

# Uma Ontologia CIDOC-CRM para o Museu Virtual da Emigração

Ricardo G. Martini<sup>1\*</sup>, Cristiana Araújo<sup>1</sup>, Pedro Rangel Henriques<sup>1</sup>

<sup>1</sup>Centro Algoritmi, Departamento de Informática,  
Universidade do Minho - Campus de Gualtar - 4710-057, Braga, Portugal

rgm@algoritmi.uminho.pt, decristianaaraujo@hotmail.com, prh@di.uminho.pt

**Abstract.** *In this article we describe the use of the international standard for museum ontologies, CIDOC-CRM, to build a concrete ontology for ME, a virtual museum for the Emigration phenomena. Although the digital document repository is a conventional relational database, more abstract data models should be used to enable conceptual navigation over the information to help end-users (the museum visitors) to extract knowledge from the navigation (the visit). Thus, we will emphasize the design of a Reduced CRM-compatible form to describe the objects of our assets (people that emigrate in some moment to some other country), and the relations holding among them that better express the emigration movements. We also discuss the choice of RDF to expose the ontology triples via Apache Jena Fuseki server, a SPARQL end-point accessible over HTTP, and an engine (Jena ARQ) to query the RDF data. Finally, it was possible to transform these data into RDF and apply SPARQL queries, returning the desired results.*

**Resumo.** *Neste artigo nós descrevemos o uso do padrão internacional para ontologias museológicas, CIDOC-CRM, para construir uma ontologia concreta para o ME, um museu virtual para o fenômeno da emigração. Embora o repositório digital de documentos seja uma base de dados relacional convencional, modelos de dados mais abstratos devem ser usados para habilitar a navegação conceitual sobre as informações para ajudar os usuários finais (visitantes do museu) a extrair conhecimento da navegação (visita). Desta forma, nós enfatizaremos o desenvolvimento de um formato compatível reduzido do CRM para descrever os objetos do nosso espólio (pessoas que emigraram em algum momento para algum outro país) e as relações entre eles que melhor expressam os movimentos de emigração. Nós também discutimos a escolha do RDF para expor as triplas da ontologia através do servidor Apache Jena Fuseki, um end-point SPARQL acessível por HTTP e um motor (Jena ARQ) para consultar os dados RDF. Por fim, foi possível transformar essas informações em RDF e aplicar consultas SPARQL, retornando os resultados desejados.*

## 1. Introdução

As instituições detentoras do patrimônio cultural - como bibliotecas, arquivos e museus - são responsáveis pelo tratamento técnico, preservação e divulgação dos acervos documentais. Estes repositórios de dados são fundamentais para a compreensão das histórias,

\*Bolsheiro CNPq - 201772/2014-0

contribuindo não só para a riqueza patrimonial do país, como também para a compreensão da própria história de um concelho.

No contexto aqui relatado, estamos particularmente interessados em reunir documentação sobre a emigração a fim de criar um Museu Virtual de Emigração.

A problemática aqui discutida, neste contexto, tem a ver com preservação digital dos documentos e a sua exploração na Web com base em ontologias. Pode-se então dizer que a ideia central passa por atender um fundo documental do Arquivo (os processos de requisição de passaportes de emigração) para “fazer” história de um determinado fenômeno social.

Em geral é sabido que as bibliotecas, arquivos e museus possuem diversos documentos em papel. No entanto, para se poder consultar e estudar a informação presente nesses documentos, torna-se uma tarefa difícil, além de que, o excesso de manuseio do papel faz com que este se deteriore rapidamente.

Perante isto, verifica-se a necessidade de conservar todos os documentos presentes nestes repositórios, ou seja, evitar excesso de manuseio do papel, tornar o conteúdo dos documentos acessível à distância, criando para isso repositórios documentais, por exemplo na Web, onde qualquer pessoa consiga pesquisar, relacionar e perceber a informação que lhe é fornecida, organizar e combinar os documentos entre si, por exemplo, biografias, almanaques, pedidos de passaporte, passaportes, entre outros, acrescentando deste modo, um valor significativo aos dados existentes, etc. Isto permite facilitar a pesquisa e melhorar a qualidade da experiência do utilizador final.

