



24° ENANCIB
Encontro Nacional de Pesquisa em Ciência da Informação
Perspectivas Contemporâneas na Ciência da Informação
• Vitória - ES • Ancib • PPGCI/UFES



XXIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXIV ENANCIB

ISSN 2177-3688

GT 2 – Organização e Representação do Conhecimento

**ORGANIZAÇÃO DO CONHECIMENTO EM ARQUIVOS: UMA ANÁLISE SEMÂNTICA
COMPARATIVA ENTRE RIC-CM E CIDOC-CRM**

***KNOWLEDGE ORGANIZATION FOR ARCHIVES: A COMPARATIVE SEMANTIC ANALYSIS
BETWEEN RIC-CM AND CIDOC-CRM***

Marieta Marks Löw – Universidade Federal de Santa Catarina (UFSC)

Caliel Cardoso de Oliveira – Universidade Federal do Rio Grande do Sul (UFRGS)

Thiago Henrique Bragato Barros – Universidade Federal do Rio Grande do Sul (UFRGS)

Modalidade: Trabalho Completo

Resumo: Ontologias e modelos conceituais podem ser analisados como representações do conhecimento de um domínio. No âmbito dos arquivos e patrimônio material destacamos dois modelos: RiC-CM e CIDOC-CRM. O primeiro modelo, RiC-CM, foi desenvolvido pelo Conselho Internacional de Arquivos e possui uma ontologia derivada. A família RiC tem como escopo a produção de documentos de arquivo. O segundo modelo, CIDOC-CRM, também uma ontologia, foi desenvolvido pelo Conselho Internacional de Museus e fornece diretrizes para modelagem formal de artefatos e patrimônio cultural. Como os arquivos podem conter documentos com valor artístico, o modelo CIDOC-CRM também pode ser usado para representar esses registros. Neste trabalho apresentamos os resultados de pesquisa que investigou como conceitos arquivísticos são definidos nos dois modelos. Identificamos aproximações e divergências semânticas entre as duas ontologias. Para isso, analisamos suas classes, atributos, relações e atributos de relações, a partir de uma análise semântica comparativa, e correspondências conceituais foram propostas, quando possível. Entendemos que as abordagens utilizadas para representar a realidade no contexto arquivístico e de patrimônio cultural são distintas, dadas as visões ontológicas com as quais cada modelo está comprometido. Assim, há alguma sobreposição entre eles, mas com especificidades próprias de cada área.

Palavras-chave: Organização do Conhecimento; Análise semântica; RiC-O; CIDOC-CRM.

Abstract: Ontologies and conceptual models can be analyzed as representations of the knowledge of a domain. In archives and material heritage domain, two models stand out: RiC-CM and CIDOC-CRM. The first model, RiC-CM, was developed by the International Council on Archives and has a derived ontology. The scope of the RiC family is the production of archival records. The second model, CIDOC-CRM, also an ontology, was developed by the International Council of Museums and provides guidelines for the formal modeling of artifacts and cultural heritage. As archives can contain documents with artistic value, the CIDOC-CRM model can also be used to represent these records. In this paper, we present the results of research that investigated how archival concepts are defined in

the two models. We identified semantic similarities and differences between the two ontologies. To do this, we analyzed their classes, attributes, relations, and attributes of relations, based on a comparative semantic analysis, and conceptual correspondences were proposed where possible. We understand that the approaches used to represent reality in the context of archives and cultural heritage are distinct, given the ontological visions to which each model is committed. Thus, there is some overlap between them, but with specificities for each area.

Keywords: Knowledge Organization; Semantic analysis; RiC-O; CIDOC-CRM.

1 INTRODUÇÃO

O desenvolvimento de modelos conceituais que representem adequadamente um determinado universo de conhecimento é fundamental para a criação de sistemas de informação eficientes e envolve várias disciplinas. Esses modelos auxiliam no processo de construção de uma visão conceitual particular. A Organização do Conhecimento (OC) dialoga com as práticas de modelagem conceitual com base em sua aplicação na construção de sistemas de organização do conhecimento (SOCs) e, no uso de ontologias, para validar a consistência de um modelo conceitual. Nesse sentido, os modelos conceituais, que fornecem artefatos ontológicos para sua aplicação, são um terreno fértil para investigar a representação adequada do conhecimento de um domínio. Assim, esta pesquisa oferece uma oportunidade de explorar semelhanças e diferenças entre duas ontologias, baseadas em dois modelos conceituais diferentes, quanto a formas de representar o conhecimento arquivístico, conforme enunciado por Barros e Tognoli (2015).

O modelo conceitual RiC-CM foi desenvolvido pelo Conselho Internacional de Arquivos (ICA), para servir de referência no domínio da produção de documentos administrativos e é complementado por sua ontologia, RiC-O. Juntos, o modelo e a ontologia permitem a identificação intelectual e a descrição de documentos arquivísticos, as pessoas que os criaram, usaram ou que estão documentados neles, e as atividades daqueles que os documentos beneficiam e documentam (International Council on Archives, 2023). O modelo foi lançado como uma versão 1.0 completa e oficial em novembro de 2023. Versões preliminares anteriores foram lançadas em 2016 e 2021. Já o CIDOC-CRM, desenvolvido pelo Conselho Internacional de Museus (ICOM), fornece diretrizes para modelar formalmente artefatos e patrimônio cultural em um instrumento ontológico. O modelo CIDOC-CRM pode ser uma opção para representar a documentação arquivística, pois os arquivos possuem também documentos com valor cultural.

