

Da ficha bibliográfica aos dados relacionados: uma revolução em curso

Rosa Maria Galvão

Biblioteca Nacional de Portugal
Campo Grande, 83
1749-081 Lisboa
Tel: 217982454
E-mail: rgalvao@bnportugal.pt

Maria Margarida Lopes

Biblioteca Nacional de Portugal
Campo Grande, 83
1749-081 Lisboa
Tel: 217982066
E-mail: mmlopes@bnportugal.pt

RESUMO

Esta comunicação aborda a questão da desadequação das funcionalidades dos catálogos bibliográficos face às exigências dos utilizadores atuais, e da necessidade de desenvolver uma nova geração de catálogos que possuam a capacidade de ir para além dos aspetos tradicionais, não ficando a perder em funcionalidades relativamente a outros interfaces da WWW. Neste contexto, é sublinhada a importância dos profundos desenvolvimentos normativos que se têm produzido à escala internacional nos últimos anos, bem como o seu impacto nos instrumentos de produção e gestão de informação bibliográfica. São referidas, nomeadamente, a publicação da família FRBR, da nova ISBD consolidada, e do código RDA. É também abordada a evolução mais recente do formato UNIMARC, e a questão da possibilidade de integração dos dados bibliográficos na web semântica. Destacam-se as implicações que os modelos conceptuais e o novo código RDA têm na catalogação e no desenho futuro de catálogos, bem como no repensar das *Regras Portuguesas de Catalogação*. Refere-se o trabalho que neste âmbito tem vindo a ser desenvolvido na Biblioteca Nacional de Portugal (BNP), enquadrado nas suas funções normativas.

PALAVRAS-CHAVE: Catálogos na WWW; Modelos conceptuais; RDA; UNIMARC

ABSTRACT

This paper addresses the issue of the inadequacy between bibliographic catalogues and user expectations, and of the need to build a new generation of catalogues with enhanced capabilities that can successfully “compete” with other WWW interfaces. It underlines the importance of changes occurring in cataloguing principles, standards and rules in this context. The publication of the FRBR, the new consolidated edition of ISBD, and the new RDA are discussed, and so are the recent changes in the UNIMARC format and the possibility of integration of bibliographic data in the semantic web. The influence of the new conceptual models and of the RDA code in future tools, including the re-design of *Regras Portuguesas de Catalogação*, is underlined. The standardization activities of the

National Library of Portugal (BNP), in its normative functions, are also presented.

KEYWORDS: Catalogues in the Web; Conceptual models; RDA; UNIMARC

INTRODUÇÃO

As bibliotecas têm uma vasta tradição na criação e desenvolvimento de normas para a organização da informação e para partilha de dados (Denton, 2007; Dunsire, 2010), tendo produzido, ao longo dos anos, gerações de códigos de catalogação. Têm também uma já longa história de abertura a um ambiente normativo alargado, utilizando não só a sua estrutura de dados tradicional mas também normas e protocolos oriundos de outras comunidades, adaptando-se à mudança, criando e desenvolvendo novos princípios, novos modelos conceptuais e novas regras [1]. Nesta abertura, onde é visível a absorção do contexto da sociedade em rede trazido pelo modelo conceptual e social da Web, a partilha de dados torna-se mais lata e mais profunda do que a mera permuta de registos entre bibliotecas (Galvão e Cordeiro, 2010).

Com a progressiva sofisticação e banalização das tecnologias da informação criou-se a noção de uma certa desadequação, e consequentemente subaproveitamento, dos dados existentes nas bases de dados bibliográficos. A questão da desadequação das funcionalidades dos catálogos bibliográficos face às exigências dos utilizadores e também do arcaísmo da sua estrutura, que não propicia uma cabal interoperabilidade no presente ambiente informacional, são um tema recorrente na comunidade biblioteconómica e na respetiva literatura (Coyle, 2010b, 2010c; Coyle e Hillmann, 2007; Weinberger, 2005). De facto, a transição dos catálogos da ficha bibliográfica em papel para o OPAC e, subsequentemente, a migração deste para a Internet e para a WWW, teve impactos significativos no modo como a informação bibliográfica é acedida e utilizada. No entanto, o interface dos recursos continuou a funcionar tendo por base as mesmas regras e códigos, cuja origem é muito anterior ao advento da moderna tecnologia computacional, mantendo assim os constrangimentos inerentes a um sistema de catálogo em fichas (Coyle, 2010a).

A WWW, em particular, veio alterar, de maneira substancial, quer a natureza das publicações, quer as características das ferramentas de mediação da informação, com novos serviços a ocupar o lugar que durante muitos anos esteve cometido à biblioteca e ao seu catálogo. Os próprios utilizadores adquiriram um conjunto de competências de informação bem diferentes das dos utilizadores de algumas décadas atrás, quando as bibliotecas iniciaram o processo de informatização dos seus catálogos: o utilizador de hoje vive num ambiente de rede e de interação e recorre aos mecanismos de pesquisa da WWW para obter resposta às suas necessidades de informação.

Contrastando com as alterações verificadas ao nível tecnológico, as regras e códigos na base da organização dos dados bibliográficos mantiveram-se relativamente estáveis ao longo dos anos, e esta estabilidade refletiu-se também nos formatos MARC. A evolução, quer de uns, quer de outros, foi resultando de alterações pontuais, incrementais, com a adição de novos elementos de dados, para responder a novas solicitações, mas não de uma verdadeira revisão em termos estruturais, não estando neste momento a responder cabalmente, ao nível da funcionalidade e interoperabilidade, às exigências suportadas pelas novas tecnologias [2]. As regras e códigos que durante anos serviram de base às práticas catalográficas estão agora a necessitar de trilhar um novo rumo, mais consentâneo com o ambiente em rede.

