

O modelo FRBR e a descoberta de informação: a experiência do projecto TELplus

Margarida Lopes, Nuno Freire

Biblioteca Nacional de Portugal

Campo Grande, 83

1749-081 Lisboa

Tel: 217982000

E-mail: {mmlopes, nuno.freire}@bnportugal.pt

*Hugo Manguinhas, Gilberto Pedrosa,
Maria Teresa Teixeira, José Borbinha*

Instituto Superior Técnico / INESC-ID

Rua Alves Redol, 9

1000-029 Lisboa

E-mail: {hugo.manguinhas, gilberto.pedrosa,
jlb}@ist.utl.pt; teixeira.mariateresa@gmail.com

RESUMO

Neste artigo descreve-se uma experiência levada a cabo pela Biblioteca Nacional de Portugal e o Instituto Superior Técnico, no âmbito do projecto TELplus, para testar a relevância da aplicação do modelo FRBR para a apresentação dos resultados no portal TEL. O modelo conceptual FRBR veio apresentar um novo paradigma resultante do reconhecimento das limitações dos modelos MARC tradicionais. Descreve-se o modelo FRBR nos seus aspectos principais, seguindo-se a descrição da experiência levada a cabo e seus resultados. Os registos utilizados estavam nos formatos MARC21 e UNIMARC, pelo que apenas se testou a relevância do modelo FRBR para apresentação de resultados de pesquisa, não tendo este estudo incidido sobre a catalogação.

PALAVRAS-CHAVE: FRBR, FRBRização, TEL, MARC21, UNIMARC, OPAC

INTRODUÇÃO

Os Requisitos Funcionais dos Registos Bibliográficos – FRBR (*Functional Requirements for Bibliographic Records*), publicados em 1998 pela IFLA, são um modelo conceptual cujo desenvolvimento resultou do reconhecimento das limitações dos catálogos bibliográficos tradicionais, e que veio apresentar um novo paradigma do tratamento bibliográfico.

Desde essa altura que temos assistido a diversas experiências de aplicação do modelo FRBR, maioritariamente em catálogos de dimensão restrita, que têm resultado em protótipos. No entanto, ainda se esperam implementações relevantes, estáveis e em escala relevante desse modelo.

Os principais desafios apontados a esta implementação resultam da conjugação de vários aspectos que procuraremos abordar, o principal dos quais é talvez a natureza teórica do estudo, não directamente relacionado com normas, estruturas ou conteúdos de dados concretos, o que faz com que seja ainda hoje percepcionado, por muitos profissionais de informação, como um modelo complexo e de compreensão difícil.

O MODELO FRBR

Génesese

O modelo FRBR teve a sua génesese no Seminário sobre Registos Bibliográficos, realizado em Estocolmo em 1990, sob a égide da IFLA. Estavam em curso profundas mudanças no ambiente das bibliotecas, e a percepção dos profissionais de informação era da existência de uma crescente desadequação entre os princípios normativos de controlo bibliográfico em vigor, baseados essencialmente nas ISBD e nos *Princípios de Paris* concebidos em 1961, e o novo ambiente tecnológico, radicalmente transformado desde essa altura. Desde então, o desenvolvimento de sistemas automatizados para processamento de dados bibliográficos permitiu o crescimento, em grande escala, de bases de dados nacionais e internacionais, com registos criados e usados por milhares de bibliotecas participantes em projectos de catalogação partilhada. A produção editorial cresceu também exponencialmente nesse período, e surgiram novas formas de edição electrónica e de redes de acesso aos recursos de informação, sendo necessário adaptar códigos e práticas de catalogação para acomodar estes recursos. Por outro lado, era crescente a pressão para reduzir os custos do processamento bibliográfico, através da catalogação a um nível mínimo e de forma cada vez mais simplificada.

Tornava-se essencial resolver o dilema da conjugação de aspectos profundamente contraditórios: por um lado, a necessidade de corresponder a expectativas cada vez maiores dos utilizadores, e de conseguir abordar diferentes tipos de materiais; por outro a pressão económica para reduzir os custos de catalogação, num momento em que a informatização retrospectiva dos catálogos pesava significativamente às bibliotecas.

Deste seminário resultaram duas conclusões relevantes: o reconhecimento de que a contínua pressão para catalogar a um nível mínimo requeria uma ponderação cuidada das relações entre determinados elementos de dados de um registo e as necessidades dos utilizadores; e o reconhecimento da necessidade de estabelecer uma norma que especificasse qual o nível mínimo ou fundamental do registo, para viabilizar, neste contexto, os programas e projectos de catalogação partilhada.

Uma das resoluções então aprovadas foi a da realização de um estudo para definir quais os requisitos funcionais de um registo bibliográfico, e recomendar um nível mínimo de funcionalidade para os registos criados pelas agências bibliográficas nacionais.

Esse estudo ficou a cargo de um grupo de trabalho da IFLA – o IFLA Study Group on the *Functional Requirements for Bibliographic Records* –, e foi publicado em 1998. Na sequência desta publicação, temos assistido nos últimos 12 anos a uma reflexão e viva discussão entre os profissionais de informação sobre o modelo, seus pressupostos, sua viabilidade, e sobre as implicações da sua implementação, num volume de produção teórica talvez sem paralelo na história da Biblioteconomia.

Aspectos essenciais do modelo

O modelo FRBR implica uma mudança radical no paradigma bibliográfico tradicional. Identifica e define claramente as entidades de interesse para os utilizadores de registos bibliográficos, os atributos de cada entidade, e os tipos de relações que se estabelecem entre as entidades. A intenção foi a de produzir um modelo conceptual para relacionar atributos e relações específicos (presentes no registo como elementos de dados separados) com as diversas operações que os utilizadores realizam quando consultam os registos bibliográficos.

