscados na *Web* enriquecendo sua contextualização.

3 Metodologia

Como vimos, os repositórios são bibliotecas digitais que armazenam diferentes expressões e manifestações de uma mesma obra, em sua grande maioria comunicações científicas, cujas datas de depósito revelam ao consulente o processo de maturação de uma pesquisa ao longo do tempo. Nos repositórios institucionais encontramos relatórios de pesquisa, que nascem nas bancadas dos laboratórios, depois são (re)formatados e apresentados como *papers* em congressos e/ou artigos em periódicos científicos. A riqueza do fluxo de informações em um repositório, onde são depositadas comunicações formais e informais, faz dessa ferramenta um campo de estudo profícuo para aplicação dos novos modelos conceituais. No entanto, os repositórios abertos só podem depositar publicações livres, ou seja, aquelas cujos direitos autorais não estão empenhados às editoras ou a outros publicadores. Por isso a necessidade de integrá-los a bancos de da-

dos que abrigam periódicos como a base de periódicos científicos *Scientific Electronic Library Online* (SCIELO)³, assim como a grandes bancos de dados de caráter internacional, agregadores da produção científica de diversos países, como o *International Nuclear Information System* (INIS)⁴, o *International System for Agricultural Science and Technology* (AGRIS)⁵ o *Medical Literature Analysis and Retrieval System Online* (MEDLINE)⁶.

Para testar as possibilidades de agregar aos atributos das entidades bibliográficas novas informações, escolheu-se como campo empírico o repositório ALICE. Sua escolha se deu em função de ser uma biblioteca digital muito prestigiada junto à comunidade científica, referência na divulgação da pesquisa agropecuária no Brasil e no mundo; suas coleções são bastante diversas e contemplam na íntegra diferentes expressões e manifestações de estudos científicos produzidos por seus pesquisadores, publicadas como teses e dissertações, notas técnicas, artigos, capítulos de livros, etc.

O repositório ALICE é um produto de informação do Sistema Embrapa de Bibliotecas (SEB) e foi desenvolvido utilizando o software DSpace⁷, tecnologia recomendada pelo Instituto Brasileiro de Ciência e Tecnologia (IBICT) para a construção de repositórios digitais no Brasil. O processo de depósito é feito pelos bibliotecários e não diretamente pelos pesquisadores. Cabe aos bibliotecários realizar a catalogação e a indexação dos documentos depositados pelos pesquisadores no sistema de bibliotecas e o carregamento do arquivo no repositório. No momento da catalogação, o bibliotecário faz o *upload* do arquivo no formato pdf e o inclui na coleção digital em se enquadra, conforme sua natureza científica ou tecnológica (CASTRO; PELUFÊ; ARRUDA, 2013).

A aplicação do modelo conceitual FRBR fez-se necessária para demonstrarmos as contribuições do novo modelo conceitual no processo de recuperação da informação em sistemas de acesso aberto, como os repositórios; as vantagens da modelagem para os usuários e as suas tarefas de navegar, encontrar, identificar, selecionar e obter os recursos, evitando perda de tempo e garantindo-lhes maior consistência na qualidade das respostas.

Para identificação dos metadados de autoridade foi feita uma pesquisa por autor no repositório. A escolha de Zanus, M. C. (figura 1) se deu em virtude

de ser um pesquisador muito produtivo, com trabalhos escritos em diferentes periódicos científicos e outras produções indexadas na base AGRIS, base de dados internacional com informações técnico-científicas da área agrícola.

Figura 1 – Busca avançada, campo autor, repositório ALICE.

Fonte: Repositório Alice.

Há no repositório cerca de vinte trabalhos do autor, todos em parceria com outros pesquisadores. A obra escolhida foi “*Características sensoriais de vinhos tropicais*”, escrita em parceria com Pereira, G. E, e Guerra, C. C., cujos metadados encontram-se descritos abaixo, sendo os mesmos relacionados às respectivas entidades e atributos do modelo conceitual FRBR.

Figura 2 – Metadados DC e entidades e atributos FRBR e FRAD.