O início do séc. 21 veio encontrar as bibliotecas em geral, e a catalogação em particular, a ter que responder a todos estes novos desafios. A distinção entre objeto físico (o “livro”) e a entidade intelectual (a “obra”), ainda não é visível nos catálogos de biblioteca, pensados para serem uma agregação de registos bibliográficos autónomos (mesmo quando implementam o controlo de autoridade para nomes, títulos e assuntos), e não como conjuntos de entidades relacionadas. A influência do catálogo em fichas ainda é forte, e muito presente nos registos bibliográficos criados em sistemas automatizados. As bibliotecas são também detentoras de um legado “pesado”, na medida em que o volume de informação criada e armazenada de acordo com as regras tradicionais é enorme, pelo que as respostas aos novos desafios terão de ser, necessariamente, respostas ao mesmo tempo fraturantes e de continuidade, isto é, que reflitam um novo paradigma de abordagem relacional aos dados, mas que permitam continuar a usar informação criada e guardada de acordo com as regras anteriores.

É este o contexto da presente comunicação, que tem como objetivo dar uma panorâmica dos desenvolvimentos que estão em curso para responder aos desafios trazidos pela evolução tecnológica, e das perspetivas que se desenham para o futuro.

CATÁLOGOS E CATALOGAÇÃO

Embora a história dos catálogos de biblioteca remonte a tempos imemoriais, é no século 19 que tem início o que se considera a catalogação moderna, i.é, a catalogação baseada em axiomas e normas. Como pioneiros da moderna catalogação destacamos Sir Anthony Panizzi (1797-1879), Charles Cutter (1837-1903), S.R. Ranganathan (1892-1972) e Seymour Lubetzky (1898-2003) cujos trabalhos constituíram a base para o

desenvolvimento de princípios e regras catalográficas (Denton, 2007; Lubetzky, 1953).

Se, no início do século 20, a emergência dos assuntos trazidos pelo sistema de catalogação cooperativo era o centro das preocupações e o aparecimento das fichas bibliográficas trazia economia, plenitude, uniformidade e legibilidade aos dados bibliográficos, já as preocupações do século 21 se centram na comunicação em ambiente WWW, na difusão de novos modos de acesso à informação e, essencialmente, na resolução de um paradoxo – o de os dados bibliográficos, ricos em informação e estrutura, serem invisíveis no ambiente WWW.

O Google, fácil e intuitivo, reflete a realidade do utilizador vulgar, mesmo que os resultados obtidos não sejam os mais relevantes nem fidedignos para o assunto pesquisado. Por outro lado, os interfaces dos OPAC refletem o que as bibliotecas consideram ser ontologicamente verdadeiro e os metadados asseguram a qualidade do que é transmitido, esperando-se que os utilizadores empreguem a terminologia adequada de acordo com os normativos apropriados. Temos, portanto, de um lado o foco na captação de utilizadores e, do outro, o foco na qualidade e precisão dos resultados de uma pesquisa. Como estabelecer o equilíbrio? Os utilizadores procuram a simplicidade, a facilidade de uso, algo que funcione bem para eles (Tischler, 2005), mas como simplificar a apresentação dos dados sem redução da qualidade dos mesmos?

A simplicidade na apresentação, face visível para o utilizador, não deve ser confundida com a estrutura de metadados que lhe dá suporte, onde uma simplificação e redução do número de elementos de metadados apenas vai despojar os objetos descritos de informação descritiva, contextual, de proveniência, técnica e administrativa associada a esses objetos, empobrecendo-os semanticamente. Por outro lado, como a interoperabilidade semântica incorpora aspetos linguísticos e culturais, uma maior simplicidade e facilidade de uso dos interfaces (sítio web e OPAC da biblioteca) também não pode passar apenas por uma solução técnica (por ex., mapeamentos de metadados).

Neste contexto colocam-se algumas perguntas para as quais é necessário encontrar respostas. Que tipo de estruturas e normas são necessárias para providenciar um controlo bibliográfico efetivo num ambiente alargado de utilizadores e de utilizações? Como pode ser feito um melhor uso das estruturas e normas existentes e/ou em desenvolvimento, criando melhores interfaces e funcionalidades para o utilizador, para ir ao encontro das suas necessidades? Que infraestruturas é necessário rever para que os dados bibliográficos passem a ser da WWW, mais do que estarem na WWW?

“We must look with cold and hard-headed rationality at our current practices and ask ourselves not what value they offer, but rather what value our patrons believe they offer. If what we offer our patrons is not perceived as valuable by them, then we have two choices: change their minds, or redirect our resources. The former is virtually impossible; the latter is enormously painful. But the latter is possible, and if we do not

undertake such a redirection ourselves, it will almost certainly be undertaken for us.” (Anderson, 2011)

Os catálogos bibliográficos são o produto da catalogação e, se há a necessidade de repensar a maneira como os dados são disponibilizados ao utilizador, é ao nível da catalogação que têm que se verificar as primeiras alterações. O primeiro desafio a vencer é o reconhecimento da necessidade de mudança. O segundo é o de compreender a natureza dessa mudança: objetivos, possibilidades e limitações naturais que decerto irão surgir e que irão influenciar a tomada de decisão (nem sempre a solução ideal é a mais realista ou viável). Por último, se queremos atingir um bom resultado final é necessário chegar a um consenso dentro da comunidade sobre o caminho a seguir.

Ao analisarmos o processamento catalográfico patente ainda nas nossas bibliotecas podemos distinguir duas grandes dimensões, a da estrutura descritiva, composta por princípios, códigos e normas que estão na base e dão suporte à descrição bibliográfica e que, ao terem sido elaboradas num determinado contexto catalográfico e cultural apresentam diversidade entre si, sendo o produto da sua época de produção; e a da estrutura de comunicação que se reflete no modo como os registos bibliográficos são comunicados ao utilizador.