As entidades constituem os objectos-chave de interesse para o utilizador de dados bibliográficos, e podem representar produtos intelectuais ou artísticos que são descritos nos registos bibliográficos: Obra, Expressão, Manifestação ou Item (grupo 1). Podem também representar os responsáveis pela criação destes produtos intelectuais: Pessoa ou Colectividade (grupo 2). Este segundo grupo pode ter responsabilidades no conteúdo intelectual ou artístico, na produção física, na difusão ou na guarda dos produtos. As entidades podem ainda representar assuntos das obras: Conceito, Objecto, Evento e Lugar (grupo 3).

Cada uma das entidades tem associado um conjunto de características ou atributos, que são usados pelos utilizadores para formular pesquisas, e interpretar respostas, quando procuram informação acerca de uma entidade em particular.

Elementos essenciais deste modelo são também as relações. As relações são formas de estabelecer vínculos entre diversas entidades, e ajudam o utilizador a navegar pelo universo que está representado numa bibliografia, catálogo, etc.

Ou seja, o utilizador formula a pergunta através de um ou mais atributos, e encontra a entidade que procura. As relações reflectidas no registo bibliográfico servem para dar informação adicional, que ajuda o utilizador a estabelecer conexões entre a entidade encontrada e outras que com ela se relacionam, de alguma forma, e que também podem ser do seu interesse.

O modelo relacional publicado em 1998 diverge radicalmente da concepção dos catálogos bibliográficos automatizados, que têm por base registos bibliográficos em formato MARC, que descrevem manifestações, numa estrutura linear que praticamente não apresenta hierarquias ou relações. Oferece-nos uma representação

do universo bibliográfico bastante mais rica e complexa, mas a sua implementação nas práticas bibliográficas implica a superação de desafios consideráveis.

Impacto potencial do modelo FRBR

O processo de informatização massiva dos catálogos que ocorreu nos anos 80 conduziu a perdas significativas na integridade das estruturas bibliográficas tradicionais e à ausência de apresentações hierárquicas e relacionais (TEIXEIRA, 2010, 51). Em resultado das condicionantes económicas e do volume de trabalho das conversões retrospectivas, os antigos catálogos manuais, altamente estruturados, foram transformados em listas de registos não relacionados. A forma mais habitual de apresentar os resultados das pesquisas nos catálogos tradicionais é listar os resultados ordenados por algoritmos de relevância simples (muitas vezes percebe-se que os resultados são apresentados apenas pela ordem dos registos na base de dados). A informação aparece assim em longas listagens de registos bibliográficos, sem uma organização com valor extra que ajude o utilizador na sua tarefa, restando-lhe assim apenas percorrer essa lista ou tentar refinar a interrogação.

Os formatos MARC foram originalmente concebidos como um sistema linear, com toda a informação sobre determinado recurso bibliográfico armazenado num registo bibliográfico, dividido em campos ou subcampos. A estrutura linear das bases actuais não permite portanto oferecer uma visualização estruturada e colocada dos resultados da pesquisa.

O modelo FRBR tem o potencial de conseguir proceder à recuperação das relações perdidas dos antigos catálogos manuais, transformando as actuais longas listas alfabéticas de resultados em visualizações OPAC altamente estruturadas, agrupadas de acordo com as entidades FRBR. As vantagens incluem uma pesquisa mais fácil, o *clustering*, e a presença de informação relativa às relações bibliográficas. O utilizador terá assim uma experiência muito mais positiva quando usar o catálogo em linha.

Uma concreta implementação do modelo que melhor potencie estas vantagens depende no entanto da superação de alguns desafios importantes, que se colocam neste momento.

O primeiro desafio consiste na necessidade de realizar trabalho adicional de reflexão, para transformar as definições conceptuais das entidades FRBR em directrizes de implementação mais claras e específicas (BENNETT et al., 2003). Há algumas áreas em particular, identificadas pelos investigadores, em que esta necessidade se faz sentir com maior premência, porque experiências de implementação já efectuadas mostram a necessidade de orientações mais concretas. São os casos, entre outros, da abordagem às publicações em série, às relações todo-parte, ou a questão da definição operacional da fronteira entre expressões diferentes de uma mesma obra.

Outro desafio que se coloca neste momento prende-se com o desenvolvimento de ferramentas que ofereçam suporte aos FRBR. Por um lado os sistemas informáticos têm de ser redesenhados, para estarem em consonância com esta nova forma de descrever e

organizar a informação. Por outro lado, é necessário disponibilizar ferramentas normativas que enquadrem a sua aplicação. Este caminho tem vindo a ser percorrido paulatinamente nos últimos anos, com a integração progressiva dos princípios e conceitos do modelo nos instrumentos normativos mais recentes. Os *Princípios Internacionais de Catalogação* e a *ISBD consolidated edition* reflectem já muitas destas preocupações de integração. A recente publicação da zona 0 da ISBD é reconhecidamente um esforço de aproximação ao conceito de expressão, por exemplo. Ao nível dos vários códigos de catalogação nacionais, esse caminho tem também vindo a ser percorrido, com a publicação das *Resource Description and Access (RDA)* em substituição das AACR2, ou de outros códigos, como as novas *Regole italiane di catalogazione (REICAT)*.

Mas não é, de todo, fácil fazer a passagem dos sistemas informáticos tradicionais para o novo desenho conceptual. Há um legado de milhões de registos criados e armazenados pelos sistemas tradicionais. A nossa herança bibliográfica é dos velhos registos MARC, pelo que a passagem ao modelo FRBR só pode ser feita se for encontrada uma forma de eficaz de reestruturar a informação existente em MARC, de acordo com o novo paradigma FRBR. Este processo chama-se correntemente “FRBRização”: um conjunto de operações que consiste em identificar e extrair entidades, atributos e relações FRBR que estão presentes nos registos bibliográficos, e apresentá-los de acordo com o modelo.