Considerando esses dois modelos conceituais, nossa pesquisa foi orientada pela seguinte pergunta: Como os conceitos do domínio arquivístico podem ser representados no CIDOC-CRM, no RiC-CM e no RiC-O, e quais aproximações e divergências semânticas podem ser observadas entre os dois modelos como estruturas conceituais para gestão de documentos e arquivos? Vale ressaltar que investigações semelhantes têm sido realizadas desde 2017, relacionando modelos ontológicos com sistemas de gestão de documentos e organização do conhecimento arquivístico (Barros; Gomes, 2018; Barros; Sousa, 2019; Gomes *et al.*, 2019).

2 MODELOS CONCEITUAIS E ONTOLOGIAS: O CIDOC-CRM E A FAMÍLIA RiC

O campo da modelagem conceitual visa capturar e representar um determinado aspecto da realidade dentro dos limites da percepção humana (Wand; Storey; Weber, 1999). A Ontologia Aplicada oferece diferentes princípios, de acordo com a visão ontológica que está sendo adotada, para que um determinado domínio possa ser representado formalmente, capturando as principais características das entidades que compõem essa porção da Realidade e representando-as de forma clara e sem ambiguidades (Almeida, 2020). Diante disso, é natural que existam aproximações entre as atividades de modelagem conceitual e a aplicação de seus princípios para a construção de artefatos ontológicos. Historicamente, as ontologias têm sido inseridas na modelagem conceitual principalmente como um mecanismo de avaliação da consistência das linguagens de modelagem conceitual (Signore, 2009). Verdonck *et al.* (2019) explicam que os dois critérios de avaliação utilizados para validar a eficácia dos modelos conceituais, Adequação de Domínio e Adequação de Compreensão, passaram pelo crivo das ontologias, onde possíveis erros ou ambiguidades na representação puderam ser mais facilmente identificados, bem como inconsistências na compreensão do domínio derivadas dessas falhas.

A segunda interação entre as ontologias e a modelagem conceitual baseou-se no entendimento de que as ontologias poderiam oferecer uma base teórica para a construção de um modelo conceitual com base em sua capacidade de capturar e representar os principais elementos do universo que elas buscam formalizar (Verdonck *et al.*, 2019). Embora os modelos conceituais por si só representem um determinado segmento da realidade, as ontologias fundacionais permitem categorizar os elementos que fazem parte da Realidade em um nível conceitual, orientando como a própria natureza do Universo será compreendida

pelos modelos conceituais que as utilizam como base teórica (Carvalho *et al.*, 2015). Em resumo, as ontologias de topo fornecem uma maneira de entender o todo. Essas propriedades gerais compreendem todas as coisas que podem existir no Universo e que devem ser representadas nessa ontologia (Almeida, 2013).

Por fim, a terceira interface entre os modelos conceituais e as ontologias é a modelagem conceitual orientada por ontologia ou ODCM. Verdonck *et al.* (2019) afirmam que a ODCM difere da modelagem conceitual tradicional porque as ontologias não são mais ferramentas separadas que influenciam a modelagem conceitual, mas se tornam técnicas de modelagem, tornando os modelos conceituais mais robustos e, ao mesmo tempo, sua criação ainda mais complexa. Assim, podemos observar que a interação entre a modelagem conceitual e a ontologia se manifesta na aplicação de teorias ontológicas em atividades de modelagem tradicionais e no desenvolvimento de artefatos ontológicos, como os modelos conceituais em si.

Em desenvolvimento contínuo desde 1996, o CIDOC-CRM é amplamente reconhecido como um modelo ontológico de alta qualidade e serve como padrão para vários outros projetos de modelagem e representação no domínio do patrimônio cultural e da museologia, abrangendo desde sítios arqueológicos a documentos bibliográficos e até mesmo documentação digital e conjuntos de dados (Biagetti, 2021). O CIDOC-CRM tem como objetivo fornecer recursos que facilitem a troca e a integração de informações sobre elementos da cultura e do patrimônio da humanidade, considerando as diferentes fontes dessas informações e as diferentes maneiras pelas quais elas são representadas por instituições que lidam com um domínio maior do Patrimônio Cultural, como museus e bibliotecas (CIDOC, 2024). Dada a complexidade envolvida na representação desse domínio, as entidades que compõem o modelo variam de classes de nível superior, esperadas para ontologias básicas que visam representar a Realidade como um todo, até classes mais específicas que permitem a representação de entidades altamente individualizadas.

Uma das principais características do CIDOC-CRM é seu compromisso ontológico baseado em uma perspectiva centrada em eventos, que permite a modelagem de processos ou evoluções de diferentes indivíduos, sejam eles, na linguagem do modelo, Entidades Temporais (classe E2 Entidade Temporal) ou Itens Persistentes (classe E77 Item Persistente). Aqui, E2 corresponde à noção filosófica de entidades perdurantes, enquanto E77 é equivalente à noção de entidades endurantes. A perspectiva centrada em eventos determina

que as entidades de E2 podem ser eventos específicos que também se relacionam com indivíduos específicos por meio de relacionamentos, chamados de propriedades no modelo. Essas propriedades também são bastante específicas, projetadas para representar a Realidade por meio de um jogo de interações entre as entidades temporais derivadas de E2 e as entidades persistentes derivadas de E77, de modo que seja possível representar os diferentes eventos temporais que atravessam a existência de um indivíduo persistente, bem como os diferentes indivíduos que têm alguma influência sobre os eventos que se desenrolam ao longo do tempo (CIDOC, 2024). Essa redução da realidade a eventos que envolvem indivíduos e ocorrem em contextos históricos e geográficos mais amplos é o principal ponto forte do CIDOC-CRM em seu esforço para padronizar as informações de museus e patrimônios de diferentes instituições (Biagetti, 2021). Dado o papel das ontologias no desenvolvimento de ambientes da Web semanticamente ricos, essa especificação também abre novas possibilidades para a operação de agentes inteligentes que exploram modelos com a marcação semântica proposta pelo CIDOC-CRM, permitindo que eles façam inferências diferentes de acordo com o tipo de entidade ou relacionamento para o qual suas pesquisas são direcionadas, sejam elas entidades temporais, entidades persistentes ou cadeias de relacionamentos entre indivíduos de ambos os tipos (Signore, 2009).

