



Projeto ICON - Integração de CONteúdos: para um novo modelo de descrição

Francisco Barbedo^a, Lucília Runa^b, Maria José de Almeida^c

^aDGLAB, DSIAE, Portugal, francisco.barbedo@dglab.gov.pt

^bDGLAB, DSIAE, Portugal, lucilia.runa@dglab.gov.pt

^cDGLAB, DSIAE, Portugal, m-jose.almeida@dglab.gov.pt

Resumo

A descrição arquivística percorreu um longo caminho desde a publicação da *ISAD(G)* pelo Conselho Internacional de Arquivos, nem sempre isento de problemas, como a compatibilização com novos conceitos de sociedade tecnologicamente mediada.

A massificação da Internet e a web semântica determinam a representação de conhecimento num universo de discurso cada vez mais alargado, esbatendo barreiras entre comunidades de prática: o mundo tornou-se relacional, e do estabelecimento e contextualização dessas relações depende a eficácia da descoberta/reutilização da informação.

Assim surgiu, na DGLAB, o projeto *ICON*, que analisou modelos alternativos aos da *ISAD(G)* e normas conexas que permitissem ampliar a capacidade de representação da realidade de forma conectada e pluralmente relacionada.

A *ISO 21127:2014 - CIDOC-CRM*, surgiu como muito promissora, com potentes capacidades de representação do universo de discurso cultural, e em consonância como novo modelo de descrição, ainda em fase de desenvolvimento, do ICA - RiC-CM.

O projeto incluiu a elaboração de modelos de dados baseados no CIDOC-CRM, adaptados à realidade da DGLAB. Relata-se o processo de análise e translação de um modelo para outro, respeitando os princípios arquivísticos, sem comprometer a capacidade relacional inerente a uma maior compreensão do mundo no qual os arquivos se integram.

Palavras-chave

arquivos, descrição arquivística, modelos de dados, normas de descrição

Introdução

O projeto *ICON* (Integração de CONteúdos), desenvolvido pela Direção Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB) desde 2017, tem como objetivo proceder à renovação da arquitetura dos seus sistemas de informação que suportam a descrição arquivística e serviços conexas. O ponto de partida assenta em sistemas distintos que se pretende agregar:

1. *DigitArq-CRAV* – descrição da documentação de arquivo, de ações de preservação e restauro, incluindo serviços, tradicionalmente presenciais, em modo de balcão eletrónico (pesquisa e consulta de documentos digitalizados, pedido de reproduções digitais, reserva de documentos para leitura presencial, pedido de certificados, etc.);

2. *Portal Português de Arquivos* (PPA) – disponibilização, num único ponto centralizado, de metainformação descritiva e imagens da documentação detida pelos aderentes à Rede Portuguesa de Arquivos (RPA), facilitando a sua pesquisa e recuperação;
3. *Ficheiro Nacional de Autoridades Arquivísticas* (FNAA) – descrição dos produtores/ detentores da documentação de arquivo e respetivas funções.

Desenvolvidos ao longo das últimas décadas, apresentam diferentes arquiteturas, baseadas nas normas de descrição arquivística emanadas pelo International Council on Archives (ICA): ISAD(G), ISAAR(CPF), ISDIAH e ISDF (CIA/ICA 2002, 2004, 2007, 2008). Esta circunstância condicionou a recolha e processamento de informação, já que as normas apresentam áreas de sobreposição e redundâncias, criando algumas dificuldades de articulação da informação registada em sistemas distintos.

A análise das normas do ICA tinha já sido efetuada no âmbito da então Direção-Geral de Arquivos (DGARQ), que produziu um documento de problematização como ponto de partida para a elaboração da 3.ª versão das *Orientações para a Descrição Arquivística* (DGARQ, 2011). Dessa análise destaca-se a constatação da ausência de definição prévia de um modelo de descrição por parte do ICA, facto do qual decorrem limitações do atual modelo, como a delimitação vaga do objeto de cada uma das normas, ou a necessidade de desenvolvimento de formas de articulação explícitas entre elas.

A aposta na representação relacional encontra-se expressa em todas elas, à exceção da ISAD(G), que não apresenta zona de relações. A sua ausência na ISAD(G) limita a criação e caracterização de relações entre os diferentes objetos de descrição. Os elementos *Unidades de descrição relacionadas* e *Nome do produtor*, permitem identificar relações entre agregações de documentos e/ou documentos, ou entre estes e o respetivo produtor, mas não permitem estabelecer relações autónomas, dotadas de atributos próprios. O mesmo se verifica para a associação da documentação a entidades geográficas, eventos ou funções/atividades, por exemplo.