CRIAÇÃO DE UM REPOSITÓRIO FRBR PARA THE EUROPEAN LIBRARY

Actualmente, é incontestável que os FRBR são um modelo a ter em conta na abordagem da realidade bibliográfica. Apesar da complexidade da sua implementação, decorrente da necessidade de “eivar” o nível da catalogação, os FRBR podem oferecer ao utilizador significativas vantagens no acesso à informação: uma organização do catálogo mais próxima dos modelos mentais de quem pesquisa, um catálogo com mais potencialidades, mas que permanece simples e apresenta resultados rigorosos. Por isso os conceitos e princípios dos FRBR têm vindo a ser gradualmente absorvidos pelos desenvolvimentos normativos mais recentes.

Acompanhando as experiências de FRBRização que se têm vindo a realizar internacionalmente, a Biblioteca Nacional de Portugal, em parceria com o Instituto Superior Técnico, participaram num projecto para testar a relevância do modelo FRBR para a apresentação dos resultados no portal de pesquisa do TEL – The European Library [1].

O TEL foi criado em 2005 como serviço da Conference of the European National Libraries (CENL), tendo como objectivo constituir um ponto de acesso comum às colecções distribuídas de 48 bibliotecas nacionais da Europa em 35 línguas. Actualmente o TEL dá acesso a 150 milhões de entradas, relativas a recursos que podem ser bibliográficos e digitais, e é um serviço gratuito.

O sucesso do TEL tem conduzido à realização de projectos associados, tendo em vista desenvolvimentos e melhorias no serviço, como o caso do projecto TELplus, levado a cabo por um consórcio de 26 parceiros, incluindo bibliotecas nacionais e centros de

investigação. Tem como objectivo fornecer serviços e produtos que constituam uma mais-valia para o TEL, especialmente no que diz respeito ao desempenho da arquitectura do serviço, à melhoria das funcionalidades do acesso multilingue, e ao acesso a textos integrais criados por processos de OCR. Um das linhas de acção é a da investigação de soluções que suportem serviços alternativos de pesquisa e navegação no TEL, de acordo com o paradigma FRBR.

É nesta linha de acção - investigar possíveis melhorias para a pesquisa e acessibilidade dos conteúdos do TEL - que se enquadrou a Task 3.3, tendo como objectivo a criação de um repositório e interface de pesquisa FRBR para a European Library. Este projecto foi desenvolvido pelo Instituto Superior Técnico, a Biblioteca Nacional e o TEL Office. Para ele contribuíram também várias outras bibliotecas nacionais europeias, na qualidade de consultores. O objectivo concreto foi desenvolver um sistema de apoio ao OPAC tradicional do TEL, que oferecesse uma nova funcionalidade de apresentação dos resultados de pesquisa agrupados com inspiração nos princípios FRBR, alternativa à visualização corrente.

Considerando que os conteúdos deste serviço atingem actualmente 150 milhões de entradas no catálogo em 35 línguas diferentes, pretendemos demonstrar que o modelo FRBR pode melhorar significativamente a usabilidade do catálogo, através do agrupamento lógico (*cluster*) dos resultados sob os conceitos de “obra”, “expressão”, “manifestação” e “item”. Um serviço alinhado com os princípios FRBR no TEL reduzirá significativamente o número de resultados da pesquisa, fazendo a sua colocação de forma lógica e imediatamente perceptível para o utilizador. Dando o exemplo de um autor consagrado, como Steinbeck, os resultados da pesquisa do seu romance *The grapes of wrath* serão agrupados num único resultado, correspondente ao conceito de “obra”, na qual o utilizador pode depois navegar, em vez de ser forçado a percorrer cada colecção de cada biblioteca nacional para aceder aos resultados.

O “estado da arte”: revisão de literatura sobre experiências FRBR

A implementação do modelo FRBR é um trabalho complexo, para o qual não existem muitas aplicações já desenvolvidas e utilizáveis de forma directa. A nossa primeira tarefa foi, por isso, estudar projectos, serviços e tecnologia relevantes, que reutilizam a informação bibliográfica tradicional para construir repositórios, ou sistemas OPAC associados, alinhados com os princípios dos FRBR. Esta análise do “estado da arte” fundamentou as decisões relativas ao “desenho” da solução técnica que foi posteriormente adoptada (TEIXEIRA et al., 2008).

Verificámos que foram já testadas várias implementações do modelo FRBR, mas geralmente com colecções pequenas. O nosso projecto de implementação difere significativamente dos que estudámos em três aspectos: lida com um conjunto de informação muito grande, proveniente de diversas bibliotecas e de diferentes catálogos; lida com registos bibliográficos tal como eles existem, em vários formatos MARC criados com base em diferentes regras de catalogação; e desenvolve-se num contexto

multilingue. O contexto do TEL é portanto extremamente heterogéneo no formato dos dados e nas práticas bibliográficas (baseadas em diferentes códigos de catalogação nacionais ou nas AACR2, todos anteriores à publicação dos FRBR). A diversidade linguística é também um factor de complexidade, que se verifica quer nos recursos descritos, quer na língua da catalogação.

No que diz respeito à identificação de ferramentas informáticas de apoio à FRBRização já desenvolvidas, impõe-se destacar a ferramenta de conversão para FRBR da BIBSYS, desenvolvida por Trond Aalberg, por ter sido bastante relevante para o nosso trabalho (AALBERG, 2007). Esta aplicação é um auxiliar no processo de extracção das entidades e relações FRBR a partir da informação bibliográfica existente, estruturada em formato MARC. Foi desenvolvida e aplicada a 4 milhões de registos da base bibliográfica BIBSYS – o catálogo colectivo da Noruega. Apesar de algumas questões por resolver, relacionadas com informação inconsistente e escalabilidade, esta ferramenta provou ser valiosa na criação de visualizações FRBRizadas para os catálogos.