A família de modelos RiC (*Records in Contexts*), voltada para a descrição arquivística, foi desenvolvida por um grupo com membros de 15 países, formado pelo ICA, para discutir a integração entre os modelos de descrição existentes. Uma primeira versão preliminar do modelo conceitual foi lançada para discussão pública em 2016, e uma segunda versão ampliada foi lançada em 2019, ainda como uma versão preliminar, chamada de versão 0.2. Enquanto isso, uma ontologia OWL foi desenvolvida a partir do modelo conceitual, e uma versão preliminar da ontologia RiC-O foi apresentada em julho de 2021. As versões completas do RiC-CM (v. 1.0) e do RiC-O (v. 1.0) foram lançados em novembro de 2023. Há, ainda, um documento chamado *Foundations of Archival Descriptions* (Fundamentos da descrição arquivística) onde os princípios da descrição arquivística são apresentados. E está em desenvolvimento uma quarta e última parte do RiC, chamada RiC-AG ou Diretrizes de Aplicação. O modelo traz conceitos fundamentais para descrever o contexto da produção de documentos de arquivo, e sua ontologia apresenta classes e propriedades de entidades, juntamente com suas relações. Desde seu lançamento, a comunidade arquivística tem discutido o modelo e a ontologia do ICA (Lów; Rocha; Barros, 2023; Bianchini, 2022;

Mikhaylova; Metilli, 2023; Cé; Flores; Campos, 2023). Eles foram implementados nos Arquivos Nacionais da França, dos Países Baixos e nos Arquivos da Cidade de Amsterdã, entre outros, e algumas ferramentas foram desenvolvidas.

Algumas comparações foram feitas com outras ontologias ou modelos conceituais. Feliciati (2022) analisa a integração de metadados de controle de autoridade em arquivos e bibliotecas, comparando os modelos RiC-CM e RiC-O com o modelo IFLA-LRMs desenvolvido para biblioteconomia. O autor conclui que a comparação de classes e propriedades é necessária para qualquer perspectiva de interoperabilidade entre modelos semânticos. Bianchini (2022) compara o IFLA-LRM, o RiC-CM e o CIDOC-CRM sob a perspectiva da Web Semântica. Koch, Lopes e Ribeiro (2023) usam o CIDOC-CRM como modelo de dados para um aplicativo chamado ArchOnto, baseado em ISAD(G), para descrever documentos dos arquivos portugueses.

A principal característica do modelo e da ontologia é propor os conceitos mais importantes para descrever um ambiente de arquivamento e produção de documentos e suas relações em um modelo relativamente simples. De acordo com o ICA, o RiC-CM é um modelo conceitual de alto nível que descreve e identifica recursos documentais e todos os seus componentes e entidades relacionadas, como o agente que criou ou usou o documento, os agentes e as atividades documentadas nele e seus contextos.

Os conceitos centrais do modelo estão todos relacionados a documentos de arquivos e sua produção. Estes conceitos, na perspectiva teórica da Arquivologia, se vinculam ao chamado Princípio da Proveniência, segundo o qual os documentos de um determinado produtor devem ser mantidos reunidos, com a ordem original em que foram produzidos. O conceito de Proveniência é fundamental na teoria arquivística, e dele decorrem diversos outros conceitos que sustentam o modelo e ontologia RiC.

O Recurso de Registro (*RiC-E02 Record Resource*) é o principal deles, um tipo de Coisa (*RiC-E01 Thing*) definido como informação produzida ou recebida e mantida por um agente no decorrer da vida ou da atividade de trabalho. Trata-se tanto de uma produção inicial de informações quanto de uma reutilização de informações existentes anteriormente. Aqui estão presentes os conceitos de documento e fundo arquivístico, ambos são resultados das ações do produtor. O documento é o registro de uma ação ou atividade realizada por esse produtor. Já o fundo é o resultado da acumulação dos documentos produzidos e recebidos, e mantidos pelo produtor para fins de prova ou consulta.

A Instanciação (*RiC-E06 Instantiation*) também é um tipo de Coisa, intimamente relacionada ao *Record Resource* (*RiC-E02*), como a existência material de um registro. Ela é definida como a inscrição de informações por um agente em um suporte em qualquer forma persistente e recuperável, como um meio de comunicar informações através do tempo e do espaço.

Agente e Atividade são os outros conceitos centrais do modelo. O Agente é tanto o produtor como a pessoa ou equipamento que registra a informação. A Atividade é a ação do produtor que resultou no registro documental.