Unidade da Embrapa/Coleção: Embrapa Semiárido - Resumo em anais de congresso (ALICE)	
Atributos do item	Identificador: 43209
Atributos da expressão	Data de Envio: 1-Jul-2010
Entidades FRAD	Tipo do Material: Resumo em anais de congresso (ALICE)
Atributos FRAD	Autoria: GUERRA, C. C. PEREIRA, G. E. ZANUS, M. C.
Atributos a manifestação	Informações Adicionais: CELITO CRIVELLARO GUERRA, CNPQV; GIULIANO ELIAS PEREIRA, CPATSA; MAURO CELSO ZANUS, CNPQV.
Entidades FRAD	Título: Características sensoriais de vinhos tropicais.
Atributos a manifestação	Edição: 2010
	Fonte/Imprensa: In: INTERNATIONAL SYMPOSIUM OF TROPICAL WINES, 2., 2010, Petrolina. Abstracts... Petrolina: Embrapa Semiárido: Bento Gonçalves: Embrapa Uva e Vinho. 2010.
	Descrição Física: 1 CD-ROM.

Fonte: Elaboração dos autores.

Os atributos são as características apresentadas pelas entidades que auxiliam o usuário a encontrar, identificar, selecionar e obter um recurso, ou, simplesmente, a navegar pelo universo bibliográfico de forma dinâmica. Os atributos podem ser intrínsecos, quando se referem aos aspectos físicos das entidades,

como as dimensões, a data de publicação, o título, entre outros, e podem ser extrínsecos quando se referem aos aspectos contextuais, como um identificador da entidade – nesse caso, necessita-se de uma fonte de referência para se estabelecer tal atributo. Os atributos são importantes por ajudar na recuperação da informação, por atribuir às entidades características que levam o usuário a alcançar suas respostas. A figura abaixo mostra exemplos de atributos referentes às entidades FRAD - pessoa e “entidade coletiva”.

Figura 3 – Principais atributos FRAD.

Grupo 2	
Atributos de uma pessoa	Nome da pessoa Datas da pessoa Título da pessoa Outra designação relacionada à pessoa
Atributos de uma entidade coletiva	Nome da entidade coletiva Número relacionado à entidade coletiva Lugar relacionado à entidade coletiva Etc.

Fonte: Elaboração dos autores. Adequação: modelos de Oliver (2011) e Mey; Silveira (2009).

4 Análise e resultados

Conforme ressalta a IFLA (2009b) na Declaração de Princípios Internacionais de Catalogação, o modelo FRBR é um modelo conceitual extensível a todas as ferramentas, podendo ser “[...] aplicado a bibliografias e a outros ficheiros (arquivos) de dados criados por bibliotecas, arquivos, museus e outras comunidades.” (IFLA, 2009b, p. 2). A fim de demonstrar as possibilidades de aplicação do modelo conceitual FRBR à estrutura descritiva de um repositório, e aferir quais vantagens o processo de modelagem traz para a recuperação da informação, foi realizada uma simulação de modo a adequar a estrutura de metadados Dublin Core às entidades FRBR, sendo a elas associados outros atributos.

Figura 4 – Metadados DC, dados de autoridade.

Formato Dublin Core	Conteúdo
dc.contributor.author:	GUERRA, C. C.
dc.contributor.author:	PEREIRA, G. E.
dc.contributor.author:	ZANUS, M. C.

Entidade FRAD (pessoa)

Fonte: Elaboração dos autores.

Dado o caráter de amostragem desse estudo, aplicamos a modelagem das entidades do modelo conceitual FRBR apenas aos dados de autoridade. O objetivo foi ampliar a arquitetura de metadados DC do repositório ALICE associando os novos metadados aos atributos de autoridade do modelo conceitual FRAD, além dos recomendados, como a forma completa do nome da pessoa e, valendo-nos do *open link data*, acrescentamos outros metadados, todos relacionados à produção científica dos autores, mas não depositados no repositório em virtude do *copyright*.

Apesar da política de coleção do Repositório ALICE prever a inclusão de artigos publicados em periódicos que adotam uma licença que permita disponibilização do texto eletronicamente, como SciELO e outros, a pesquisa demonstrou que nem todos os artigos livres são depositados. Por isso, aproximar essas obras de seus autores através de *links* semânticos mostrou-se ser uma opção viável.

A extensão do modelo à estrutura DC permitiu-nos acrescentar atributos que completem as informações sobre a entidade pessoa (forma completa do nome), além de permitir o mapeamento dos vínculos associando a entidade pessoa a outras informações sobre suas produções (artigos científicos) depositadas em outros repositórios, revistas científicas e a base AGRIS.