No que se refere às dificuldades de implementação dos FRBR em catálogos bibliográficos, referidas recorrentemente, a revisão de literatura permitiu-nos elencar um conjunto de dificuldades e também alguns novos desafios, identificados em projectos anteriores, e que sabíamos desde logo que iríamos enfrentar no nosso trabalho.

Muitas das questões identificadas prendem-se com limitações decorrentes da utilização das estruturas de dados em MARC. Essas limitações decorrem da linearidade das estruturas, mas sobretudo da não exploração das possibilidades oferecidas pelos formatos (ausência de conteúdos relacionais e de contexto nos registos bibliográficos).

A estrutura das bases de dados em MARC torna difícil a extracção de entidades FRBR dos registos bibliográficos actuais: é particularmente difícil extrair e separar as quatro entidades do grupo 1 a partir de um registo MARC único. Diversos elementos do registo MARC aplicam-se à obra, expressão, manifestação e item, mas definir explicitamente que um determinado campo ou subcampo se aplica a uma entidade em particular do grupo é uma tarefa bastante complexa. Além disso, estas entidades não são representadas de forma consistente em diferentes bibliotecas, e em alguns casos estão representadas em campos não estruturados (como campos de notas), o que torna difícil o processamento automático. Conseguir fazer estas operações de modo rigoroso – separar o registo MARC em obra, expressão, manifestação e item, e depois ligar registos de diferentes manifestações ao registo correcto da expressão, e registos de diferentes expressões ao registo correcto da obra – é, no entanto, essencial para transformar bases de dados MARC em bases de dados FRBR.

Outra dificuldade apontada de forma sistemática é o facto de os campos de título disponíveis serem muitas vezes insuficientes para identificar a obra: o campo de título uniforme é o mais apropriado para esta identificação, porque permite agrupar expressões de

uma mesma obra, que surgem sob diferentes títulos (em línguas diferentes, por exemplo). No entanto, estatisticamente, só uma pequena percentagem dos registos bibliográficos dos catálogos contém pontos de acesso a partir do título uniforme. No caso da experiência do catálogo BIBSYS, verificou-se que apenas 5% dos seus registos têm título uniforme. No catálogo da Biblioteca Nacional de Portugal, por exemplo, essa percentagem é de 6%. Ou seja, em cerca de 95% dos casos a identificação da obra tem de ser feita a partir do campo de título próprio, que nem sempre é uma boa fonte de identificação: é um campo susceptível de erros tipográficos, pode conter abreviaturas, etc. No caso da obra traduzida em línguas diferentes, se não existir título uniforme, a identificação de que estamos perante expressões de uma mesma obra terá de ser feita a partir de um campo de notas, que relaciona o registo catalogado com o registo da obra na língua original.

De uma forma geral, as diversas experiências mostram que a FRBRização pode ser feita, mas que é difícil atingir resultados com um grau elevado de qualidade. A conversão, com qualidade, dos velhos catálogos MARC depende da existência de dados praticamente perfeitos. É por isso que os resultados obtidos em bibliografias nacionais são potencialmente melhores do que em catálogos individuais (PISANSKI et al., 2009). Os erros e inconsistências nas práticas de catalogação dificultam grandemente o processo de FRBRização, e é por isso que alguns autores argumentem que o desafio é construir um algoritmo que seja eficaz com dados que não sejam perfeitos.

Por outro lado, a literatura revista permitiu também identificar novos desafios já lançados entretanto: a necessidade de definir entidades a um nível superior ao das obras (“*superworks*”), que permitam uma melhor colocação de obras relacionadas nos catálogos e bases de dados; e também a necessidade de especificar a natureza das relações entre entidades FRBR. As regras de catalogação prevêm códigos de função para especificar a natureza dessas relações, e os registos MARC estão preparados para acomodar esta informação. No entanto, a inclusão das funções dos responsáveis foi tornada opcional pelas AACR2, e consequentemente a sua utilização caiu em desuso em diversos países, fruto da pressão para reduzir os custos de catalogação. Uma decisão que se reconhece, retrospectivamente, não ter sido a mais acertada, porque torna impossível a identificação da natureza das relações para os futuros utilizadores. O desafio lançado é o de reverter esta situação, de modo a permitir flexibilidade na manipulação dos registos e oferecer ao utilizador melhor informação na navegação pelos catálogos (TILLET, 2004, 3).

Aspectos metodológicos: controlo dos resultados

O desenho da experiência desenvolvida teve em consideração a necessidade de avaliação rigorosa dos resultados, uma vez que o contexto heterogéneo do TEL (em línguas, formatos, regras e práticas de catalogação) coloca dificuldades não enfrentadas em trabalhos realizados anteriormente.

Foi por isso seleccionada uma colecção com características específicas, que constitui o “grupo de controlo” durante o trabalho experimental no protótipo,

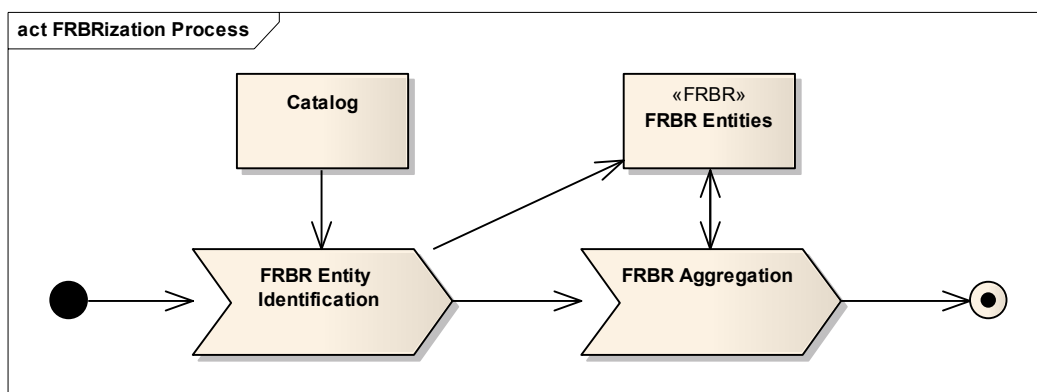


Figura 1: Fluxo de trabalho de um processo de FRBRização típico

tornando mais fácil a verificação dos resultados. Esta colecção foi constituída por obras cujos autores foram vencedores do Prémio Nobel da Literatura. A sua relevância advém da existência de numerosas expressões de cada obra, consistindo em traduções para um número de línguas muito alargado. Com a análise desta colecção pretendemos abordar as questões decorrentes da representação de um conjunto de obras do mesmo autor, com múltiplas traduções em diferentes línguas [2].

