

XIX encontro nacional
de pesquisa em
ENANCIB ciência da informação

// SUJEITO INFORMACIONAL E AS
PERSPECTIVAS ATUAIS EM CIÊNCIA
DA INFORMAÇÃO. //

22-26
OUTUBRO
2018
LONDRINA/PR



XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018

GT-2 – Organização e Representação do Conhecimento

MODELOS CONCEITUAIS NA CIÊNCIA DA INFORMAÇÃO: UMA REVISÃO DE LITERATURA

Marcos Fragomeni Padron (Universidade de Brasília – UnB, Senado Federal)

Fernando William Cruz (Universidade de Brasília – UnB)

Juliana Rocha de Faria Silva (Instituto Federal de Brasília – IFB)

CONCEPTUAL MODELS IN INFORMATION SCIENCE: A LITERATURE REVIEW

Modalidade da Apresentação: Comunicação Oral

Resumo: O artigo faz uma revisão de literatura sobre os principais modelos conceituais bibliográficos que surgiram desde o final dos anos 1990, com destaque especial aos modelos que seguiram os Requisitos Funcionais para Registros Bibliográficos e culminaram na publicação do Modelo de Referência para Dados Bibliográficos em 2017. É apresentado também o Modelo Conceitual de Referência do Conselho Internacional de Museus e sua extensão para integração com o domínio bibliográfico. O trabalho tem o objetivo de prover uma visão unificada das principais características desses modelos conceituais, obtida a partir do estudo de suas especificações oficiais e de textos complementares selecionados por meio de um levantamento bibliográfico exploratório. Conclui-se que os modelos conceituais são um importante ferramental para lidar com os problemas de organização da informação no contexto da *Web Semântica*, pois (i) ao introduzirem as entidades Obra, Expressão, Manifestação e Item ao tratamento da informação bibliográfica concebem uma visão física do objeto informacional (publicação e exemplares de recursos bibliográficos) e outra abstrata (criação intelectual e suas expressões); (ii) reconhecem a importância entre os relacionamentos permitindo que o modelo possa ser pensado para contextos não bibliográficos; e (iii) são mais genéricos e permitem a integração com outros modelos no atendimento de uma maior gama de aplicações.

Palavras-Chave: Modelos conceituais; FRBR; IFLA LRM; CIDOC CRM; FRBRoo.

Abstract: The article is a literature review on the main bibliographic conceptual models that have arisen since the late 1990s, giving special emphasis to the models that followed the Functional Requirements for Bibliographic Records which culminated in the publication of the Library Reference Model in 2017. It is also presented the International Council of Museums' Conceptual Reference

Model and its extension for integrating with the bibliographic domain. This work aims to provide a unified vision of the main characteristics of those conceptual models, obtained by the analysis of their official specifications and complementary texts selected through an exploratory bibliographic research. It is concluded that conceptual models are an important tool to deal with problems of information organization in the context of the *Semantic Web*, because, (i) by introducing the entities Work, Expression, Manifestation and Item to the treatment of bibliographic information, they provide a physical view of the informational object (publication and copies of bibliographic resources) and an abstract one (intellectual creation and its expressions), (ii) they recognize the importance of relationships, allowing the model to be used in non-bibliographic contexts, and (iii) they are more generic and allow integration with other models of a wider range of applications.

Keywords: Conceptual models; FRBR; IFLA LRM; CIDOC CRM; FRBROO.

1 INTRODUÇÃO

Os avanços das Tecnologias de Comunicação e Informação (TIC) promoveram melhorias tanto na infraestrutura de conexão quanto nas formas de representação de recursos de informação. A grande quantidade de documentos textuais e não textuais disponibilizados nas mais variadas formas provocou mudanças paradigmáticas no tratamento desses objetos. Dessa forma, percebe-se uma tendência em se repensar a arquitetura dos objetos de informação para que possam ser submetidas aos processos de seleção, armazenamento, indexação e recuperação, típicos da área de organização da informação.

Nesse contexto, muitos pesquisadores (COYLE, 2010; SEEMAN e GODDARD, 2015; MURRAY, 2008; MILLER, 2011) vislumbram novas formas de catalogação e trato informacional com base no uso de modelos conceituais e ontologias e, vão mais além, propondo que a descrição de recursos informacionais seja feita com base em considerações como: (i) um processo de registro catalográfico colaborativo; (ii) com priorização das ligações entre os recursos de informação; (iii) realizado não apenas por bibliotecários; e (iv) apoiado por serviços de curadoria digital mais qualificados.

De fato, desde os mais antigos catálogos bibliográficos até os tempos modernos, as ferramentas de organização bibliográfica têm buscado prover aos seus usuários acesso ao acervo e ao agrupamento de materiais relacionados. Desde a década de 1950, foram desenvolvidas teorias para diferenciar uma obra do item ou itens físicos que a incorporam. A distinção entre obra e item foi, pela primeira vez, formalmente apresentada à comunidade de catalogação durante a Conferência Internacional de Paris sobre Princípios Catalográficos de 1961, todavia a publicação da segunda edição das Regras Anglo-Americanas de Catalogação (AACR2) em 1988 continuou a demonstrar confusão em relação à natureza das obras (DICKEY, 2008).

Dentro dessa perspectiva, os modelos conceituais bibliográficos propostos a partir do final da década de 1990 se tornaram um elemento de organização importante face aos desafios impostos pela explosão informacional e aos princípios subjacentes à *Web Semântica*. A partir dessa iniciativa, surgiram desdobramentos que desencadearam em uma série de modelos conceituais bibliográficos, com diferentes características.

Nas próximas sessões é feita uma revisão de literatura com o objetivo de prover uma visão unificada das principais características desses modelos, obtida a partir do estudo de suas especificações em manuais oficiais e de textos complementares selecionados por meio de um levantamento bibliográfico exploratório.

2 FAMÍLIA FRBR

O impulso para a definição dos Requisitos Funcionais para Registros Bibliográficos (FRBR) surgiu no Seminário de Catalogação de Estocolmo em 1990 patrocinado pela Federação Internacional de Associações e Instituições Bibliográficas (IFLA). De 1992 a 1997 o Grupo de Estudos da IFLA desenvolveu um modelo Entidade-Relacionamento (ER) que representa uma visão geral do universo bibliográfico, com o objetivo de ser independente de código de catalogação ou de detalhes de implementação. O relatório final do FRBR (IFLA, 1998) contém: *(i)* a especificação do modelo conceitual (entidades, relacionamentos e atributos); *(ii)* tarefas de usuário relacionadas a recursos bibliográficos descritos em catálogos ou outras ferramentas bibliográficas; e *(iii)* uma proposta de registro bibliográfico de nível nacional para todos os tipos de materiais.

A família de modelos FRBR também inclui o modelo Requisitos Funcionais para Dados de Autoridade (FRAD), publicado em 2009, e o modelo Requisitos Funcionais para Dados de Autoridade de Assunto (FRSAD). O modelo conceitual FRAD é voltado para entidades descritas em registros de autoridade, provimento de uma estrutura para a análise de requisitos funcionais para o tipo de dados de autoridade requerido para apoiar a tarefa de controle de autoridades e compartilhamento internacional de dados de autoridade. O modelo tem foco nos dados, independentemente de como são agrupados (p. ex., em registros de autoridade) (IFLA, 2013).

Por outro lado, o modelo conceitual FRSAD endereça questões relacionadas a dados de autoridade de assunto. Seu principal propósito é prover uma estrutura para o entendimento

compartilhado e explícito sobre dados, registros e arquivos de autoridade de assunto (IFLA, 2010).