Na primeira, a da estrutura descritiva, vamos encontrar as regras, códigos de catalogação e princípios orientadores – desde as regras de Cutter (1904) [3], onde pela primeira vez foram estabelecidas, explicitamente, as funções a que o catálogo devia responder, passando pelos *Princípios de Paris* (1961), marco histórico para a normalização e cooperação internacional, pelas *International Standard Bibliographic Description* (ISBD), pelas *Anglo American Cataloguing Rules* (AACR) e, mais recentemente, pelos *Princípios Internacionais de Catalogação* (ICP) (2009) e pelo *Resource Description and Access* (RDA) (2010) –, e vamos encontrar também os formatos para codificar os dados bibliográficos, compostos pela família de formatos MARC, onde o UNIMARC se integra.

Resultante desta estrutura, os registos bibliográficos (os metadados bibliográficos) que têm vindo a ser criados pelas bibliotecas são compostos por texto. É certo que o texto está codificado em campos e subcampos, constituindo um conjunto de dados legível e manipulável por computador, mas a base física do registo continua a ser essencialmente texto, cujo principal objetivo é o de ser legível por humanos, permitindo-lhes encontrar e fazer uso da informação.

O modelo de catalogação tradicional pode ser representado do seguinte modo:

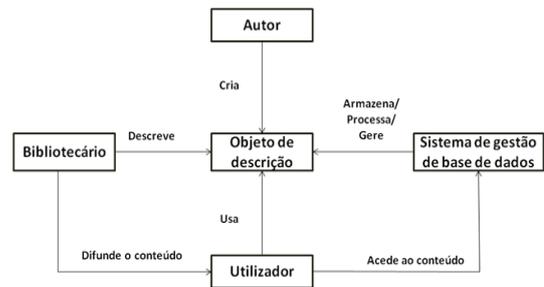


Figura 1: Modelo de catalogação tradicional [4]

Nos dias de hoje, a conceção dos dados para a descrição de um recurso tem que obter o equilíbrio entre dois tipos de funcionalidade - as necessárias para o processamento e inteligibilidade por máquina, e as necessárias para a legibilidade e compreensão humana.

Por outro lado, a catalogação não pode permanecer agarrada meramente ao conceito de descrição da publicação, pensada para produzir catálogos constituídos por uma agregação de registos autónomos, ignorando toda a riqueza relacional que as entidades bibliográficas possuem e as possibilidades de visualização de dados que essas relações permitem. Os conceitos e vocabulário trazidos pelo modelo conceptual FRBR e pelos Princípios Internacionais de catalogação correspondem à consciencialização de novas exigências e também à reformulação de necessidades antigas. O novo código RDA – *Resource description and access* já constitui uma explicitação e aplicação destes conceitos e princípios.

Na segunda dimensão, a da estrutura de comunicação, vamos encontrar uma panóplia de apresentações e visualizações que vão desde o catálogo em livro, passando pela ficha bibliográfica (e pela sua miniaturização e disponibilização em microficha), pelos catálogos em linha e finalmente pelos catálogos disponíveis na WWW. Verificamos, no entanto, que todos estes sistemas convencionais de recuperação de informação bibliográfica desenvolvem de uma forma modesta as inter-relações, pensados que estão para relacionar apenas recursos bibliográficos, através de termos e descrições bibliográficas individualizadas de uma publicação, e não através de ligações efetivas entre registos bibliográficos que compartilhem elementos de dados ou descrições comuns.

Considerando que o ambiente envolvente das bibliotecas está a evoluir rapidamente, é necessário que estas adquiram mais flexibilidade e optem por soluções mais dinâmicas de apresentação da informação sem, no entanto, perderem a qualidade e riqueza dos seus dados. É necessário também começar a olhar para os dados bibliográficos do ponto de vista da tecnologia da web semântica e perceber de que modo se pode transformar a vasta experiência em trabalhar com metadados bibliográficos numa experiência em trabalhar com ontologias ou modelações de conhecimento, tornando estes dados também significantes e acionáveis por máquinas.

NOVOS ENQUADRAMENTOS NORMATIVOS

Uma melhor adaptação ao ambiente da WWW implicou a introdução de alterações nos normativos existentes.

As primeiras, que constituíram a base para posteriores desenvolvimentos, foram trazidas pelos FRBR (*Functional Requirements for Bibliographic Records*) com a introdução no “universo biblioteconómico” do modelo Entidade-Relação.

A reanálise dos objetivos dos catálogos e a revisão da filosofia de base de organização da informação, visando uma exploração mais consentânea com a riqueza dos dados e os interesses dos utilizadores, trazida pelos FRBR, influenciou a preparação de revisões ou reconcepções de códigos de catalogação, estando na base do novo código de catalogação RDA.

Os modelos conceptuais

Os modelos FRBR (*Requisitos funcionais dos registos bibliográficos*), FRAD (*Requisitos funcionais para dados de autoridade*) e FRSAD (*Requisitos funcionais para dados de autoridade de assunto*) [5], constituem, no seu conjunto, uma família de estudos onde são desenvolvidos vários aspetos e níveis de aprofundamento teórico sobre o universo bibliográfico, cuja publicação trouxe uma perspetiva diferente à maneira como o podemos olhar, com implicações profundas na forma como encaramos os catálogos das bibliotecas.

Os modelos têm sido apresentados como representando o retorno a uma tradição de catalogação na qual o catálogo, mais do que um somatório de registos, é visto como um produto final, uma representação do universo bibliográfico.

De facto, nos catálogos em livro, as descrições bibliográficas assumiam tradicionalmente a forma de entradas apresentadas hierarquicamente onde, sob a entrada de cada autor eram dispostos, alfabeticamente, os títulos das respetivas obras. A primeira edição de uma obra existente na biblioteca era objeto de uma descrição completa; se a biblioteca possuía outras edições dessa obra, a entrada era abreviada e feita a seguir à primeira edição, indicando “outra edição”.

Esta organização do catálogo tinha como objetivo principal, para além da economia, “colocar” numa estrutura hierárquica entidades como obras, edições e exemplares. Esta estrutura organizativa manteve-se, de uma forma geral, embora com ajustamentos, quando os catálogos em livro foram substituídos por catálogos de fichas (Svenonius, 2000).