A ampliação do elenco de atributos FRAD só foi possível em função das características do padrão DC, que além da interoperabilidade semântica, simplicidade e escopo internacional, são passíveis de customização/extensibilidade, o que permite que as comunidades que o adotam possam adequá-lo às necessidades dos usuários ao descrever um recurso.

Figura 5 – Modelagem da estrutura DC aplicando o modelo conceitual FRAD.

ENTIDADE FRAD	dc.contributor.author:	ZANUS, M. C.
Atributo FRAD	Forma completa do nome	Mauro Celso Zanús
Atributo FRAD	Curriculo Lattes	http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4784107Y6
Atributo FRAD	Produção base AgriS	http://agris.fao.org/agris-search/searchIndex.do?query=ZANUS&x=0&y=0
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0100-29452010005000014
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0100-29452008000300018
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0103-84782009005000242
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0103-84782011000600027
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/1984-70332014v14n4c42
ENTIDADE FRAD	dc.contributor.author:	PEREIRA, E.G.
Atributo FRAD	Forma completa do nome	Giuliano Elias Pereira
Atributo FRAD	Curriculo Lattes	http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4703464Z3
Atributo FRAD	Produção base AgriS	http://agris.fao.org/agris-search/searchIndex.do?query=GIULIANO+ELIAS+PEREIRA&x=9&y=18
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0100-29452006000200024
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/01000683rbs20150284
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S0100-40422010000500032
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/01000683rbs20150284
Atributo FRAD	Produção Sicelo	http://dx.doi.org/10.1590/S1413-70542008000500026

Fonte: Elaboração dos autores.

A estratégia de complementar o elenco de atributos FRAD através do entrelaçamento de dados disponibilizados na *Web*, como o Currículo Lattes, produções submetidas a periódicos qualificados de acesso aberto (via dourada) ou depositadas em bases de dados temáticas, como a AGRIS, amplia as possibilidades de se conhecer melhor um cientista e de se ter acesso às suas obras. Isso faz com que um repositório institucional ultrapasse os estreitos limites de armazenar basicamente literatura cinzenta⁸⁹, abrindo-se à *Web* de dados e interligando-a às suas coleções.

Conforme ressalta Marcondes (2012), a proposta de dados abertos interligados oferece grande potencial e agrega valor aos produtos de informação, como os Repositórios Institucionais (RI), abrindo a possibilidade de conectá-los a outros recursos informacionais através de *links* semânticos significativos, que, ao contrário dos *links* convencionais, são “[...] meios para que programas navegadores, a partir de um recurso, acessem outros [...]” e acrescenta: “Sendo significativos para programas, *links* semânticos podem ser processados enriquecendo cognitivamente o significado (legível por máquina) da ligação entre ambos os recursos.” (MARCONDES, 2012, p. 173).

Sabe-se que, por conta da cessão dos direitos autorais (*copyright*) a editores, o autor abre mão da sua titularidade. Ou seja, continua a ser o produtor intelectual da obra, mas cede à editora o seu direito de utilizar, fruir e dispor da

obra (direito autoral patrimonial). A associação de atributos FRAD às informações disponibilizadas na *Web* por meio de identificadores estáveis (URI) permite-nos minimizar esse problema no âmbito dos repositórios.

Marcondes (2012, p. 179) esclarece que os URIs são “[...] *links* permanentes e estáveis, além de terem um significado preciso, ao contrário dos URL, meros endereços de páginas, instáveis e sempre sujeitos ao frustrante Erro 404 *page not found*”. Finaliza sua análise ressaltando que os *links* semânticos devem ser usados como recursos, direcionando o usuário para catálogos de bibliotecas, verbetes de enciclopédias, e não para meras páginas hipertextuais, cujos *links* são, como já frisado, na maioria das vezes, instáveis.

Na aplicação que sugerimos (figura 5), fez-se um movimento inverso (de dentro para fora), apontando os metadados correspondentes a um determinado atributo FRAD a *links* estáveis e significativos, contextualizando de forma segura a entidade pessoa.