Os três modelos adotaram diferentes pontos de vista e propuseram diferentes soluções para questões comuns entre eles, mesmo que tenham sido desenvolvidos utilizando a mesma estrutura de modelagem Entidade-Relacionamento. A tentativa de adotar os três modelos em um mesmo sistema requereria a solução de questões complexas e de forma *ad hoc*. Ficou claro que seria necessário combinar ou consolidar a família FRBR em um modelo único e coerente para esclarecer o entendimento do modelo geral e remover as barreiras à sua adoção (IFLA, 2017, p. 5).

O contexto de utilização dos modelos evoluiu desde a publicação do modelo FRBR em 1998 e os Dados Ligados se tornaram parte do panorama bibliográfico, requerendo maior estruturação e clareza para que se pudesse aproveitar as técnicas da *Web Semântica* para navegação nos dados bibliográficos (RIVA, 2016b).

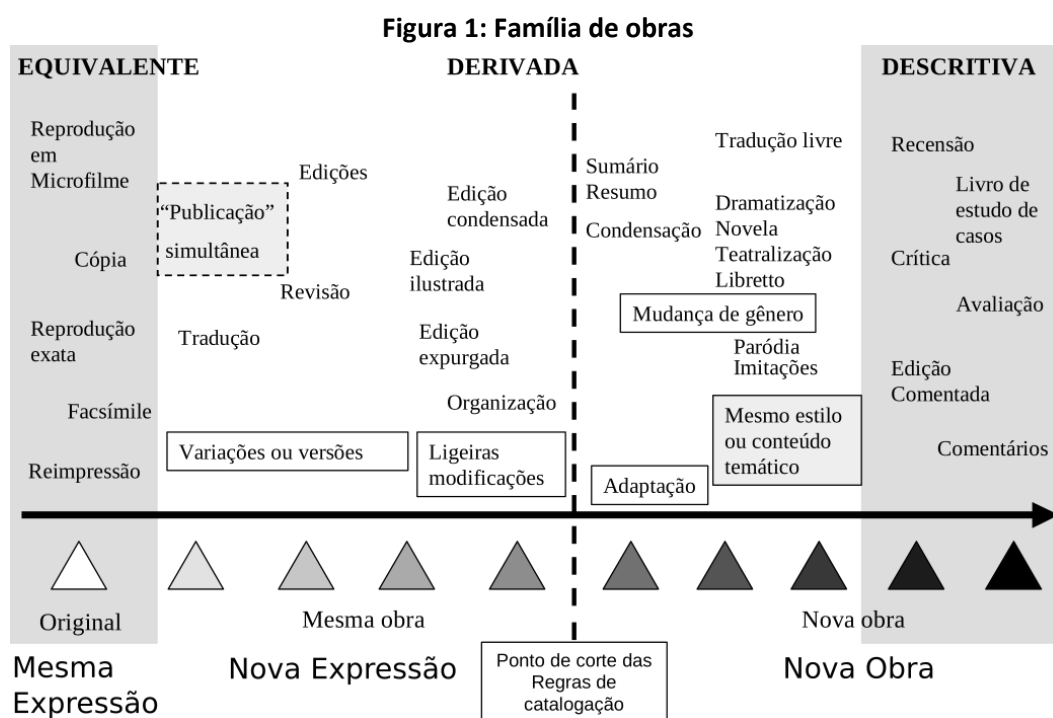
Em agosto de 2017, foi aprovado pelo comitê de padrões da IFLA o relatório final do Modelo de Referência para Dados Bibliográficos (IFLA LRM), que se propôs a ser um modelo conceitual único apoiado em uma estrutura Entidade-Relacionamento de alto nível que cobre todos os aspectos do universo bibliográfico unindo os modelos FRBR, FRAD e FRASAD (RIVA, 2016b).

O modelo FRBR possui dez entidades divididas em três grupos: (i) o grupo 1 compreende os produtos intelectuais ou artísticos. Suas entidades são conhecidas pela sigla OEMI (Obra, Expressão, Manifestação e Item); (ii) o grupo 2 compreende as entidades responsáveis pelo conteúdo intelectual ou artístico, pela produção física e disseminação, ou pela guarda das entidades do primeiro grupo (Pessoa e Entidade Coletiva); e (iii) o terceiro grupo, que compreende um conjunto adicional de entidades que servem como assunto de obras (Conceito, Objeto, Evento e Lugar).

O FRBR oferece uma nova perspectiva sobre a estrutura e os relacionamentos dos registros bibliográficos e de autoridade, além de estabelecer um vocabulário mais preciso para as regras de catalogação (TILLET, 2003). De uma maneira simplificada, uma obra é entendida como uma criação intelectual ou artística de natureza distinta. Uma expressão refere-se à realização intelectual ou artística de uma obra na forma de notação alfanumérica, musical ou coreográfica, imagem, movimento, ou qualquer combinação dessas formas. Por sua vez, a

manifestação é a incorporação física de uma expressão de uma obra e o item é um exemplar de uma manifestação.

Uma obra é realizada por meio de uma ou várias expressões, uma expressão é incorporada em uma ou mais manifestações, e uma manifestação é exemplificada por um ou mais itens. Esta leitura dos relacionamentos primários das entidades OEMI se alinha a uma visão de usuário. Uma leitura no sentido inverso se adéqua mais à visão do catalogador. Observa-se que o único relacionamento primário com cardinalidade múltipla nos dois sentidos ocorre entre as entidades Expressão e Manifestação, indicando que uma manifestação pode incorporar mais de uma expressão. Tillett (2003) enfatiza a importância de outros relacionamentos entre as entidades do grupo 1, além dos relacionamentos primários apresentados anteriormente. Esses relacionamentos também permitem o agrupamento de itens relacionados e a navegação em uma rede, por vezes complexa, do universo bibliográfico. A autora utiliza a Figura 1 para descrever os relacionamentos de conteúdo como um contínuo de obras, expressões, manifestações e itens.



Fonte: Tillett (2001 apud TILLET, 2003, p. 4).

O eixo horizontal indica o quanto o conteúdo de uma obra derivada se diferencia do conteúdo da obra original. O FRBR, assim como o LRM, não especifica regras para a delimitação de fronteiras entre as obras, deixando essa responsabilidade para as regras de

catalogação¹. O diagrama apresenta uma linha de corte tracejada que representa uma possível fronteira. De acordo com essa visão, uma edição condensada ainda seria considerada a mesma obra, ao passo que uma adaptação já seria considerada uma nova obra derivada.

Relacionamentos do tipo todo-parte e parte-parte também são endereçados pelo FRBR. Para Tillett (2003), relacionamentos do tipo todo-parte têm especial importância para o controle bibliográfico de recursos digitais, citando o exemplo de um sítio *Web* que pode ser visto como um “todo” composto de seus componentes, tratados como “partes”. Como exemplos de relacionamento parte-parte pode-se citar os relacionamentos de “sequência”, “suplemento” e “apêndice”. O nível de granularidade para a catalogação desses recursos deve ser feito pelo catalogador, com base em políticas locais que reflitam as necessidades de seus usuários.

As entidades OEMI e seus relacionamentos primários foram mantidos no modelo IFLA LRM e permanecem sendo o núcleo do modelo. Entretanto, ao prover um nível comum de conhecimento dos três modelos, o LRM apresenta um maior nível de abstração. O LRM é um modelo lógico com foco na organização estrutural e, por essa razão, está mais longe da implementação que os outros modelos. Ao mesmo tempo, ele se adianta ao adotar a forma necessária para sua utilização em aplicações de dados abertos ligados (*Linked Open Data*) (RODRÍGUEZ *et al.*, 2017).

O LRM objetiva tornar explícitos os princípios gerais que governam a estrutura lógica da informação bibliográfica sem fazer pressuposições sobre como os dados serão armazenados em uma aplicação em particular. Os dados bibliográficos são tratados pelo LRM de uma maneira ampla, sem distinção entre dados armazenados em registros bibliográficos, de acervo e dados armazenados em registros de autoridades. Para seu propósito, todos esses dados são tratados genericamente como informação bibliográfica (IFLA, 2017, p. 9).