Frequentemente, nas quatro normas do ICA, as relações estão implícitas, em elementos cuja informação é registada em texto narrativo. É o caso de *História administrativa/biográfica/familiar*, *História custodial e arquivística*, *Fonte imediata de aquisição ou transferência*, *Âmbito e conteúdo*. A atomização desta informação permite, para além de uma maior normalização no registo e de uma mais eficaz recuperação, explicitar e caracterizar relações atualmente apenas implícitas.

Um outro aspeto identificado prende-se com o facto de a aplicação estrita da grelha de elementos de descrição das normas do ICA, nomeadamente da ISAD(G), sobretudo nos níveis mais baixos de descrição, levar frequentemente a que, num mesmo elemento, seja reunida informação de diferente natureza, com as consequências que daí decorrem para a sua normalização, recuperação e coerência global do modelo de descrição.

Por outro lado, partindo do pressuposto da pertinência de definir elementos específicos para documentos com características especiais, põe-se a questão de como os seleccionar, já que a ISAD(G) é uma norma geral. Coloca-se a questão de como articular a realidade dos documentos convencionais com a dos documentos eletrónicos, mantendo, em simultâneo, a coerência entre as diferentes grelhas de elementos de informação propostos pelas normas gerais e específicas.

Esta reflexão surge no contexto das análises levadas a cabo por diferentes países (Comisión de Normas Españolas de Descripción Archivística, 2017; National Archives of Finland, 2013), e também pelo ICA (ICA, 2012), que mais recentemente colocou em consulta pública o RiC-CM (ICA, 2016), que introduz um conceito de modelação de dados que “pode ser definido como ‘descrição multidimensional’”. Mais do que uma hierarquia, a descrição pode ser representada como um grafo ou uma rede. A modelação de

dados em forma de grafo acomoda a descrição multinível – com base no conceito de fundo –, contemplada na ISAD(G), mas também permite abordar uma compreensão mais ampla do conceito de proveniência. O modelo multidimensional possibilita assim a descrição do fundo, mas também da existência do fundo num contexto alargado, em relação com outros fundos.” (ICA, 2016, p.10 [tradução livre]).

Saliente-se que o RiC-CM mantém um apreciável número de propriedades que preveem o registo em texto narrativo: RiC-P4 General Note, RiC-P5 Authenticity and Integrity Note, RiC-P7 Content Extent, RiC-P8 Quality of Information, RiC-P9 Scope and Content, etc.

No entanto o RiC-CM, que se assume como versão preliminar do modelo conceptual de descrição, que deveria ter sido acompanhado pela divulgação de uma proposta de ontologia (RiC-O) no final de 2016, continua sem desenvolvimento, ignorando-se os resultados da discussão pública, da qual se conhecem apenas alguns contributos divulgados pelos próprios autores. Assim, as quatro normas disponíveis continuam a apontar para uma lógica vertical (quer considerando a área a que se reportam – apenas arquivos – quer os objetos a representar – documentação, produtores, detentores e funções), caracterizadas por alguma articulação, mas por pouca ou nenhuma integração.

Esta lógica encontra-se em contra-ciclo com as tendências de novos contextos sociais e tecnológicos que deveriam pautar um novo modelo de descrição arquivística - *web* semântica, interoperabilidade e complementaridade entre sistemas -, visando assegurar a pesquisa integrada de recursos, independentemente da sua natureza ou da entidade que os detém (arquivos, bibliotecas ou museus), já que os atuais utilizadores privilegiam tendencialmente o acesso remoto e não mediado aos conteúdos, em contraste com o utilizador presencial, que tradicionalmente constituía o público-alvo dos arquivos.

O projeto ICON visa responder a estes desafios, definindo um novo modelo de dados que articule e evite a redundâncias entre os atuais sistemas. Simultaneamente procura definir as melhores formas de acesso e recuperação da informação, identificar novas funcionalidades e serviços a integrar.

Método

A adoção de um novo modelo conceptual, destinado a utilizadores internos e externos da rede DGLAB, levou ao envolvimento de ambos no projeto, através da promoção da discussão e do debate no âmbito de reuniões realizadas para o efeito.