Foi a passagem dos catálogos em ficha para os catálogos em linha que levou à perda, quase por completo, da sua estrutura bibliográfica – a estrutura sindética do catálogo.

Os modernos catálogos em linha, apesar de disponibilizarem vários pontos de acesso, não expressam o conjunto do universo bibliográfico (ou a parte do universo bibliográfico que é representado nas existências da biblioteca) numa estrutura bibliográfica completa, levando à existência de um conflito, que anteriormente não existia, entre duas funções essenciais do catálogo: i) a de encontrar o recurso (pesquisa específica); ii) e a de o “colocar” (pesquisa de materiais

relacionados/com características comuns) (Bianchini e Guerrini, 2009).

Ou seja, com a informatização, a estrutura sindética dos antigos catálogos em livro não foi implementada nos catálogos em linha, pelo menos completamente, pelo que este cumpre apenas a primeira função – ajudar o leitor a encontrar o recurso específico que procura.

Os modelos relacionais, ao apresentarem o universo bibliográfico como um conjunto de entidades relacionadas entre si, vêm colocar ênfase nas relações e não nas entidades consideradas individualmente, e chamam por isso a atenção para a contradição patente entre a riqueza e complexidade do universo bibliográfico, por um lado, e a linearidade das estruturas apresentadas pelo catálogo, por outro.

No entanto, como modelos teóricos que são, de alto nível, fornecem o enquadramento, mas não substituem nem acarretam, em si próprios, princípios e regras de catalogação, ou novos modelos de registo de dados. Para a concretização dos objetivos de melhoramento dos catálogos é fundamental que existam regras de catalogação e formatos de registo de dados que os implementem.

A ISBD consolidada

Publicada pela primeira vez em 1971, a ISBD - *International Standard Bibliographic Description*, ao normalizar a forma de recolha e a estrutura de apresentação dos elementos de dados, constituiu a base sobre a qual muitos códigos de catalogação desenvolveram as regras para a descrição bibliográfica [6].

O movimento de refundação de princípios, normas e códigos de catalogação, verificado internacionalmente na sequência da publicação dos FRBR, também se refletiu na ISBD. Tendo até então conhecido várias fases evolutivas, que passaram por processos de especialização, com a criação de ISBD específicas por tipo de material, e por revisões regulares, com o surgimento dos FRBR a IFLA deu início a um processo de revisão mais profunda.

Deste processo de revisão resultou a evolução da ISBD, de oito textos individualizados para um texto único, num esforço para integrar os textos e alinhar a descrição de todos os tipos de materiais ao mesmo grau de conformidade com o FRBR. Como principais alterações, estendeu-se a descrição bibliográfica a novos tipos de recursos, criou-se uma nova zona 0, para indicar a forma do conteúdo e suporte do recurso, e tornaram-se opcionais os elementos de dados que tinham a taxa de pertinência mais baixa relativamente às operações efetuadas pelos utilizadores e definidas no modelo FRBR (pesquisar, identificar, selecionar e obter). O novo texto oferece as bases para descrever todos os aspetos do recurso, incluindo o seu conteúdo, o seu suporte e o modo de publicação.

Apesar da profundidade das alterações deste novo texto da ISBD, a sua evolução recente para um texto integrado único, publicado como edição consolidada, constitui um exemplo de como é difícil “dar o salto” para o novo paradigma: se em termos de terminologia e conceitos utilizados, o texto da nova ISBD vai ao encontro dos modelos conceptuais, ela continua, no

entanto a centrar-se na publicação, que agora designa por “recurso”, e não numa abordagem relacional ao universo bibliográfico.

A ISBD continua, efetivamente, centrada na descrição da manifestação. Uma vez que não está no âmbito desta norma estabelecer prescrições relativamente à construção dos pontos de acesso, ela não é a norma onde, por excelência, poderão ser determinadas regras relativamente à explicitação dos diversos tipos de relações que se podem estabelecer entre as entidades do universo bibliográfico.

No entanto, embora a descrição bibliográfica e o acesso sejam aspetos distintos - a descrição dizendo respeito aos recursos bibliográficos; e o acesso (e a estrutura sindética) dizendo respeito às obras existentes no universo bibliográfico, dos quais os recursos bibliográficos constituem uma manifestação, - na verdade estes dois aspetos estão intimamente relacionados.

A ISBD continua por isso a ter um papel a desempenhar, não só no que diz respeito à função de descrição, mas também, ainda que indiretamente, de organização dos catálogos: as regras para descrição bibliográfica aceites e aplicadas internacionalmente possibilitam a existência de visualizações consistentes e previsíveis dos dados, e colocam uma descrição bibliográfica de uma manifestação num determinado contexto do catálogo e de outras descrições de outros recursos, numa lista mais ou menos complexa (Bianchini e Guerrini, 2009).

RDA - Resource description and access

As alterações até agora referidas, ao nível da informatização, da diversificação dos suportes de informação, da crescente complexidade dos conteúdos, da própria publicação dos modelos relacionais, constituem o contexto geral em que podemos enquadrar o surgimento de RDA - *Resource description and access*.

De facto, tanto no caso das *Anglo-American Cataloguing Rules* como de diversas regras de catalogação nacionais se previa, desde a década de 90, uma renovação inevitável, dadas as limitações das regras vigentes.

O novo código RDA nasce da necessidade de refundar os códigos de catalogação, para os alinhar com os modelos relacionais, e vem substituir as AACR2, mas não se resume a uma simples atualização ou revisão destas. Constitui-se como uma explicitação e aplicação dos conceitos e princípios dos modelos relacionais, e apresenta-se como uma nova forma de abordagem ao paradigma bibliográfico, estabelecendo novas regras de descrição e propondo formas concretas de proceder à articulação entre os vários níveis das entidades bibliográficas.