5 Considerações finais

Um Repositório Institucional tem o objetivo de preservar o conhecimento acumulado nas instituições, apresentando como principais vantagens o armazenamento de grande quantidade de informações, facilidade de busca e recuperação e autonomia para o criador fazer a descrição e o autoarquivamento de seus documentos. Ao lado desses objetivos somam-se outros. No caso do repositório institucional ALICE, dar a visibilidade à produção técnico-científica da Embrapa e mensurar o impacto dos resultados das pesquisas por meio da maximização do acesso e do uso da informação produzida na instituição, avaliando de que forma ela vem contribuindo para o progresso da ciência.

O processo de simulação da aplicação dos atributos do FRAD na estruturação descritiva demonstrou que tal implementação melhorará consideravelmente as atividades de busca e recuperação da informação, qualificando os registros com o acréscimo de atributos à entidade pessoa, tais como a forma completa do nome, o Currículo Lattes e a produção científica, inclusões estas permitidas graças ao entrelaçamento de dados abertos (*Linked Open Data*) em que identifica-

dores estáveis, como o *Digital Object Identifier* (DOI) dos artigos submetidos a revistas científicas, substituíram os metadados textuais.

Conforme vimos (figura 5), *links* semânticos associados ao metadado *dc:contributor.author*, na qualidade de atributos, conforme determina o vocabulário FRBR, ampliam as possibilidades de conhecermos a produção intelectual de uma entidade (no exemplo, entidade pessoa) à medida que se permite associá-la a *links* significativos como o seu Currículo Lattes, assim como outras obras depositadas pelo autor na Web, os artigos científicos enviados para periódicos de acesso aberto, ou capturados pela base AGRIS. Outra possibilidade que pode ser aventada diz respeito ao ponto de acesso de autoridade, caso o autor possuísse um registro no arquivo de autoridades do VIAF (*Virtual International Authority File*) poderíamos substituir a descrição textual de seu metadado (uma assertiva RDF) pelo URI do registro VIAF. Infelizmente, no âmbito das instituições de pesquisa ainda não há um serviço de curadoria que compartilhe esse tipo de dado.

Pelo fato de o repositório ALICE ter sido desenvolvido no padrão Dublin Core, cujas características primam pela interoperabilidade e extensibilidade, foi possível explorarmos sua estrutura de metadados, não só estendendo-a, mas agregando novos metadados referentes a atributos que qualificam e complementam as informações da entidade pessoa (criadores das obras). Isso permitiu que, por exemplo, um atributo referente à produção científica pudesse ser repetido quantas vezes fossem necessárias, pois seu objetivo era mapear as diferentes obras produzidas por uma pessoa, depositadas no ambiente *Web*, e que tivessem a ela associado um identificador único, URI, ou seja, um *link* semântico. A experiência mostrou que a criação de novos campos de metadados abre as ferramentas bibliográficas digitais, como os repositórios, ao diálogo com outras: como os repositórios de periódicos científicos e as bases de dados internacionais, como a AGRIS.

No que diz respeito à recuperação da informação, dois aspectos merecem destaque: a) o detalhamento de um registro de autoridade com acréscimos de atributos amplia as possibilidades de relacionamentos entre as entidades pessoas e as demais entidades bibliográficas que compõem o registro DC, como, por

exemplo, obra; b) a incorporação de *links* permanentes e estáveis como o Currículo Lattes e o DOI inaugura a possibilidade de associar as ferramentas bibliográficas aos dados abertos disponíveis na *Web*, cujos *links* expressam ligações com recursos armazenados de forma estável.

Referências

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of *web* content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, [S.l.], v. 284, n. 5, p. 34-43, 2001.

BERNERS-LEE, T. **Linked Data**: Design issues. 2006. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: 12 jul. 2015.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. In: HEATH, T.; HEPP, M.; BIZER, C. (Ed.). **International Journal on Semantic Web and Information Systems**, [S.l.], v. 5, n. 3, p. 1-22, 2009. Special Issue on Linked Data.

BIZER, C.; CYGANIAK, R.; HEATH, T. **How to publish Linked Data on the Web**. 2007. Disponível em: <<http://www4.wiwiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>>. Acesso em: 12 jul. 2015.

CAFÉ, L. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: ENCONTRO NACIONAL DE CENTROS DE INFORMAÇÃO E BIBLIOTECAS DA ÁREA DE COMUNICAÇÃO, 13, Belo Horizonte, MG, 2003. **Anais...** Belo Horizonte: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2003 Disponível em: <http://dspace.ibict.br/dmdocuments/ENDOCOM_CAFE.pdf>. Acesso em: 10 jul. 2015.