O escopo do modelo é restrito às tarefas de usuário (encontrar, identificar, selecionar, obter e explorar), definidas do ponto de vista do usuário final e suas necessidades. Metadados administrativos estão fora do escopo e metadados relacionados aos direitos de uso só estão no escopo caso estejam relacionados diretamente à tarefa de obtenção do recurso pelo usuário.

¹ Acredita-se que o LRM ao prover um modelo genérico que atenda a todos os tipos de recursos bibliográficos não poderia definir critérios fixos para o estabelecimento de fronteiras entre as entidades Obra e Expressão sem se especializar nos diferentes tipos de recursos informacionais, e, por esse motivo, deixa a responsabilidade para as aplicações e regras de catalogação.

Uma estrutura de modelagem Entidade-Relacionamento Estendida (EER) foi utilizada para a confecção do modelo, que difere da estrutura Entidade-Relacionamento usada pelos modelos anteriores, principalmente pela introdução do conceito de especialização de classes, relacionamentos e atributos.

O LRM não se atém a nenhum tipo de recurso especificamente, procurando assim revelar as similaridades e a estrutura subjacente aos recursos bibliográficos. São utilizados, portanto, termos e definições aplicáveis de maneira genérica a todos os tipos de recurso. Ainda assim, são apresentados no modelo alguns exemplos de atributos relacionados a tipos de recursos específicos.

Definido na forma de um modelo conceitual de alto nível², o IFLA LRM deve ser utilizado como guia ou base para a formulação de regras de catalogação e para a implementação de sistemas bibliográficos. Qualquer aplicação prática terá que determinar o nível requerido de precisão, podendo expandir o contexto do modelo ou até omitir alguns de seus elementos. Exige-se, entretanto, que a aplicação respeite ao menos os relacionamentos essenciais entre as entidades OEMI, consideradas como o núcleo do modelo, para que ainda seja considerada uma implementação do IFLA LRM (IFLA, 2017, p. 10).

Ao todo, o LRM é composto por 11 entidades, 37 atributos e 36 relacionamentos. Cada elemento do modelo recebe uma identificação única para uma referência não ambígua, com utilização do prefixo “LRM” seguido de uma letra indicativa do tipo do elemento (‘E’ para entidades, ‘A’ para atributos e ‘R’ para relacionamentos) e de um número sequencial (p. ex., LRM-E2 e LRM-R12).

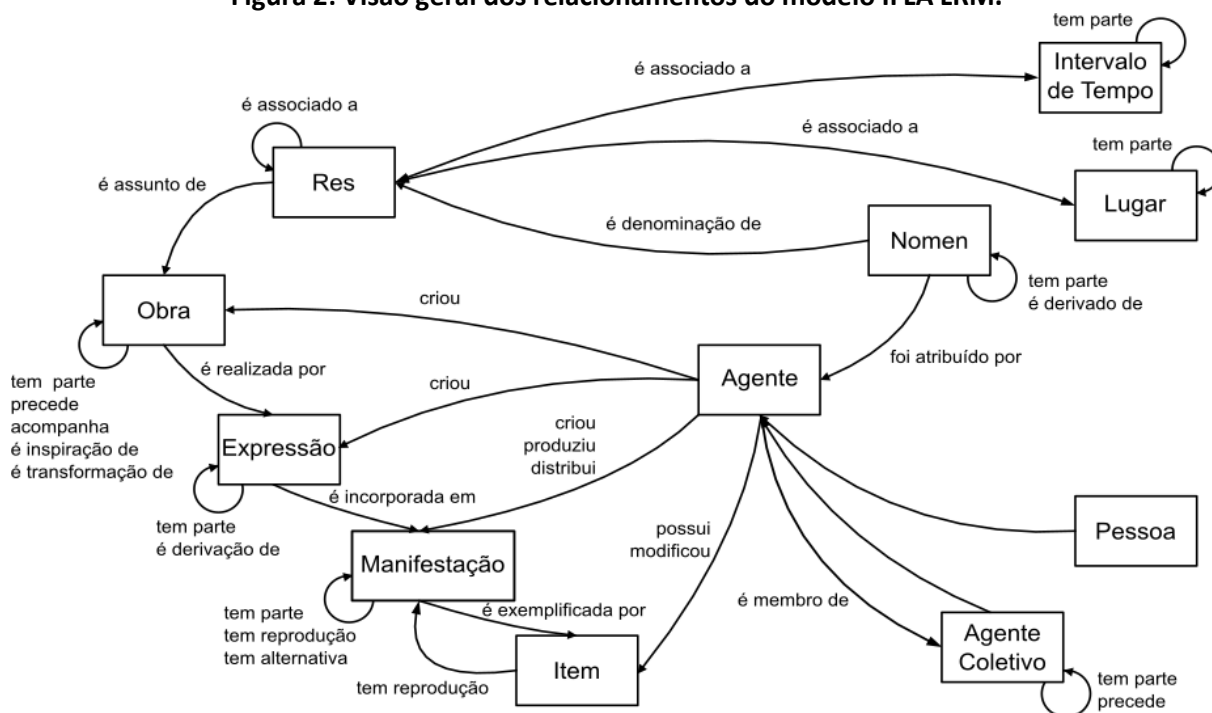
O modelo prevê vários mecanismos de extensão que provavelmente serão necessários em qualquer implementação. Por exemplo, a definição do atributo LRM-E1-A1 (Categoria da entidade *Res*) permite às implementações criarem subclasses que acharem úteis a qualquer entidade e atributos especializados podem ser adicionados a todas as entidades. Além disso, várias das entidades são definidas em um nível geral, novamente com a intenção de serem especializadas durante a implementação (IFLA, 2017, p. 10).

² Um modelo conceitual de alto nível descreve conceitos do domínio em um nível de generalidade que normalmente não é suficiente para a sua aplicação direta. Por exemplo, o relacionamento ‘LRM-6 Expressão foi criada por Agente’ não indica a responsabilidade do Agente em relação à Expressão. Uma aplicação poderia criar um relacionamento ‘Expressão foi traduzida por Agente’, que especializa o relacionamento LRM-6, para indicar o tradutor de uma Expressão.

No caso de atributos, é utilizado como prefixo o identificador da entidade, antes da letra 'A', e a sequência numérica é reiniciada para cada entidade (p. ex., LRM-E2-A1). Cada elemento também recebe um nome curto, porém o significado de cada elemento é definido nas notas de escopo. Para cada elemento é apresentado um quadro contendo: identificador, nome, definição, notas de escopo e exemplos. Para classes, são apresentadas também algumas restrições. Para os atributos, é identificada a classe à qual pertence. Para cada relacionamento, são definidos ainda as classes do domínio e da imagem (origem e destino), sua cardinalidade e o nome no sentido inverso. Vale observar que as 11 entidades do modelo são apresentadas numa estrutura hierárquica e são relacionadas aos diversos conceitos provenientes dos modelos tradicionais do FRBR, como ilustrado na Figura 2.

No FRBR, as entidades Expressão, Manifestação e Item são definidas utilizando outras entidades do grupo 1 em suas definições, de modo que as definições das quatro entidades dependem cada uma das outras. No LRM, as definições dessas três entidades buscam permanecer independentes, sem referenciar uma a outra. Por outro lado, a função da entidade obra é ser um ponto de agrupamento daquelas expressões e manifestações relacionadas que, aparentemente, deveriam estar juntas para a sua recuperação (RIVA, 2016b)

Figura 2: Visão geral dos relacionamentos do modelo IFLA LRM.



Fonte: IFLA (2017, p. 86)

Uma obra não existe independentemente de suas expressões, e só poderá vir a existir quando sua primeira expressão é criada. Uma visão geral das classes e relacionamentos do modelo IFLA LRM é apresentada na Figura 2.

A partir dessa figura, pode-se ressaltar algumas mudanças do LRM em relação aos modelos anteriores: (i) no LRM, a entidade *Res* substitui a entidade *Thema* do modelo FRAD; (ii) foi adicionada a entidade Agente como superclasse de Pessoa e Agente Coletivo, que, por sua vez, representa as entidades Entidade Coletiva e Família do FRBR; (iii) a entidade *Nomen* é uma fusão de *Nomen* do FRAD, Nome do FRAD e de ponto de acesso controlado; e (iv) foi adicionada ainda a entidade Intervalo de Tempo, inexistente nos modelos anteriores (BIANCHINI, 2017).

As entidades Lugar e Intervalo de Tempo podem ser associadas a qualquer outra entidade, seguindo a lógica da *Web Semântica*. Atributos do tipo “local de ...” ou “data de ...” foram generalizados nos relacionamentos entre *Res* e estas entidades, reduzindo de forma drástica a quantidade de atributos do novo modelo (37 no total).

A maioria dos atributos de uma manifestação FRBR consiste em declarações transcritas de um exemplar representativo de toda a manifestação. Todos esses atributos previstos no FRBR são representados no LRM pelo atributo LRM-E4-A4 declaração de manifestação, que deve ser normalmente implementado como uma série de atributos especializados, de acordo com as regras relevantes de catalogação. O Padrão Internacional para Descrição Bibliográfica (ISBD) detalha o consenso internacional sobre as declarações mais relevantes que devem ser incluídas na descrição de recursos. Ao tratar essas declarações de manifestação de uma forma genérica, o LRM evita tomar o lugar do ISBD ou de regras especializadas de catalogação, ao mesmo tempo que provê um ponto de extensão do modelo (BIANCHINI, 2017; RIVA, 2016b).

A definição de alguns elementos-chave do IFLA LRM tem a intenção de facilitar sua operacionalização por meio de uma variedade de códigos de catalogação. Um desses casos é o atributo LRM-E2-A2, atributo da expressão representativa da entidade Obra, que registra os valores desses atributos de expressão considerados essenciais para a caracterização da obra, sem predeterminar o critério a ser usado por um código de catalogação em particular (IFLA, 2017, p. 10). Diferentemente da visão original do FRBR, no qual identificadores, nomes e termos eram modelados como atributos das entidades, o LRM adota a visão do modelo FRAD, uma entidade *Nomen* única para associar as entidades às suas denominações. Qualquer entidade referenciada no universo do discurso é nomeada por meio de pelo menos

um *Nomen*. O modelo LRM define vários atributos para a entidade *Nomen*, tais como *script*, conversão de *script* e esquema, que são básicos para a utilização de um *Nomen* na função de ponto de acesso controlado (RIVA, 2016a). Há uma grande diferença entre a entidade *Nomen* e o atributo *string* do *Nomen*. Essa entidade é muito mais que uma simples *string* e trata de uma distinção que se aplica cotidianamente no âmbito bibliográfico quando se desenvolve a atividade de controle de autoridades. O objetivo do novo modelo lógico é reforçar a estrutura de relacionamentos dos dados bibliográficos, tornando-a mais adaptada à estrutura dos grafos no RDF³ e favorizar a integração de dados bibliográficos na *Web Semântica*. A escolha do nome *Res*, por exemplo, se relaciona ao termo *Thing* em inglês, que representa qualquer entidade definida na *Web Semântica* (BIANCHINI, 2017).

O LRM prevê duas modalidades de representação de um atributo: uma sequência de símbolos ou uma URI⁴. Sendo preferido o estabelecimento de um relacionamento entre entidades, sempre que um atributo possa ser representado como URI. Essa abordagem deriva diretamente da *Web Semântica* e dos Dados Ligados.

3 CIDOC CRM

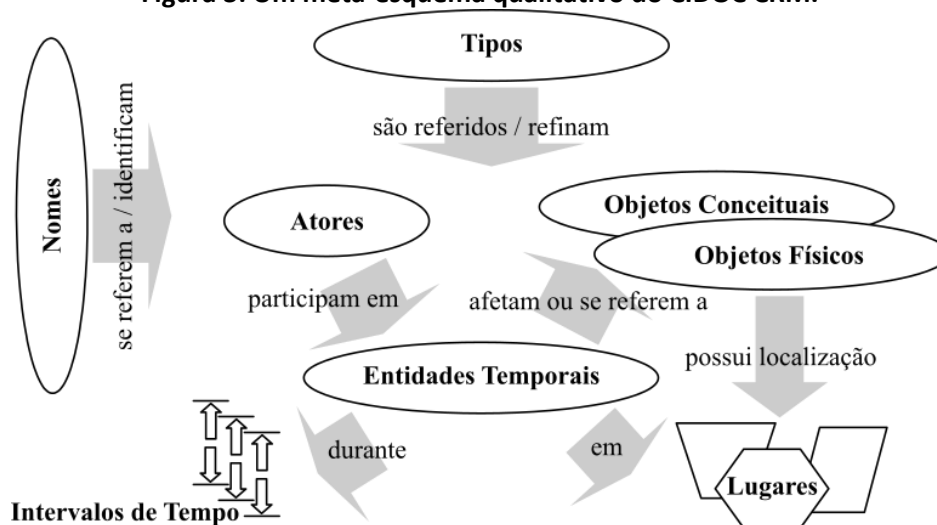
O modelo CIDOC CRM (Modelo Conceitual de Referência) representa uma ontologia para a informação de memória cultural, ou seja, descreve em uma linguagem formal os conceitos implícitos e explícitos relevantes à documentação da memória cultural. Sua principal função é servir como base para a mediação da informação histórica e prover a “cola semântica” necessária para a integração entre diferentes fontes de informação de memória cultural, tais como as publicadas em museus, bibliotecas e arquivos (CIDOC CRM SIG, 2016).

O CRM surgiu a partir do Grupo de Documentação do Comitê Internacional de Documentação (CIDOC) do Conselho Internacional de Museus (ICOM) visando alcançar a interoperabilidade semântica de dados museológicos. Em setembro de 2006, o CRM foi aceito como norma ISO 21127 e em dezembro de 2015 foi publicada a nova versão ISO 21127:2014.

³ O RDF (*Resource Description Framework*) é o modelo básico de dados para a publicação de metadados na *Web Semântica*.

⁴ A URI (*Uniform Resource Identifier*) é o padrão para identificação de recursos na *Web Semântica*.

Figura 3: Um meta-esquema qualitativo do CIDOC CRM.