Trata-se de uma norma de metadados baseada num quadro teórico, que tem como propósito essencial descrever o conteúdo dos recursos, melhorando a identificação e o acesso a esses mesmos conteúdos. É construída sobre a experiência do desenvolvimento das AACR2, mas com um âmbito internacional mais alargado, a dois níveis: por um lado, pretende-se que venha a ser usada internacionalmente, estendendo-se

para além do conjunto de países anglo-saxónicos que já utilizavam as AACR. Por outro lado, embora se destine em primeiro lugar à aplicação aos recursos existentes nas bibliotecas, procura estender-se também para além dessa fronteira mais restrita, num esforço para ser também efetivamente compatível com as abordagens aos metadados que caracterizam comunidades relacionadas, como arquivos, museus e editores.

Tal como as AACR2, o RDA consiste num conjunto de instruções, mas com um novo enquadramento teórico que ao definir a sua forma, estrutura e conteúdo, lhe dá uma organização diferente. A chave para compreender o RDA é o seu alinhamento com os modelos relacionais FRBR e FRAD.

Outro aspeto relevante do RDA é a sua independência relativamente a qualquer estrutura ou formato específicos para visualizar ou armazenar metadados. De facto, o RDA apresenta-se deliberadamente como uma norma de conteúdo, independente de qualquer esquema de codificação ou visualização dos dados, aspeto relativamente ao qual é quase completamente omissa. Apenas em apêndice é feita referência à ISBD, considerada meramente como uma norma de apresentação dos dados; e é também em apêndice que são estabelecidas as formas para os pontos de acesso.

Ou seja, o RDA define os elementos necessários para a descrição e o acesso, e dá instruções sobre o modo como a informação deve ser registada em cada um dos elementos. Assim, a informação é segmentada em elementos definidos de forma clara, e cada elemento é definido sem ambiguidade, contendo um tipo específico de informação.

Esta forma de registar a informação permite que esta possa ser codificada independentemente do esquema de metadados. Podem ser usados esquemas de metadados já existentes – MARC21, UNIMARC, Dublin Core, MODS – ou outros esquemas atuais ou futuros. Torna-se assim possível assegurar um aspeto fundamental – o da integração dos dados criados com base neste novo código RDA com os dados já armazenados, que foram criados de acordo com regras de catalogação pré-existentes, sejam elas as AACR ou outras.

Assim, o RDA pode ser aplicado usando a tecnologia e as bases de dados atuais, por exemplo em registos MARC armazenados nos catálogos “tradicionais”. Mas pode também ser usado em ambiente Web e em novos tipos de estruturas de bases de dados, como base para um conjunto de elementos de metadados que tornam a informação visível e utilizável num ambiente Web.

OS DESENVOLVIMENTOS NORMATIVOS E O UNIMARC

Os FRBR modelam o universo bibliográfico de um catálogo como um diagrama de entidades distintas mas interrelacionadas entre si. Cada registo bibliográfico é representativo de uma parte desse universo.

O formato UNIMARC bibliográfico permite, desde a sua origem, dar resposta a uma estrutura relacional a vários níveis [7], através de uma gama de relações muito vasta que vão desde as relações entre manifestações quase idênticas às totalmente distintas, como manifestações de outras expressões da obra ou

manifestações de expressões de obras novas resultantes de adaptações ou modificações.

Ao analisarmos os registos em UNIMARC efetivamente existentes nos catálogos bibliográficos verificamos que, na sua maioria, já apresentam uma estrutura que pode ser considerado como uma estrutura FRBR simplificada, onde as entidades, implícita ou explicitamente descritas, podem ser inferidas quer através das etiquetas e códigos quer pela interpretação do próprio modelo.

Esta abordagem representa, no entanto, uma visão simplificada do modelo. Partindo da premissa que a maior parte dos registos descrevem uma única publicação/edição que, no modelo FRBR, é designada por “manifestação”, materializando, pelo menos, uma expressão, que é a realização de uma obra, está a ser excluída a amplitude de utilização do modelo FRBR tal como ele foi desenhado.

As manifestações podem materializar várias expressões que podem ser outras entidades, com as quais estão relacionadas mas nas quais podem, ou não, estar contidas. Um registo UNIMARC pode ou não identificar explicitamente a expressão relacionada, tornando impossível providenciar uma descrição adequada do recurso relacionado.

Um registo bibliográfico em UNIMARC é para ser usado como um todo, providencia o contexto para cada elemento de dados e congrega todos os campos que descrevem uma manifestação em particular. Pode, no entanto, conter metadados que pertençam a mais do que uma entidade “obra” ou a mais do que uma entidade “manifestação”. Quando isto ocorre é necessário fazer uma análise dos dados que permita a sua separação lógica [8].

O problema da correta identificação das entidades FRBR num registo bibliográfico tradicional é bem conhecido [9] e resulta da prática catalográfica corrente, que não está alinhada com o modelo. Por outro lado também é necessário que o UNIMARC possua uma estrutura que permita a correta representação de estruturas complexas de entidade relação.

Ao longo dos seus 35 anos de existência o UNIMARC evoluiu de modo a adaptar-se às necessidades emergentes para responder não só à descrição de diferentes tipos de recursos, mas também para se manter alinhado com as alterações verificadas na ISBD e com os conceitos e terminologia trazidos pelos novos Princípios Internacionais de Catalogação e, mais recentemente, para responder aos desafios trazidos pela implementação dos FRBR e FRAD nas normas bibliográficas.

A publicação do RDA implementando, na prática, o modelo FRBR, motivou a adaptação do UNIMARC para descrever as entidades obra e expressão sem, no entanto, deixar de manter a coerência e usabilidade da informação criada e armazenado de acordo com as regras “tradicionais”. O movimento que levou à introdução destas alterações, criando novos campos, especialmente ao nível do UNIMARC Autoridades, teve início em 2008 e atingiu o seu expoente máximo em 2011 [10], onde metade das propostas apresentadas ao PUC (Permanent UNIMARC Committee) se

destinavam a atingir este objetivo.