Nesse sentido, procedeu-se à análise de modelos alternativos às normas do ICA, que permitissem ampliar a capacidade de representação da realidade de forma conectada e pluralmente relacionada. A ontologia CIDOC-CRM (ISO 21127:2014) emergiu como uma alternativa viável, planeada para representar informação e conhecimento num amplo universo de discurso, particularmente vocacionado para a cultura. A este respeito será interessante reter as referências a património cultural expressas na respetiva Lei n.º 107/2001 de 8 de setembro, que claramente apontam para uma dimensão muito abrangente e inclusiva, refletindo uma realidade holística, em que a cultura é ubíqua a todas as áreas da atividade humana.

Assim, este modelo conceptual potenciará a interoperabilidade semântica dos sistemas de informação da DGLAB, permitindo aos utilizadores aceder, a partir de um único ponto centralizado, à informação disponibilizada por diferentes entidades e domínios culturais. A hiperconectividade que caracteriza a sociedade atual contraria a visão tradicional de separação de conteúdos consoante áreas disciplinares ou entidades responsáveis pela custódia da informação. Uma ontologia tão abrangente como o CIDOC-CRM permitirá articular o sistema de informação da DGLAB com outros domínios culturais – arte, arqueologia, história natural, arquivos e literatura, ciência e investigação, etc. – e enriquecer a representação das múltiplas relações entre os respetivos objetos de informação.

Definida a ontologia CIDOC-CRM como base para o novo modelo de dados, foram realizados mapeamentos exaustivos para o conjunto das quatro normas de descrição arquivística do ICA, bem como do RiC. Foram igualmente consideradas nos mapeamentos as normas EAD (SAA, 2018); EAC-CPF (SSA, SBB, 2014); EAG (APEX, 2015) bem como outras mais abrangentes, usados por várias comunidades de prática no âmbito do património cultural: EDM (EUROPEANA, 2017); DC (ISO 15836-1:2017).

Este trabalho de mapeamento esteve na base das experiências práticas de transposição, para CIDOC-CRM, de descrições arquivísticas de diferentes tipos de objetos de informação: documentos, entidades produtoras e/ou detentoras, entidades geográficas, períodos cronológicos, funções e respetivas relações. Os casos práticos foram escolhidos no âmbito da discussão do projeto com os técnicos da Torre do Tombo e dos Arquivos Distritais, para demonstrar a exequibilidade de denotar, do ponto de vista descritivo, os elementos utilizados no modelo vigente.

Procurou-se ainda exemplificar a lógica de aplicação da entidade “evento” e a sua integração com os tradicionais elementos de descrição. Este exercício permitiu identificar um conjunto de questões decorrentes da sua aplicação prática, cuja análise se revelou fundamental para a definição do novo modelo.

Foram igualmente alvo de mapeamento registos descritivos produzidos por outras comunidades de prática, nomeadamente bibliotecas e museus, tendo-se confirmado a capacidade de resposta do CIDOC-CRM às necessidades específicas dos diferentes domínios culturais.

Importa salientar que, se do ponto de vista conceptual, esta escolha se revelou a mais interessante, a alteração do atual modelo de dados está condicionada à sua exequibilidade tecnológica e à possibilidade de acomodar os dados residentes no atual sistema Digitalq, que terão que migrar para a nova solução a adotar. Neste âmbito, a DGLAB candidatou o desenvolvimento do projeto ICON a financiamento no âmbito do sistema de apoio à transformação digital da Administração Pública (SAMA2020). Como entidade parceira, integra também um projeto de Investigação Científica e Desenvolvimento Tecnológico em Ciência dos Dados e Inteligência Artificial na Administração Pública (EPISA) promovido pelo Instituto de Engenharia de Sistemas e Computadores do Porto (INESC Porto/FE/UP), candidatado a financiamento através da Fundação para a Ciência e a Tecnologia. Tem como objetivo a construção de um protótipo que suporte o novo modelo de dados e o desenvolvimento de ferramentas de migração semântica, a testar com os dados do sistema Digitalq. De momento, ambos os projetos aguardam decisão formal quanto a financiamento.