Para constituir esta colecção, várias bibliotecas nacionais que participam no TEL enviaram os respectivos registos existentes nos seus catálogos, em formatos MARC diversos, codificados em ISO2709, MARCXML ou MarcXchange. O número de registos de cada biblioteca, bem como o total, é apresentado na Tabela 1 e na Tabela 2.

Biblioteca	Registos Nobel
Biblioteca Nacional de Portugal	3.130
British Library	18.838
Royal Library of Belgium	1.084
National Library of Latvia	1.786
National Library of Lithuania	341
Russian State Library	20
National Library of the Czech Republic	2.945
Bibliothèque nationale de France	28.302
German National Library	22.598
National Library of Spain	12.444
National Library of Servia	1.047
	92.535

Tabela 1: Registos enviados por cada biblioteca

Formatos MARC	Registos Nobel
UNIMARC (4 bibliotecas)	31.793
MARC21 (7 bibliotecas)	60.742
	92.535

Tabela 2: Registos enviados, por formato

DESENVOLVIMENTO DO SISTEMA FRBR

O desenvolvimento do sistema assumiu um cenário de duas componentes: FRBRização e criação do interface do utilizador. No final do projecto, estas duas

componentes foram integradas.

FRBRização

No âmbito desta componente ocorre o processo de FRBRização, esquematizado na Figura 1: numa primeira etapa este processo é responsável pela identificação e extracção de entidades, de atributos e das relações FRBR presentes nos registos bibliográficos das bibliotecas. É também responsável, num segundo momento, pela respectiva consolidação, isto é, pela identificação dos casos em que duas entidades aparentemente distintas são, na verdade, uma mesma. Por exemplo, quando o nome de um autor se escreve de forma diferente em diferentes línguas, ou quando ocorrem erros tipográficos no nome de um autor ou num título.

O primeiro momento – de extracção de elementos FRBR – é relativamente simples. No entanto, a consolidação das entidades é um processo bastante complexo, uma vez que o modelo apresenta um nível de abstracção superior ao do registo bibliográfico: os registos bibliográficos encontram-se, conceptualmente, ao nível da manifestação, deixando os níveis da obra e expressão sem uma correspondência directa. Quando lidamos com diversos catálogos em simultâneo a questão torna-se ainda mais complexa, por não poderem ser usados, como elementos auxiliares, os identificadores primários de registos, atribuídos no contexto dos catálogos individuais.

No âmbito do nosso trabalho em concreto, estas duas etapas foram antecidas de outra tarefa prévia: a análise dos registos em presença e dos requisitos necessários para a correcta FRBRização da informação bibliográfica. Esta análise prévia permitiu identificar desde logo alguns problemas, e procurar a melhor abordagem para os ultrapassar. A maior parte deles foi resolvido logo na fase inicial da FRBRização, por se tratar de questões específicas de cada conjunto de registos e de cada prática de catalogação. Esta avaliação e correcção prévias permitiram que praticamente todos os registos provenientes dos diversos catálogos fossem aceites para FRBRização.

Ou seja, no desenvolvimento do nosso sistema, a FRBRização foi concretizada em três fases: uma fase de pré-processamento, em que os registos MARC foram corrigidos e enriquecidos de acordo com os requisitos da FRBRização; uma segunda fase, em que as entidades FRBR foram extraídas dos registos MARC; e uma fase final, que consistiu na sua agregação (fusão das entidades duplicadas).

No final deste processo foi constituído um novo repositório, acessível por um serviço *web*, contendo os registos bibliográficos originais, mas agora organizados de acordo com uma estrutura FRBR.

Interface do utilizador

Na segunda componente do sistema, o portal do TEL utiliza a estrutura FRBR criada no passo anterior para auxiliar o utilizador final na descoberta de informação. Assim, o motor de busca pode oferecer ao utilizador final uma forma alternativa de navegar nos resultados de uma pesquisa, organizada segundo a estrutura e as relações FRBR, como apresentado na Figura 2.

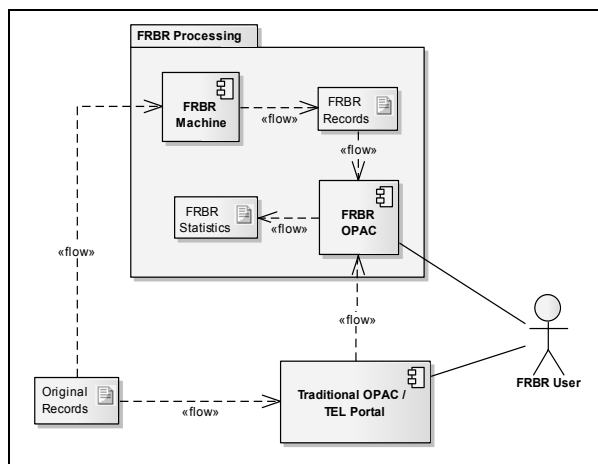


Figura 2: O sistema final

O protótipo foi concebido para avaliação do utilizador,

e não como produto final. Pretendeu-se disponibilizar uma ferramenta que permitisse ao utilizador comparar facilmente os resultados obtidos a partir de uma pesquisa num catálogo ou portal em linha tradicional, com os de uma pesquisa usando o sistema inspirado nos FRBR.

