

INTEROPERABILIDADE SEMÂNTICA EM PATRIMÔNIO CULTURAL: ESTRATÉGIAS E FERRAMENTAS PARA A CONCEPÇÃO DE UM CENTRO DE MEMÓRIA

SEMANTIC INTEROPERABILITY IN CULTURAL HERITAGE: STRATEGIES AND TOOLS FOR DESIGNING A MEMORY CENTER

Alison Carlos Filgueiras

Universidade Fernando Pessoa/ Universidade Estadual de Goiás
alisoncf@gmail.com

Feliz Ribeiro Gouveia

Universidade Fernando Pessoa
fribeiro@ufp.edu.pt

Paula Roberta Chagas

Universidade Fernando Pessoa/ Universidade Estadual de Goiás
paula.chagas@ueg.br

Resumo: A interoperabilidade semântica entre repositórios de patrimônio cultural tem se mostrado uma necessidade em face ao aumento de acervos digitais de galerias de arte, bibliotecas digitais, arquivos e museus na *Web*. Esse aumento, aliado às características da Sociedade em Rede (CASTELLS, 2016), gerou uma demanda de intercâmbio de informações que requer um grande esforço no que tange organização, armazenamento e recuperação da informação para produção do conhecimento. Iniciativas de coletas de e padrões de metadados têm um papel fundamental na disseminação das informações, no entanto, apresentam limitações no que diz respeito à semântica e tem perdido a motivação para seu uso (FREIRE et al., 2018b). No tocante ao patrimônio cultural há um aparato tecnológico semântico disponível para a integração desses repositórios que sugere um futuro promissor para a integração de espaços como os Centros de Memória, que envolve ontologias, além de linguagens e outras ferramentas para armazenamento e recuperação da informação em triplas. O objetivo deste trabalho é trazer à discussão os aspectos tecnológicos e estratégias possíveis para criação de um centro de memória à luz dos Dados Abertos Conectados (*Linked Open Data*). A metodologia utilizada consiste de um estudo descritivo e exploratório de fontes bibliográficas e tecnologias disponíveis. Como resultado apresenta-se uma proposta inicial de arquitetura para o Centro de Memória de Goiás.

Palavras-Chave: Interoperabilidade; *Web* Semântica; Patrimônio Cultural; Arquivos Abertos Conectados; Centro de Memória.

Abstract: Semantic interoperability among cultural heritage repositories has been shown to be a necessity in the face of increasing digital collections of art galleries, digital libraries, archives, museums (GLAMs) on the *Web*. demand for information exchange that requires a great deal of effort regarding the organization, storage, availability, access and retrieval of information for the production of knowledge. Metadata collection initiatives and metadata standards such as OAI-PMH and Dublin Core have a key role in disseminating information, however, they have limitations with regard to semantics. As regards cultural heritage, there is a semantic technological apparatus available for the integration of these repositories, which suggests a promising future for

the integration of so-called GLAMs, which involves ontologies such as CIDOC-CRM, ICOM and EDM of Europeana, as well as languages and APIs and others tools for storing and retrieving RDF information. The purpose of this article is to discuss this technological apparatus by analyzing the characteristics of each of the technologies involved, and in the light of open data connected (*Linked Open Data*) to discuss questions related to the design of a Memory Center. The methodology used consists of a descriptive and exploratory study of bibliographical sources and available technologies. As a result, an architecture proposal is presented for the Memory Center of Goiás.

Keywords: Interoperability; semantic *Web*; cultural heritage; *Linked Open Data*; memory center.

INTRODUÇÃO

A *Web* permitiu um aumento significativo na quantidade de repositórios digitais relacionados ao patrimônio cultural. Museus, bibliotecas, arquivos, galerias e outros espaços detentores de informações de patrimônio material e imaterial (GLAMs) utilizam, cada vez mais, as tecnologias da *World Wide Web* para se integrarem à sociedade em rede. Em uma sociedade em rede (CASTELLS, 2016), a informação, que nas palavras de Le Coadic (1994), “é um conhecimento inscrito (gravado), sob a forma escrita (impressa ou numérica), oral ou audiovisual”, torna-se mais rica quando ela adquire capacidade de se conectar com outras informações relevantes. Essa busca incessante pela informação é reforçada por Le Coadic (1994) que destaca que:

Seja pelo simples prazer de conhecer (Freud), de estar informado sobre os acontecimentos políticos, os progressos da ciência e tecnologia, ou pelo prazer menos simples de estar a par dos últimos temas e resultados das pesquisas, de acompanhar a vanguarda do científico, o objetivo da informação permanece sendo a apreensão de sentidos ou seres em sua significação, ou seja, continua sendo o conhecimento; e o meio é a transmissão do suporte, da estrutura. (LE COADIC, 1994, p. 5).

Para além de documentos, essa condição de efetiva utilidade e conectividade das informações expostas na *Web* requer uma estruturação que possibilite que as informações sejam recursos interoperáveis entre os diversos serviços e fins (BERNERS-LEE, 2006). Os dados esparsos, sejam em forma de metadados, planilhas ou qualquer outro artefato descritivo de objetos, nato-digitais ou não, tendem a se dissolver com o passar do tempo, uma vez que na *Web* dos dados não há espaço para recursos desconectados. Garantir essa interoperação, no entanto, não é uma tarefa trivial e encontra duas dificuldades principais: a primeira, sintática diz respeito à possibilidade conectividade em primeira ordem (MOURA, 2011). Em garantir uma forma comum para comunicação. A segunda, é a semântica. Mais importante e mais

profunda, tem a ver com o significado. Para além da forma, é necessário permitir que a informação seja compreendida de acordo com um contexto apresentado .

O problema da interoperabilidade semântica está epistemologicamente ancorado no campo de estudos da representação e recuperação da informação (ARAÚJO, 2009). Esse problema tem suas raízes nos primórdios da *Web*, que conforme Lima, Santos e Santarem Segundo (2016) afirmam, nasceu descentralizada, funcionando como um grande depósito de documentos, por isso, há imensa dificuldade para a recuperação das informações, tornando impossível que agentes automatizados compreendam o sentido que os dados possuem (LIMA; SANTOS; SANTARÉM SEGUNDO, 2016).

As informações de patrimônio material e imaterial sugerem múltiplas conexões e integrações entre repositórios digitais. Um objeto relativo à cultura de um povo possui potencial possibilidade de se conectar com muitos outros. Por exemplo, elementos da música popular brasileira que está envolvida em eventos históricos, com pessoas e em locais distintos e com forte influência sobre movimentos culturais, e que também podem ser vistos como artefatos em algum repositório digital na *Web*. Não é difícil encontrar documentos em arquivos ou museus que estão relacionados a outros documentos externos e fatos históricos, conhecidos ou não do grande público.