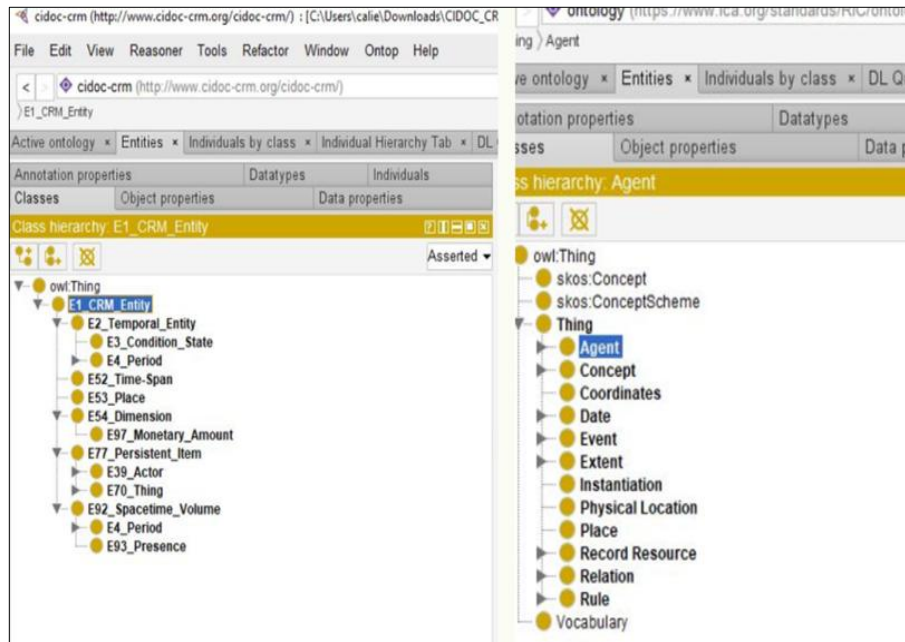
3 METODOLOGIA

O primeiro passo em nosso trabalho comparativo foi analisar as classes que compõem as ontologias RiC-O e CIDOC-CRM, com etapas subsequentes para análise das outras entidades. Em todas as etapas, nosso trabalho consistiu em realizar o *download* dos arquivos das ontologias e executá-los no software Protegé, amplamente considerado como o principal software para construção de ontologias. Para esta etapa de nossa pesquisa, usamos a versão 1.0.1 da ontologia RiC-O, lançada em maio de 2024, e a versão 7.1.3 do CIDOC-CRM, lançada em fevereiro de 2024. A Figura 1, abaixo, apresenta a taxonomia geral de ambas as ontologias visualizadas no Protegé.

Apesar de haver algumas semelhanças entre as duas taxonomias de classes, é possível observar que ambos os modelos têm prioridades diferentes. Enquanto o RiC-O oferece conceitos mais técnicos e específicos, adequados para representar o domínio especializado dos recursos arquivísticos, o CIDOC-CRM inicia sua taxonomia com conceitos de nível superior projetados para representar o escopo mais geral em torno do trabalho e da pesquisa envolvendo Patrimônio Cultural. Pode-se perceber, por exemplo, como alguns dos conceitos centrais do RiC-O aparecem apenas como subclasses de segundo nível no CIDOC-CRM, subordinados às entidades centrais mais generalistas, por exemplo, *Agent* do RiC-O versus *Actor* do CIDOC-CRM.

O trabalho comparativo nesta etapa de estudo usou a taxonomia RiC-O como objeto referencial para encontrar equivalentes dentro do CIDOC-CRM. As entidades RiC-O usadas na comparação pertenciam ao primeiro e segundo níveis da taxonomia.

Figura 1 - Taxonomia geral do CIDOC-CRM e RiC-O no Protegé



Fonte: Imagem do autor

Algumas entidades de RiC-O se mostraram difíceis de equiparar a entidades CIDOC-CRM durante nossa análise. Alguns, como *Instanciação*, podem ser equiparados não a uma única classe CIDOC-CRM, mas sim a uma relação específica entre classes por meio de uma propriedade de objeto específica. Em outros casos, como *Rule*, uma única classe CIDOC-CRM foi suficiente para abranger a maior parte do significado de sua contraparte RiC-O, embora não todo, impedindo uma correspondência completa entre os dois modelos.

As etapas seguintes deste estudo foram conduzidas de maneira ligeiramente diferente. Em vez de usar apenas as ontologias RiC-O e CIDOC-CRM, nosso ponto de partida foi a lista de entidades RiC-CM. Em seguida, mapeamos entidades do RiC-CM para seus equivalentes no RiC-O, um processo que se tornou mais simples graças à página da web do RiC-O, que informa com facilidade quais elementos listados são implementações adaptadas de entidades do modelo conceitual. Em seguida, pegamos as duas listagens do RiC-CM e do RiC-O e repetimos o processo analítico de encontrar equivalentes no CIDOC-CRM.

Quantitativamente, foram analisados os 45 atributos de entidade descritos no RiC-CM, 86 relações e 06 atributos de relação. Em muitos casos, uma única entidade do RiC-CM foi dividida em duas ou mais contrapartes no RiC-O; os atributos geralmente se tornaram propriedades de dados ou classes, enquanto as relações se tornaram propriedades de objetos ou classes. Em todos os casos, tentamos encontrar equivalentes simples no modelo CIDOC-

CRM. Nosso trabalho comparativo visou a fidelidade, sempre traduzindo entidades RiC-O para CIDOC-CRM de acordo com sua natureza: classes por classes, propriedades por propriedades. Em alguns casos, a equivalência pode ser encontrada entre um tipo de entidade, mas não o outro; isso foi visto principalmente nos casos em que uma relação RiC-CM se tornou uma propriedade de classe e objeto em RiC-O. Nesses casos, a classe resultante representava alguma relação genérica e aberta, de uma natureza que o CIDOC-CRM geralmente não se destina a modelar; nesses casos, encontrar uma propriedade equivalente no CIDOC-CRM pôde ser feito, no geral, sem problemas, mas nenhuma classe equivalente foi nomeada. Ao contrário da etapa anterior deste estudo, em que uma cadeia de relacionamento representaria entidades sem um único equivalente CIDOC-CRM, decidimos não apresentar nenhuma cadeia elaborada, deixando apenas comparações um-para-um (há algumas exceções em que duas entidades CIDOC-CRM distintas, em conjunto, abrangeriam todo o significado de sua contraparte RiC-O; nesses casos, ambas as entidades são apresentadas como dados discretos em nossas tabelas comparativas). Finalmente, conforme explicado na seção Resultados, as diferenças de propósito e escopo entre os modelos fizeram com que muitas entidades de RiC não tivessem um equivalente apropriado no CIDOC-CRM, levando a espaços em branco em nosso esforço comparativo.