O protótipo, disponível em linha, oferece por isso duas interfaces de pesquisa ao utilizador: “Library search” e “FRBR clustering search”[3]. O utilizador pode assim comparar os resultados da pesquisa clicando nas ligações disponíveis: a primeira opção dá acesso à pesquisa pelo meio “tradicional” do OPAC, na segunda os resultados de pesquisa são apresentados organizados de acordo com alguns princípios do modelo FRBR, nomeadamente o conceito de “obra” e a língua (como atributo da expressão). Os agrupamentos (*clusters*) de registos são ordenados segundo a relevância para a pesquisa efectuada.

Na Figura 3 e na Figura 4 apresentam-se exemplos do resultado de uma pesquisa de acordo com cada uma das opções atrás explicitadas.

Avaliação dos resultados

O protótipo foi publicitado em diversas listas de discussão internacionais, cujos participantes são potenciais interessados nas questões do modelo FRBR ou no TEL. Foi também divulgado na Biblioteca Nacional de Portugal, junto de profissionais e também de utilizadores da biblioteca. No âmbito desta divulgação, os utilizadores do protótipo foram convidados a responder, após a utilização do protótipo, a um questionário de avaliação disponível em linha.

Search Library	FRBR CLUSTERING
... Try this search with the FRBR clustering	
Results: 0 - 20 of 586 for "Jungle Book" ... all 20 Jungle Book Search	
Results for 'Jungle Book':	
<p>Title Het jungleboek The jungle book in the second jungle book illustraties van C. Detmold # # The jungle book in the second jungle book illustraties van C. Detmold # # The jungle book in the second jungle book</p> <p>Author Kipling, Rudyard, 1865-1936</p> <p>Author Reader's Digest</p> <p>Author Detmold, C.</p> <p>Date 1993</p> <p>Summary - Het jungleboek The jungle book in the second jungle book illustraties van C. Detmold # # The jungle book in the second jungle book illustraties</p>	View Detailed Record
<p>Title The jungle book</p> <p>Summary - The jungle book</p>	View Detailed Record
<p>Title The jungle book</p> <p>Summary - The jungle book</p>	View Detailed Record
<p>Title Z indických chrestí The Jungle Book [or rather eight tales from "The Jungle Book" and "The Second Jungle Book"]. Z angličtiny preložila: Margita Pauliny-Tóthová.</p> <p>Author Kipling, Rudyard,</p> <p>Date 1865-1936</p> <p>Summary - Z indických chrestí The Jungle Book [or rather eight tales from "The Jungle Book" and "The Second Jungle Book"]. Z angličtiny preložila: Margita</p>	View Detailed Record
<p>Title Het jungleboek The Jungle Book vertaling : Ernst van Altena # # The Jungle Book vertaling : Ernst van Altena # # The Jungle Book</p> <p>Author Kipling, Rudyard, 1865-1936</p> <p>Author van Altena, Ernst, 1933-1999</p> <p>Author Facet Internationaal</p> <p>Date 1995</p> <p>Summary - Het jungleboek The Jungle Book vertaling : Ernst van Altena # # The Jungle Book vertaling : Ernst van Altena # # The Jungle Book Kipling</p>	View Detailed Record
<p>Title Het jungleboek The jungle book vertaling : Ernst van Altena # # The jungle book vertaling : Ernst van Altena # # The jungle book</p> <p>Author van Altena, Ernst, 1933-1999</p> <p>Author Kipling, Rudyard, 1865-1936</p> <p>Author Facet Internationaal</p> <p>Date 2004</p> <p>Summary - Het jungleboek The jungle book vertaling : Ernst van Altena # # The jungle book vertaling : Ernst van Altena # # The jungle book Kipling</p>	View Detailed Record
<p>Title Jungle-Bogen. Avtoriseret Oversættelse ved A. Halling. Tredie Oplag. [Eight tales from "The Jungle Book" and "The Second Jungle Book."]</p> <p>Author Kipling, Rudyard,</p> <p>Date 1865-1936</p>	View Detailed Record

Figura 3: Resultados de pesquisa em “Library Search”