CASTRO, R. L. de; PELUFÊ, M. S.; ARRUDA, R. G. Repositórios digitais da Embrapa: acesso livre a produção técnico-científica. In: CONGRESSO BRASILEIRO DE BIBLIOTECOLOGIA, DOCUMENTAÇÃO E CIÊNCIAS DA INFORMAÇÃO, 25, Florianópolis, SC, 2013. **Anais...** Disponível em: <<http://portal.febab.org.br/anais/article/view/1442>>. Acesso em: 2 set. 2015.

CROW, R. **The case for institutional repositories**: a SPARC position paper. Washington, DC, Scholarly Publishing: Academic Resources Coalition, 2002. Disponível em: <http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf>. Acesso em: 12 jul. 2015.

FUSCO, E. **Modelos conceituais de dados como parte do processo da catalogação**: perspectiva de uso dos FRBR no desenvolvimento de catálogos biblio-

gráficos digitais. 2010. Tese de Doutorado, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2010.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

IFLA. Study Group on the Functional Requirements for Bibliographic Records. Functional requirements for bibliographic records. **Final report**. München: K. G. Saur, 1998.

IFLA. Grupo de Trabajo sobre los Requisitos Funcionales y Numeración de Registros de Autoridad. **Informe final**. München: K.G. Saur, 2009a. Disponível em: <http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-es.pdf>. Acesso em: 2 set. 2015.

IFLA. **Declaração de princípios internacionais de catalogação**. 2009b. Disponível em: <http://www.ifla.org/files/cataloguing/icp/icp_2009-pt.pdf>. Acesso em: 20 maio 2015.

KURAMOTO, H. **IBICT estimula a adoção do Open Archives no Brasil**. 2005. Disponível em: <<http://www.ibict.br/noticia.php?id=148>>. Acesso em: 10 fev. 2015.

MARCONDES, C. H. Metadados: descrição e recuperação na web. In: MARCONDES, C. H. et al. (Org.). **Bibliotecas digitais: saberes e práticas**. Salvador: EDUFBA; Brasília; IBICT, 2005. p. 77-143.

MARCONDES, C. H. “Linked data”: dados interligados: e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli**, Florianópolis, v. 17, n. 34, p. 171-192, maio/ago. 2012.

MEY, E. S. A. **Não brigue com a catalogação!** Brasília: Briquet de Lemos/Livros, 2003.

MEY, E. S. A.; SILVEIRA, N. C. **Catalogação no plural**. Brasília: Briquet de Lemos/Livros, 2009.

MORENO, F. P. **Requisitos funcionais para registros bibliográficos – FRBR**: um estudo no catálogo da Rede Bibliodata. 2006. Dissertação de Mestrado, Universidade de Brasília, 2006.

MORENO, F. P.; BRASCHER, M. MARC, MARCXML e FRBR: relações encontradas na literatura. **Inf & Soc. Est.**, João Pessoa, v. 17, n.3, p. 13-25, set./dez. 2007.

OLIVER, C. **Introdução à RDA**: um guia básico. Brasília: Briquet de Lemos/Livros, 2011.

ROSETTO, M. **Metadados e formatos de metadados em sistemas de informação: caracterização e definição**. 2003. Dissertação de Mestrado, Escola de Comunicações e Artes, Universidade de São Paulo, 2003.

SAORÍN, T. Cómo linked open data impactará en las bibliotecas a través de La innovación abierta. **Anuario ThinkEPI**, [S.l.], v. 6, p. 288-292, 2012.

SILVA, R. E. da; SANTOS, P. L. V. A. da C. Requisitos Funcionais para Registros Bibliográficos (FRBR): considerações sobre o modelo e sua implementabilidade. **RBBB**, São Paulo, v. 8, n. 2, p. 116-129, 2012.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 93-102, jan. /abr. 2000.

TILLET, B. **O que é FRBR?**: um modelo conceitual para o universo bibliográfico. 2003. Disponível em: <<http://www.Snake.ne.jp/~yama/nph-docomo.cgi/010000A/http/www.loc.gov/catdir/cpsa/o-que-e-frbr.pdf>>. Acesso em: 9 fev. 2015.