Como permitir que esses artefatos sejam de fato conectados de forma múltipla e viva entre os diversos repositórios é um desafio que se estende desde a definição mais básica dos conceitos, como as ontologias de domínio (GUARINO, 1998), até as mais avançadas estratégias de integração e compartilhamento de conteúdo, como os Dados Abertos Conectados - *Linked Open Data* (OWAIS; AL-FEEL; HASSANEIN, 2017).

Na perspectiva da preservação e acesso de patrimônio cultural, há diversas iniciativas, muitas delas de forma isolada, que atuam em vários níveis e vão desde a digitalização de acervos a disponibilização de repositórios digitais sob várias tecnologias. Na Universidade Estadual de Goiás tem sido desenvolvida uma dessas iniciativas, em formato de programa de extensão universitária, o Centro de Memória de Goiás – CMG, que dentre seus objetivos, pretende disponibilizar para a comunidade informações sobre o patrimônio material e imaterial do estado de Goiás.

O objetivo deste trabalho é trazer à discussão os aspectos tecnológicos e estratégias possíveis para criação de um Centro de Memória à luz dos Dados Abertos Conectados - *Linked Open Data* (BERNERS-LEE, 2006), e, conseqüentemente, orientado à *Web Semântica*. A metodologia utilizada consiste de um estudo descritivo e exploratório de fontes bibliográficas e tecnologias disponíveis, e, com base em fundamentos da Engenharia de

Software, apresenta-se uma proposta inicial de arquitetura de um sistema de informação, orientado à *Web Semântica* e sob as diretrizes das iniciativas dos Dados Abertos Conectados – *Linked Open Data*, para Centro de Memória de Goiás.

Apresenta-se, além deste tópico introdutório, a seguinte organização: a seção 2 fornece uma fundamentação teórica sobre interoperabilidade semântica; a seção 3 apresenta uma discussão sobre os Dados Aberto Conectados; na seção 4 há elementos estruturantes para a visão dos centros de memória discutidos neste trabalho, especificamente, sobre o Centro de Memória de Goiás; e, por fim, na seção 5 são apresentadas as considerações finais e questões em aberto.

INTEROPERABILIDADE SEMÂNTICA

A grande quantidade de documentos e dados disponíveis na Internet requer um esforço tecnológico para que, para além das pessoas, as informações sejam “inteligíveis” para as máquinas. Difícil pensar em interoperabilidade sem levar em consideração que as informações transitadas entre os diversos repositórios possam ser consumidas e agregam um maior poder informacional através de processamentos entre sistemas de acordo com algum contexto de utilização.

Tome-se como exemplo museus da imagem e do som que possuam algum objeto relacionado à música popular brasileira. Questões relativas ao período de composição, aos instrumentos utilizados ou mesmo ao autor são fundamentais para a geração de conhecimento. Entretanto, essas informações não precisam estar fisicamente no mesmo repositório especializado em imagem e som. Essa necessidade de conectividade informacional e colaboração entre os repositórios, sendo possível, tende a potencializar e enriquecer a informação para diversas formas de utilização.

Há um esforço da comunidade científica em gerir vocabulários, tesouros, taxonomias e outras formas de classificação e descrição de objetos digitais ou não em diversos contextos do discurso. Tudo isso gera uma quantidade estrondosa de informação em vários formatos, em especial de metadados (dados sobre dados).

Sobre esse crescimento acelerado na produção de informação de toda natureza e formato, Pierre Lèvy (2014) cognomina “caos digital”, onde este caos se origina das multiplicidades linguísticas naturais, dos sistemas de classificação e das ontologias. Paradoxalmente, o aumento massivo de dados (incluindo-se os metadados) experimentado nos

últimos anos de Internet tem contribuído para que o caos digital também aumente (LÈVY, 2014).

Em outras palavras, mesmo se tratando de fontes de informação com potencial capacidade de inter-relacionamentos, os diferentes contextos criam barreiras para a conectividade. Este cenário envolve inúmeros desafios que vão além da simples associação de artefatos ou a descrição em metadados (CELKO, 2014).

Há uma necessidade crescente de trabalhar os significados e os contextos relacionados à informação, e o contexto, mesmo estando associado a um universo específico do discurso, é composto de um extenso, e muitas vezes subjetivo, conjunto de classes, propriedades e relacionamentos, o que torna as tarefas relativamente simples em sistemas de bancos de dados convencionais potencialmente complexas (HOU et al., 2011).

Se de um lado o barateamento de determinadas tecnologias de hardware permitiu que se digitalizassem, em grande escala, documentos históricos, obras de arte e outros recursos de memória e patrimônio, de outro lado há grande confusão e dificuldade de permitir que esses recursos sejam acessados, mantidos e distribuídos de forma efetiva. Em outras palavras, há dados sobre tudo, mas há imensa dificuldade de conectividade entre as partes do todo. Partes estas que estão inseridas em um contexto próprio composto de linguagens, formas e conteúdos distintos que são, de alguma forma, organizados em sistemas de informação que, justificado pelo que discuti até aqui, clamam por mais semântica.

Interoperabilidade semântica pode ser vista como a capacidade que os sistemas de informação possuem de se integrarem por meios de padrões e tecnologias semânticas. No contexto dos repositórios digitais de patrimônio cultural, pressupõem-se o emprego de padrões e protocolos específicos de museus, arquivos, galerias de arte e outras instituições de preservação do patrimônio cultural material e imaterial (PADILHA; CAFÉ, 2016).

Esse cenário, segundo Doerr (2009), apresenta uma proliferação de terminologias especializadas que nem sempre são utilizadas como meio de concordância entre especialistas, mas muitas vezes apenas como uma ferramenta intelectual para construção de hipóteses ou descrever fenômenos (DOERR, 2009).

Entretanto, a despeito da dicotomia entre técnica e academia, há uma complexidade importante que envolve os assuntos, as terminologias e as estratégias em sistemas especializados, e que se desdobra em um complexo problema de integração de esquemas a fim de associar informações variadas e quase sempre complementares.

Dessa forma, para a colaboração entre os repositórios ocorra de forma que os dados possam ser conectáveis, abertos e, conseqüentemente, possam contribuir para a produção do

conhecimento, para além da interligação é necessário haver um cuidados processo de interoperação semântica (SABOU; EKAPUTRA; BIFFL, 2017) e (PADILHA; CAFÉ, 2016).

DADOS ABERTOS CONECTADOS E OS REPOSITÓRIOS DE PATRIMÔNIO CULTURAL

Segundo Berners-Lee (1998) a *Web* foi projetada para ser um espaço universal de informações, e portanto, ela perde seu poder se houver certos tipos de coisas aos quais você não pode possa vínculos (BERNERS-LEE, 1998).