As tabelas que mostram os resultados de nosso experimento comparativo estão disponíveis como conjuntos de dados no repositório Zenodo. Nelas, apresentamos cada uma das entidades utilizadas em nosso trabalho comparativo e, no caso da primeira tabela (comparando as classes RiC-CM com as classes CIDOC-CRM), o raciocínio por trás das escolhas feitas nos casos em que não foi possível encontrar equivalência um-para-um. Os dados utilizados neste estudo constam anexos a este documento, como arquivos no formato .csv. Nossas conclusões sobre o experimento são apresentadas a seguir.

4 RESULTADOS

Como resultado preliminar, que ocorreu com todos os tipos de entidades, nosso experimento destacou algumas diferenças importantes em como a Realidade é percebida tanto no CIDOC-CRM quanto no RiC-O. Geralmente, as entidades apresentadas pelo RiC-O permitem uma representação muito mais direta que permite uma troca eficiente de informações relevantes entre as instituições. As entidades centrais no primeiro e no segundo níveis da taxonomia permitem a modelagem de informações relevantes para o trabalho

arquivístico: definindo o que pode ser entendido como um *Record Resource* (RiC-E02), bem como os elementos que, juntos, constituem as informações de proveniência desse recurso, como o agente por trás de sua produção ou o mandato que justifica sua custódia por uma determinada organização. Embora isso possa levar a relações taxonômicas aparentemente contraintuitivas (como o *Record Resource* estar no mesmo nível hierárquico que o *Concept*), é consistente com a preocupação do RiC-O em descrever o *Record Resource* em si mesmo, bem como os elementos imediatos que compõem o contexto em que esse recurso está localizado.

O CIDOC-CRM, por outro lado, oferece classes mais gerais em um nível mais alto de abstração. Entendendo a Realidade como um jogo de interações entre entidades temporais e persistentes, o CIDOC-CRM permite modelar o histórico de um item até e além do momento de sua inclusão em uma coleção e vários eventos temporais à medida que se desenvolvem uns com os outros. Essa característica permite uma contextualização mais rica dos processos e atividades que levaram à criação de um artefato ou outro recurso e de todas as partes envolvidas ao longo de sua existência. No entanto, em comparação com o RiC-O, o CIDOC-CRM apresenta algumas lacunas perceptíveis em sua representação, principalmente quando surge a necessidade de modelar ambientes corporativos. Apesar do alto nível de abstração de suas classes principais, o CIDOC-CRM ainda é concebido para uso em instituições de Patrimônio Cultural.

As tabelas de comparação foram criadas usando classes, atributos, relações e atributos de relação RiC e suas definições, mapeando seus equivalentes CIDOC-CRM. As correspondências eram tão precisas e semanticamente semelhantes quanto possível, mas nem sempre idênticas. Ao identificar equivalências em classes, pudemos fazer algumas cadeias de relações onde não havia correspondência semântica direta. Para outras tabelas (atributos e relações), decidimos não fazer tal cadeia de relações. Nesse caso, a cadeia relacional não era tão simples, então decidimos usar correspondência direta sempre que possível. Em relação aos atributos de relação, não houve correspondência relacionada entre RiC e CIDOC-CRM.

Em relação às entidades de Atributo do RiC-CM, observou-se que o modelo as trata como adjacentes às suas entidades, resultando em serem representadas como propriedades de tipo de dados no RiC-O. Na prática, isso significa que a maioria deles se comporta como notas anexadas às entidades que caracterizam, tendo dados do tipo literal ou mesmo string como seu possível intervalo. Isso se deve à natureza menos estruturada que esses dados

podem ter: a propriedade de tipo de dados *Rico:history*, por exemplo, deve ser preenchida com um texto de forma livre resumindo o desenvolvimento de uma determinada entidade; sendo uma propriedade de dados, ela não aponta para nenhuma outra entidade no modelo para servir como seu intervalo, como o CIDOC-CRM faria, trabalhando em um nível tão alto de granularidade e expressando informações tão específicas que se torna orientada para o usuário humano que trabalha com o modelo. Essa abertura em relação a muitas propriedades de tipo de dados causa problemas ao tentar mapear esses atributos para o CIDOC-CRM, pois esse modelo (pelo menos idealmente) faria uso de cadeias de relações mais complexas para expressar o histórico anterior de um determinado item, tornando-o o mais próximo possível para facilitar a automação e a integração de dados entre diferentes instituições. A solução adotada para criar uma equivalência foi fazer uso da propriedade de dados do CIDOC-CRM *P3 has note*, que permite a inclusão de dados de forma livre, do tipo literal (conceitualmente, esses dados seriam entendidos como uma instância da classe *E62 String*, embora essa entidade não seja usada nos arquivos de ontologia do CIDOC-CRM).