WEITZEL, S. R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 51-71, jan./jun. 2006.

Description of resources in a metadata structure inspired in the FRBR model

Abstract: Analyzes the semantic web impacts on organizational processes and access to information and perspectives brought by the FRBR conceptual model for the field of descriptive cataloging. Simulates the application of FRAD attributes in a repository in order to ascertain the limits and possibilities of use of conceptual models in the descriptive structure of this type of digital library. It demonstrates how the use of conceptual models, when combined with other technologies such as best practices indicated by the use of linked data, can add value to the search process and information retrieval. It concludes that the possibility of replacing textual metadata by universal identifiers will allow software agents to accurately identify the key entities for the recovery of different resources, such as people (creators of scientific works), thus linking the repository, logically and securely, to other data available on the Web.

Keywords: Descriptive Representation. FRBR conceptual model. Digital repository. Semantic Web.

Recebido: 28/09/2015

Aceito: 23/02/2016



-
- ¹ Das tarefas dos usuários especificadas no modelo FRAD, as duas primeiras também estão presentes no modelo FRBR, são elas: **encontrar** uma ou um grupo de entidades que correspondam a um critério de busca do usuário, ou seja, encontrar uma única entidade ou um conjunto delas usando, como critérios de pesquisa, um de seus atributos ou combinações de atributos ou relações entre elas; **identificar** uma entidade, confirmando se a entidade recuperada corresponde à entidade desejada, distinguir entre duas ou mais entidades semelhantes para validar a forma do nome a ser utilizada como ponto de acesso controlado; **contextualizar** uma entidade, esclarecendo a relação entre uma ou mais entidades e conhecer a relação entre duas ou mais pessoas, entidades coletivas, obras, etc.; a forma pela qual uma pessoa, corporação, etc., é conhecida, como, por exemplo, a forma do nome usada por uma pessoa em uma determinada região; **justificar** a criação dos dados de autoridade, fornecendo a justificativa da escolha e da forma do nome atribuído a uma entidade e em que se baseia o ponto de acesso controlado. (IFLA, 2009a, tradução nossa).
- ² Para Tim Berners-Lee, James Hendler e Ora Lassila (2001, p. 37, tradução nossa) “A Web Semântica é uma extensão da Web atual, na qual é dada à informação um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação [...] em um futuro próximo estes desenvolvimentos darão uma nova funcionalidade às máquinas: estas serão mais eficazes e terão maior capacidade de processar e ‘entender’ os dados do que simplesmente exibí-los”.
- ³ A *Scientific Electronic Library Online* - SciELO é uma biblioteca eletrônica que abrange uma coleção selecionada de periódicos científicos brasileiros. (<http://pt.wikipedia.org/wiki/SciELO>)
- ⁴ O INIS é um sistema cooperativo da IAEA (*International Atomic Energy Agency*) com mais de 45 anos de experiência que conta com a participação de mais de 140 países e organizações internacionais. (<http://www.cnem.gov.br/centro-de-informacoes-nucleares/inis-base-de-dados>)
- ⁵ AGRIS é um banco de dados público que fornece acesso a informações bibliográficas sobre ciência e tecnologia agrícola. (<http://agris.fao.org/content/about>)
- ⁶ MEDLINE é a base de dados bibliográficos da Biblioteca Nacional de Medicina dos Estados Unidos da América (*US National Library of Medicine's* - NLM). (<http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&base=MEDLINE&lang=p&form=F>)
- ⁷ Dspace - software de código de fonte aberto que fornece facilidades para o gerenciamento de acervos digitais. segundo dados do IBICT, esse software tem sido o mais usado para a implementação de repositórios de acesso aberto.
- ⁹ O termo Literatura Cinzenta ou *Grey Literature* surgiu em 1978 em uma reunião organizada pela *British Library Lending Division* (BLLD) para designar uma ampla categoria de documentos que não eram objeto de depósito legal, tais como: os relatórios técnicos e de pesquisa, publicações governamentais, traduções avulsas, *preprints*, dissertações, teses e literatura originada de encontros científicos como os anais de congressos, etc. Esses recursos, em sua maioria livres dos compromissos que envolvem os direitos autorais, constituem os acervos das coleções dos repositórios, tal como ocorre como o Repositório ALICE.