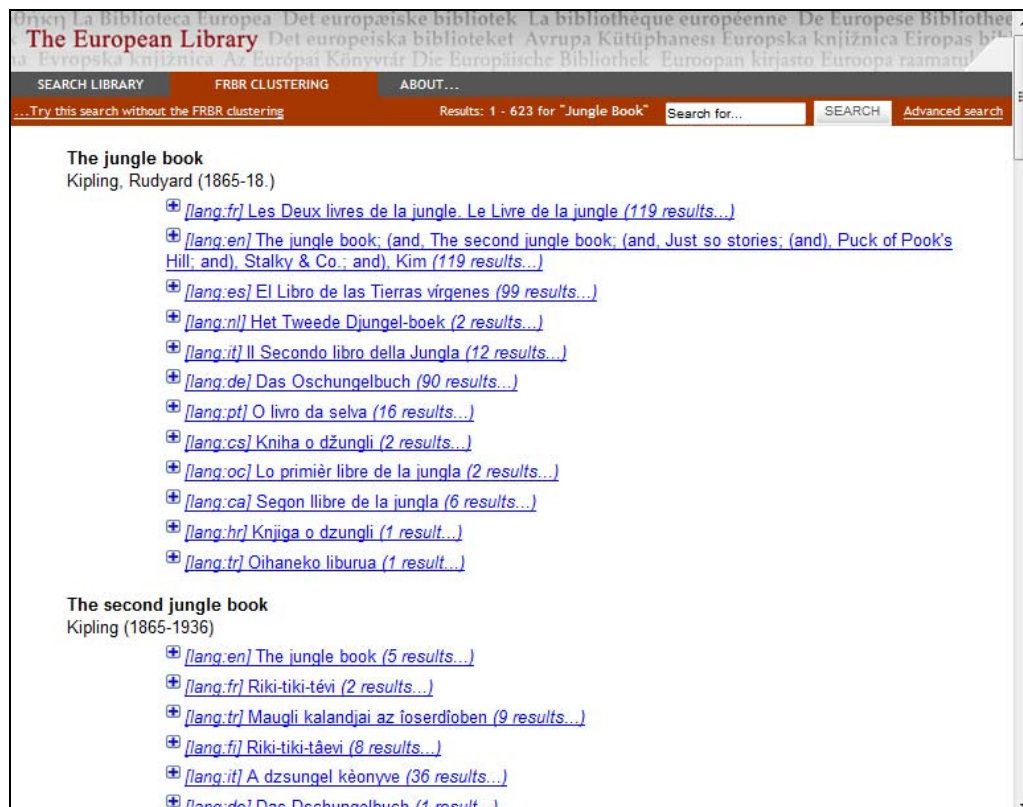


Figura 4: Resultados de pesquisa em “FRBR clustering”

Foi recebido um total de 31 respostas, as quais se encontram resumidas na Tabela 3. O número de respostas obtidas foi suficiente para suportar as conclusões com dados estatisticamente relevantes.

Das respostas recebidas a esse questionário, podemos concluir que a apreciação geral dos utilizadores relativamente ao protótipo foi francamente positiva. A maior parte dos comentários negativos recebidos referia-se a problemas relacionados com a informação original em MARC, pequenos erros de *software*, ou funcionalidades esperadas num sistema final, mas que não foram implementadas, deliberadamente, no protótipo. Podemos por isso deduzir, dos resultados do questionário, que as funcionalidades de “colocação” dos FRBR são implementáveis com a informação

actualmente existente, e que os seus resultados serão de utilidade aos utilizadores da European Library.

CONCLUSÕES E DESAFIOS FUTUROS

A experiência permitiu-nos identificar um conjunto de técnicas de processamento FRBR que pode ser aplicado de forma automática a dados com uma estrutura tradicional em MARC, e disponibilizar resultados satisfatórios para o utilizador, ou seja, resultados que ajudam o utilizador final na recuperação da informação, organizada segundo o paradigma FRBR.

A complexidade do projecto decorreu da necessidade de criar uma estrutura para FRBRização dos registos num universo de informação com características muito diversas, a todos os níveis: na informação a tratar, nos

Questão	Sim, sempre		Sim, por vezes		Não		Sem resposta
	%	# resp.	%	# resp.	%	# resp.	
1 - A agregação dos resultados é útil para encontrar os recursos relevantes para a sua pesquisa?	50,0%	15	40,0%	12	10,0%	3	1
2 - A agregação dos resultados é útil para descobrir recursos relevantes para além dos que procurava?	36,7%	11	53,3%	16	10,0%	3	1
3 - A agregação dos resultados é útil para descobrir recursos em diferentes línguas?	67,7%	21	25,8%	8	6,5%	2	0
4 - A agregação dos resultados é útil para encontrar a primeira publicação de um recurso?	16,1%	5	54,8%	17	29,0%	9	0
5 - A agregação dos resultados é útil para descobrir quais as bibliotecas que têm o mesmo recurso?	41,9%	13	48,4%	15	9,7%	3	0

Tabela 3: Resultados da avaliação do protótipo

formatos usados para a registar, e nas práticas de catalogação seguidas, mesmo quando é usado um formato comum. Os maiores desafios resultaram, aliás, não da concepção de regras para identificar entidades FRBR num registo num formato MARC específico, mas da dificuldade em estabelecer regras válidas e consistentes com as várias práticas de catalogação, quando é utilizado o mesmo formato MARC.

Esta questão só pode ser ultrapassada por intervenção humana em momentos específicos e cruciais no processo de FRBRização: é hoje largamente reconhecido que a implementação do modelo FRBR eleva a catalogação a um nível semântico mais elevado, mais complexo e mais rico. Para esse fim foi ainda desenvolvida uma interface profissional, representada na Figura 5, que permite a intervenção dos profissionais de informação, quando as regras automáticas se revelam insuficientes ou originam agrupamentos incorrectos.

Uma particularidade do sistema desenvolvido é que as bibliotecas envolvidas neste trabalho podem, em vários momentos ao longo das etapas, avaliar os resultados do processamento FRBR, indicar pontos em que considerem que o processamento pode ser melhorado, e verificar a qualidade global da estrutura FRBR resultante. A aplicação inclui esta funcionalidade, disponibilizando uma interface de navegação para “profissionais”: o bibliotecário pode navegar, verificando e corrigindo os agrupamentos das obras de cada autor, e também verificar a proveniência dos registos e respectivos detalhes. Pode assim melhorar a qualidade da estrutura FRBR resultante. Consideramos que só por si este serviço já é relevante, independentemente dos restantes resultados do projecto.

Apesar dos resultados alcançados serem francamente positivos, há ainda muito trabalho a ser feito na melhoria geral do processo, especialmente no que diz respeito às obras agregadas e às séries.

Há também aspectos que ainda podem ser melhorados,

nas diversas etapas da FRBRização. Por exemplo, na fase de pré-processamento é possível melhorar alguns aspectos relativos à identificação da natureza das relações, mesmo quando não existem códigos de função, através dos elementos constantes na descrição bibliográfica (“trad.”, “coord.”). Na fase da extracção das entidades, é necessário fazer uma nova avaliação com um universo mais abrangente de registos, especialmente para validar aspectos relacionados com os atributos. A fase de agregação pode também ser ainda melhorada, em especial no caso das entidades “expressão” e nas entidades do grupo 3 (assuntos).