O objetivo deste artigo incide na criação de uma ontologia específica, a partir de um padrão definido para os museus (CIDOC-CRM) e na exploração da ontologia, onde se pode realizar pesquisas através de *queries* SPARQL.

Neste artigo apresentamos, na Seção 2 o Museu da Emigração e os documentos arquivados no mesmo. Na Seção 3 descortina-se o significado e o objetivo de uma ontologia CIDOC-CRM, e como se construiu uma ontologia baseada na mesma, mas totalmente compatível com a original. Já na Seção 4, são abordadas as formas de representação dos dados para facilitar o armazenamento e consultas de triplas (sujeito, predicado e objeto). Na Seção 5, descreve-se o armazenamento e consulta das triplas em formato RDF. Por fim, na Seção 6, é realizada uma síntese do trabalho realizado, e também são abertas questões de trabalhos futuros.

## 2. O Museu da Emigração e seus documentos

De acordo com o Concelho Internacional de Museus (ICOM), um museu é uma instituição permanente e sem fins lucrativos servindo a sociedade e seu desenvolvimento, aberta para o público. Um museu tem o papel de adquirir, conservar, pesquisar, comunicar e exibir o patrimônio cultural tangível e intangível da humanidade e é um ambiente que propõe a educação, o estudo e a apreciação<sup>1</sup>.

O ICOM ainda define Museus Virtuais como uma coleção logicamente relacionada de objetos digitais compostos por uma variedade de meios de comunicação, por isso, um museu virtual não tem lugar ou espaço real, seus objetos e as informações relaciona-

<sup>1</sup><http://icom.museum/the-vision/museum-definition/>

das podem ser disseminadas por todo o mundo, independente de localização geográfica [Schweibenz 2004].

O Museu de Emigração incorpora a dimensão comunicacional e simbólica, sendo, por isso, um projeto organizado para ser entendido como plataforma informativa e de dinamização de atividades de pesquisa e divulgação. Este está organizado nas seguintes vertentes: salas temáticas, casa museu, Arquivo, núcleos museológicos e sítios históricos, tendo em conta o município de origem e a estrutura de serviços [Monteiro 2007].

Neste caso concreto, o foco é o Arquivo histórico. O Arquivo, no sentido da descoberta dos indivíduos e dos quotidianos, procura recuperar documentos e objetos usados pelos emigrantes e descendentes, solicitando a doação ou depósito à guarda do museu, contribuindo, deste modo, para a investigação e estimulando a preservação e estudo da história da emigração e do emigrante [Monteiro 2007].

Assim, a coleção dos documentos de emigração tratada neste projeto abrange a época entre 1960 e 1970 e pertence ao Arquivo municipal de Fafe (cidade situada ao norte de Portugal), no entanto, este projeto pode ser aplicado a qualquer período de tempo.

Os documentos de emigração caracterizados neste projeto restringem-se aos que descrevem os dados necessários dos emigrantes para solicitar um passaporte. Estes contêm informações como dados gerais do próprio documento (Documento de identificação), dados gerais sobre o emigrante, lista dos documentos contidos em anexo, os membros da família que acompanham o emigrante, tipo de transporte desejado, qualificações (acadêmicas e profissionais) e antecedentes criminais, familiares que ficam no país à responsabilidade do emigrante, emigrações anteriores (caso tivesse realizado), detalhes da pessoa que “chamou” o emigrante, o contrato de trabalho, a ajuda no país de destino, entre outras informações. Todos esses itens (mais de 80) devem ser fornecidos para obter um passaporte [Martini et al. 2015].

Este conjunto de informações possui um grande potencial para descrever cada indivíduo integrado na sociedade da sua época, mas também fornece o conhecimento sobre a sociedade num contexto preciso da história do país [Martini et al. 2015].