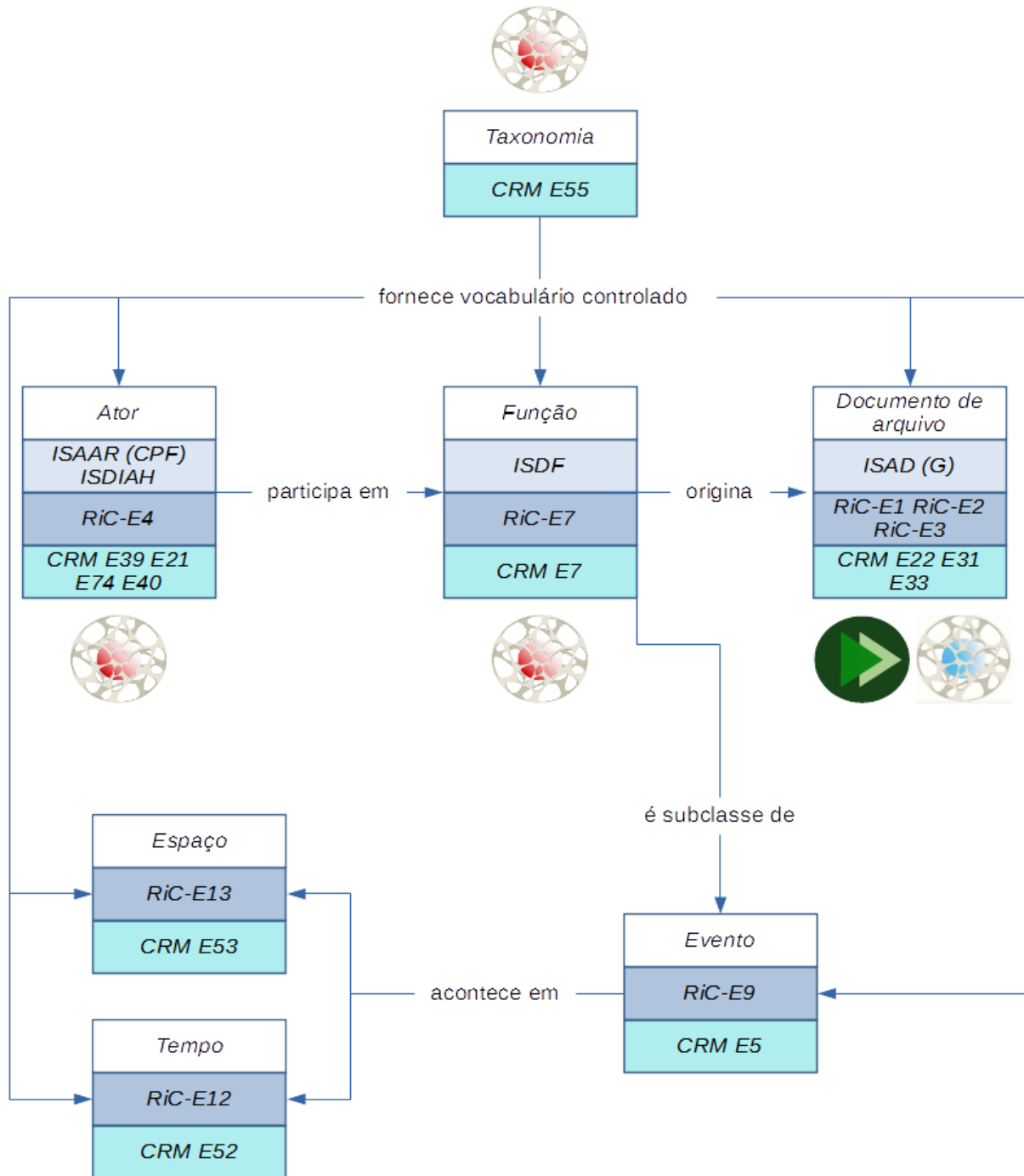
Resultados e Discussão

No decurso da análise da aplicação da ontologia CIDOC-CRM à descrição arquivística, uma das questões discutidas passou pelo facto de, para os diferentes objetos a descrever, existir um conjunto apreciável de elementos de informação cujos conteúdos terão que ser modelados através do recurso à classe CRM E55 *Type* e da respetiva propriedade CRM P2 *has type (is type of)*.

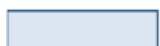
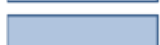

Um outro aspeto a considerar passa pelo facto de deixar de ser possível associar, de forma direta, um período cronológico, um local, um produtor (ou outro tipo de ator), diretamente a um documento e/ou a uma agregação documental (*fundo, secção, série, documento simples ou composto*). Uma vez que o CIDOC-CRM se centra no conceito de evento, que ocorre num determinado período cronológico e num determinado local, é da classe CRM E5 *Event* que decorre a produção de documentos e/ou a constituição de agregações documentais. Assim, será ao evento que irão associar-se um ou vários atores (CRM E39), pessoas coletivas (CRM E40), singulares (CRM E21) ou famílias (CRM E74), por nele terem, de alguma forma, participado (na qualidade de produtores, autores, detentores, etc.). Por sua vez os atores também

serão associados entre si não de forma direta, mas através dos eventos que moldam as suas relações: o nascimento (CRM E67), por exemplo, determina o estabelecimento de relações de maternidade e paternidade através das propriedades associadas (CRM P96 *by mother [gave birth]* ou P97 *from father [was father for]*)