Em alguns casos, as *Data Properties* de RiC-O foram mapeadas para *Object Properties* do CIDOC-CRM sempre que este oferecia propriedades que poderiam cumprir a mesma função semântica que as do modelo anterior. Um exemplo seria a propriedade de tipo de dados *rico:productionTechnique*, que foi comparada com a propriedade de objeto *P32 used general technique (was the technique of)*. Ainda assim, em relação aos atributos, era comum que os modelos RiC-CM fossem traduzidos para a ontologia não apenas como propriedades de tipo de dados, mas também como classes em si. Essas classes também funcionavam em um nível muito alto de granularidade, representando informações específicas úteis durante o trabalho de arquivamento. Nesses casos, pesquisamos classes CIDOC-CRM que fossem suficientes para representar o significado pretendido. Na maior parte dos casos, isso foi alcançado sem problemas, embora o referido grau de abertura dos atributos RiC-O, combinada com a especificidade das informações transmitidas, tenha causado que algumas dessas classes derivadas do atributo RiC-O tivessem que ser mapeadas para classes CIDOC-CRM mais genéricas. O *classRico: OccupationType*, por exemplo, representa a "Categorização de uma profissão, comércio ou ofício exercido por uma Pessoa no cumprimento de uma Atividade". Essas informações seriam melhor representadas no CIDOC-CRM como uma instância da classe *E55 Type* vinculada a uma instância de *E7 Activity* por meio da propriedade *P2i is type of (has type)*. Como mencionado, não expressamos cadeias de relacionamento ao

trabalhar com entidades RiC-CM que não sejam as classes primárias. Devido a essa escolha no procedimento, *Rico: OccupationType* foi comparado à classe *E55 Type*; outros casos semelhantes aconteceram durante o trabalho. Em suma, os atributos no RiC-CM são entendidos menos como entidades em si mesmos e mais como “adjetivos” para instâncias de classe. Mesmo quando implementadas como classes no RiC-O, seu papel semântico é condicional ao de outra instância, ilustrando as características dessa instância por estar relacionada a ela. Isso ajuda a ilustrar como o RiC-O visa atender aos usuários que lidam com o trabalho arquivístico, pois os atributos são implementados para enriquecer outros objetos, transmitindo informações pertinentes para melhor caracterizá-los. Isso cria um contraste com o CIDOC-CRM, onde o objetivo de facilitar a interoperabilidade entre instituições de patrimônio cultural significa que não há muitas opções para representar o conhecimento não estruturado de forma acessível, e a tipificação de entidades é melhor implementada por meio do uso de uma classe específica e das relações que podem envolvê-la.

Uma situação semelhante envolveu as entidades do tipo Relação da RiC-CM em relação à sua implementação na RiC-O. Eles foram em geral traduzidos para a RiC-O como propriedades de objeto, embora muitos também tiveram uma implementação dupla, tornando-se classes. As relações nos modelos Ric seguem a lógica mencionada no primeiro parágrafo desta seção: visam modelar, particularmente, informações pertinentes à proveniência de recursos de documentos, bem como informações organizacionais relevantes para a guarda de tais itens. Outra característica mencionada acima é o uso de entidades um pouco abertas para expressar informações, como o *rico:class WorkRelation*, que representa qualquer relação entre dois agentes que trabalharam juntos; outro exemplo é a primeira relação de RiC-O, *rico:isRelatedTo*, que poderia representar qualquer interação ocorrida entre duas entidades (essas entidades de fato possuem sub-relações instanciadas, mas sua inclusão chama atenção). Assim como antes, tentamos mapear classes para classes e propriedades de objetos para propriedades de objetos, embora não sem dificuldades.

Em muitos casos onde ocorreu a implementação tanto de uma propriedade quanto de uma classe na RiC-O, foi possível apontar para uma equivalência entre uma propriedade Ric-O e uma propriedade CIDOC-CRM, mas não houve êxito em encontrar equivalências de classes. Nesta fase, a ênfase dos modelos Ric na representação das relações organizacionais tornou-se muito evidente, e a falta de representação do CIDOC-CRM para esse tipo de entidade se fez aparente; não conseguimos encontrar uma propriedade equivalente suficiente

no CIDOC-CRM para *rico:isOrWasLeaderOf*, que vincula um agente a uma organização da qual ele era chefe, por exemplo. Algumas particularidades também ilustram de maneira mais sutil as diferenças de finalidade entre os modelos RiC e o CIDOC-CRM: Ric-O tem diferentes entidades implementadas para expressar diferentes formas de custódia de um determinado *Record Resource*, como *rico:hasCollector* ou *rico:hasAccumulator*, permitindo maior clareza sobre como esse registro veio a estar em sua posse, o que é informação muito pertinente para uma instituição de arquivo. Por outro lado, o CIDOC-CRM desenvolveu-se voltado a instituições de patrimônio cultural, nas quais a questão de como determinado item passou a estar sob custódia desta instituição não é tão relevante, e não oferece o mesmo nível de detalhamento para expressar tais informações. Da mesma forma, RiC-O oferece mais propriedades destinadas a representar aspectos legais do trabalho de arquivo, como a propriedade *Rico:issued*, que liga um agente a uma norma específica que ele emitiu ou publicou, um detalhe que não se consegue, ao menos dessa maneira, representar tão facilmente utilizando o CIDOC-CRM.