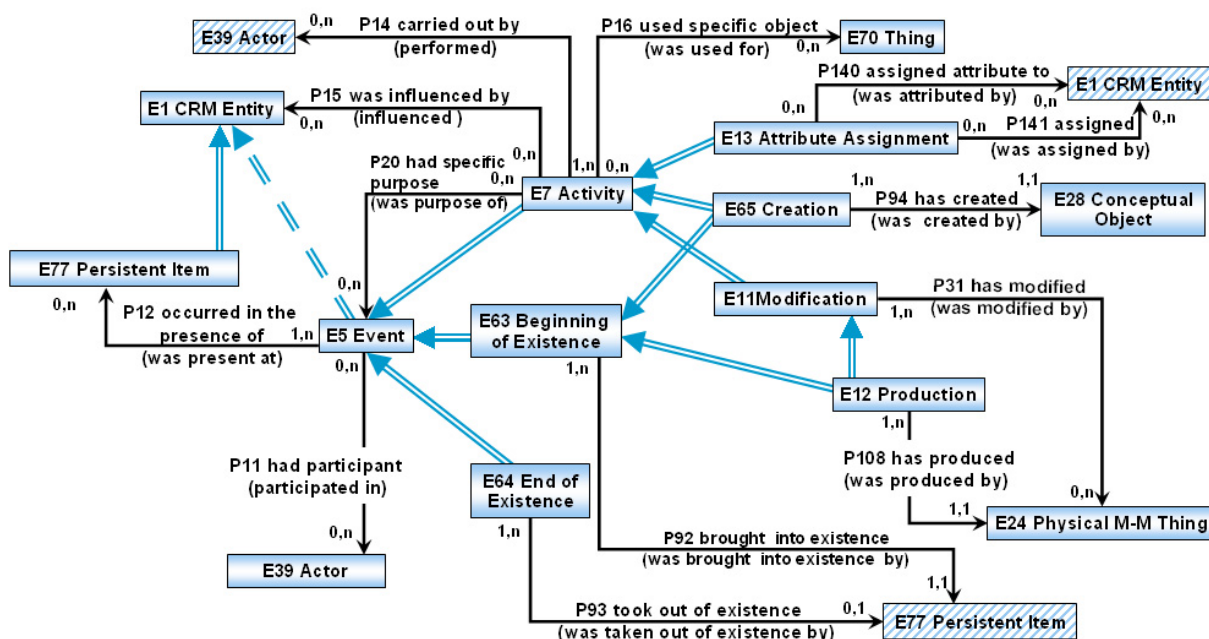


Fonte: Doerr (2003, p. 85).

Doerr (2003) caracteriza o CRM como uma ontologia de alto nível para habilitar a integração de informação para dados de memória cultural e sua correlação com a informação bibliotecária e arquivística. Explica, ainda, que o modelo analisa os conceitos comuns por trás das estruturas de dados e metadados para suportar a transformação, mediação e mesclagem de dados. Segundo Lima (2008), uma das inovações do CRM é a estruturação das informações em torno dos eventos temporais, em oposição à maioria dos modelos de metadados que têm o recurso como objeto central de interesse. Nesta abordagem, os eventos são definidos como entidades que agregam atores, fatos e objetos (físicos e abstratos), localidades e duração de intervalo de tempo.

Em sua versão 5.0.4, o CIDOC CRM contém 86 classes e 137 propriedades únicas (CIDOC CRM SIG, 2011). Essas propriedades puseram as entidades temporais e os eventos em uma posição central, como apresentado na Figura 3. Entidades temporais, atores, objetos conceituais, objetos físicos, lugares e intervalos temporais são organizados em uma hierarquia de tipos, bem como os relacionamentos entre eles, representados por propriedades binárias. Múltiplos nomes e identificadores podem ser atribuídos a todas as entidades do modelo.

Figura 4: Principais eventos do CIDOC CRM.



Fonte: CIDOC CRM SIG (2016).

A Figura 4 contempla uma representação gráfica das classes e propriedades que modelam os principais eventos do modelo. As classes são identificadas por um código iniciado pela letra “E” seguida de seu número e nome em inglês enquanto as propriedades são identificadas por um código iniciado pela letra “P” seguido de seu número, nome e nome para leitura no sentido inverso entre parênteses. As propriedades representam relacionamentos binários entre classes. No diagrama da Figura 4, pode-se observar que a criação de um objeto conceitual, por exemplo a pintura de um quadro ou uma música, é modelada como um evento de criação (E65), que é uma atividade (E7), que foi realizada por (P14) atores (E39), podendo ter sido influenciado por (P15) qualquer outra entidade do modelo (E1), e assim por diante.

Doerr (2003) nos dá um exemplo de como o CIDOC CRM pode ser utilizado para atingir seu objetivo primário de habilitar a troca de informações e a integração entre fontes heterogêneas de informação de memória cultural. A conferência de Yalta, que ocorreu em fevereiro de 1945, é o evento que registra oficialmente o fim da Segunda Guerra Mundial.

Figura 5: Possível registro *Dublin Core* para o Tratado de Yalta.

Type:	<i>Text</i>
Title:	<i>Protocol of Proceedings of Crimea Conference</i>
Title.Subtitle:	<i>II. Declaration of Liberated Europe</i>
Date:	<i>February 11, 1945.</i>
Creator:	<i>The Premier of the Union of Soviet Socialist Republics The Prime Minister of the United Kingdom The President of the United States of America</i>
Publisher:	<i>State Department</i>
Subject:	<i>Postwar Division of Europe and Japan</i>

Fonte: Doerr (2003, p. 78)

Durante esse evento, foi assinado um tratado de divisão pós-guerra da Europa e do Japão e foi tirada uma foto dos líderes aliados, Winston Churchill, Franklin Delano Roosevelt e Josef Stalin. A Figura 5 representa um possível registro *Dublin Core* de uma cópia do Tratado de Yalta no Departamento de Estado dos Estados Unidos e a Figura 6 apresenta um registro semelhante para a fotografia exposta no *Bettmann Archive* em Nova York.

Figura 6: Possível registro *Dublin Core* para a foto dos líderes aliados em Yalta.

Type:	<i>Image</i>
Title:	<i>Allied Leaders at Yalta</i>
Date:	<i>1945</i>
Publisher:	<i>United Press International (UPI)</i>
Source:	<i>The Bettmann Archive</i>
Copyright:	<i>Corbis</i>
References:	<i>Churchill, Roosevelt, Stalin</i>

Fonte: Doerr (2003, p. 79)

Surpreendentemente, os dois registros de metadados não têm nada em comum a não ser a data de 1945, um atributo que praticamente não caracteriza os documentos. A Figura 7 apresenta um registro do *Thesaurus of Geographic Names* (TGN), que poderia servir como peça de integração entre os dois documentos.

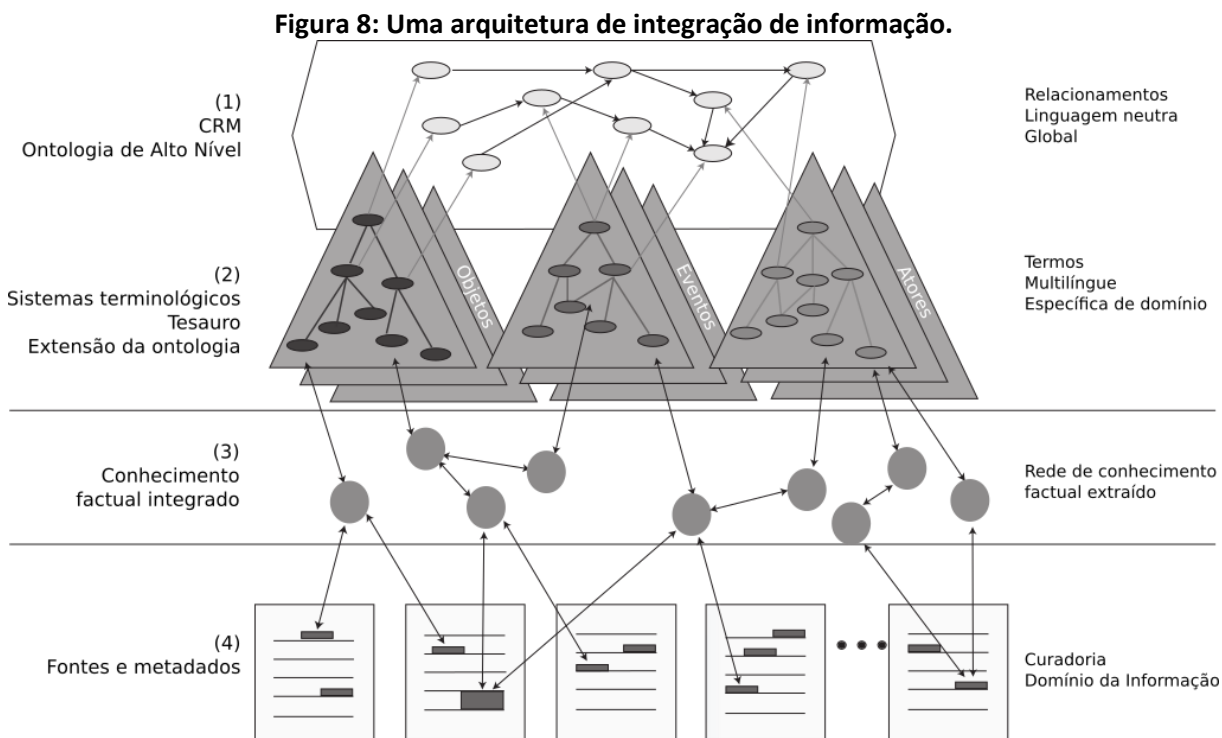
Surgem então os problemas de relacionar “*Crimea*” com “*Krym*” e então com “*Yalta*”, relacionar o “*Premier of the Union of Soviet Socialist Republics*” com “*Joseph Stalin*”, e assim por diante. Os artefatos sozinhos não são suficientes para responder à questão sobre a conferência de Yalta. O problema fundamental é que, para recuperar informação relacionada a um assunto específico, é necessária a integração de diversas fontes de informação. A unificação de vocabulário e de estruturas de dados somente não é suficiente para resolver o problema.