Na figura 1 podem observar-se as classes CIDOC-CRM (e correspondentes entidades RiC) a utilizar para modelar a descrição dos objetos contemplados nas ISAAR(CPF)/ISDIAH, ISDF e ISAD(G), normas que constituem a base dos sistemas de informação da DGLAB a renovar. Para modelar, através do recurso ao CIDOC-CRM, a totalidade da informação prevista em cada uma dessas normas, é necessário recorrer a outras classes e propriedades do CIDOC-CRM, já identificadas de forma exaustiva, mas cuja análise não cabe no âmbito da presente comunicação.



Normas

-  ICA normas
-  ICA – RiC (draft)
-  ICOM CIDOC-CRM

Sistemas de informação a renovar




-  FNAA
-  Digitarq
-  PPA

Figura 1: Transição do modelo de dados subjacente aos atuais sistemas de informação DGLAB para o novo modelo baseado no CIDOC-CRM

Considerando o papel assumido pelo evento, tornou-se premente clarificar o conceito subjacente a esta classe CRM, sendo de destacar que tanto pode ter origem natural (o terremoto de 1755, por exemplo) como humana (o ato que deu origem à produção de um documento, por exemplo). Hierarquiza-se em subclasses, das quais a que melhor se adequa à realidade dos objetos de descrição arquivística é a *atividade* (CRM E7). Procurou-se exemplificar a lógica da sua aplicação e a sua integração com os tradicionais elementos de descrição.

A definição do CIDOC-CRM de *atividade* (E7) enquadra as noções de função, processo de negócio e atividade, considerada como subunidade do processo de negócio. A autonomização do evento enquanto *atividade*, no sentido que lhe é atribuído pelo CIDOC-CRM, permite o seu ulterior relacionamento com os demais objetos de informação e o enriquecimento semântico da descrição. Neste último caso, a maior dificuldade reside na constituição de uma estrutura vocabular normalizada para designar eventos, dada a sua natureza virtualmente infinita, ao contrário do que já acontece para o universo de discurso funcional ou tematicamente delimitado.

Na perspetiva do arquivo, a aplicabilidade da modelação da descrição centrada no evento ficou demonstrada para os níveis mais altos, mas colocou problemas nos mais baixos. Para documentos como a fotografia, por exemplo, o problema reside na seleção de um evento significativo e pertinente que contextualize, de forma singular, a produção do documento, pois seria pouco útil, porque universalmente aplicável, considerar o disparo do obturador o evento contextualizador do documento fotográfico.

A utilização das classes e propriedades do CIDOC-CRM mostrou-se eficaz na modelação das relações hierárquicas que caracterizam o atual modelo de descrição. Ficou igualmente demonstrada a possibilidade de, para um mesmo documento, produzir e gerir diversas descrições, correspondentes às suas diferentes manifestações físicas (caso dos documentos fotográficos, ou registados em diferentes tipos de suporte).

O novo modelo, caracterizado pela flexibilidade no relacionamento dos diferentes objetos, incluindo o evento, permite integrar outras valências, como o registo de ações de conservação e restauro, preservação digital ou balcão de serviços.

A utilização do CRM implica uma muito maior atomização da informação, aspeto determinante para a normalização de conteúdos, por diminuir, embora não eliminando, a subjetividade no registo; para aumentar as possibilidades de relacionamento entre a informação registada de forma atomizada; para possibilitar a constituição de uma rede (*linked data*) semanticamente rica, já que para cada “átomo” podem ser estabelecidas várias relações autónomas.

Destaca-se igualmente a possibilidade do modelo relacional permitir a representação autonomizada de partes ou componentes de documentos - um selo ou uma iluminura, por exemplo - sem a condicionante do modelo hierárquico, que prevê um número limitado de níveis de descrição. Esta flexibilidade permite representar qualquer unidade lógica ou física de um objeto, independentemente do nível intelectual que ocupe na hierarquia descritiva. No atual modelo hierárquico, o nível mais baixo utilizável é o de *documento*, pelo que a descrição de uma das suas componentes não se processa autonomamente, sendo remetida para o elemento genérico *Âmbito e conteúdo*, que acomoda diferentes tipologias de informação.