É também importante realçar que os resultados foram em grande medida condicionados pela qualidade dos registos bibliográficos. Isto verificou-se essencialmente devido às características diferenciadas das práticas de catalogação adoptadas por cada agência bibliográfica nacional. Esta disparidade nas práticas só poderá vir a ser significativamente minorada quando o modelo FRBR dispuser de um suporte normativo mais consistente e específico. Tal permitirá ultrapassar muitas das dificuldades com que nos batemos actualmente, de uma forma muito mais consistente. O caso mais óbvio que nos ocorre é o do título original, cuja informação terá de estar presente quando se cataloga cada uma das expressões de uma obra, seja num campo de título uniforme ou num registo de autoridade.

Estamos em crer que a integração do paradigma FRBR nas práticas biblioteconómicas é uma evolução que pode ser feita em várias frentes, a melhor das quais terá de ser encontrada em conjunto pela comunidade de profissionais, sendo certo que se trata de um caminho que temos vantagem em percorrer o mais rapidamente possível. Julgamos que esta experiência terá contribuído para mostrar não só a relevância desse caminho, mas também a sua praticabilidade.

NOTAS

[1] <http://www.theeuropeanlibrary.org/>

The European Library						
JOIN TITLES						
Author: Saramago, José (The Nobel Prize in Literature: 1998) b. 1922						
Work	Title	Language	Main Authors	Other Authors	Source	Original Record
<input type="checkbox"/>	<input type="checkbox"/> Baltasar and Blimunda /		Saramago, José		bl	XML
	<input type="checkbox"/> Baltazar i blimunda /		Saramago, José		bl	XML
	<input type="checkbox"/> Baltasar & Blimunda	eng	Saramago, José	Pontiero, Giovanni	bnp	XML
	<input type="checkbox"/> Baltasar & Blimunda /		Saramago, José		bl	XML
	<input type="checkbox"/> Baltasar och Blimunda	swe	Saramago, José	Eyre, Marianne	bnp	XML
	<input type="checkbox"/> Baltazar i Blimunda	pol	Saramago, José	Milewska, Elżbieta	bnp	XML
	<input type="checkbox"/> Baltasar a blimunda	cze	Saramago, José	Havliková, Marie	bnp	XML
<input type="checkbox"/>	<input type="checkbox"/> Cuadernos de Lanzarote	por	Saramago, José		bnp	XML
	<input type="checkbox"/> Cuadernos de Lanzarote (1993-1995)	spa	Saramago, José	Naval, Eduardo	bnp	XML
	<input type="checkbox"/> Cuadernos de Lanzarote		Saramago, José		spa	XML
	<input type="checkbox"/> Cuadernos de Lanzarote	por	Saramago, José		bnp	XML
	<input type="checkbox"/> Cuadernos de Lanzarote		Saramago, José		spa	XML
<input type="checkbox"/>	<input type="checkbox"/> Candida Hofer :		Saramago, José		bl	XML
<input type="checkbox"/>	<input type="checkbox"/> Casi un objeto		Saramago, José		spa	XML
	<input type="checkbox"/> Casi un objeto		Saramago, José		spa	XML

Figura 5: Interface de edição para profissionais de informação

- [2] Chegou a ser constituída uma segunda colecção à volta da obra conceptual “Bíblia”, um grupo considerado relevante pelo seu reconhecimento global e pelas relações muito alargadas. Posteriormente, os planos de trabalho sobre esta segunda colecção acabaram por ser postos de parte, dados os desafios que foram surgindo e o calendário apertado do projecto, que terminou no final de 2009. Estamos no entanto convictos de que os resultados obtidos são extensíveis a transferíveis para esta, ou outra qualquer colecção.
- [3] <http://digimap2.ist.utl.pt:8080/lgteFrbr/indexFrbrClustering.jsp>

REFERÊNCIAS

AALBERG, Trond – MARC and FRBR: match or mismatch? In *Cataloguing 2007, Reikjavik, Iceland, 1-2 February 2007: Back to basics – and flying into the future* [on line]. Reikjavík: the National and Univ. Library of Iceland, Nordic Research Library Associations Network (NFBN), Icelandic Library and Information Science Association (Upplysing), the University of Iceland – Department of Library and Information Science, 2007. Disponível em: <URL: <http://www.congress.is/cataloguing2007/>>

BENNETT, R. ; Lavoie, B. ; O’Neill, E. – The concept of a work in WorldCat. *Library Collections, Acquisitions and Technical Services*, 27(1), 2003

IFLA. Study Group on the Functional Requirements for Bibliographic Records - *Requisitos funcionais dos registos bibliográficos : relatório final* . trad. Fernanda Campos. Lisboa : Biblioteca Nacional de Portugal, 2008. Disponível em linha, na versão inglesa, em: <URL: www.ifla.org/VII/s13/wgfrbr/finalreport.htm>

PISANSKI, Jan ; Zumer, Maja ; Aalberg, Trond – Frbrisation: towards a bright new future for national bibliographies. In *World Library and Information Congress: 75th IFLA General Conference and Council*. Milan, 2009. Disponível em: <URL: <http://www.ifla.org/files/hq/papers/ifla75/77-pisanski-en.pdf>>

TEIXEIRA, Teresa [et al.] – *Report on FRBR experiments: TELplus deliverable n. D-3.6*. July 2008. 23 p.

TEIXEIRA, Teresa – *Contributions to FRBRization. The library catalogue in a new digital environment: how FRBR can help produce better displays in OPACs*. Lisboa : [s.n.], 2010. Dissert. Mestr. Estudos de Informação e Bibl. Digitais, ISCTE

TILLET, Barbara – *What is FRBR? A Conceptual Model for the Bibliographic Universe*. Washington : Library of Congress, Cataloging Distribution Service, 2004. Disponível em: <URL: <http://www.loc.gov/cds/downloads/FRBR.PDF>>