Figura 7: Partes do registro TGN para Yalta. TGN Id: 7012124

Names:	<i>Yalta (C, V), Jalta (C, V)</i>
Types:	<i>inhabited place(C), city (C)</i>
Position:	<i>Lat: 44 30 N, Long: 034 10 E</i>
Hierarchy:	<i>Europe (continent) <- Ukryaina (nation) <- Krym (autonomous republic)</i>
Note:	<i>Located on the south shore of the Crimean Peninsula, site of the conference between Allied powers during World War II in 1945. It is a vacation resort noted for pleasant climate and coastal and mountain scenery; it produces wine, canned fruit, and tobacco products.</i>
Source:	<i>Thesaurus of Geographic Names.</i>

Fonte: Doerr (2003, p. 79)

Na Figura 8, é apresentada uma possível arquitetura que integra o CIDOC CRM, como ontologia de alto nível (1) que provê a semântica para as propriedades dos sistemas terminológicos subordinados (2), e uma camada de conhecimento factual (3) construída a partir de fontes de dados, metadados e conhecimento existente (4) (DOERR, 2003). As relações entre (1) e (2) posicionam na ontologia do CRM os objetos, eventos, atores e lugares do domínio. A camada (3) pode dessa forma estabelecer os relacionamentos entre as entidades do domínio, entre si e com os objetos informacionais a partir de seus metadados (4), refletindo a estrutura de relacionamentos provida pelo CRM (1).



Fonte: Adaptado de Doerr (2003, p. 81)

A descrição de problemas semânticos em linguagem natural é passível de erro, imprecisão e certo grau de ambiguidade apesar do esforço empenhado na precisão da

argumentação (SANTOS, 2016). O CRM é uma ontologia formal, que reduz a imprecisão na descrição dos conceitos do domínio, facilitando uma comunicação com menos ambiguidade. Dessa forma o modelo conceitual permite o entendimento unificado dos conceitos do domínio entre os especialistas e profissionais de memória cultural, cientistas da computação e implementadores de sistemas.

Além da interoperabilidade semântica, o CRM pode ser utilizado também como ontologia para modelagem de aplicações e para o processamento de consultas complexas que necessitam de inferências (SANTOS, 2016).

4 FRBRoo

O FRBRoo é um modelo conceitual que expressa os conceitos do FRBR utilizando a metodologia e estrutura orientada a objetos do CIDOC CRM. Deve ser visto como uma alternativa sobre a conceitualização bibliográfica e não como substituição ao FRBR. O modelo oferece uma visão comum da documentação bibliográfica e museológica, necessária ao desenvolvimento de sistemas integrados de informação desses dois domínios. Seu desenvolvimento é fruto de um esforço conjunto entre a IFLA e o ICOM, por meio do *International FRBR/CIDOC CRM Harmonization Group*, formado em 2003 (RIVA; DOERR; ŽUMER, 2008).

O objetivo do FRBRoo é capturar e representar a semântica da informação bibliográfica e facilitar a integração e o intercâmbio de informações bibliográficas e museológicas (CHOFFÉ; LERESCHE, 2016).

De acordo com Riva, Doerr e Žumer (2008), o trabalho de harmonização dos modelos se tornou uma oportunidade de enriquecimento dos modelos FRBR e CIDOC CRM, como por exemplo: (i) adição ao FRBR da modelagem de tempo e eventos; (ii) esclarecimento da entidade Manifestação; (iii) modelagem explícita de performances e gravações no FRBR; (iv) adição da entidade Obra ao CRM; e (v) adição do processo de atribuição de identificadores ao CRM.

Enquanto o FRBR modela os resultados de processos (p. ex., criação, realização e planejamento), o FRBRoo é focado em processos. Os criadores do FRBRoo modelaram todos os conceitos implícitos do FRBR necessários para justificar a existência de certos atributos no FRBR. O documento do FRBRoo dedica uma seção ao mapeamento entre o FRBR e o FRBRoo, que explica como esses conceitos foram endereçados.

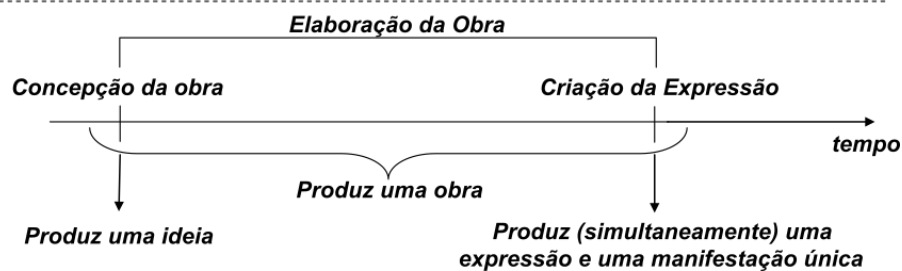
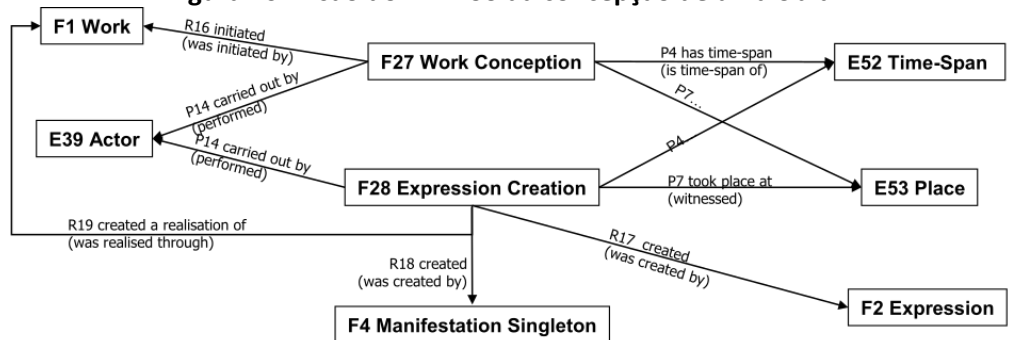
As classes e propriedades do FRBRoo recebem um nome e um identificador construídos de acordo com as convenções do CIDOC CRM. O identificador das classes é formado pela letra “F” seguida de um número sequencial. No caso das propriedades, o identificador é formado pela letra “R” seguida de um número sequencial e recebem um sufixo “i” quando a propriedade é referenciada em seu sentido reverso. Algumas propriedades, ditas Propriedades de Classe, são identificadas pelo prefixo “CLP” seguido de um número sequencial. Estas propriedades têm como domínio a classe F3 Manifestação Tipo de Produto e indicam que todos os exemplares de uma determinada publicação devem apresentar (ou espera-se que apresentem) as características da publicação à qual pertencem (IFLA, 2015, p. 35).