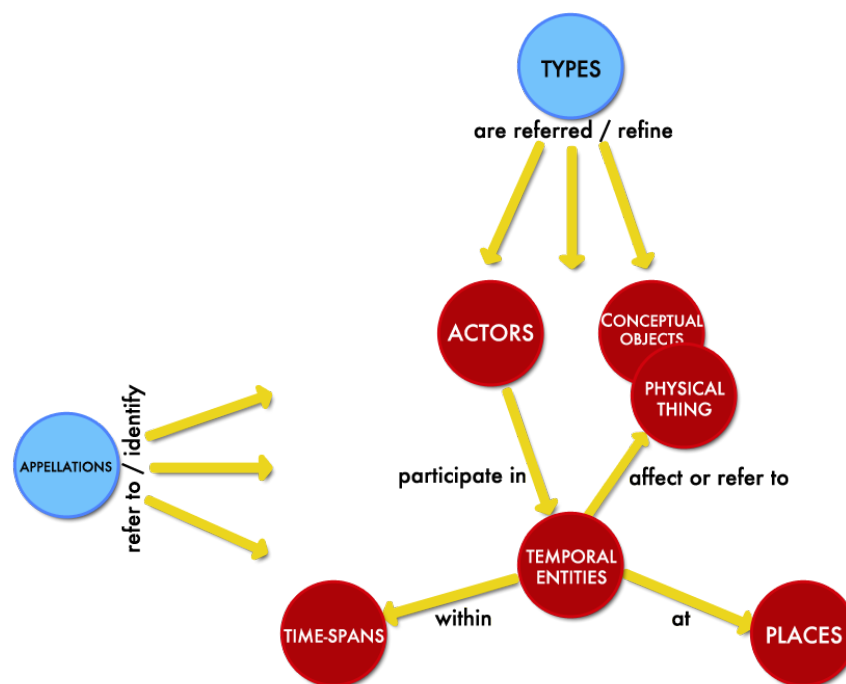
### 3. A Ontologia CIDOC-CRM

O uso de ontologias na computação já vem sendo estudado há vários anos. Inicialmente na área de Inteligência artificial e mais tarde em outras áreas da computação, as quais desejam manter conceitos relativos ao domínio da aplicação [Gava and Menezes 2003]. Uma das importâncias de seu uso é devido à capacidade de representar hierarquias de classes de objetos (taxonomias) e seus relacionamentos [Librelotto et al. 2008].

Como dito anteriormente na Seção 2, o domínio da aplicação tratada nesse trabalho é voltada ao patrimônio cultural, mais especificadamente sobre o fenômeno da emigração em Portugal nos anos 1960 até 1970. Para descrever e trabalhar esse domínio, uma ontologia padrão específica para a descrição de museus chamada CIDOC-CRM foi utilizada.

O objetivo da ontologia CIDOC-CRM é promover um entendimento compartilhado do domínio de herança cultural, fornecendo um *framework* semanticamente extensível e comum onde qualquer informação sobre o domínio possa ser mapeado [ICOM/CIDOC 2013][Oldman and Labs 2014].

Para entender como a ontologia CIDOC-CRM é organizada, a Figura 1 ilustra seu núcleo principal.



**Figura 1. Estrutura principal da ontologia CIDOC-CRM (adaptado de [Oldman and Labs 2014])**

A ontologia CIDOC-CRM é baseada em eventos, como pode-se notar na Figura 1, onde as entidades principais são chamadas *Temporal Entities*, ou seja, classes que tratam eventos que aconteceram no passado. Como as classes *Temporal Entities* são relacionadas ao tempo, elas podem ter datas e tempos relacionados aos *Time-Spans*. Já as classes *Actors*, *Conceptual Objects*, *Physical Things* and *Places* não podem ser diretamente relacionadas ao tempo (*Time-Spans*). Assim, elas precisam ser associadas a um evento (*Temporal Entity*).

Um lugar (*Place*) pode ser qualquer coisa que descreva uma localização, seja ela geográfica ou por exemplo dentro de um carro ou no topo do monte Everest.

Atores (*Actors*) são entidades que possuem uma responsabilidade legal. Um ator pode ser um indivíduo ou um grupo. Um indivíduo pode ser uma pessoa e um grupo pode ser uma empresa, por exemplo. Atores interagem com coisas (*Conceptual Objects* and *Physical Things*) através de eventos.

Uma coisa física (*Physical Thing*) é algo que pode ser fisicamente destruída e transformada (criada) em algo novo, se preservada alguma parte do todo. Já objetos conceituais (*Conceptual Objects*) não podem ser destruídos. Por exemplo, um objeto físico como um livro, uma pintura, pode ser destruído, mas o conteúdo (informação) não pode. Para destruir um objeto conceitual, é necessário destruir tudo que retrata aquele conceito, inclusive as pessoas.

Coisas no CIDOC-CRM podem ter designações, ou seja, *Appellations*. Uma designação pode ser um nome, um número de identificação, etc. Além disso, diferen-

tes organizações possuem diferentes tipos de classificação. Na ontologia CIDOC-CRM, classificações são chamadas de tipos (*Types*) e elas servem para a classificação de coisas. São exemplos de tipos de eventos: casamento, terremoto, guerra, aniversário, nascimento, etc.

Além disso, a ontologia CIDOC-CRM possui notações que devem ser seguidas, como por exemplo: uma entidade sempre inicia seu nome com a letra maiúscula “E” seguida de um código numérico (ex: E21 Person). As relações são iniciadas por “P” maiúsculo seguido de um código numérico (ex: P131 is identified by).

Após descrever a ontologia CIDOC-CRM e o conceito de museus, nós podemos descrever como foi desenvolvida uma ontologia para o Museu da Emigração baseada na ontologia padrão CIDOC-CRM. Essa ontologia será apresentada na Seção 3.1.

### 3.1. Uma ontologia para o Museu da Emigração

Após uma análise intensiva do CIDOC-CRM e de todas as suas classes e propriedades, conseguiu-se perceber quais as entidades compatíveis com os dados presentes no fundo documental do Arquivo Municipal de Fafe.

Deste modo, construiu-se uma ontologia baseada na CIDOC-CRM, mas totalmente compatível com a original. Quando uma ontologia é criada baseada na ontologia CIDOC-CRM e não fere algumas regras impostas, ela é chamada de uma ontologia reduzida, ou o termo dado pelos criadores, em inglês, *Reduced CRM-compatible form*. Sendo assim, nós criamos uma ontologia baseada na CIDOC-CRM sem modificar absolutamente nada (entidades e relacionamentos), ou seja, nenhuma classe ou relacionamento da ontologia original foi modificado, nem foi criado algo novo. Deste modo, temos uma ontologia considerada *Reduced CRM-compatible form*.

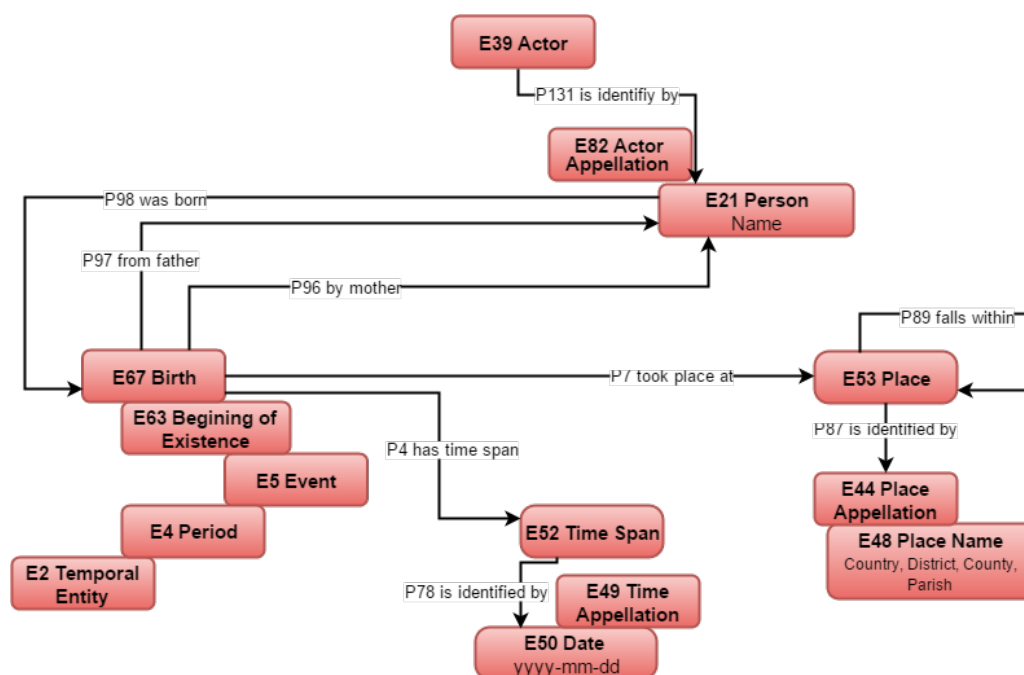
Por motivos de espaço, na Figura 2 é apresentada uma versão reduzida da ontologia criada. Mostramos um exemplo descrevendo o evento de nascimento de uma pessoa, a qual pode ser, em nosso domínio, um emigrante, um contratante, um chamante, um acompanhante, etc.

Como pode-se ver na Figura 2, o evento nascimento (*E67 Birth*), o qual é uma subclasse de *E5 Event*, é composto pelo local de nascimento (*E53 Place*) através do relacionamento *P7 took place at*; data de nascimento (*E52 Time-Span*) através do relacionamento *P4 has time-span*; nome do pai (*E21 Person*) através do relacionamento *P97 from father*; e nome da mãe (*E21 Person*) através do relacionamento *P96 by mother*.

Após definirmos a ontologia para o Museu da Emigração, notamos a necessidade de representar os dados reais do fundo documental do Arquivo municipal de Fafe e avaliar se a forma como criamos a ontologia estava correta. A Seção 4 apresenta as formas de representação dos dados reais que utilizamos para comprovar que a ontologia para o Museu da Emigração está de acordo com os documentos reais e nossos objetivos.

## 4. Formas de representação dos dados

Para definir e utilizar uma ontologia é necessário uma representação concreta. Há uma variedade de linguagens de representação de ontologias que podem ser usadas para este propósito, como *eXtensible Markup Language (XML)*, *Resource Description Framework*



**Figura 2. Entidades e relacionamentos atribuídos à descrição do evento de nascimento de uma pessoa**

(RDF), *Web Ontology Language (OWL)*, entre outras. Elas variam em termos de expressividade [McGuinness and Harmelen 2004].

Sabendo que o CIDOC-CRM é descrito, em sua forma original, em RDF, nós aqui neste trabalho também o utilizamos para descrever todas as informações reais do fundo documental do Arquivo municipal de Fafe para o pedido de passaporte por parte do emigrante. Porém, antes de escolhermos e descrevermos as informações em formato RDF, foi necessário um estudo de qual formato de representação de triplas utilizar.

A fim de colocar as informações em uma forma organizada já pensando nas instâncias de nossa ontologia já criada, como primeiro passo, descrevemos as informações do fundo documental em formato texto (TXT) para posteriormente analisarmos e definirmos a sintaxe a ser utilizada para as triplas.

#### 4.1. Triplas TXT

Após a criação da ontologia, era crucial verificar se esta era apropriada para os dados existentes no fundo documental.

Para resolver o problema, foram criadas, de forma manual, instâncias de triplas com informação real, num documento TXT.

A fim de clarificar e demonstrar a forma de organização das triplas TXT, em *Listing 4-1* é retratado um trecho das triplas a partir da informação real dos documentos de pedidos de passaporte.

##### Listing 4-1. Triplas TXT - Evento Nascimento Emigrante

```
1 E21 Person: Emigrante 1; P131 is identified by; E82 Actor
  Appellation: Manuel Vaz
```

```

2 E21 Person : Emigrante 1 ; P98 was born ; E67 Birth : B1
3 E67 Birth : B1; P97 from father ; E21 Person : Pai Emigrante 1
4 E21 Person: Pai Emigrante 1; P131 is identified by; E82 Actor
  Appellation: Julio Vaz
5 E67 Birth: B1; P96 by mother; E21 Person: Mae Emigrante 1
6 E21 Person: Mae Emigrante 1; P131 is identified by; E82 Actor
  Appellation: Virginia Delgado
7 E67 Birth: B1; P4 has time span; E52 Time-Span: TS1
8 E52 Time-Span: TS1; P78 is identified by; E52 Time-Span:
  1924.03.26
9 E67 Birth: B3; P7 took place at; E53 Place: PL1
10 E53 Place: PL1; P87 is identified by; E48 Place Name: Monte
11 E53 Place: PL1; P89 falls within; E53 Place: PL2
12 E53 Place: PL2; P87 is identified by; E48 Place Name: Fafe
13 E53 Place: PL2; P89 falls within; E53 Place: PL3
14 E53 Place: PL3; P87 is identified by; E48 Place Name: Braga

```

Neste trecho de código estão listados, em forma de triplas, todas as informações relacionadas ao nascimento do Emigrante 1, o qual possui o nome de Manuel Vaz. Cada linha do *Listing 4-1* representa uma tripla.

Somente a tripla da linha 1 será explicada (para todas as demais, segue a mesma lógica). Pode-se ver na linha 1 (E21 Person: Emigrante 1; P131 is identified by; E82 Actor Appellation: Manuel Vaz) que o sujeito, predicado e objeto, separados por ponto e vírgula, são:

- Sujeito: a instância *Emigrante 1* que pertence a classe *E21 Person*;
- Predicado: o relacionamento *P131 is identified by* que relaciona o Sujeito com o Objeto;
- Objeto: a instância *Manuel Vaz* que pertence a classe *E82 Actor Appellation*.

Sendo assim, podemos dizer que o *Emigrante 1*, o qual é uma pessoa (*E21 Person*) é identificado por (*P131 is identified by*) *Manuel Vaz*, o qual é uma designação de nome para o emigrante e que pertence à classe *E82 Actor Appellation*.

A partir dessa organização de ideias no formato TXT e comprovação de que é possível descrever as informações reais do fundo documental em triplas, verificou-se a necessidade de representar as triplas em formato RDF a fim de permitir a realização de consultas sobre as mesmas, utilizando SPARQL. A Seção 4.2 descreve como esse trabalho foi realizado.

## 4.2. Triplas RDF

A partir da necessidade de descrever as triplas do fundo documental da emigração em um formato que ajudasse no propósito do nosso trabalho, foram criadas e armazenadas um conjunto de triplas em uma base de dados para posterior análise e consulta.

Um trecho das triplas em formato RDF/XML reduzido é mostrado no *Listing 4-2*. Note que todos os relacionamentos referentes às instâncias *Emigrante\_1* e *B1* estão descritos de acordo com o *Listing 4-1*.

#### Listing 4-2. Triplas RDF - Evento Nascimento Emigrante

```
1 <rdf:Description rdf:about="Emigrante_1">
2     <rdf:type rdf:resource="E21_Person"/>
3     <P131_is_identified_by rdf:resource="Manuel_Vaz"/>
4     <P98i_was_born rdf:resource="B1"/>
5 </rdf:Description>
6
7 <rdf:Description rdf:about="B1">
8     <rdf:type rdf:resource="E67_Birth"/>
9     <P98_brought_into_life rdf:resource="Emigrante_1"/>
10    <P97_from_father rdf:resource="Pai_Emigrante_1"/>
11    <P96_by_mother rdf:resource="Mae_Emigrante_1"/>
12    <P4_has_time-span rdf:resource="TS1"/>
13    <P7_took_place_at rdf:resource="PL1"/>
14 </rdf:Description>
```

O próximo passo a ser feito, depois da descrição das informações reais do fundo documental em RDF/XML, para atingir nossos objetivos, é trabalhar sobre essas informações formatadas, ou seja, ter uma forma de armazená-las e consultá-las. A Seção 5 descreve essa importante etapa do nosso trabalho.

### 5. Armazenamento e consulta das triplas RDF

Como já foi referido anteriormente, utilizou-se RDF para descrever as informações do nosso fundo documental. No entanto, para ter acesso a estes dados e poder realizar algumas consultas sobre eles, foi necessário criar e armazenar triplas em RDF. Para armazenar as triplas em um conjunto de dados, utilizou-se uma base de dados RDF (Apache Jena TDB).

Após as informações necessárias estarem armazenadas no Apache Jena TDB, o próximo passo era executar certas consultas SPARQL sobre as triplas lá armazenadas. Para isso foi necessário utilizar um *end-point*, novamente da Apache, chamado Fuseki, em sua versão 2.0.

Algumas consultas, com o apoio de uma arquivista do Arquivo Municipal de Fafe, foram elaboradas e testadas, todas executadas com sucesso. As consultas respondem questões como: (1) emigrantes e o respectivo país de emigração; (2) emigrantes e seu destino, por sexo; (3) emigrantes por destino e seus acompanhantes; (4) emigrantes por sexo, destino e data; (5) emigrantes cujo chamante é o cônjuge e respetivo destino; (6) total de emigrantes por país; (7) emigrantes e idade de emigração; (8) emigrantes e seus acompanhantes e idade de emigração.

Por questão de espaço, somente uma consulta é explicada. No *Listing 5-3* é apresentada a consulta SPARQL referente à *query* número 8.

#### Listing 5-3. Consulta SPARQL: emigrantes e seus acompanhantes e idade de emigração

```
1 PREFIX ecrm: <http://erlangen-crm.org/140617/>
2 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
3 prefix xsd: <http://www.w3.org/2001/XMLSchema#>
```



```

4
5 SELECT DISTINCT ?nomeEmigrante ?dtMOVE ?dtNascEMIGRANTE ?
   nomeAcompanhante ?dtNascACOMPANHANTE
6 WHERE {
7   ?pessoa a ecrm:E21_Person ;
8     ecrm:P131_is_identified_by ?ActorApellation ;
9     ecrm:P11i_participated_in ?move .
10  ?ActorApellation ecrm:P3_has_note ?nomeEmigrante.
11  ?pessoa ecrm:P98i_was_born ?dtNascimento.
12  ?dtNascimento ecrm:P4_has_time-span ?dt.
13  ?dt ecrm:P78_is_identified_by ?dtNascEMIGRANTE.
14  ?move ecrm:P4_has_time-span ?dtMove.
15  ?dtMove ecrm:P78_is_identified_by ?dtMOVE.
16 OPTIONAL{
17   ?move ecrm:P14_carried_out_by ?pessoa2 .
18   ?pessoa2 ecrm:P131_is_identified_by ?ap2 .
19   ?pessoa2 ecrm:P3_has_note "Acompanhante"^^xsd:string .
20   ?ap2 ecrm:P3_has_note ?nomeAcompanhante .
21 OPTIONAL{
22   ?pessoa2 ecrm:P98i_was_born ?dtNascimentoAcompanhante.
23   ?dtNascimentoAcompanhante ecrm:P4_has_time-span ?
     dtAcompanhante.
24   ?dtAcompanhante ecrm:P78_is_identified_by ?
     dtNascACOMPANHANTE.
25 }
26 }
27 }

```

Entre as linhas 7 e 10 do *Listing 5-3* pretende-se procurar uma ou mais pessoas (E21 Person) que tenham participado num evento de Emigração (P11i participated in) e que é identificada por uma denominação (E82 Actor Appellation). A classe E82 Actor Appellation possui uma nota (P3 has note) onde descreve o nome da pessoa. Como se pode observar na linha 5 'nomeEmigrante' é uma das variáveis que desejamos como resposta.

Para saber a idade da pessoa é necessário obter a data de nascimento da pessoa e a data de emigração, isso é realizado entre as linhas 11 e 15 (dtNascEMIGRANTE e dtMOVE, respectivamente).

Como também existem emigrantes que levaram acompanhantes aquando a sua emigração, torna-se necessário repetir o mesmo processo descrito acima, para identificar o acompanhante (nomeAcompanhante), linhas 17 a 20.

Alguns desses acompanhantes possuem data de nascimento (dtNascACOMPANHANTE) e o processo para a identificar é semelhante a data de nascimento do emigrante e está descrito entre as linhas 22 e 24.

É importante salientar, nesta consulta, a utilização do *OPTIONAL*. Este é utilizado porque no fundo documental em estudo, existem emigrantes que levam acompanhantes e outros não. Para os que emigram sozinhos, a *query* encerra na linha 15. Já para os que

possuem acompanhantes, a *query* executa o primeiro *OPTIONAL* que identifica o nome do acompanhante. Existem ainda acompanhantes que não possuem identificadas, no fundo documental, a sua data de nascimento. Dessa forma, a *query* retorna somente o nome do emigrante, sua idade no momento da emigração e, caso possua, o nome do(s) acompanhante(s). Para obter mais informações sobre os resultados e consultas executadas, favor consultar a página <http://www4.di.uminho.pt/~gepl/coming2015>.

## 6. Conclusão

O principal objetivo deste artigo foi descrever um fundo documental em um formato entendível tanto por máquina quanto por humanos, a fim de extrair informações sobre determinado domínio com o intuito de estudar e repassar a história que os documentos do Arquivo possuem.

Contudo, para ser processado por máquina, este fundo documental deve ser descrito em um formato estruturado. Sendo assim, utilizou-se RDF para a descrição do fundo documental no momento de criação da ontologia para o Museu da Emigração.

A criação e armazenamento das triplas RDF habilitou a realização de consultas e extração de informações reais sobre o fenômeno de emigração em Portugal entre os anos 1960 e 1970.

O trabalho aqui descrito é de suma importância para a finalização do projeto principal, visto que existe um projeto focado na criação de ambientes virtuais de aprendizagem (salas de exposição virtuais) que necessita dos resultados desse trabalho prévio para atingir os objetivos principais.

Como trabalho futuro, idealizou-se criar uma ferramenta que através dos dados do fundo documental, crie o RDF de forma automática. Isto porque a experiência com as triplas TXT e a criação do RDF foi efetuada manualmente, no entanto, torna-se difícil de o fazer para grandes quantidades de informação.

Uma outra tarefa para o futuro seria reunir informação não só dos pedidos de passaporte, como também de outras fontes (biografias, cartas, atestados de residência, etc.), a fim de enriquecer a pesquisa e tornar a história de um determinado fenômeno social disponível nesses ambientes virtuais de aprendizagem como museus virtuais.

**Agradecimentos:** ao CNPq, por ter possibilitado e financiado esta pesquisa. Ao Arquivo Municipal de Fafe, na pessoa da sua diretora Dr<sup>a</sup> Mónica Guimarães, pela informação disponibilizada e apoio prestado.

## Referências

- Gava, T. B. S. and Menezes, C. S. D. (2003). Uma ontologia de domínio para a aprendizagem cooperativa. *XIV Simpósio Brasileiro de Informática na Educação NCE IMUFRJ*, pages 336–345.
- ICOM/CIDOC (2013). Definition of the cidoc conceptual reference model. Technical report, ICOM/CIDOC.
- Librelotto, G. R., Freitas, L. O., Gassen, J. B., Copetti, M., Turchetti, R. C., da Silva, F. L., and Augustin, I. (2008). Uma ferramenta para o processamento da representação do domínio de atividades médicas. *Hifen*, pages 25–32.

- Martini, R., Guimarães, M., Librelotto, G., and Henriques, P. (2015). Storing archival emigration documents to create virtual exhibition rooms. In Rocha, A., Correia, A. M., Costanzo, S., and Reis, L. P., editors, *New Contributions in Information Systems and Technologies*, volume 353 of *Advances in Intelligent Systems and Computing*, pages 403–409. Springer International Publishing.
- McGuinness, D. L. and Harmelen, F. V. (2004). Owl web ontology language overview.
- Monteiro, M. (2007). O museu da emigração e os “brasileiros” do rio: o público e o privado na construção de modernidade em Portugal. *Revista da Faculdade de Letras História*, 8.
- Oldman, D. and Labs, C. (2014). The CIDOC Conceptual Reference Model (CIDOC-CRM): Primer. *International Council of Museums (ICOM)*, 1.
- Schweibenz, W. (2004). The development of virtual museums. In *Virtual Museums*, volume 57(3). ICOM.