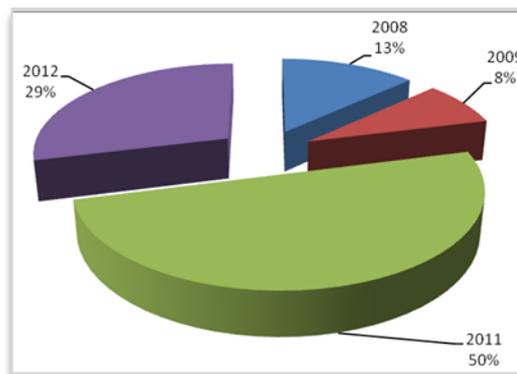


Figura 2: Propostas de alteração para acomodar entidades FRBR

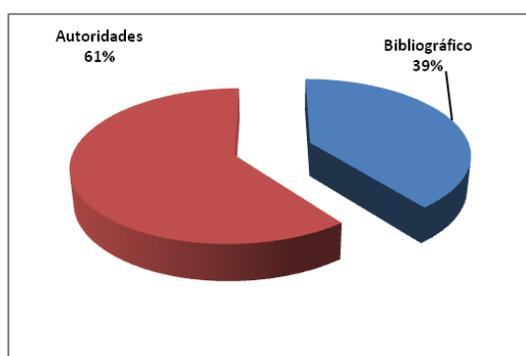


Figura 3: Alterações por formato para acomodar entidades FRBR

OS FORMATOS DE DADOS E A WEB SEMÂNTICA

Os catálogos bibliográficos podem ser a conexão entre a biblioteca e o espaço do conhecimento na web, ao possibilitarem a conexão entre os recursos da biblioteca e os recursos que lhe são externos. No entanto, para que isto aconteça, não basta distribuir os metadados bibliográficos na rede, eles também têm que pertencer e interagir na web, só assim será possível incrementar não só as oportunidades para os utilizadores os descobrirem, mas também o seu valor, ao permitir que sejam utilizados numa grande variedade de contextos (Coyle, 2010c). É, portanto, essencial que os dados bibliográficos venham a ter a capacidade de serem integrados no espaço virtual dos utilizadores.

Num catálogo tradicional, onde a gestão predominante assenta em conjuntos de informação ligada para formar registos - armazenados em ficheiros com uma determinada estrutura específica, e tradicionalmente a operarem em ambientes fechados e estáticos -, um campo UNIMARC que seja retirado do contexto pode não ser significativo, a menos que a informação seja concebida para ser significativa tanto dentro de um registo como fora dele.

As tecnologias da web semântica conceptualizam os dados de uma forma diferente desta conceptualização na base dos formatos de dados do século 20, estruturam os dados como diagramas (construções que, em princípio, podem ser ilimitadas), colocam a ênfase na relação

entre dados oriundos de fontes diversas e operam em sistemas abertos, dinâmicos e interativos.

Assim, é necessário que as bibliotecas sejam capazes de preparar os seus metadados para este ambiente, criando identificações unívocas de coisas e relações, de modo a assegurar que os dados não perdem o significado específico quando misturados com outros dados no ambiente WWW (Coyle, 2010c). Isto é conseguido providenciando relações e identificadores explícitos para o assunto da descrição que permitam processar elementos de dados individuais dentro do seu contexto de significância, dando à máquina uma maior possibilidade de poder atuar sobre eles. É necessário, portanto, que as bibliotecas sejam capazes de preparar os seus metadados de modo a que possam ser utilizados como *linked data* [11].

Os *linked data* são um modelo de dados que identifica, descreve, conecta e relaciona elementos de dados estruturados e que têm como principal objetivo facilitar a reutilização, a interligação, a integração e a partilha de dados em WWW [12]. Como meta-modelo fornece uma estrutura para a definição, conceção, desenvolvimento e manutenção de esquemas e vocabulários de qualquer tipo e tamanho de um determinado domínio, o que permite que instituições, como as bibliotecas, não tenham, obrigatoriamente, de abandonar as normas de metadados existentes e todo o seu legado de vocabulário e listas de autoridade, mas sim transformá-los em dados que possam ser lidos e interpretados por máquinas.

Para que isto seja possível é necessário: i) conceber um modelo de dados base – as bibliotecas ainda não o têm mas já têm um modelo de informação na forma de requisitos funcionais (FRBR, FRAD e FRSAD); ii) definir as propriedades dos metadados bibliográficos, parte do processo passa por tornar as definições disponíveis na Web num formato acionável por máquina; iii) definir toda a lista de vocabulário controlado, num formato compatível com os *linked data*, e disponibilizá-la na Web de modo a que qualquer um a possa utilizar (Coyle, 2010c).

Os movimentos relacionados com a passagem dos dados bibliográficos para a web semântica começam a fazer-se com mais rapidez com o registo de conjuntos de elementos de metadados e vocabulário de FRBR (Dunsire, 2011), de RDA (Hillmann e outros, 2010), da ISBD e, mais recentemente, também o UNIMARC tem em desenvolvimento um projeto para o registo dos seus elementos e vocabulário.

Perspetivas para os catálogos

Não há ainda uma definição clara que caracterize a realidade a que nos referimos quando falamos em nova geração de catálogos. Em termos gerais, o que se pretende é que os catálogos bibliográficos possuam a capacidade de ir para além dos aspetos que tradicionalmente possuem, e que, se comparados com outros interfaces da WWW, não percam em funcionalidades.

As tecnologias da web semântica colocam a ênfase na relação entre dados oriundos de fontes diversas mas, o que ainda não está bem clarificado é o modo como podemos vir a partilhar estas entidades. No entanto, sem um corpo de dados bibliográficos convertidos para

linked data também não é possível criar *software* e novas aplicações que tirem proveito dos *linked data* ou que os implementem num ambiente existente.

No que diz respeito aos formatos de dados bibliográficos, no momento presente ainda não se pode dizer que existam evoluções muito significativas. O seu desenvolvimento tem sido essencialmente de manutenção, acrescentando ou alterando elementos pontuais para novas necessidades específicas numa estrutura base que, em geral, se tem mantido.