Por fim, a última etapa do nosso estudo foi mapear equivalentes no CIDOC-CRM para as entidades do tipo Atributo de Relação existentes nos dois instrumentos de RiC. Não foi possível localizar equivalentes adequados a essas entidades no CIDOC-CRM. Isso pode ser explicado pelo propósito de cada uma dessas três entidades, a saber, *rico:relationCertainty*, *rico:relationSource* e *rico:isEvidencedBy*, tem na Ric-O. Em todos os três casos, seu papel no modelo é vincular uma relação (expresso através do uso de propriedades) a uma dada entidade que ofereça validação àquela relação. Assim, a falta de entidades equivalentes no CIDOC-CRM se deve à forma como as questões de validade e atestação das propriedades de um documento são tratadas nesse modelo. Sendo especificamente destinado a representar artefatos culturais já sob custódia de uma determinada instituição (e informações contextuais em torno desses artefatos), o objetivo final do CIDOC-CRM dificilmente tem preocupações de usar um determinado artefato como um recurso documentário, cuja informação pode inclusive ainda ter vigência legal ou normativa. Isso está em oposição direta com os modelos RiC e seu escopo, que visa facilitar as atividades nas instituições de arquivo, onde os recursos de documentos frequentemente são usados como documentos com informações legais ou administrativas relevantes. Este valor informativo, que é muitas vezes de natureza legal ou jurídica, deve ser validado através do escrutínio do recurso de documento. Este aspecto do trabalho arquivístico explica a necessidade de tais entidades na Ric-O e oferece um argumento

convicente sobre por que as relações de equivalência entre elas e entidades do CIDOC-CRM não puderam ser alcançadas.

5 CONCLUSÕES

A análise comparativa das ontologias Ric-O e CIDOC-CRM revela abordagens distintas para representar a Realidade no contexto dos recursos arquivísticos e do patrimônio cultural, dadas as distintas visões ontológicas às quais cada modelo está comprometido. A escolha por uma ou outra forma de observar a Realidade necessariamente acontece de acordo com a aplicação e objetivo final de cada modelo. Podemos observar que a estrutura taxonômica tem algumas semelhanças, à primeira vista, porém, após uma observação mais aprofundada no nível semântico, diferenças importantes emergiram. Por exemplo, a representação de eventos é mais simplificada no RiC e mais complexa e detalhada no CIDOC-CRM. Além disso, o CIDOC-CRM faz uma distinção importante entre conteúdo intelectual e objetos materiais. Isso se deve à natureza do material do patrimônio cultural, no qual os fenômenos culturais podem ser expressos de forma imaterial, e sua influência pode ser sentida em características como representações nos artefatos físicos. Nos modelos RiC esta distinção não é essencial já que todos os registros são, de alguma forma, também um objeto material. Assim, o problema de distinguir entre conteúdo e portador é resolvido pela classe de *Instanciação*; o mesmo conteúdo pode ser instanciado em um portador diferente através do tempo e do espaço.

A Ric-O proporciona uma representação mais direta focada em possibilitar a troca eficiente de informações entre instituições arquivísticas, com entidades centrais que permitem a modelagem de informações relevantes para o trabalho de arquivo. Ele descreve o recurso de documento em si e os elementos imediatos que compõem o contexto em que ele está localizado. O CIDOC-CRM, por outro lado, oferece classes mais gerais em um nível mais alto de abstração, entendendo a Realidade como interações entre entidades temporais e persistentes. Essa visão ontológica permite modelar a história de um item, múltiplos eventos temporais e as partes envolvidas. No entanto, o CIDOC-CRM apresenta algumas lacunas perceptíveis em sua representação, principalmente na modelagem de ambientes corporativos/institucionais. Apesar do alto nível de abstração de suas classes principais, ele ainda é concebido para uso dentro das instituições do Patrimônio Cultural. O CIDOC-CRM oferece um compromisso ontológico com a divisão entre entidades temporais e físicas com um nível de granularidade que a Ric-O não possui. Propriedades de ambos os modelos

permitem a representação de relacionamentos entre eventos, tipo de informação que pode ser relevante para coleções documentais.

Conclui-se, portanto, que apesar de haver alguma sobreposição entre ambos os domínios, eles ainda são distintos, com suas especificidades. É possível que essa sobreposição entre os domínios do patrimônio cultural e do trabalho arquivístico explique justamente o porquê de não haver uma quantidade altamente expressiva de estudos e esforços para comparar esses dois modelos. Em resumo, ambos os modelos conceituais têm seus pontos fortes e limitações, e a escolha entre eles deve ser baseada nas necessidades específicas de cada projeto arquivístico ou de patrimônio cultural, ou de cada profissional que lida com registros arquivísticos ou objetos patrimoniais.

REFERÊNCIAS

ALMEIDA, M. B. **Ontologia em Ciência da Informação: teoria e método**. Curitiba: CRV, 2020.

ALMEIDA, M. B. Revisiting ontologies: A necessary clarification. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 64, n. 8, p. 1682-1693, 2013. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/asi.22861>. Acesso em: 10 jul. 2024.

BARROS, T. H. B.; GOMES, D. L. Classification and Knowledge Organization Systems: ontologies and archival classification. **Challenges and Opportunities for Knowledge Organization in the Digital Age**, [S. l.], p. 103-111, 2018. Disponível em: <http://dx.doi.org/10.5771/9783956504211-103>. Acesso em: 10 jul. 2024.

BARROS, T. H. B.; SOUSA, R. T. B. de. Organização do Conhecimento e Arquivologia: abordagens metodológicas. **Informação & Informação**, Londrina, v. 24, n. 2, p. 76-92, 2019. Disponível em: <http://hdl.handle.net/10183/257121>. Acesso em: 10 jul. 2024.

BARROS, T. H. B.; TOGNOLI, N. B. Os processos de representação do conhecimento arquivístico: elementos históricos e conceituais da classificação e descrição. In: GUIMARÃES, J. A. C.; DODEBEL, V. (org.). **Estudos avançados em organização do conhecimento: organização do conhecimento e diversidade cultural**. Marília: Fundepe, 2015. v. 3. p. 94-99.