Acontece que em meados dos anos 2000 percebeu-se que as ligações entre documentos (links e hipertexto) podem se apresentar um tanto quanto frágeis enquanto possibilidade de produção do conhecimento. Saltar através do hipertexto começou a representar pouco, uma vez que diante da crescente quantidade de informação, escancarava-se a necessidade de melhor aproveitamento dela não em forma de simples dados processados expostos.

Esse conhecimento, conforme destaca Le Coadic(1994), é representado por uma estrutura de conceitos e suas relações de determinados assuntos em determinados momentos (LE COADIC, 1994). Tais estruturas, relações e contextos são, de forma estrita, a semântica.

A proposta de Berners-Lee pretende fornecer uma solução para essa fragilidade semântica e desconexão dos dados na *Web*. A gênese da proposta trazia como pano de fundo, dentre outros objetivos, permitir que as pessoas pudessem trabalhar de forma colaborativa e que elas pudessem compartilhar conhecimento, universalmente (BERNERS-LEE, 1998).

A dicotomia *Web* dos Documentos (hiperlinks) versus *Web* dos Dados (semântica) é abordada em seu texto *Linked Data*:

Como a *Web* de hipertexto, a *Web* de dados é construída com documentos na *Web*. No entanto, ao contrário da *Web* de hipertexto, onde os links são âncoras de relacionamentos em documentos de escritos em HTML, para os dados eles ligam coisas descritas pelo RDF (BERNERS-LEE, 2006 P.1) .

Berners-Lee (2006) ainda fornece quatro expectativas para que a *Web* dos dados possa crescer: 1) uso de URIs para nomear coisas; 2) uso de HTTP URIs para que pessoas possam procurar por esses nomes; 3) fornecer informações úteis usando padrões estabelecidos (RDF *, SPARQL), quando alguém buscar um URI; e 4) incluir ligações para outros URIs para que eles possam descobrir mais coisas.

Acervos culturais digitais podem se apresentar como grandes facilitadores de acesso sob várias perspectivas. Os repositórios dos ditos GLAMs (galerias, bibliotecas, arquivos e museus) oferecem, além das possibilidades de disponibilização de públicos distintos, uma enorme potencialidade de colaboração entre si para geração de conhecimento.

Nota-se aqui oportunidades de ferramentas para integração entre várias camadas dessas fontes de informação heterogênea que vão desde definições de vocabulário até serviços de trocas de informação qualificada para geração de sistemas de conhecimento.

Carvalho Junior; Martins; Germani (2019) enfatizam que estamos diante de um novo ambiente informacional, uma vez que a digitalização de acervos das instituições culturais, aliada às vantagens em se promover interoperabilidade entre suas bases, e a necessidade logística de compartilhar de infraestrutura de software e hardware, promovem uma emergência de infraestruturas que permitem a operação em ambientes digitais contíguos e até permeáveis (CARVALHO JUNIOR; MARTINS; GERMANI, 2019).

Sistemas de bibliotecas digitais, como o *DSpace* (VECHIATO et al., 2017) , e iniciativas de exposição e consumo de metadados como o OAI-PMH (FREIRE et al., 2018a) aliados a padrões de metadados como o Dublin Core (XIE; MATUSIAK, 2016) foram responsáveis por um aumento acentuado de repositórios digitais nos últimos anos, em especial no campo das bibliotecas digitais.

Outras iniciativas como o DigitArq (FERROS; FERREIRA; RAMALHO, 2010) e, mais recentemente no Brasil, o Projeto Tainacan (MARTINS; CARMO; GERMANI, 2018) ofereceram alternativas mais voltadas para museus, galerias de arte e arquivos. Entretanto, quando invocam-se os requisitos de efetividade e conectividade da *Web Semântica* aplicados ao patrimônio cultural, o projeto Europeana (EUROPEANA, 2017) destaca-se pela complexidade e quantidade de recursos envolvidos – inclusive financeiros.

O projeto Europeana é uma biblioteca virtual que integra mais de 50 milhões de objetos culturais de bibliotecas, museus e arquivos. Na perspectiva de Arquivos Abertos Conectados, a Europeana permite-se recuperar grafos RDF a partir de consultas em um endpoint SPARQL. A estruturação dos dados da Europeana é orientada pelo modelo conceitual EDM – Europeana Data Model (CONEGLIAN; SANTAREM SEGUNDO, 2016).

Os Dados Abertos Conectados (*Linked Open Data*), são no entanto, uma mistura de Dados Conectados (*Linked Data*) e Dados Abertos (*Open Data*). Defende-se aqui, que para uma colaboração ampla entre os repositórios, que além de conectáveis, os dados sejam abertos. Iniciativas como *DBPedia*, *GeoNames*, e outros serviços, constituem o projeto LOD de dados abertos conectados.

Berners-Lee (2006) propõe um espectro de desenvolvimento dos dados conectados em 5 estrelas: 1) disponibilizar com licença aberta na *Web*; 2) disponibilizar os dados em formato estruturado; 3) disponibilizar os dados em um formato aberto não proprietário; 4) utilizar URIs para identificar recursos; e 5) conectar com outros dados para prover contexto.

O modelo conceitual de referência CIDOC-CRM

Tudhope et. al (2013) afirmam que para promover interoperabilidade semântica se faz necessário mais do que um modelo comum de dados semântico e sintático (TUDHOPE et al., 2013) .

A CIDOC-CRM (ISO 21127:2014) é uma ontologia formal para integração e intercâmbio de informações heterogêneas de patrimônio cultural para redes de conhecimento (DOERR; ORE; STEAD, 2007). Ela atua fornecendo definições semânticas que visam clarificar informações locais para que elas sejam melhor aproveitadas de forma global.

Segundo Doerr (2009), na CIDOC-CRM são seguidos princípios estreitos para a integração global e restrições filosóficas sobre o tipo de discurso, sendo a ontologia descrita na linguagem TELOS. Conforme apresentado na Figura 1, distingue classes individuais de propriedades, que são bidirecionais com rótulos distintos para cada direção permitindo herança múltipla tanto pra classes quanto para propriedades (DOERR, 2009).