Os nomes das classes e das propriedades são apenas um indicativo do seu significado. O conjunto do identificador seguido do nome existe para caracterizar unicamente estes elementos na documentação. O significado real de uma classe ou propriedade é dado em sua nota de escopo. No FRBRoo, as entidades Obra, Expressão e Manifestação foram divididas em várias classes, cada uma com suas propriedades específicas. A classe Obra (F1) é especializada em Obra Individual (F14), Obra Complexa (F15), Obra Contêiner (F16), Obra Agregadora (F17), Obra Serial (F18), Obra de Publicação (F19), Obra de Performance (F20) e Obra de Gravação (F21).

A classe Expressão (F2), por sua vez, é especializada nas subclasses Expressão Autocontida (F22), Fragmento de Expressão (F23), Expressão de Publicação (F24), Plano de Performance (F25) e Gravação (F26). Em relação às manifestações, o FRBRoo trata separadamente as Manifestações Tipo de Produto (F3) que são produtos de um processo de publicação, das Manifestações Originais (F4) produzidas como objetos únicos.

De acordo com Choffé e Leresche (2016), a Obra Complexa é a classe que melhor representa o conceito de obra do FRBR. A Figura 10 exemplifica o papel central exercido pelas entidades temporais no FRBRoo. As classes Concepção de Obra (F27) e Criação de Expressão (F28) são atividades, exercidas por um ou mais atores (E39) e que ocorrem em um intervalo de tempo (E52) e em um lugar (E53). Inicialmente surge uma ideia durante a concepção da obra, seguida de um processo de criação de expressão, que produz simultaneamente uma expressão e sua primeira manifestação, uma manifestação original (p. ex., manuscrito do autor), ambos realizando a obra.

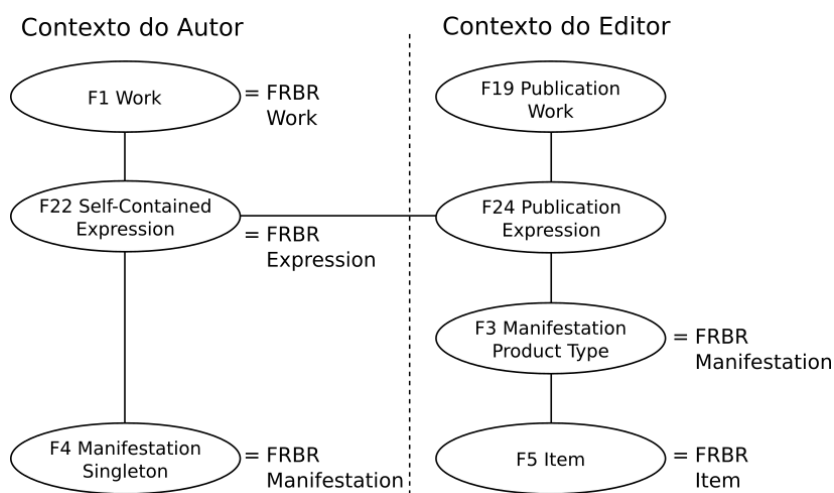
Figura 10: Visão do FRBRoo da concepção de uma obra.



Fonte: IFLA (2015, p. 17)

O FRBRoo torna explícita a contribuição intelectual do editor, não modelada no FRBR. A Figura 11 demonstra como é entendida a relação entre a obra do autor (F1) e a obra do editor (F19). A manifestação tipo de produto (F3) incorpora uma expressão de publicação (F24), que, por sua vez, compreende a expressão autocontida (F22) do autor e realiza a obra do editor.

Figura 11: Relação entre entidades do FRBR e classes do FRBRoo.



Fonte: IFLA (2015, p. 19)

Apesar de reconhecer performances gravadas como novas expressões de obras, o FRBR não apresenta relacionamentos entre a performance e o material utilizado como fonte para a performance. Na visão do FRBRoo, intérpretes, ao participarem de uma performance (F31) seguindo um plano de performance (F25), criam uma expressão, que pode ser

incorporada em uma gravação (F26) durante um evento de gravação (F29) (RIVA; DOERR; ŽUMER, 2008).

A análise de processos bibliográficos no FRBRoo trouxe diversos refinamentos ao CRM que beneficiam a comunidade da Museologia, tais como: *(i)* um melhor entendimento da produção em série de objetos (p. ex., impressão de gravuras); *(ii)* o estabelecimento de relações entre o conteúdo imaterial de criações e suporte físico; e *(iii)* a introdução de um modelo básico de concepção intelectual e derivação aplicável a todos os produtos artísticos.

Um apêndice do FRBRoo apresenta um processo de atribuição de identificadores que pode ser aplicado de forma geral. A atribuição de identificadores inclui a construção de pontos de acesso controlados com base nos nomes associados a entidades específicas, por meio da aplicação de regras específicas de uma agência. Nesse processo, deve-se ressaltar a importância da identificação da manifestação “representativa” de uma expressão e da expressão “representativa” de uma obra. Características das manifestações ou expressões representativas podem ser abstraídas para a expressão ou obra e serem utilizadas na criação de identificadores.

5 CONSIDERAÇÕES FINAIS

Tradicionalmente, os recursos têm sido descritos ou representados com o suporte de um modelo conceitual de entidade única (também conhecido como mono-entity ou single-entity), no qual uma entidade principal, seja ela um livro, um artigo ou uma música é caracterizada por um conjunto de atributos, materializados como metadados (BAKER, COYLE e PETIYA, 2014). No decorrer das últimas duas décadas, entretanto, o mundo bibliográfico tem desenvolvido modelos de recursos bibliográficos que não veem um livro apenas como uma entidade singular, mas como um conjunto de entidades variadas que refletem o significado, expressão e aspectos físicos desse recurso e que possui relacionamentos com outros recursos.

Entende-se como sendo a grande contribuição do modelo FRBR, a introdução das entidades OEMI (Obra, Expressão, Manifestação e Item) ao tratamento da informação bibliográfica. A distinção entre conceitos físicos e conceitos abstratos, além do reconhecimento da importância entre os relacionamentos existentes entre objetos culturais, teve grande impacto nas regras de catalogação e em outras aplicações não bibliográficas.

Coyle (2016, p. 157) declara que o “FRBR representa um grande salto na evolução da teoria de descrição bibliográfica”, e abre uma “nova era de regras de catalogação

bibliográfica”. É importante ressaltar o pioneirismo do FRBR ao utilizar a modelagem Entidade-Relacionamento para a informação bibliográfica em um trabalho iniciado em 1992, quase uma década antes do início da Web Semântica, utilizando o que existia de mais moderno no campo da modelagem conceitual. A publicação de dados estruturados no contexto da Web Semântica e dos Dados Ligados depende do reconhecimento da existência de entidades relacionadas, algo não contemplado pela prática tradicional da catalogação anglo-americana - AACR2.

O modelo IFLA LRM, sucessor do FRBR, já nasce com o objetivo de facilitar a publicação de dados bibliográficos na Web de dados. Além de compatibilizar os modelos anteriores da família FRBR, o LRM simplificou os modelos e empregou um maior nível de generalidade e abstração. O novo modelo manteve as entidades OEMI e suas relações, e removeu as delimitações entre os grupos de entidades, tratando recurso, assunto e autoridades como dados bibliográficos de uma forma geral.