As maiores inovações têm-se dado ao nível da interoperabilidade com outros esquemas de metadados, como o Dublin Core [13], reformulação ou integração com outros dados para novas finalidades de gestão, como os esquemas METS, MODS e MADS [14], ou simplesmente o desenvolvimento de novas formas de transmissão, como o MARCXchange [15] e o MARCXML [16] para o alargamento da utilização de dados MARC no ambiente Web.

A abordagem trazida pelos FRBR, ao contrário da consignada na ISBD e nas regras nacionais que nela se baseiam, que é a da descrição e transcrição da informação, traz-nos já uma perspetiva mais consentânea com as novas necessidades, dando ênfase à identificação das entidades e às relações existentes entre essas entidades. Cada registo descreve uma entidade, criando assim a base para que o “registo bibliográfico” criado não seja visto como uma peça única mas sim como um conjunto de registos de entidade ligados entre si através da WWW, o que se traduz num modo mais eficiente de armazenar dados na medida em que cada entidade é registada apenas uma vez, evitando a redundância.

Esta abordagem abriu espaço para o redesenhar de novas normas para o registo de dados, estando a emergir o movimento *Bibliographic Framework Transition Initiative* [17], coordenado pela Biblioteca do Congresso. O objetivo principal é o de encontrar uma nova estrutura de dados que permita a substituição do formato MARC21, tirando o melhor partido das tecnologias, e garantido ao mesmo tempo que é preservada a solidez da troca de dados que nas últimas décadas permitiu às bibliotecas partilhar recursos e obter economias significativas nos custos da catalogação.

PERSPECTIVAS NACIONAIS E CONSIDERAÇÕES FINAIS

Também na realidade portuguesa é necessário considerar todas estas transformações, avaliar as suas implicações, e encontrar um caminho para avançar, num momento em que este processo de refundação ainda está em curso e sem fim à vista. Em especial, é necessário refletir sobre as implicações que poderão ter no repensar o futuro das *Regras Portuguesas de Catalogação*, cuja reformulação é uma necessidade há muito sentida.

Presentemente, a BNP, na sua estratégia de reforço às atividades normativas, tem acompanhado os desenvolvimentos internacionais e realizado a sua divulgação a nível nacional essencialmente através da tradução e publicação de documentos normativos da IFLA e de normas da ISO (esta última decorrente de ser Organismo de Normalização Setorial para a Informação

e Documentação).

Dos resultados práticos já alcançados destacam-se as edições portuguesas de documentação da IFLA como os FRBR e os FRAD, as últimas edições de manuais UNIMARC, bibliográfico e de Autoridades, a Declaração dos Princípios Internacionais de Catalogação e, já no presente ano, a edição consolidada da ISBD.

No âmbito do RDA, a BNP tem atualmente em preparação a tradução para português do código, e tem acompanhado os trabalhos do EURIG – *European RDA Interest Group*, que se tem debruçado, discutido e partilhado experiências relativamente à implementação do RDA na Europa, onde a ISBD continua a constituir a norma-base da descrição.

Uma vez terminada a fase de tradução, encetar-se-á o trabalho conducente à preparação de novas Regras Portuguesas de Catalogação, um trabalho que terá naturalmente como fontes os novos Princípios Internacionais de catalogação, os modelos conceptuais FRBR e FRAD e o código RDA.

Num momento em que assistimos a transformações profundas, que implicarão um trabalho significativo no repensar das assunções e práticas nacionais, é mais do que nunca essencial que a comunidade de profissionais se mantenha atualizada e se envolva ativamente neste movimento, para que dele possa resultar uma mudança em que toda a comunidade se reveja.

NOTAS

(Todos os URL válidos em 2012.09.20)

[1] Ver: *Declaração de Princípios Internacionais de Catalogação* (2009); o texto em português pode ser consultado em http://www.ifla.org/files/cataloguing/icp/icp_2009-pt.pdf. A família de modelos conceptuais iniciada com os *Requisitos Funcionais de Registos Bibliográficos* (FRBR) (IFLA, 1998, act. 2009; edição portuguesa, 2008), seguidos de *Requisitos Funcionais para Dados de Autoridade* (FRAD) (IFLA, 2009; edição portuguesa, 2011) e de *Functional Requirements for Subject Authority Data* (FRSAD) (IFLA, 2011) e dos trabalhos em curso para os *Functional Requirements for Bibliographic Records object-oriented* (FRBRoo).

[2] Ver, por ex., Thomale (2010) para o MARC, em <http://journal.code4lib.org/articles/3832>, e Galvão e Cordeiro (2012) para o UNIMARC, em <http://conference.ifla.org/sites/default/files/files/papers/wlic2012/92-galvao-en.pdf>.

[3] Data da 4ª ed., publicada postumamente, que introduz alterações à primeira edição de 1876.

[4] Figura adaptada de Alemu, Stevens e Ross, 2012.

[5] Ver também a comunicação de Galvão e Cordeiro, 2010.

[6] As *Anglo American Cataloguing Rules*, 2nd ed. (AACR2), cuja estrutura descritiva é baseada na ISBD(G), são um exemplo desta aplicação.

[7] Permite relações do tipo descendente/ascendente e de todo/parte (as designadas relações verticais); anterior/posterior (as relações cronológicas); relações entre manifestações ou expressões ou obras diferentes, ou outras relações, (as relações horizontais).

[8] Por ex., registos bibliográficos com campos do bloco 4xx do tipo “publicado com” ou com o campo 327 identificando vários conteúdos, como por ex., os títulos das várias faixas de um disco.

[9] Consultar, por ex.: Hegna e Murtooma, 2002, em <http://archive.ifla.org/IV/ifla68/papers/053-133e.pdf>; Hickey e O’Neill, 2005, em *Cataloging & Classification Quarterly*. Vol. 39, no. 3-4, p. 239-251; Pisanski, Žumer e Aalberg, 2009, em <http://conference.ifla.org/past/ifla75/77-pisanski-en.pdf>; Galvão, Freire e Lopes, 2009, em <http://conference.ifla.org/past/ifla75/135-freire-en.pdf>.