BIAGETTI, M. T. Ontologies as Knowledge Organization Systems. **Knowledge Organization**, [S. l.], v. 48, n. 2, p. 152-176, 2021. Disponível em: https://www.researchgate.net/publication/344070697_Ontologies_as_Knowledge_Organization_Systems. Acesso em: 10 jul. 2024.

BIANCHINI, C. The Entities of the IFLA-LRM, RiC-CM, and CIDOC-CRM Models in the Semantic Web. **JLIS.it**, v. 13, n. 3, p. 63-75, 2022. Disponível em: <https://doi.org/10.36253/jlis.it-482>. Acesso em: 10 jul. 2024.

CARVALHO, V. A.; ALMEIDA, J. P. A.; FONSECA, C. M.; GUIZZARDI, G. Extending the Foundations of Ontology-Based Conceptual Modeling with a Multi-level Theory. In: JOHANNESON, P.; LEE, M.; LITTLE, S.; OPDAHL, A.; PASTOR LÓPEZ, Ó. (ed.) **Conceptual Modeling**. Lecture Notes in Computer Science. v. 9381, p. 119-133, 2015. Disponível em: https://doi.org/10.1007/978-3-319-25264-3_9. Acesso em: 10 jul. 2024.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

CÉ, G.; FLORES, D.; CAMPOS, L. M. A Visão do Records in Contexts (RIC) como Objeto de Fronteira de um Grupo de Profissionais da Arquivologia no Brasil e do EGAD/ICA. **Revista Brasileira De Biblioteconomia e Documentação**, v. 19, p. 1-26. 2023. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1874>. Acesso em: 12 jul. 2024.

FELICIATI, P. Call Me by Your Name: Towards an Authority Data Control Shared Between Archives and Libraries. *In*: BERGAMIN, G.; GUERRINI, M. (ed.) **Bibliographic Control in the Digital Ecosystem**. Roma: Associazione italiana biblioteche; Macerata: Edizioni Università di Macerata; Firenze: Firenze University Press, 2022, p. 203-214.

GOMES, D. L.; BARROS, T. H. B.; SOUSA, R. T. B.; SANTOS JUNIOR, R. L. Proposta de uma ferramenta para classificação arquivística com base em ontologias. **Em Questão**, Porto Alegre, v. 26, n. 1, p. 351-374, 2019. Disponível em: <https://doi.org/10.19132/1808-5245261.351-374>. Acesso em: 13 jul. 2024.

INTERNATIONAL COMMITTEE FOR DOCUMENTATION (CIDOC). **Definition of the CIDOC Conceptual Reference Model**. Version 7.1.3. [S. l.]: ICOM-CIDOC, 2024. Disponível em: <https://www.cidoc-crm.org/Version/version-7.1.3>. Acesso em: 10 jul. 2024.

INTERNATIONAL COUNCIL ON ARCHIVES (ICA). **International Council on Archives Records in Contexts Conceptual Model (ICA RiC-CM)**. Versão 1.0. [S. l.]: ICA, 2023. Disponível em: <https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf>. Acesso em: 10 jul. 2024.

KOCH, I.; LOPES, C. T.; RIBEIRO, C. **Moving from ISAD(G) to a CIDOC-CRM-based Linked Data Model in the Portuguese Archives**. [S. l.]: [S. n.], 2023. Disponível em: <https://doi.org/10.48550/arXiv.2310.11140>. Acesso em: 10 jul. 2024.

LÖW, M.; ROCHA, R. P.; BARROS, T. H. B. Representação do Conhecimento Arquivístico: uma Análise Comparativa entre Entidades de Records in Contexts Ontology e Entidades Derivadas de Basic Formal Ontology. *In*: TOGNOLI, N. B.; ALBUQUERQUE, A. C.; CERVANTES, B. M. N. (ed.). **Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na Era da Datificação**: Londrina: ISKOBrasil; PPGCI-UEL, 2023. Disponível em: https://isko.org.br/wp-content/uploads/2023/06/livro-isko-Brasil_23.pdf. Acesso em: 10 jul. 2024.

MIKHAYLOVA, D.; METILLI, D. Extending RiC-O to Model Historical Architectural Archives: the ITDT Ontology. **Journal on Computing and Cultural Heritage**, [S. l.], v. 16, n. 4, p. 1-15, 2023. Disponível em: <https://doi.org/10.1145/3606706>. Acesso em: 10 jul. 2024.

SIGNORE, O. Representing knowledge in archaeology: from cataloguing cards to Semantic Web. **Archeologia e Calcolatori**, [S. l.], v. 20, p. 111-128, 2009. Disponível em: <http://www.archcalc.cnr.it/journal/id.php?id=508>. Acesso em: 10 jul. 2024.

VERDONCK, M.; GAILLY, F.; PERGL, R.; GUIZZARDI, G.; SOUZA, B. F. M.; PASTOR, O. Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. **Information Systems**, v. 81, p. 92-103, Mar. 2019. Disponível em: <https://doi.org/10.1016/j.is.2018.11.009>. Acesso em: 10 jul. 2024.

WAND, Y.; STOREY, V. C.; WEBER, R. An Ontological Analysis of the Relationship Construct in Conceptual Modeling. **ACM Transactions on Database Systems**, [S. l.], v. 24, n. 4, p. 494-528, 1999. Disponível em: <https://doi.org/10.1145/331983.331989>. Acesso em: 10 jul. 2024.