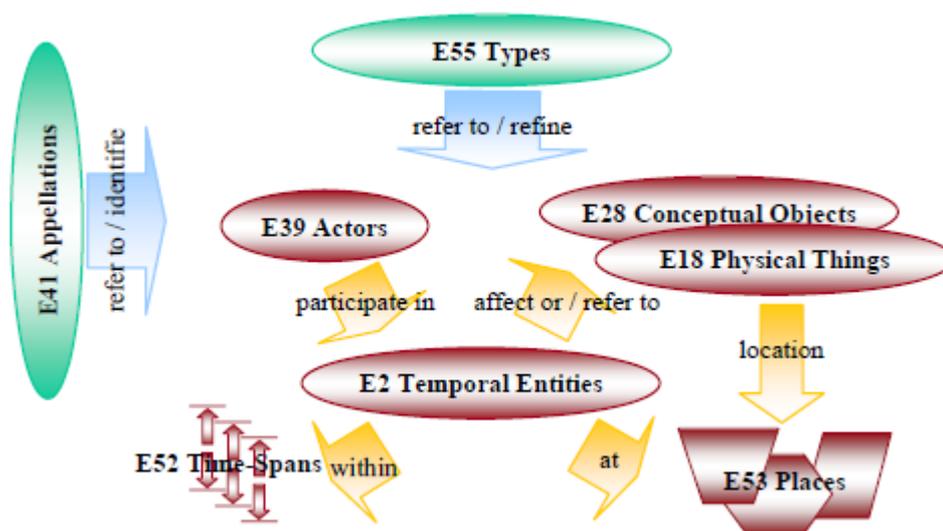


Figura 1: Conceitos fundamentais do CIDOC-CRM

Fonte: Extraído de Doerr (2009)

Pensada como um guia para o processo análise de requisitos e fase de modelagem conceitual para os sistemas de informação cultural, a CIDOC-CRM fornece uma codificação das principais conceituações de domínio interdisciplinar da cultura, de forma que permita uma funcionalidade extensível (SANTOS, 2016).

A CIDOC-CRM pode ser visualizada, de uma maneira geral, através das funções suportadas pelos seus relacionamento que são 1) identificação de itens do mundo real por através de nomes do mundo real; 2) classificação de itens do mundo real; 3) decomposição parcial de coisas materiais e imateriais, entidades temporais, grupo de pessoas, lugares e tempos; 4) participação de itens persistentes em unidades temporais; 5) localização de entidades temporais no espaço-tempo e coisas materiais no espaço; 6) influência de objetos em atividades e produtos e vice-versa; e 7) referências de objetos de informação a qualquer item do mundo real (BOEUF et al., 2018).

O modelo EDM

O projeto Europeana surgiu com a intenção de unir os acervos digitais das bibliotecas, museus, arquivos e outros centro culturais espalhados por toda Europa, na tentativa de disponibilizar em um único ambiente digital o acesso a essa grande massa informacional (CONEGLIAN; SANTAREM SEGUNDO, 2016)

O modelo EDM (Europeana Data Model) foi desenvolvido para suceder a ESE (Europeana Semantic Elements). O modelo ESE é simples e permite registrar metadados sobre os objetos culturais de uma forma não-hierárquica (ou seja, não é possível registrar informação sobre partes ou componentes de um objeto), recorrendo em parte aos metadados Dublin Core. A sua simplicidade, face à necessidade de registrar informação mais coleta sobre os objetos, levou à proposta do EDM, baseado em várias propostas da *Web* semântica e em normas existentes como LIDO (museus), EAD (arquivos) e METS (bibliotecas digitais). A Figura 2 apresenta uma visão hierárquica das classes do modelo EDM, onde as classes representadas por retângulos brancos são reutilizadas a partir de esquemas externos, e as em azul claro definidas propriamente pelo EDM.

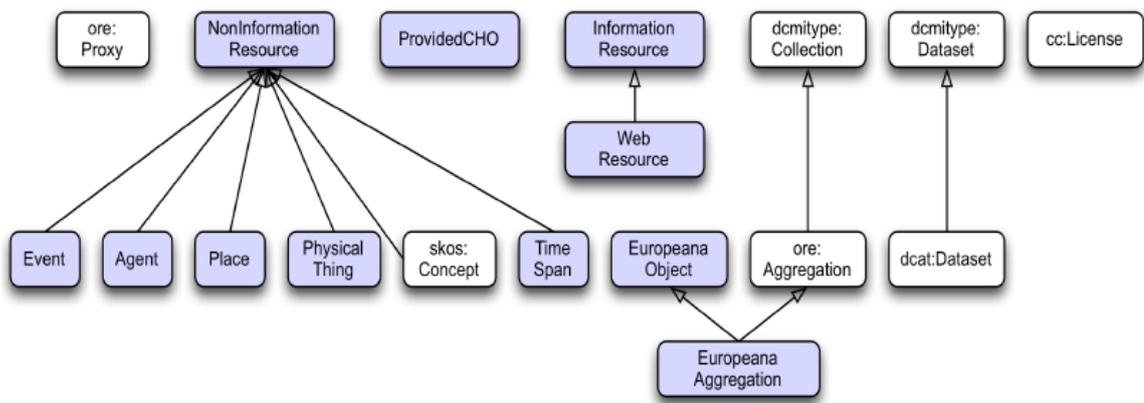


Figura 2: Classes do modelo EDM

Fonte: Extraído de Specification of Europeana Data Model (EUROPEANA, 2017)

A classe *ProvidedCHO* inclui os objetos de Patrimônio Cultural que são descritos na Europeana. Funciona como uma espécie de nó central no EDM.

A classe *NonInformation Resource* se refere a todos os recursos que não são de natureza da informação, por exemplo, pessoas, lugares e objetos tangíveis. Todos esses recursos estão associados a eventos (*Event*), agentes (*Agent*), lugares (*Places*), intervalo de tempo (*Time Span*), que são classes próprias da Europeana, e referem-se a conceitos, definidos pela classe externa *skos:Concept*. Tais conceitos são utilizados para contextualização dos recursos a partir de ideias ou significados específicos de acordo a descrição de sua estrutura conceitual dentro um sistema de organização do conhecimento.

A classe *EuropeanaAggregation* reúne em uma única unidade conceitual todas as informações sobre os objetos de patrimônio cultural. É partir dela que ocorre a agregação de fato da descrição dos objetos que estão em diferentes provedores de conteúdo.

O modelo *Linked.Art*

Uma das vantagens de se usar LOD é que podem-se usar várias ontologias para modelar os dados. O *Linked Art*, criado pela American Art Collaborative - AAC, é uma extensão simplificada do CIDOC-CRM que pretende fornecer padrões e modelos para que instituições de cultura possam publicar seus dados para usuários especializados, pesquisadores e desenvolvedores não especializados.

O *Linked Art*, além de utilizar a CIDOC-CRM como a ontologia central, apoia-se no vocabulário Getty como as principais fontes de identidade e na notação de objeto JavaScript para dados vinculados (JSON-LD) como a serialização. O foco principal deste modelo de

destino são a facilidade de padronização entre os repositórios e a facilidade de uso (NEELY; LUTHER; WEINARD, 2019).

O conjunto básico das classes é composto por objetos, atividades, pessoas, lugares e tempo em quatro questões: o quê, quando, quem e onde. A busca por simplificação do modelo CIDOC-CRM é justificada pela equipe *Linked Art* que afirma que o custo humano de entender um sistema com centenas de classes e relacionamentos é extremamente alto e aumenta a probabilidade de que diferentes abordagens sejam adotadas, reduzindo assim a interoperabilidade.

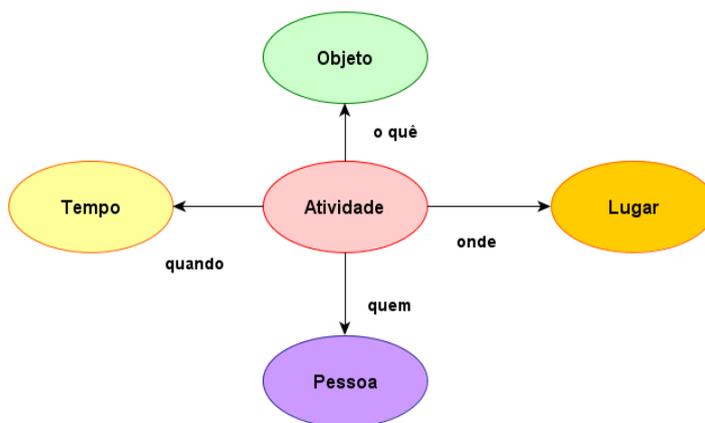


Figura 3: Núcleo de classes do *Linked Art*

Fonte: Adaptado de (FINK, 2018)

O CENTRO DE MEMÓRIA DE GOIÁS

Para Camargo e Goulart (2015) os centros de memória, enquanto espaço de formação e difusão do saber, preveem a coleta, preservação e transmissão do patrimônio cultural com a missão principal de tornar acessível a informação em diversos contextos (CAMARGO; GOULART, 2015).

De forma complementar Dodebei (2011), afirma que historicamente, a interseção entre museus, bibliotecas e arquivos dividiu seu foco entre valor documental dos objetos o processo de identificação, descrição e comunicação dos conteúdos, e o fato de haver conflitos na conceituação sobre objetos, documentação e memória revela uma interface entre as práticas disciplinares (DODEBEI, 2011).

Um dos desafios impostos pela própria essência dos centros de memória está em garantir esse acesso à população de forma interativa e com um certo grau de

interconectividade com outras fontes de conhecimento disponíveis. O que não pode ser realizado sem levar em consideração os padrões estabelecidos pelas comunidades e as tecnologias experimentadas para esse fim.

O Centro de Memória de Goiás (CMG), caracteriza-se como unidade institucional da Política de Extensão da Universidade Estadual de Goiás (UEG), conforme Resolução CsA nº 1045/2017 (UEG, 2017) e tem como princípio a construção de uma rede de estudos, pesquisas e trabalhos, envolvendo instituições públicas e privadas, voltada para ações de preservação da memória do Estado de Goiás e na formação continuada de profissionais da educação nas mais diversas áreas de conhecimento no estado de Goiás. As ações vinculadas ao CMG estão agregadas nos seguintes 4 núcleos, conforme esquema apresentado na Figura 4.

I – Núcleo de Documentação: responsável pela higienização, catalogação e/ou digitalização de arquivos dos poderes executivo, legislativo e judiciário; e arquivos cartoriais, eclesiásticos e privados. Esse núcleo não realiza a salvaguarda dos documentos e funciona como apoio a outros centros, arquivos e instituições que contenham material de interesse histórico e cultural para o Centro de Memória de Goiás.

II – Núcleo de História Oral e Imagem: se encarrega de ações que tem como objetivo o registro e catalogação de depoimentos orais e imagens das mais variadas formas de manifestação artística e cultural do povo goiano. O núcleo também é responsável pela produção de material audiovisual das atividades deste e dos outros eixos do Centro de Memória para realização de documentários, exposições e divulgação.

III – Núcleo de Arqueologia: Responsável pelas ações de arqueologia, que envolvem a coleta, catalogação e/ou organização de acervo arqueológico, bem como a sua disponibilização e apresentação em exposições ao público em geral.

IV – Educação Patrimonial e Temas específicos da docência – refere-se às atividades relacionadas especificamente ao campo educacional, desde a produção de material didático até a formação continuada dos docentes estaduais e municipais a partir dos acervos e materiais dos outros três eixos.

Na perspectiva de espaço integrado de educação e difusão do conhecimento, o Centro de Memória de Goiás (CMG) é uma proposta de iniciativa ambiciosa da Universidade Estadual de Goiás, que tem, dentre outros objetivos, promover ações de digitalização, preservação e acesso ao patrimônio material e imaterial do Estado de Goiás.

O CMG possui 3 ações em andamento nos municípios de Morrinhos, Goiás e Uruaçu. Apesar da crise enfrentada em vários setores da cultura, e também da profunda crise financeira institucional na Universidade, o CMG tem conseguido manter atividades nas

três frentes, muito pelo seu formato extensionista que conta com professores, estudantes e membros da comunidade como atores e entusiastas da preservação do patrimônio cultural, em especial, do povo goiano.

A partir das primeiras ações do CMG, fica explícita, através da relação universidade-comunidade, que a extensão universitária se apresenta como um dos caminhos para a sustentabilidade de espaços de cultura e patrimônio. Um dado que reforça essa potencialidade é que, no ano de 2018, somente na Universidade Estadual de Goiás, foram realizadas 884 ações de extensão, contando com a participação de 821 professores, 8614 estudantes e 1747 membros da comunidade, beneficiando pouco mais de 380 mil pessoas. Outro dado essencial é que a UEG oferece 179 cursos de licenciatura em todas as regiões do estado.

A Figura 4 apresenta a estruturação acadêmica do CMG, que, de acordo com os objetivos de cada núcleo, tem como eixo articulador a educação patrimonial. Reforça, portanto, o caráter e a preocupação com professores, estudantes e outros atores da educação a partir da preservação da cultura e patrimônio como identidade do povo.



Figura 4: Esquema dos núcleos do Centro de Memória de Goiás.

Fonte: Extraído da PrE/UEG

Uma proposta de arquitetura para o Centro de Memória de Goiás

Os repositórios de patrimônio cultural são sistemas de informação, que segundo, Doerr(2009) podem ser distinguidos pelas seguintes funções: a) Gestão de coleções (aquisição, registro, “desativação”, inventário, empréstimos, exposições, seguros, direitos, zonas de proteção) b) Conservação (diagnóstico de deterioração, medidas preventivas, intervenções, tratamentos e agentes químicos); c) Pesquisa (investigação, descrição, interpretação); e d) Apresentação (portais, ensino, publicação).

Do ponto de vista dos sistemas de informação para memória e patrimônio, uma questão fundamental é, como suportar os vários modelos de organização da informação? A diversidade dos dados de memória aliadas às facetas exigidas pelas funções e as expectativas de utilização geram um cenário de alta complexidade.

Por outro lado, há de se levar em consideração que a criação de um sistema de informação não semântico pode levar a um grande retrabalho no futuro, sem falar na possível perda de informações pela ausência de um modelo que leve em consideração o contexto dos objetos digitais, da utilização dos profissionais e da expectativa dos usuários. Com base nessas premissas, a buscou-se, a partir da identificação dos requisitos e interação com colaboradores, construir caminhos possíveis para a criação de uma arquitetura para o sistema.

Essa arquitetura foi modelada em grandes ações: 1) definição conceitual; 2) definição tecnológica; e 3) definição de utilização. Os conceitos devem receber especial atenção. A organização da informação é fundamental para a recuperação da informação e pode garantir melhor usabilidade dos recursos.

E por fim, qual o propósito de um sistema de informação que não é utilizado? daí a preocupação com o reuso e a conectividade com outras plataformas possíveis, para que pesquisadores, professores, estudantes e comunidade em geral possam utilizá-lo de forma efetiva. As definições de utilização requerem, no futuro, um amplo estudo das necessidade da interface com o usuário, com o fim de tornar as informações o mais amigável e facilmente navegável possível para um grupo heterogêneo de usuários.

Sendo o CMG pensado em três eixos e quatro núcleos, um ponto de partida para o conhecimento produzido no sistema é enquadrado em quatro dimensões: 1) tempo; 2) espaço; 3) pessoas; e 4) fenômeno.

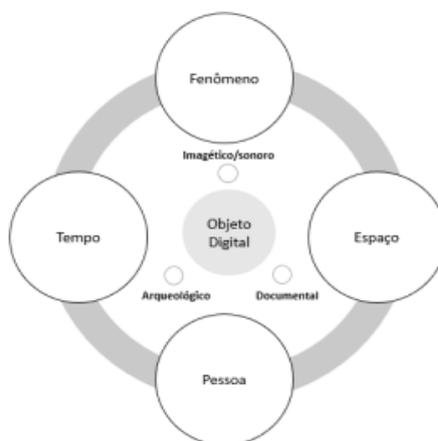


Figura 5: Dimensões do CMG

O modelo é focado no objeto digital, seja ela nato-digital ou digitalizado, que encontra-se cercado nas quatro dimensões. Enquanto Fenômeno, o objeto gera ou se refere a qualquer ação, evento, acontecimento ou manifestação, que ocorreu (ou ocorre) em um Tempo, que pode envolver período ou data; em algum Espaço (qualquer local geográfico) e envolve, de alguma maneira, alguma Pessoa (pessoas, povo, grupo populacional).

A proposta de sistema de informação para o CMG passa pela divisão em módulos. A Figura 6 apresenta um ensaio inicial de termos utilizado para a organização da informação com base nos eixos e núcleos de arqueologia, documental e história oral e imagem. Esse ensaio inicial mostra a complexidade das relações entre os objetos e propriedades possíveis em três entidades principais: eventos, pessoas e objetos digitais (recursos). Os recursos podem ser, em uma primeira visão, arqueológicos, documentais ou imagético/sonoro. Nada impede que tal objeto seja relacionado aos três. Isto sugere a ideia de herança múltipla (como no CIDOC-CRM) e requer um cuidado especial.

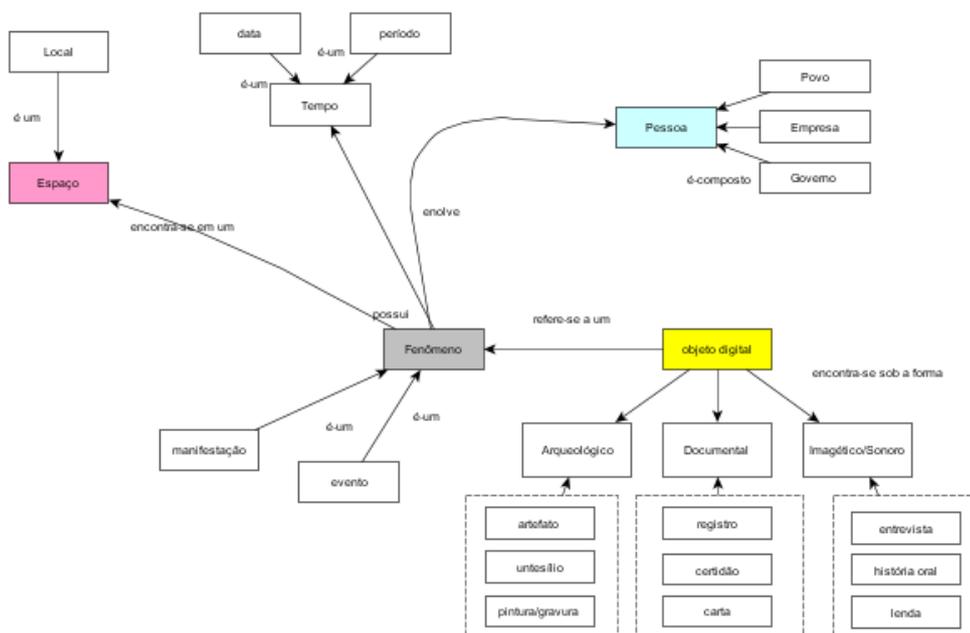


Figura 6: Um ensaio de organização dos conceitos do CMG

Um objeto digital pode ser encontrado sob várias formas. Referem-se a fenômenos. Os fenômenos podem envolver pessoas em um determinado tempo, de acordo com as quatro dimensões apresentadas. Uma ontologia pode ser utilizada para fornecer um domínio nas quatro dimensões e nos três eixos fundamentais.

A Figura 7 traz um modelo simplificado arquitetural do sistema proposta. Nesta visão arquitetural são recortados do sistema apenas os aspectos internos de produção da informação.

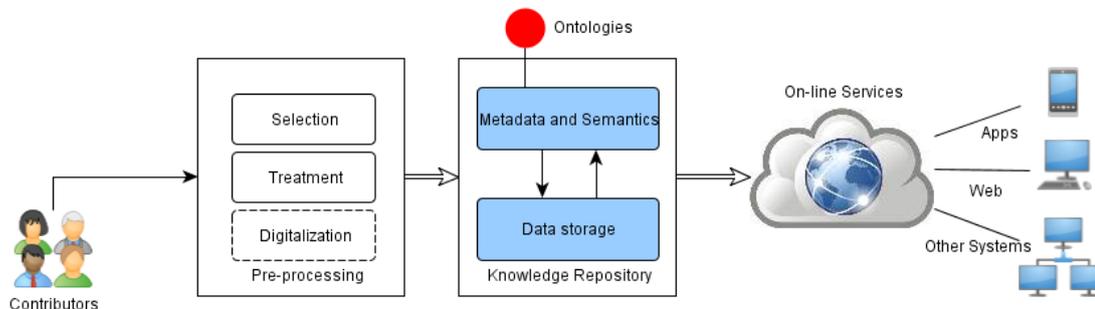


Figura 7: Modelo arquitetural simplificado para o CMG

Fonte: (CHAGAS et al., 2018)

Os “contribuidores” representam os diversos atores do sistema: estudantes, professores, técnicos e membros da comunidade que colaboram para essa produção. Os processos de seleção e tratamento também são fenômenos (eventos) que precisam ser levados em consideração para a produção do conhecimento. O processo de digitalização de objetos não nato-digitais também precisa receber atenção enquanto evento. Toda essa fase chamada de pré-processamento deve alimentar/gerar um fluxo de armazenamento dos objetos digitais e disparar o início da composição semântica nas quatro dimensões (tempo, espaço, pessoas e fenômeno).

A nuvem de serviços on-line apresentada na Figura 7 representa os *Web Services*, SPARQL endpoints e outras possibilidades de colaboração externa que o sistema possa oferecer.

Há de se reafirmar o caráter de “estado momentâneo” de cada objeto. Um objeto não é fechado em si. As propriedades e relacionamentos precisam ser vivas e em cada momento pode-se haver um novo estado a partir de um novo conjunto de relações.

CONSIDERAÇÕES FINAIS

Muitos projetos têm sido desenvolvidos com foco nos dados abertos conectados. Apesar dos esforços humanos, científicos e tecnológicos, e em muitos casos financeiros, ainda há um longo caminho a ser seguido para que os espaços de memória possam, em maior

escala, colaborar para gerar mais conhecimento e, não menos importante, tornar todo esse conhecimento facilmente acessível.

Para que os dados sejam conectados (e explorados) em todo o seu potencial, nota-se que há necessidade de uma mudança cultural na forma como os dados são produzidos, gerenciados, disseminados para que sejam, para além de conectáveis, facilmente utilizáveis pelos usuários finais.

Para o Centro de Memória de Goiás, um projeto que ainda dá seus primeiros passos em direção aos imensos desafios das humanidades digitais, percebe-se um longo caminho e desafios próprios que permeiam a gestão da informação. Uma faceta é importante desse desafio é a escolha das tecnologias a serem utilizadas.

Tanto a experiência da AAC em criar seu LOD quanto o projeto Europeana apresentam possibilidades que podem ser utilizadas em diversos graus de tecnologia e metodologia. Há de se levar em consideração, no entanto, que ambas tratam de agregadores de informações semânticas dos GLAMs com diferenças estratégicas justificadas em cada um dos casos.

Do ponto de vista da utilização de ontologias, por exemplo, percebe-se que a CIDOC-CRM apresenta-se como um modelo que pretende ser completo e padrão para, sobretudo, os museus, mas, conforme a experiência da AAC (*Linked Art*), possui dezenas de classes e centenas de relacionamentos, (alguns desnecessários), tornando-a em ontologia complexa e difícil de utilizar.

Já pelo viés financeiro, há de se levar em consideração as limitações de investimento que praticamente inviabilizam tecnologias que não sejam abertas e colaborativas.

Partindo-se pela visão da diversidade da informação, pode-se dizer que há grande aproximação dos modelos apresentados em ambos os casos com aquilo que se pretende para o CMG, sendo notável que o modelo simplificado do *Linked.Art* parece corresponder bem em termos de ontologia e facilidade.

Por fim, há de se reconhecer a importância e também a complexidade das tarefas que envolvem a concepção do sistema de informação proposto. Estando claro que qualquer que seja o formato, sua construção, deve estar alinhada com o princípio básico de qualquer ação extensionista: a colaboração entre universidade e sociedade.

REFERÊNCIAS

ARAÚJO, C. A. Á. Correntes teóricas da ciência da informação. **Ciência da Informação**, [s. l.], v. 38, n. 3, p. 192–204, 2009. Disponível em:

- <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652009000300013&lng=pt&tlng=pt>
- BERNERS-LEE, T. **Realising the Full Potential of the Web**. 1998. Disponível em: <<https://www.w3.org/1998/02/Potential.html>>.
- BERNERS-LEE, T. **Linked Data**. 2006. Disponível em: <<https://www.w3.org/DesignIssues/LinkedData.html>>.
- BOEUF, P. Le; DOERR, M.; ORE, C. E.; STEAD, S. **Definition of the CIDOC Conceptual Reference Model**. [s.l.: s.n.].
- CAMARGO, A. M.; GOULART, S. **Centros de Memória: Uma proposta de definição**. 1a. ed. São Paulo: Sesc-SP, 2015.
- CARVALHO JUNIOR, J. M. C.; MARTINS, D. L.; GERMANI, L. B. GLAM e Instituições de Memória em Rede: uma 'Infosfera' de Florida? **pragMatizes**, [s. l.], 2019.
- CASTELLS, M. **A Sociedade em Rede**. 17a Edição ed. São Paulo.
- CELKO, J. **Joe Celko's Complete Guide to NoSQL**. 1a. Ed. ed. [s.l.] : Joe Celko's, 2014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/B9780124071926000030>>
- CHAGAS, P. R.; GOUVEIA, F. A. R.; FILGUEIRAS, A. C.; TORRES, M. A. Da. Goiás State University as locus for heritage preservation: genesis of a memory center. In: 6TH INTERNATIONAL CONFERENCE ON HERITAGE AND SUSTAINABLE DEVELOPMENT 2018, Granada-ES. **Anais...** Granada-ES
- CONEGLIAN, C. S.; SANTAREM SEGUNDO, J. E. Europeana no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [s. l.], v. 22, n. 48, p. 88, 2016. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n48p88>>
- DODEBEI, V. Cultura Digital: novo sentido e significado de documento para a memória social? **DataGramZero**, [s. l.], v. 12, n. 2, 2011.
- DOERR, M. Ontologies for Cultural Heritage. In: STAAF, S.; STUDER, R. (Eds.). **Handbook on Ontologies**. Second Edi ed. New York, NY, USA: Springer, 2009. p. 463–486.
- DOERR, M.; ORE, C.-E.; STEAD, S. The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing. **Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual Modeling - Volume 83**, [s. l.], v. 83, n. Er, p. 51–56, 2007. Disponível em: <<http://dl.acm.org/citation.cfm?id=1386957.1386963>>
- EUROPEANA. **Definition of the Europeana Data Model**. 2017. Disponível em: <<https://pro.europeana.eu/resources/standardization-tools/edm-documentation>>. Acesso em: 24 jul. 2018.
- FERROS, L. M.; FERREIRA, M.; RAMALHO, J. C. DigitArq e o novo módulo de interoperabilidade. [s. l.], 2010.
- FINK, E. E. American Art Collaborative (AAC) Linked Open Data (LOD) Initiative: overview and recommendations for good practices. **The American Art Collaborative (AAC)**, [s. l.], 2018. Disponível em: <http://americanartcollaborative.org/wp-content/uploads/2018/03/AAC_LOD_Overview_Recommendations.pdf>
- FREIRE, N.; ISAAC, A.; ROBSON, G.; HOWARD, J. B.; MANGUINHAS, H. A survey of Web technology for metadata aggregation in cultural heritage. **Information Services and Use**, [s. l.], v. 37, n. 4, p. 425–436, 2018. a.
- FREIRE, N.; MEIJERS, E.; VOORBURG, R.; ISAAC, A. Aggregation of cultural heritage datasets through the Web of Data. **Procedia Computer Science**, [s. l.], v. 137, p. 120–126, 2018. b. Disponível em: <<https://doi.org/10.1016/j.procs.2018.09.012>>
- GUARINO, N. **Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy**. 1st. ed. Amsterdam, The Netherlands, The Netherlands: IOS Press, 1998.