Em termos de metodologia de modelagem, a principal diferença em relação aos modelos anteriores está na utilização das relações de especialização, presentes no modelo Entidade-Relacionamento Estendido. Entidades, atributos e relacionamentos são estruturados em hierarquias que refletem relações taxonômicas entre esses elementos. A utilização desse tipo de relação permitiu (i) a simplificação do modelo em termos de redução de elementos, (ii) a aplicação de um maior nível de abstração, (iii) melhor entendimento do relacionamento entre as entidades e (iv) maior facilidade na extensão do modelo. O LRM é mais genérico que seus predecessores e deixa claro no seu documento que “qualquer aplicação prática necessitará determinar o nível apropriado de precisão”, requerendo expansões ou omissões do modelo (IFLA, 2017, p 10). Esse nível de generalidade é bem-vindo em se tratando de um modelo conceitual de alto-nível. Modelos mais genéricos facilitam a integração entre modelos derivados e atendem a uma gama maior de aplicações. Em contrapartida, se posicionam mais longe da implementação, exigindo que o implementador tome mais decisões em relação ao domínio da aplicação. Em função das características citadas, é factível se imaginar que tais modelos possam ter uma aplicabilidade ampla pelas comunidades interessadas em trabalhar com dados bibliográficos no contexto da Web Semântica.

REFERÊNCIAS

BAKER, T.; COYLE, K.; PETIYA, S. Multi-entity models of resource description in the Semantic Web: a comparison of FRBR, RDA and BIBFRAME. **Library hi tech**, v.32, n.4, p.562–582, 2014.

Disponível em: <<https://www.emeraldinsight.com/doi/abs/10.1108/LHT-08-2014-0081>>. Acesso em: 10 jul. 2018.

BIANCHINI, C. Osservazioni sul modello IFLA Library Reference Model. **Italian Journal of Library, Archives, and Information Science**, v.8, n.3, p.86–99, 2017. Disponível em: <<https://www.jlis.it/article/view/12416>>. Acesso em: 10 jul. 2018.

CHOFFÉ, P.; LERESCHE, F. Doremus : Connecting sources, enriching catalogues and user experience. In: IFLA WORD LIBRARY AND INFORMATION CONGRESS, 82., 2016, Columbus, Ohio. **Anais eletrônicos...** Columbus, Ohio: IFLA, 2016. p.1-20. Disponível em: <<http://library.ifla.org/id/eprint/1322>>. Acesso em: 20 abr. 2018.

CIDOC CRM SIG. **CIDOC Conceptual Reference Model**, Sítio Web. [S.l.], 2016. Disponível em: <<http://www.cidoc-crm.org>>. Acesso em: 14 out. 2016.

CIDOC CRM SIG. **Definition of the CIDOC Conceptual Reference Model**, Version 5.0.4. [S.l.], nov. 2011. Disponível em: <http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_5.0.4.pdf>. Acesso em: 14 abr. 2018.

COYLE, K. **FRBR, Before and After**: a look at our bibliographic models. Chicago: American Library Association, 2016. 179p. Disponível em: <<http://kcoyle.net/beforeAndAfter>>. Acesso em: 14 mar. 2017.

COYLE, K. **Yes we can! Libraries and the Semantic Web**. 2010. Disponível em: <<http://www.kcoyle.net/presentations>>. Acesso em: 7 out. 2017.

DICKEY, T. J. FRBRization of a Library Catalog: Better Collocation of Records, Leading to Enhanced Search, Retrieval, and Display. **Information Technology and Libraries**, v.27, n.1, p.23–32, 2008. Disponível em: <<https://ejournals.bc.edu/ojs/index.php/ital/article/view/3260>>. Acesso em: 29 abr. 2018.

DOERR, M. The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. **AI Magazine**, v.24, n.3, p.75–92, 2003. Disponível em: <<https://www.aaai.org/ojs/index.php/aimagazine/article/view/1720>>. Acesso em: 14 abr. 2018.

IFLA. **Definition of FRBROO**: A Conceptual Model for Bibliographic Information in Object-Oriented Formalism. Netherlands, 2015. Disponível em: <https://www.ifla.org/files/assets/cataloguing/FRBROO/frbroo_v_2.4.pdf>. Acesso em: 2 jul. 2018.

IFLA. **Functional Requirements for Authority Data**. [S.l.], 2013. Disponível em: <https://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf>. Acesso em: 2 jul. 2018.

IFLA. **Functional Requirements for Bibliographic Records**. [S.l.], 1998. Disponível em: <<https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>>. Acesso em: 2 jul. 2018.

IFLA. **Functional Requirements for Subject Authority Data**. [S.l.], 2010. Disponível em: <<http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>>. Acesso em: 2 jul. 2018.

IFLA. **IFLA Library Reference Model: A Conceptual Model for Bibliographic Informaion**. Netherlands, 2017. Disponível em: <https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf>. Acesso em: 2 jul. 2018.

LIMA, J. A. O. **Modelo genérico de relacionamentos na organização da informação legislativa e jurídica**. Tese (Doutorado) — Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília (UnB), Brasília, 2008. Disponível em: <<http://repositorio.unb.br/handle/10482/3398>>. Acesso em: 14 out. 2016.

MILLER, S. J. **Metadata for Digital Collections: a how-to-do-it manual**. Chicago: Ed. Neal-Schuman, 2011.

MURRAY, R. J. The FRBR-Theoretic Library: The Role of Conceptual Data Modeling in Cultural Heritage Information System Design. In: INTERNATIONAL CONFERENCE ON PRESERVATION OF DIGITAL OBJECTS - IPRES. 1. **Anais eletrônicos...** London, 2008. Disponível em: <www.bl.uk/ipres2008/presentations_day1/26_Murray.pdf>. Acesso em: 12 mai. 2016.

NEWMAN, R. **Expression of Core FRBR Concepts in RDF**. [S.l.], 2005. Disponível em: <<http://vocab.org/frbr/core>>. Acesso em: 10 abr. 2018.

RIVA, P. Building RDA using the FRBR Library Reference Model. In: IFLA SATELLITE CONFERENCE, RDA IN THE WIDER WORLD, 2016a, Columbus, Ohio. **Anais eletrônicos...** Columbus, Ohio: IFLA, 2016. P.1-12. Disponível em: <<https://spectrum.library.concordia.ca/982198/1/S17-2016-riva-en.pdf>>. Acesso em: 10 mar. 2018.

RIVA, P. On the new conceptual model of the bibliographic universe: the FRBR Library Reference Model. **AIB Studi**, v.56, n.2, p.265–274, 2016b. Disponível em: <<http://aibstudi.aib.it/article/view/11480>>. Acesso em: 10 mar. 2018.

RIVA, P.; DOERR, M.; ŽUMER, M. FRBRoo: enabling a common view of information from memory institutions. In: IFLA WORLD LIBRARY AND INFORMATION CONGRESS, 74., 2008, Québec, Canada. **Anais eletrônicos...** Québec, Canada: IFLA, 2008, p.1-11. Disponível em: <http://www.ifla.org/IV/ifla74/papers/156-Riva_Doerr_Zumer-en.pdf>. Acesso em: 15 abr. 2018.

RODRÍGUEZ, E. E. *et al.* Advocating for a change of mentality in the development of metadata standards: historical celestial cartography as a specialization example. **JLIS.it**, v.8, n.3, p.39–57, 2017. Disponível em: <<https://www.jlis.it/article/view/12398>>. Acesso em: 15 abr. 2018.

SANTOS, H. P. Modelo CIDOC CRM: interoperabilidade semântica de informações culturais. **Brazilian Journal of Information Studies: Research Trends**, v.10, n.1, p.56–62, 2016. Disponível em: <<https://dialnet.unirioja.es/descarga/articulo/5376141.pdf>>. Acesso em: 14 out. 2016.

SEEMAN, D.; GODDARD, L. Preparing the way: creating future compatible cataloging data in a transitional environment. **Cataloging & Classification Quarterly**, v.53, n.3-4, p.331-340, 2015.

TILLETT, B. O que é FRBR? Um modelo conceitual para o universo bibliográfico. [S.l.]: **Technicalities**, v.25, n.5, set./out. 2003. Disponível em: <<http://www.loc.gov/catdir/cpsd/o-que-e-frbr.pdf>>. Acesso em: 2 jul. 2018.