A micro-contextualização de um documento no processo organizacional em que se integra é igualmente difícil de representar no atual modelo hierárquico. Tome-se como exemplo um projeto de arquitetura, normalmente compartimentado em eventos sequenciais: integração urbanística, anteprojecto, orçamento, peças desenhadas, etc. A descrição de cada atividade que o compõe não é fácil de representar hierarquicamente, já que a sua expressão documental pode não ser enquadrável nos níveis descritivos admitidos pela ISAD(G). A sua contextualização nas diversas fases do processo é facilitada recorrendo

ao modelo relacional, que permite ainda o relacionamento autónomo de cada parte do processo com outros objetos ou eventos.

Outra questão identificada passa pela dificuldade de assegurar, por escassez de recursos e de tempo, a profundidade e exaustividade das descrições. O CIDOC-CRM, tal como o modelo hierárquico, não condiciona estes aspetos. A profundidade (o termo *expansão* será eventualmente mais apropriado para um modelo relacional) e a exaustividade serão determinadas pelas políticas adotadas, não dependendo do modelo conceptual aplicado. Por outro lado, o sistema ganhará em estar preparado para receber descrições muito especializadas, coligida segundo uma lógica não exclusivamente arquivísticas, por utilizadores externos ou no âmbito de projetos de investigação. O processo de *crowdsourcing* é independente do modelo CRM, no entanto a sua operacionalidade é facilitada pela flexibilidade do modelo em integrar descrições elaboradas com base noutra modelo conceptual, ou sem qualquer modelo conceptual subjacente.

A utilização de thesaurus e/ou vocabulários controlados, considerada desejável no âmbito do atual modelo hierárquico, foi considerada condição necessária para a plena utilização de um modelo baseado no CIDOC-CRM. O *Ficheiro Nacional de Autoridades Arquivísticas* (FNAA) constitui um exemplo, incluindo registos de autoridade para produtores/detentores e funções. Neste domínio, seria interessante localizar ou construir repositórios cujos recursos, independentemente da sua tipologia, pudessem ser relacionados com os arquivísticos.

O evento (CRM E5) e particularmente a atividade (CRM E7) – que compreende a função, a atividade e o processo de negócio –, constituem provavelmente a estrutura mais complexa. Neste âmbito, foi realizado um teste de aplicação da *Macro Estrutura Funcional* (MEF) (DGLAB, 2013) à realidade de fundos documentais produzidos por entidades extintas, com o objetivo de avaliar:

- A adequação da MEF enquanto repositório de referência para descrição de funções e processos de negócio em contextos histórico;
- O grau de dificuldade na associação de funções a fundos documentais, bem como as respetivas segmentações hierárquicas;
- A dificuldade na eventual construção de uma estrutura de funções adequada à descrição de conjuntos documentais históricos.

Os resultados destes testes permitiram concluir que:

- As classes contidas na MEF são aplicáveis a contextos históricos, tendo sido reconhecida a existência de funções transversais que perduram na longa duração;
- Existem *Âmbitos e conteúdos* de funções e sub funções que necessitam de adaptação para se adequarem ao âmbito das funções em contexto histórico;
- O processo de negócio – terceiro e/ou quarto níveis da *Avaliação Supra institucional da Informação Arquivística* (ASIA) (DGLAB, 2016)–, não parece ter aplicabilidade em contextos históricos, já que o fluxo e a forma de um processo no âmbito de uma organização extinta (histórica) pode diferir dos atuais. Tal não exclui, no entanto, a pertinência do enquadramento funcional, caso se considere estritamente a função e não a forma como era exercida;
- A MEF não cobre funções desempenhadas por entidades privadas, já que foi elaborada para a Administração Pública. Existem, no entanto, áreas de sobreposição, tal como há funções atualmente exercidas pelo Estado que já foram desempenhadas por privados, podendo existir afinidade;

- A identificação de funções é mais complexa aos níveis mais baixos de descrição, pelo que a flexibilidade de associação deverá ser definida prevendo essa realidade. É possível considerar a identificação de temáticas, assuntos ou tópicos que, posteriormente, possam conduzir à criação de uma função e/ou de um evento, mediante a identificação de um padrão reconhecível. Note-se que o CIDOC-CRM é transitivo, o que significa que as funções associadas aos níveis superiores transitam para os níveis dependentes, exceto mediante a especialização funcional dos sub níveis. A título de exemplo, mencione-se que ao relacionar um fundo com as funções de “administração da justiça” e “formalização de atos e registos”, estas são automaticamente herdadas pelos níveis inferiores, mas é possível proceder-se à especialização de um subnível, vinculando-o apenas a uma das funções acima mencionadas.

Do exposto decorre que a MEF pode ser utilizada em contextos históricos enquanto repositório de funções, desde que adaptada e aumentada de forma incremental. Uma alternativa passa pelo recurso a termos de indexação, pese embora a ausência de estruturas vocabulares disponíveis para esse efeito.

A questão do anacronismo, subjacente à utilização da MEF em contexto histórico, tem que ser considerada numa dupla vertente:

- terminológica - apenas varia o termo utilizado para referir uma função, sendo possível associar o conceito atual aos conceitos antigos e identificar uma forma preferencial de nome;
- conceptual - a conceção coeva da função difere da atual, o que configura uma situação mais complexa e implica uma análise cuidadosa. É o caso, por exemplo, de determinar se a função *Administração da justiça* é aplicável ao Tribunal do Santo Ofício.

Existem vários repositórios contendo terminologia e conceitos autorizados ou, pelo menos, controlados. É, internamente, o caso do FNAA. Mas será possível encontrar, fora da DGLAB, em contexto nacional e internacional, outros repositórios passíveis de utilização. Os pontos de acesso geográficos constituem um exemplo. Será ainda possível criar regras para a definição de termos de indexação para assuntos, eventualmente utilizáveis aos níveis mais baixos de descrição.

Conclusões

O projeto ICON pretende responder aos desafios enunciados recorrendo a uma norma reconhecida internacionalmente. A maleabilidade do novo modelo de dados permitirá uma abordagem transversal às diferentes comunidades de prática centradas na área do património cultural. Sendo uma ontologia de base relacional, o CIDOC-CRM é compatível com tecnologias web, nomeadamente web semântica e *linked data*, fatores que vão ao encontro do interesse dos utilizadores, que procuram recursos informativos independentemente da comunidade de prática que os detém, preferencialmente com recurso à web e através de um único ponto centralizado, acessível através de qualquer dispositivo, fixo ou móvel.

A utilização do CIDOC-CRM implica uma abordagem completamente diferente da descrição arquivística, o que se refletirá na nova arquitetura do sistema, bem como na exigência dos testes a que este deverá ser submetido, para garantir a capacidade de resposta a utilizadores internos e externos.

Um dos principais objetivos do projeto reside no aumento da eficiência do funcionamento das unidades orgânicas da DGLAB, nos serviços centrais e desconcentrados. A simplificação de processos, o reforço da articulação e colaboração entre serviços, a rentabilização de recursos, evitará a duplicação de tarefas, promovendo o registo normalizado da informação a produzir, a sua melhor contextualização e divulgação seletiva, a sua rápida recuperação, partilha e reutilização.

Referências

- APEX PROJECT. WORK PACKAGES 2 & 4 (2015) - *EAG 2012: Encoded Archival Guide*. [Em linha]. [Den Haag]: APEX. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<http://www.apex-project.eu/index.php/en/outcomes/standards/eag-2012>>.
- CONSELHO INTERNACIONAL DE ARQUIVOS (2002) - *ISAD(G): Norma geral internacional de descrição arquivística: adoptada pelo Comité de Normas de Descrição, Estocolmo: Suécia, 19-22 de Setembro de 1999*. Trad. Grupo de Trabalho para a Normalização da Descrição em Arquivo. 2.ª ed. [Em linha]. Lisboa: IAN/TT. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<http://arquivos.dglab.gov.pt/servicos/documentos-tecnicos-e-normativos/lista-de-documentos/>>.
- CONSELHO INTERNACIONAL DE ARQUIVOS (2004) – *ISAAR(CPF): Norma internacional de registos de autoridade arquivística para pessoas colectivas, pessoas singulares e famílias*. Trad. Grupo de Trabalho para a Normalização da Descrição em Arquivo. 2.ª ed. [Em linha]. Lisboa: IAN/TT. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<http://arquivos.dglab.gov.pt/servicos/documentos-tecnicos-e-normativos/lista-de-documentos/>>.
- DIREÇÃO-GERAL DE ARQUIVOS. GRUPO DE TRABALHO DE NORMALIZAÇÃO DA DESCRIÇÃO EM ARQUIVO (2011) - *Orientações para a Descrição Arquivística*. 3.ª versão. Lisboa: DGARQ. ISBN 978-972-8107-91-8.
- DIREÇÃO GERAL DO LIVRO, DOS ARQUIVOS E DAS BIBLIOTECAS (2013) - *Macroestrutura funcional (MEF)*. Versão 2.0. [Em linha]. Lisboa: DGLAB. [Consult. 9 jul. 2018]. Disponível na Internet: <URL:http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2013/10/2013-03-28_MEF-v2_0.pdf>.
- DIREÇÃO GERAL DO LIVRO, DOS ARQUIVOS E DAS BIBLIOTECAS (2016) - *Avaliação Supra institucional da Informação Arquivística (ASIA): documento metodológico*. Versão 1. [Em linha]. Lisboa: DGLAB. [Consult. 9 jul. 2018]. Disponível na Internet: <URL:http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2016/03/ASIA_Doc-metodologico2016-03-10.pdf>.
- EUROPEANA (2017) - *Definition of the Europeana Data Model*. [Em linha]. V. 5.2.8. Den Haag: Europeana Foundation. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<https://pro.europeana.eu/resources/standardization-tools/edm-documentation>>.
- INTERNATIONAL COUNCIL ON ARCHIVES (2007) – *ISDF: International Standard for Describing Functions: developed by the Committee on Best Practices and Standards, Dresden, Germany, 2-4 May 2007* [Em linha]. 1.ª ed. Paris: ICA. [Consult. 7 mar. 2018]. Disponível na Internet: <URL:<https://www.ica.org/en/isdf-international-standard-describing-functions>>.
- INTERNATIONAL COUNCIL ON ARCHIVES (2008) – *ISDIAH: International Standard for Describing Institutions with Archival Holdings: Developed by the Committee on Best Practices and Standards, London, United Kingdom, 10-11 March 2008*. [Em linha]. Paris: ICA. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<https://www.ica.org/en/public-resources/standards>>.
- INTERNATIONAL COUNCIL ON ARCHIVES. COMMITTEE ON BEST PRACTICES AND STANDARDS (2012) - *Progress report for revising and harmonising ICA descriptive standards*. [Em linha]. S.l., ICA. [Consult. 6 jul. 2018]. Disponível na Internet: <URL:<https://www.ica.org/en/cbps-progress-report-revising-and-harmonising-ica-descriptive-standards>>.
- INTERNATIONAL COUNCIL ON ARCHIVES. EXPERTS GROUP ON ARCHIVAL DESCRIPTION (2016) - *Records in Context: a conceptual model for archival description*. Consultation Draft, v 0.1. [Em linha]. [S.l.]: ICA. [Consult. 28 out. 2016]. Disponível na Internet: <URL:<https://www.ica.org/en/egad-ric-conceptual-mode>>.

ISO 15836-1:2017, Information and documentation - *The Dublin Core metadata element set*. Genebra: International Standard Organization.

ISO 21127:2014, Information and documentation - *A reference ontology for the interchange of cultural heritage information*. 2. ed. Genebra: International Standard Organization.

Lei n.º 107/2001. Diário da República n.º 209/2001, Série I-A de 2001-09-08. Estabelece as bases da política e do regime de proteção e valorização do património cultural.

MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE. COMISIÓN DE NORMAS ESPAÑOLAS DE DESCRIPCIÓN ARCHIVÍSTICA (2017) - *Neda-MC: Modelo conceptual de descripción archivística: entidades, relaciones y atributos*. [Em linha]. [Madrid]: CNEDA. [Consult. 9 jul. 2018]. Disponível na Internet: <URL:http://www.mecd.gob.es/dam/jcr:f8d73e08-dc13-4dcf-85ed-41f940f0d602/20170505_NEDA-MC.pdf>.

NATIONAL ARCHIVES OF FINLAND (2013) - *Finnish conceptual model for archival description. Draft version 0.1*. Helsinki: National Archives of Finland.

SOCIETY OF AMERICAN ARCHIVISTS; STAATSBIBLIOTHEK ZU BERLIN. TECHNICAL SUBCOMMITTEE FOR ENCODED ARCHIVAL CONTEXT (2014) - *Encoded Archival Context: Corporate Bodies, Persons, and Families (EAC-CPF)*. [Em linha]. Berlin: Staatsbibliothek zu Berlin. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<http://eac.staatsbibliothek-berlin.de/schemata-and-tag-library/>>.

SOCIETY OF AMERICAN ARCHIVISTS. TECHNICAL SUBCOMMITTEE FOR ENCODED ARCHIVAL STANDARDS (2018) - *Encoded Archival Description*. [Em linha]. Version EAD3 1.1.0. Chicago: SAA. [Consult. 3 jul. 2018]. Disponível na Internet: <URL:<https://www.loc.gov/ead/>>.