- HOU, X.; ONG, S. K.; NEE, A. Y. C.; ZHANG, X. T.; LIU, W. J. GRAONTO: A Graph-based Approach for Automatic Construction of Domain Ontology. **Expert Syst. Appl.**, Tarrytown, NY, USA, v. 38, n. 9, p. 11958–11975, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2011.03.090>>
- LE COADIC, Y.-F. **A Ciência da Informação**. [s.l.: s.n.].
- LÈVY, P. **A Esfera Semântica Tomo 1: Computação, Cognição e Economia da Informação**. 1a Ed. ed. São Paulo: Annablume, 2014.
- LIMA, F. R. B.; SANTOS, P. L. V. A. C.; SANTARÉM SEGUNDO, J. E. Padrão de metadados no domínio museológico. **Perspectivas em Ciência da Informação**, [s. l.], v. 21, n. 3, p. 50–69, 2016.
- MARTINS, D. L.; CARMO, D. Do; GERMANI, L. B. MUSEU DO ÍNDIO : ESTUDO DE CASO DO PROCESSO DE MIGRAÇÃO E ABERTURA DOS DADOS LIGADOS SEMÂNTICOS DO ACERVO MUSEOLÓGICO COM O SOFTWARE LIVRE TAINACAN. In: XIX ENANCIB 2018, **Anais...** [s.l.: s.n.]
- MOURA, M. A. interoperabilidade Semântica e Ontologia Semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. **Informação & Informação**, [s. l.], v. 16, n. supl, p. 165–179, 2011.
- NEELY, L.; LUTHER, A.; WEINARD, C. Cultural Collections as Data: Aiming for Digital Data Literacy and Tool Development. In: MW19: MW 2019 2019, Boston, MA-USA. **Anais...** Boston, MA-USA Disponível em: <<https://mw19.mwconf.org/paper/cultural-collections-as-data-aiming-for-digital-data-literacy-and-tool-development/>>
- OWAIS, N. N.; AL-FEEL, H. T.; HASSANEIN, E. E. A step towards the linking of the Culture Heritage in the Egyptian museums. In: 2017 IEEE 7TH ANNUAL COMPUTING AND COMMUNICATION WORKSHOP AND CONFERENCE (CCWC) 2017, **Anais...** [s.l.: s.n.]
- PADILHA, R. C.; CAFÉ, L. M. A. A interoperabilidade semântica entre acervos de museus: discutindo o caso dos Museus da Imagem e do Som. **Em Questão**, [s. l.], v. 23, n. 1, p. 113, 2016.
- SABOU, M.; EKAPUTRA, F. J.; BIFFL, S. **Semantic Web technologies for data integration in multi-disciplinary engineering**. [s.l.: s.n.]. cited By 4. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85052076728&doi=10.1007%2F978-3-319-56345-9_12&partnerID=40&md5=81d4b221fdad0690f4ee484e2020dcf5>
- SANTOS, H. P. Modelo CIDOC CRM: interoperabilidade semântica de informações culturais. **Brazilian Journal of Information Science: Research Trends**, [s. l.], v. 10, n. 1, p. 56–62, 2016. Disponível em: <<http://revistas.marilia.unesp.br/index.php/bjis/article/view/5060/3975>>
- TUDHOPE, D.; BINDING, C.; MAY, K.; CHARNO, M. Pattern based mapping and extraction via CIDOC CRM. **CEUR Workshop Proceedings**, [s. l.], v. 1117, p. 23–36, 2013.
- UEG, U. E. de G. Política de Extensão da Universidade Estadual de Goiás 2017.
- VECHIATO, F. L. (Org. .; MARQUES, C. de A. G. (Org. .; KOSHIYAMA, D. C. A. D. G. (Org. .; MOURA, E. A. de (Org. .; TORINO, E. (Org. .; MAIA, M. A. Q. (Org. .; MARQUES, T. M. S. de M. (Org. . **Repositórios digitais: teoria e prática**. Curitiba-PR.
- XIE, I.; MATUSIAK, K. K. Chapter 5 - Metadata. In: XIE, I.; MATUSIAK, K. K. B. T.-D. D. L. (Eds.). Oxford: Elsevier, 2016. p. 129–170.

SOBRE OS AUTORES E A AUTORA

Alison Carlos Filgueiras

Graduação em Processamento de Dados pela Universidade Estadual de Goiás (UEG). Mestrado em Ciência da Computação pela Universidade Federal de Goiás. Gestor Governamental de Tecnologia da Informação do Estado de Goiás. Docente do Curso de Sistemas de Informação da UEG Câmpus Itaberaí-GO. Estudante de doutoramento em Ciência da Informação na Universidade Fernando Pessoa (UFP-PT). Áreas de interesse: extensão universitária, engenharia de software, web-semântica, sistemas de informação de patrimônio cultural.

Lattes: <http://lattes.cnpq.br/2264806467504828>

Feliz Ribeiro Gouveia

É Professor Associado com Agregação na Universidade Fernando Pessoa, Porto-PT. Orientador de doutoramento em Ciência da Informação na Linha de Sistemas, Tecnologia e Gestão da Informação.

Paula Roberta Chagas

Possui graduação em História pela Universidade Federal do Paraná (2007), Especialização em Educação Especial pela UNINTER e Mestrado em História pela Universidade Federal do Paraná (2010). Atualmente é doutoranda do programa de pós graduação em Ciências da Informação pela Universidade Fernando Pessoa (UFP/Portugal), Docente de ensino superior da Universidade Estadual de Goiás, Assessora de Projetos e Programas da Pró-Reitoria de Extensão, Cultura e Assuntos Estudantis e Coordenadora do Centro de Memória de Goiás. Tem experiência na área de História, com ênfase em História do Brasil Colônia, Ensino de História, Preservação de patrimônio material e imaterial, Diversidades e Gênero, atuando principalmente nos seguintes temas: fontes paroquiais, compadrio, centros de documentação, produção de material didático, memória, didática, gênero e diversidades.

Lattes: <http://lattes.cnpq.br/2412954484947403>

Recebido para publicação em março de 2020

Aprovado para publicação em julho de 2020