[10] Entre 2008 e 2012 foram aprovadas 38 propostas de alteração para acomodar as entidades FRBR.

[11] Yee, num artigo disponível em: <http://escholarship.org/uc/item/91b1830k>, explorou o modo como a catalogação tradicional pode ser transposta para o domínio da web semântica com a modelação dos dados em RDF.

[12] Ver informação relativa aos *linked data* especialmente em: W3C INCUBATOR GROUP (2011). *Library Linked Data Incubator Group Final Report*. Disponível em: <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>; e em BERNERS-LEE, T. (2006-2009). *Linked data*. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>.

[13] *Dublin Core Metadata Element Set*, Version 1.1, disponível em <http://dublincore.org/documents/dces/>.

[14] METS – *Metadata Encoding Transmisson Standard*, disponível em <http://www.loc.gov/standards/mets/>. MODS – *Metadata Object Description Schema*, disponível em <http://www.loc.gov/standards/mods/>. MADS – *Metadata Authority Description Schema*, disponível em <http://www.loc.gov/standards/mads/>.

[15] ISO 25577:2008 – *Information and Documentation – MarcXchange*.

[16] Ver informação relativa a esta iniciativa em <http://www.loc.gov/marc/transition/>.

REFERÊNCIAS

(Todos os URL válidos em 2012.09.20)

ALEMU, Getaneh ; STEVENS, Bret ; ROSS, Penny (2012). Towards a conceptual framework for user-driven semantic metadata interoperability in digital libraries: a social constructivist approach. *New library world*, ISSN 0307-4803. Vol. 113, no1/2, p. 38-54

ANDERSON, Rick (2011). The crisis in research librarianship. *Journal of Academic Librarianship*. Artigo em pré-impressão. Doi:10.1016/j.acalib.2011.04.001

BIANCHINI, Carlo ; GUERRINI, Mauro (2009).” From bibliographic models to cataloguing rules : remarks on FRBR, ICP, ISBD, and RDA and the relationships between them”. *Cataloging & Classification quarterly*. Vol. 47, no. 2, p.105-124

COYLE, Karen (2010a). Changing the nature of library data. *Library Technology Reports*, ISSN 0024-2586. Vol. 46, no 1 (Jan.), p.14-31

COYLE, Karen (2010b). Library data in a modern context. *Library Technology Reports*, ISSN 0024-2586. Vol. 46, no 1 (Jan.), p.5-13

COYLE, Karen (2010c). Library data in the web world. *Library Technology Reports*, ISSN 0024-2586. Vol. 46, no 2 (Fev. /Mar.), p.5-11

COYLE, Karen ; HILLMANN, Diane (2007). Resource Description and Access (RDA): cataloging rules for the 20th century. *D-Lib Magazine*. ISSN 1082-9873. Vol. 13, nr 1/2 (Jan.-Feb. 2007). Disponível em: <http://www.dlib.org/dlib/january07/coyle/01coyle.html>

DENTON, William (2007). FRBR and the History of Cataloging. In: Understanding FRBR : what it is and how it will affect our retrieval tools. ed. Arlene G. Taylor. Westport : Libraries Unlimited, 2007. p. 35-57

DUNSIRE, Gordon (2010). UNIMARC, RDA and the semantic web. *International cataloguing and Bibliographic control*. ISSN 1011-8829. Vol. 39, no 2 (Apr.-Jun.), p.37-40

DUNSIRE, Gordon (2011). Linked data and the implications for library cataloguing: metadata models and structure in the Semantic Web. In: CLA 2011 National Conference and Trade Show (2011). Halifax, Nova Scotia : The World Trade & Convention Centre: May 25 - 28, 2011. F40. Disponível em: <http://www.cla.ca/conference/2011/postcon.html>

GALVÃO, Rosa Maria ; CORDEIRO, Maria Inês (2010). Novos princípios, modelos e normas para o futuro dos serviços de informação bibliográfica. In: Políticas de informação na sociedade em rede : actas. 10º Congresso Nacional [dos] Bibliotecários, Arquivistas e Documentalistas. Guimarães : BAD

HILLMANN, Diane, e outros (2010). RDA vocabularies: process, outcome, use. *D-Lib Magazine*. ISSN 1082-9873. Vol. 16, nr 1/2 (Jan.-Feb. 2010). Disponível em: <http://www.dlib.org/dlib/january10/hillmann/01hillmann.print.html>

IFLA (1998). *Functional requirements for bibliographic records: final report*. IFLA Study Group on the functional requirements for bibliographic records. München: Saur (UBCIM Publications, New Series, 19). Versão em linha, actualizada em 2009, disponível em: http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf.

IFLA WGFRANAR (2009) *Functional Requirements for Authority Data : a conceptual model*. IFLA Working Group on Functional Requirements and Numbering of Authority Records ; ed. by Glenn E. Patton. Munchen : K. G. Saur.

IFLA WGSAR (2009) *Functional Requirements for Subject Authority Data (FRSAD) : a conceptual model* . IFLA Working Group for Subject Authority Records. 2nd Draft 2009-06-10. Disponível em: <http://nkos.slis.kent.edu/FRSAR/report090623.pdf>.

LUBETZKY, Seymour (1953, ed. 2001). Cataloging Rules and Principles: A Critique of the A.L.A. Rules for Entry and a Proposed Design for their Revision, 1953. In: SVENONIUS, Elaine, MCGARRY, Dorothy, eds. *Seymour Lubetzky: writings on the classical art of cataloging*. Englewood: Libraries Unlimited, 2001. p. 78-139

SVENONIUS, Elaine (2000). *The intellectual foundation of information organization*. Cambridge (Massachusetts): The MIT Press. ISBN 0-262-19433-3

TISCHLER, Linda (2005). The Beauty of Simplicity. *The fast Company*. Nr 100, p. 52-58

WEINBERGER, D. (2005). Tagging and Why it Matters. Disponível em: <http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf>