

MESTRADO EM CIÊNCIA DA INFORMAÇÃO

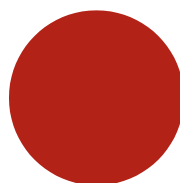
ENTIDADES EM DOCUMENTOS DE ARQUIVO E SUA EXPANSÃO COM FONTES DE DADOS LIGADOS NO PROJETO EPISA

Camilla Oliveira da Silveira

M
2022

UNIDADES ORGÂNICAS ENVOLVIDAS

FACULDADE DE ENGENHARIA
FACULDADE DE LETRAS



Camilla Oliveira da Silveira

ENTIDADES EM DOCUMENTOS DE ARQUIVO E SUA EXPANSÃO
COM FONTES DE DADOS LIGADOS NO PROJETO EPISA

Dissertação realizada no âmbito do Mestrado em Ciência da Informação, orientada
pela Doutora Maria Cristina de Carvalho Alves Ribeiro.

Faculdade de Engenharia e Faculdade de Letras
Universidade do Porto

Julho de 2022

ENTIDADES EM DOCUMENTOS DE ARQUIVO E SUA EXPANSÃO COM FONTES DE DADOS LIGADOS NO PROJETO EPISA

Camilla Oliveira da Silveira

Dissertação realizada no âmbito do Mestrado em Ciência da Informação, orientada
pela Doutora Maria Cristina de Carvalho Alves Ribeiro.

Membros do Júri

Presidente: Doutor António Manuel Lucas Soares

Professor Associado da Faculdade de Engenharia da Universidade do Porto.

Arguente: Doutora Mariana Curado Malta

Professora Adjunta do Instituto Superior de Contabilidade e Administração do Porto
e do Instituto Politécnico do Porto.

Orientadora: Doutora Maria Cristina de Carvalho Alves Ribeiro

Professora Associada da Faculdade de Engenharia da Universidade do Porto.

Agradecimentos

Primeiramente, gostaria de agradecer aos meus pais pelo contínuo suporte e incentivo ao longo de todo esse percurso.

À minha irmã e cunhado pelo apoio e assistência nesses últimos anos.

À minha orientadora Cristina Ribeiro pela disponibilidade e direcionamento que foram fundamentais para a realização desta dissertação.

À minha amiga e companheira de mestrado Aline Araújo por todo o seu apoio emocional e encorajamento durante esses dois anos de mestrado.

Ao meu namorado Jakub pela paciência e incentivo tão necessários neste momento, e cujo apoio tecnológico foi essencial para a conclusão deste trabalho.

Resumo

O Arquivo Nacional da Torre do Tombo (ANTT) é um exemplo de instituição que mantém uma coleção de grande relevância para a memória nacional portuguesa. Nele estão guardados documentos originais que vão desde o século IX até os dias de hoje, muitos já digitalizados, podendo ser pesquisados tanto por investigadores como pela população em geral. O arquivo tem vindo a adaptar-se para abraçar uma perspectiva aberta em que os dados arquivísticos estão ligados a fontes de informação globais, seguindo uma tendência mundial de tornar a informação dos repositórios do património cultural mais acessível. Este trabalho analisou uma amostra de 25 registos do Arquivo Nacional da Torre do Tombo, escolhidos pelos especialistas em arquivo como representativos de diferentes grupos de fundos e níveis de descrição, para identificar entidades e propriedades e explorar relações com outros recursos não arquivísticos. O objetivo é fornecer informação adicional às interfaces de arquivo. Após a seleção das entidades dos registos, foram escolhidas as bases de dados da Wikidata, DBpedia e Europeana para o enriquecimento e ligação de dados. A análise das entidades nos registos permitiu identificar centenas de propriedades diretamente relacionadas com classes nas respectivas bases de dados. Um conjunto mais significativo de propriedades foi obtido através da seleção de propriedades para as quais existem mais dados nas bases de dados para entidades das classes correspondentes. Em seguida, essas entidades e propriedades foram submetidas a testes utilizando o Wikidata Query Service, o DBpedia SPARQL Explorer e o Virtuoso SPARQL Query Editor na Europeana. As pesquisas realizadas com as entidades no Wikidata Query Service e no DBpedia SPARQL Explorer apenas forneceram resultados para determinadas propriedades, uma vez que nem todas as entidades possuem dados nas respetivas bases de dados. Por outro lado, as propriedades testadas provaram possibilitar dados relevantes sobre as entidades, fornecendo informação adicional sobre as entidades com as quais se relacionam. Com a Europeana as pesquisas com as entidades não se mostraram produtivas, não tendo sido possível recuperar informação sobre propriedades de entidades como realizado nas bases de dados anteriores. A Europeana é uma grande agregadora de metadados, porém esses metadados ainda se encontram em processo de conversão em dados abertos ligados. Por fim, mais estudos em repositórios se mostram necessários a fim de investigar a viabilidade de explorar as ligações entre os recursos de arquivo e a informação em grandes arquivos abertos.

Palavras-chave: arquivos, dados abertos ligados, web semântica

Abstract

The Portuguese National Archives (*Arquivo Nacional da Torre do Tombo*) is an example of an institution that maintains a collection of great relevance to the Portuguese national memory. It holds original documents from the 9th century to the present day, many already digitized and available to be searched by researchers and the general public. The archive has been adapting to embrace an open perspective in which archival data is linked to global information sources, following a worldwide trend of making information from cultural heritage repositories more accessible. This work analyzed a sample of 25 records from the Portuguese National Archives, chosen by archival specialists as representative of different fonds and description levels, to identify entities and properties and explore relationships with other non-archival resources. The goal is to provide additional information to the archival interfaces. After selecting the entities in the records, Wikidata, DBpedia and Europeana databases were chosen for data enrichment and linking. The analysis of the entities in the records provided hundreds of properties directly related to classes in the corresponding databases. A more significant set of properties was obtained by selecting properties that had more data entered into the databases for entities of the corresponding classes. These entities and properties were then subjected to tests using Wikidata Query Service, DBpedia SPARQL Explorer and the Europeana Virtuoso SPARQL Query Editor. The queries performed with the entities in Wikidata Query Service and DBpedia SPARQL Explorer only provided results for certain properties, since not all entities have data in the database. On the other hand, the properties tested proved to enable relevant data about the entities, providing additional information about the entities they are connected with. However, with Europeana the entity searches did not prove productive, and it was not possible to retrieve information about entity properties as performed in the previous databases. Europeana is a great aggregator of metadata, however its metadata is still in the process of being converted into linked open data. Finally, further studies on repositories are needed in order to investigate the feasibility of exploring the links between archival resources and information in large open archives.

Keywords: archives, linked open data, semantic web

Lista de ilustrações

Figura 1 - Utilizador nos Serviços de Busca do Arquivo	13
Figura 2 - Arquivista nos Serviços de Descrição do Arquivo	13
Figura 3 - Árvore de Objetivos	15
Figura 4 - As Camadas da Web Semântica.....	18
Figura 5 - Representação de um documento RDF	20
Figura 6 - Nuvem de dados abertos ligados	24
Figura 7 - Fragmento da grelha dos 25 registos utilizados pelo Projeto EPISA	37
Figura 8 - Registo 1 - Juízo da Índia e Mina	38
Figura 9 - Registo 11 – Venda de Louça Preta.....	39
Figura 10 - Registo 12 – Mosteiro de São Bernardo de Portalegre	40
Figura 11 - Teste com a entidade Mouzinho da Silveira na Wikidata	47
Figura 12 - Teste com a entidade Adriano VI na Wikidata	48
Figura 13 - Teste com a entidade Mouzinho da Silveira na DBpedia	49
Figura 14 - Teste com a entidade Adriano VI na DBpedia	50
Figura 15 - Teste com a entidade Mouzinho da Silveira no Virtuoso SPARQL Query Editor ..	51

Lista de tabelas

Tabela 1 - Fragmento da Tabela de Entidades e Propriedades Associadas aos 25 Registos no CIDOC CRM, Schema.org e Dublin Core	41
Tabela 2 – Fragmento da Tabela de Classes das Entidades e Propriedades Encontradas Associadas ao Registo 1 na Wikidata, Europeana e DBpedia Após Primeira Seleção	42
Tabela 3 - Fragmento da Tabela Resumida de Classes das Entidades e Propriedades Associadas aos 25 Registos na Wikidata, Europeana e DBpedia.....	43
Tabela 4 – Fragmento da Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia	44
Tabela 5 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia	45
Tabela 6 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata Após Testes no Wikidata Query Service	47
Tabela 7 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na DBpedia Após Testes no DBpedia SPARQL Explorer.....	49

Lista de abreviaturas e siglas

ANTT - Arquivo Nacional da Torre do Tombo

CCo - Creative Commons Zero

CIDOC - Centro Intercultural de Documentación

CRM - Conceptual Reference Model

CRM SIG - Conceptual Reference Model Special Interest Group

CSIG - Centro de Sistemas de Informação e de Computação Gráfica

CSV - Comma Separated Values

DCAT - Data Catalog Vocabulary

DCMES - Dublin Core Metadata Element Set

DCMI - Dublin Core Metadata Initiative

DCTERMS - DCMI Metadata Terms

DGARQ - Direcção-Geral de Arquivos

DGLAB - Direcção-Geral do Livro, dos Arquivos e das Bibliotecas

EAD - Encoded Archival Description

EDM - Europeana Data Model

EPISA - Inferência de Entidades e Propriedades para Arquivos Semânticos

ESE - Europeana Semantic Elements

ETH - Eidgenössische Technische Hochschule Zürich

FCT - Fundação para a Ciência e a Tecnologia

FinnONTO - National Semantic Web Infrastructure in Finland

FOAF - Friend of a Friend Ontology

GABRIEL - Projeto Gateway and Bridge to Europe's National Libraries

GTA - Institut für Geschichte und Theorie der Architektur

HD - Humanidades Digitais

HTML - HyperText Markup Language

ICA - Conselho Internacional de Arquivos

INESC TEC - Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência

ISAD - International Standard Archival Description

ISO - International Organization for Standardization

JSON-LD - Java Script Object Notation - Lightweight Linked Data Format

LDF - Linked Data Finland

LIDO - Lightweight Information Describing Objects

LOD - Linked Open Data

LODI4DH - Linked Open Data Infrastructure for Digital Humanities in Finland

METS - Metadata Encoding and Transmission Standard

NDE - Dutch Digital Heritage Network

NISV - Netherlands Institute for Sound and Vision

NS - Namespace

OAI-ORE - Open Archives Initiative Object Reuse and Exchange

OAIPMH - Open Archives Initiative Protocol for Metadata Harvesting

OKG - Open Knowledge Graph

ONKI - Finnish Ontology Library Service

OWL - Web Ontology Language

OWL DL - Web Ontology Language Description Logic

PDF - Portable Document Format

RDF - Resource Description Framework

RiC - Records in Context

RNOD - Registo Nacional de Objetos Digitais

SARI - Swiss Art Research Infrastructure

SeCo - Semantic Computing Research Group

SPARQL - Simple Protocol and RDF Query Language

SKOS - Simple Knowledge Organization System

SQL - Structured Query Language

TEL - The European Library

TOMS - Thailand Open Museum System

UH - University of Helsinki

UNIMARC - Universal MARC Format

URI- Universal Resource Identifier

URL - Uniform Resource Locator

VIAF - Virtual International Authority File

WissKI - Wissenschaftliche KommunikationsInfrastruktur

WWW - World Wide Web

W3C - World Wide Web Consortium

XML - Extensible Markup Language

3D - Three dimensional

Índice de conteúdo

1.	Introdução	11
1.1.	Enquadramento	11
1.2.	Definição do Problema	12
1.3.	Motivação	13
1.4.	Objetivos e Resultados Esperados	14
1.4.1.	Árvore de Objetivos	15
1.5.	Metodologia	15
1.6.	Estrutura da Dissertação	17
2.	Revisão de Literatura	18
2.1.	Web Semântica	18
2.1.1.	Estrutura de Descrição de Recursos – RDF	19
2.1.2.	Simple Protocol and RDF Query Language – SPARQL	20
2.1.3.	Ontologia	21
2.2.	Bases de Dados em RDF	22
2.2.1.	DBpedia	22
2.2.2.	Wikidata	23
2.3.	Dados Ligados no Patrimônio Cultural	23
2.3.1.	Modelo CIDOC CRM	26
2.3.2.	Modelo de Dados da Europeana – EDM	27
2.3.3.	Modelo Schema.org	30
2.3.4.	Modelo Dublin Core	30
2.4.	Trabalhos Relacionados	31
3.	Entidades em Registos de Arquivo e em Bases de Dados Ligados	34
3.1.	O Projeto DigitArq	34
3.2.	Introdução ao Projeto EPISA	35
3.3.	Expansão de Entidades nos Registos de Arquivo do DigitArq	36
3.4.	Fontes Exteriores de Enriquecimento e Ligação de Dados	41
4.	Testes de Entidades e de Propriedades dos Registos	46
4.1.	Testes com Wikidata Query Service	46
4.2.	Testes com DBpedia SPARQL Explorer	48
4.3.	Testes com Europeana SPARQL API	51
5.	Conclusões e perspetivas de desenvolvimento	54
6.	Referências bibliográficas	58

Apêndices	63
APÊNDICE A – Tabela de Entidades e Propriedades Associadas aos 25 Registos no CIDOC CRM, Schema.org e Dublin Core.....	63
APÊNDICE B - Lista de Classes das Entidades e Propriedades Encontradas Associadas Apenas ao Registo 1 na Wikidata, Europeana e DBpedia.	76
APÊNDICE C - Tabela de Classes das Entidades e Propriedades Encontradas Associadas ao Registo 1 na Wikidata, Europeana e DBpedia Após Primeira Seleção...	85
APÊNDICE D - Tabela Resumida de Classes das Entidades e Propriedades Associadas aos 25 Registos na Wikidata, Europeana e DBpedia.....	89
APÊNDICE E - Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia.....	105
APÊNDICE F - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia.	109
APÊNDICE G - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata Após Testes no Wikidata Query Service.....	113
APÊNDICE H - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na DBpedia Após Testes no DBpedia SPARQL Explorer.....	116

1. Introdução

1.1. Enquadramento

O Arquivo Nacional da Torre do Tombo (ANTT) é uma instituição que mantém uma coleção de grande relevância para a memória nacional portuguesa. Nele estão guardados documentos originais que vão desde o século IX até os dias de hoje. O Arquivo também tem como responsabilidade preservar os novos arquivos eletrônicos, cumprindo a lei que estabelece as bases da política e do regime de proteção e valorização do património cultural. Para isso, a instituição possui um serviço de reprodução encarregado de realizar a cópia dos documentos do acervo, substituindo-os por novos suportes, evitando o constante manuseamento e consequente desgaste desses registos. O serviço de reprodução atua em três suportes: papel, microfilme e, mais recentemente, digital. Nele estão digitalizados documentos que podem ser pesquisados tanto por investigadores como pela população em geral. O serviço disponibiliza documentos de arquivo previamente digitalizados em diversos formatos, suportes em papel e pergaminho, livros entre outros (Arquivo Nacional Torre Do Tombo 2022).

Com a evolução dos objetos de arquivo, a demanda pela inclusão de dados digitais e por requisitos de interoperabilidade cresceu nos repositórios de património, havendo assim a necessidade de uma nova ferramenta de descrição de dados, mais flexível e processável por máquinas. Do mesmo modo, o público que faz uso dos seus serviços também mudou e procura novas experiências e possibilidades de interação, bem como intercâmbios de informação mais eficientes e produtivos. Além disso, os arquivos têm vindo a adaptar-se para abraçar uma perspectiva aberta em que os dados arquivísticos estão ligados a fontes de informação globais, seguindo uma tendência mundial de tornar a informação dos repositórios do património cultural mais acessível (INESC TEC 2022).

O acesso e a facilidade de uso dos recursos informacionais digitais estão sujeitos aos impactos de cenários tecnológicos específicos. A vulnerabilidade estrutural da informação digital consiste em um dos maiores desafios para os pesquisadores e profissionais de informação. A preservação da informação digital abrange inúmeras variáveis, gerenciamento responsável, tecnologia, além de um imenso orçamento, o que torna a tarefa especialmente difícil para os arquivos e outras instituições do património cultural. Um dos aspectos mais proeminentes é o da identificação de um conjunto de dados e informação, expostos na forma de metadados, que apoiem as ações de gestão da preservação digital (Sayão, Luís Fernando 2010). Seguindo NISO (2004), “Metadados é a informação estruturada que descreve, explica, localiza, ou possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar”. Logo, são elementos

utilizados para se descrever outros elementos, podendo ser aplicados a qualquer objeto, seja ele físico ou digital. Geralmente eles encontram-se visíveis e suas descrições são realizadas de acordo com convenções de organizações voltadas à sua padronização. Alguns exemplos simples de metadados em um registo de arquivo são: autor, título, data de criação ou modificação de um documento, tamanho (dimensões físicas ou espaço ocupado), tipo de suporte (analógico ou digital) (Ferreira 2020).

Dessa forma, surge o projeto EPISA do INESC TEC, que visa criar ferramentas de descrição apropriadas a bibliotecas, arquivos e museus. O EPISA tem como parceiros, além do INESC TEC, a DGLAB (Direção-Geral do Livro, dos Arquivos e das Bibliotecas), órgão responsável pela gestão dos Arquivos Nacionais, e a Universidade de Évora. O projeto está a definir novos modelos para os metadados do património cultural a fim de oferecer ferramentas de pesquisa que lidem com metadados em diferentes plataformas, integrando recursos administrativos e de investigação (INESC TEC 2022).

O seguinte projeto de pesquisa se enquadra em uma das tarefas do projeto EPISA. O trabalho consiste em buscar entidades de ligação entre documentos de arquivo e fontes de dados ligados com o propósito de expandir os registos tendo como foco o utilizador, melhorando assim o acesso aos arquivos e promovendo a interoperabilidade tanto com arquivos como com redes semânticas globais. O modelo conceitual usado no projeto é baseado na norma CIDOC CRM, uma ontologia formal destinada a facilitar a integração, mediação e intercâmbio de informações heterogêneas do património cultural (CIDOC CRM 2022).

Dessa forma, poderá abranger descrições não apenas de arquivos, mas de bibliotecas e museus. Um dos benefícios disto, por exemplo, é a possibilidade de obter informação sobre uma entidade específica encontrada num dos registos dos arquivos, como o político português Mouzinho da Silveira. Se houver interoperabilidade através da ligação de dados entre a entidade e outras bases de dados, o utilizador pode obter informação relevante sobre o político, como a sua data de nascimento e data de morte, mesmo que esta não esteja explicitamente disponível nos arquivos. Para a realização deste trabalho foram utilizadas as bases de dados da Wikidata, DBpedia e Europeana.

1.2. Definição do Problema

Ao longo deste trabalho buscarei responder ao seguinte problema de pesquisa: Quais entidades na ligação de documentos de arquivo seriam relevantes expandir de forma a enriquecer as respostas às necessidades de informação dos utilizadores em geral e dos profissionais de arquivo?

1.3. Motivação

A motivação para a realização deste projeto consiste na possibilidade de expandir e enriquecer os documentos de arquivo e registos de metadados com abordagens de questionamentos mais amplas e variadas, tendo em mente dois cenários: as mais frequentes interrogações feitas pelos utilizadores em geral, representado pela Figura 1, e o suporte à tarefa de descrição dos profissionais de arquivo, representado pela Figura 2.

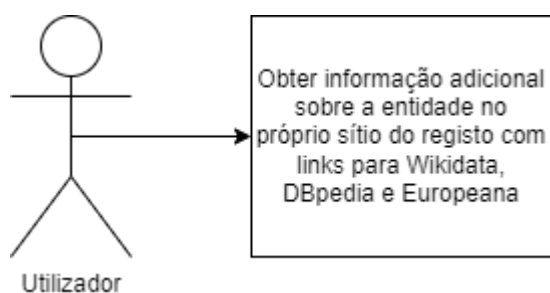


Figura 1 - Utilizador nos Serviços de Busca do Arquivo

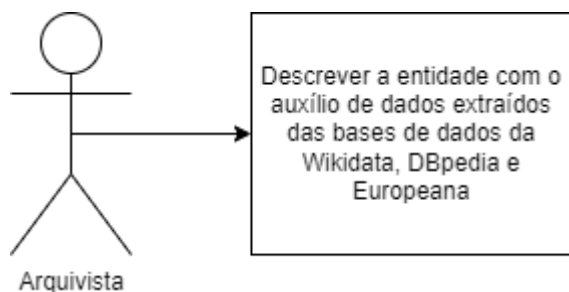


Figura 2 - Arquivista nos Serviços de Descrição do Arquivo

O trabalho realizado no âmbito da ciência da informação será o da análise da informação em documentos e em registos de metadados, buscando entidades com informação relevante em outras fontes. Espera-se também criar ligações entre fontes de informação diversas contribuindo para aprimorar a recuperação e a disseminação da informação junto do público em geral.

1.4. Objetivos e Resultados Esperados

O projeto tem como objetivo geral expandir os documentos de arquivo e registros de metadados tendo como base o conhecimento de interrogações típicas de utilizadores na plataforma de pesquisa.

Como objetivos específicos estão as pesquisas de documentos e dos 25 registros de arquivo do DigitArq utilizados no projeto, seguido de uma detalhada análise onde serão escolhidas as entidades de interesse para o trabalho. Posteriormente, serão estudadas e elaboradas questões com base nas perguntas mais frequentes dos utilizadores em plataformas de pesquisas. O próximo passo será definir as classes das entidades selecionadas de acordo com bases de dados específicas. Com base nessas conclusões serão escolhidas propriedades para a descrição dessas classes de entidades nessas futuras bases de dados. Por fim, o último objetivo será a realização de testes de buscas com os utilizadores e profissionais de arquivo.

Com isso, os resultados que se espera alcançar ao final deste trabalho são a habilidade de explorar de forma eficiente as ligações entre recursos diversos; a escolha apropriada de entidades de arquivo com informação relevante em outras fontes; a análise e seleção de propriedades capazes de enriquecer a informação de arquivo; a realização de testes que confirmem a relevância das propriedades escolhidas nas bases de dados; a avaliação dos benefícios deste trabalho para os sistemas de informação dos arquivos com arquivistas e utilizadores através de uma interface de busca; e a contribuição, por fim, para um futuro desenho de interface onde essa informação seja apresentada, vindo a auxiliar na otimização da recuperação da informação nos arquivos.

1.4.1. Árvore de Objetivos

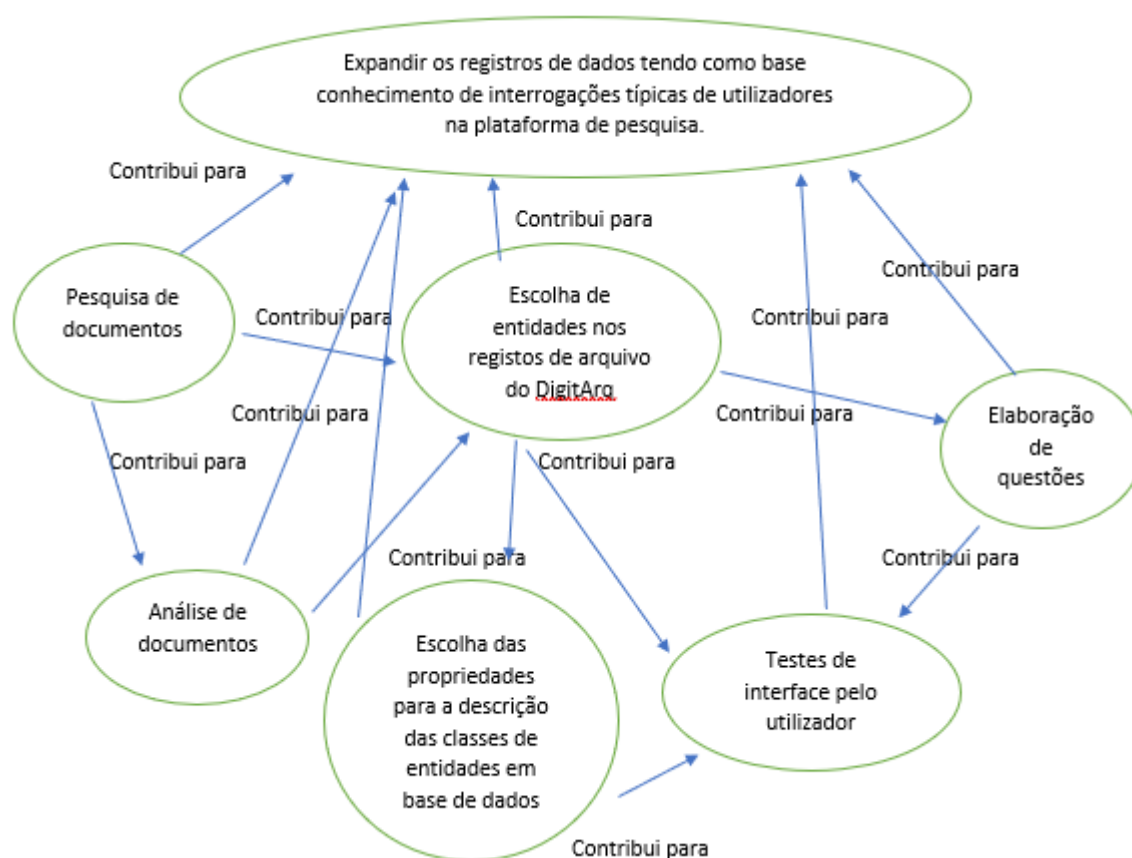


Figura 3 - Árvore de Objetivos

A Figura 3 retrata a árvore de objetivos do projeto, que apresenta o objetivo geral de expandir os documentos de arquivo e registos de metadados tendo como base o conhecimento de interrogações típicas de utilizadores na plataforma de pesquisa. Em seguida, mostra os objetivos específicos que contribuirão para o cumprimento do objetivo geral e as suas devidas relações com os outros objetivos do projeto.

1.5. Metodologia

Com relação à metodologia da pesquisa, do ponto de vista da sua natureza, ela é aplicada, pois busca produzir conhecimentos com um propósito prático e direcionado à solução de problemas específicos. A abordagem da pesquisa é qualitativa, tendo como intuito a descrição do processo. Da parte dos seus objetivos é uma pesquisa exploratória, devido ao fato de apresentar como a “principal finalidade desenvolver, esclarecer e modificar conceitos e ideias” (Gil, 2002, p. 27). Com relação aos procedimentos técnicos, na primeira fase do projeto serão utilizados artigos

científicos, documentos e livros sobre o tema com o propósito de dar respaldo ao projeto.

A pesquisa para essa primeira fase, que consiste na revisão de literatura, foi conduzida em duas bases de dados: Web of Science e no Library and Information Science Source. Além disso, foram realizadas pesquisas em sites da internet de projetos relacionados ou similares ao projeto EPISA. Abaixo, segue o método de pesquisa realizado nas duas bases de dados:

- Termo de Pesquisa Utilizado: linked data in cultural heritage em ambos os bancos de dados; sem uso de operadores booleanos.
- Pesquisa na Web of Science: não foram utilizados filtros. Foram recuperados 8 registros, sendo que apenas 5 de acesso aberto. Os 5 foram escolhidos para o trabalho.
- Pesquisa da Library and Information Science Source: Os seguintes filtros foram utilizados: texto integral e revista científica analisada por pares. Foram recuperados 23 registros, contudo, após a leitura dos resumos e conclusões, 20 não se mostraram relacionados ou interessantes para o trabalho. Logo, apenas 3 foram escolhidos para o trabalho.

Após a leitura dos artigos científicos, documentos e livros escolhidos na etapa acima, foram adicionados mais artigos e outras referências citadas nos respectivos textos consideradas pertinentes para a realização do trabalho. Para finalizar, foram adicionadas as referências do próprio projeto EPISA, assim como, trabalhos e sites de projetos referenciados por ele.

As tarefas do trabalho de dissertação serão as seguintes:

1. Análise dos 25 registros da Torre do Tombo, os quais estão sendo utilizados no projeto EPISA para a construção de ontologias;
2. Escolha de entidades de interesse dos registros;
3. Reuniões para o debate de quais dessas entidades têm relevância para a expansão;
4. Estudo das melhores bases de dados para as buscas de informação sobre as entidades escolhidas;
5. Escolha das propriedades mais relevantes para a descrição das classes de entidades em cada base de dados;
6. Testes das entidades e propriedades selecionadas nas suas respectivas bases de dados;
7. Testes de busca utilizando os 25 registros do projeto, com profissionais de arquivo, a fim de saber se as expansões das entidades sugeridas foram de fato relevantes para as suas tarefas no arquivo; e em seguida, testes de recuperação com os utilizadores.

1.6. Estrutura da Dissertação

A presente dissertação encontra-se dividida em seis capítulos. O primeiro capítulo é o da Introdução, onde é apresentado o enquadramento do trabalho, a definição do problema de pesquisa e sua motivação, o objetivo geral, os objetivos específicos do trabalho e os resultados esperados, a árvore de objetivos, a metodologia e, por fim, a estrutura da dissertação.

O segundo capítulo é o da Revisão de Literatura, onde são discutidos os principais conceitos e projetos na área da web semântica e de dados abertos ligados. Os temas tratados são: web semântica, estrutura de descrição de recursos – RDF e SPARQL, ontologia, bases de dados em RDF, dados ligados no patrimônio cultural, o modelo CIDOC CRM, o modelo EDM da Europeana, o modelo Schema.org, o modelo Dublin Core e, por último, trabalhos relacionados já realizados na área.

O terceiro capítulo é sobre Entidades em Registos de Arquivo e em Bases de Dados Ligados. Nele é abordada a plataforma DigitArq, o projeto EPISA, a expansão de entidades nos registos de arquivo do DigitArq e as fontes exteriores de enriquecimento e ligação de dados.

O quarto capítulo é o dos Testes de Entidades e de Propriedades dos Registos. Nele são explicados os testes feitos com o Wikidata Query Service, com o DBpedia SPARQL Explorer e com o Europeana SPARQL API.

O quinto capítulo é o das conclusões e perspetivas de desenvolvimento, e discutem-se as limitações e problemas encontrados ao longo do trabalho, fazendo considerações finais e notas para trabalhos futuros.

Por fim, o sexto capítulo das referências, seguido dos apêndices.

2. Revisão de Literatura

2.1. Web Semântica

Com o avanço da internet, espera-se que os computadores venham a cumprir tarefas cada vez mais complexas, principalmente com relação às buscas de informação. Um dos propósitos é fazer com que os computadores ligados a rede entendam e associem dados fazendo uso da semântica. Através da semântica é possível dar significado a palavras, frases e textos de uma língua, tornando as informações compreensíveis para o computador através de ontologias e regras de inferências. Além disso, colabora para que os utilizadores encontrem respostas mais precisas para as consultas, excluindo o que é irrelevante (Lima e Carvalho 2004). Segundo o diretor do World Wide Web Consortium (W3C) e fundador da World Wide Web Foundation, Tim Berners-Lee (2001) “A web semântica não é uma web separada, mas uma extensão da atual. Nela, a informação é dada com um significado bem definido, permitindo melhor interação entre os computadores e as pessoas”. A web semântica significou uma enorme mudança na forma de se processar, adquirir e organizar a informação na rede. Para isso, os recursos fornecidos devem ser significativos o suficiente para que máquinas sejam capazes de realizar o processamento e extrair o significado pretendido dos dados de uma forma automática e assertiva (Lima e Carvalho 2004).

Berners-Lee desenvolveu um modelo de arquitetura em camadas, onde cada uma dessas camadas teria uma determinada função e interagiriam entre si. Essas camadas podem ser observadas na Figura 4.

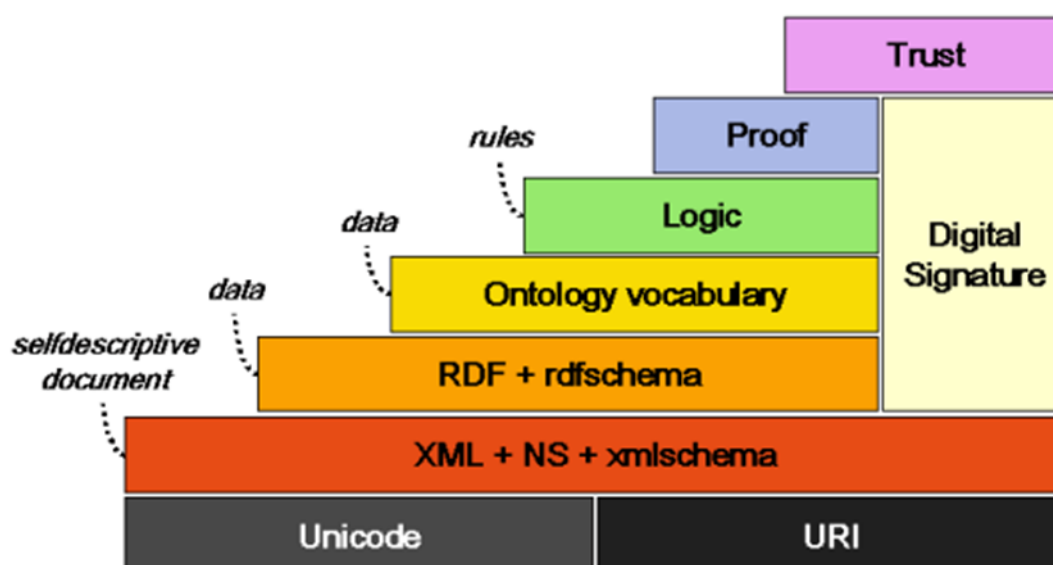


Figura 4 - As Camadas da Web Semântica (Koivunen e Miller 2001).

Conforme explica Minichiello (2017):

- Camada URL e Unicode – onde serão definidos os símbolos para a representação dos dados e suas referências;
- Camada XML, XMLSchema e NS – onde serão determinadas as normas para o uso da linguagem e suas estruturas;
- Camada RDF e RDFSSchema – onde serão descritos os dados segundo um modelo lógico e um vocabulário apropriado;
- Camada da Ontologia – onde serão estabelecidas as relações entre as palavras e os seus conceitos;
- Camada Lógica: onde serão definidas as regras que lidam com a informação;
- Camada de Prova e Confiança: onde serão feitas as provas, checagens e autenticação dos dados.

2.1.1. Estrutura de Descrição de Recursos – RDF

O modelo RDF e a linguagem XML são fundamentais no desenvolvimento da web semântica. Com o uso do XML, é possível a utilizadores criar tags próprias que permitem que scripts ou programas as empreguem, incorporando uma determinada estrutura em documentos sem a necessidade de mencionar os seus significados. Isso permite a criação de uma base tecnológica para a formação de uma web global (Lima e Carvalho 2004).

O W3C recomenda a RDF como uma tecnologia fundamental para a elaboração de projetos semânticos, sendo utilizada para a descrição dos recursos e para o acesso de dados em esfera global. Ele expressa o significado das etiquetas, as quais possibilitam a interoperabilidade dos dados, ajudando na recuperação de informação de recursos na internet (Berners-Lee, Lassila e Hendler 2001). O modelo RDF possibilita uma estrutura de triplas com sujeito, predicado e objeto, com o fim de retratar a relação entre recursos e de fazer a codificação do sentido das etiquetas. Para a construção da web semântica, em geral, se faz uso dos URIs (Universal Resource Identifier) para identificar os recursos. Os recursos no sujeito e predicado podem ser qualquer coisa acessível na web identificada por URIs (Silva, Santarem Segundo e Freire 2021). Essa estrutura de tripla da codificação RDF pode ser vista na Figura 5, com um exemplo em que o objeto é um literal.

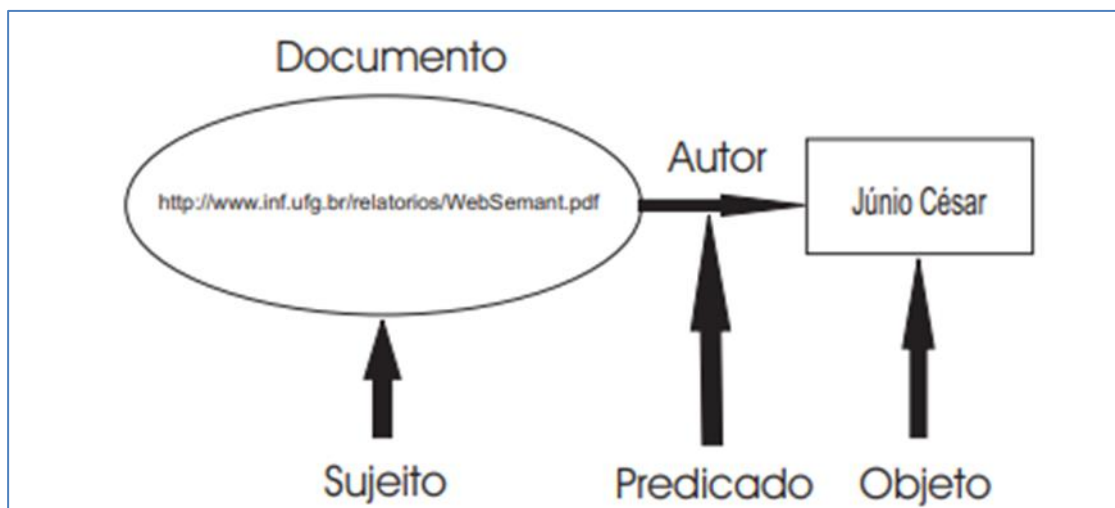


Figura 5 - Representação de um documento RDF (Lima e Carvalho 2004).

Com o uso das URIs há a possibilidade de se identificar recursos de uma forma única, assim como as propriedades que os relacionam. O conhecimento representado em triplas pode ser mantido em bancos de dados. Cada tripla é uma afirmação a respeito de um recurso e usa uma propriedade para lhe associar um valor ou outro recurso, sendo uma maneira de representar dados a serem processados por máquinas (Lima e Carvalho 2004).

O modelo RDF pode ser exibido em diversos formatos, entre eles estão: o RDF in Attributes - RDFa, RDF eXtensible Markup Language - RDF/XML, Turtle, N-Triples, JSON-LD, TRING e N-Quads. Além disso, esses formatos permitem uma leitura em grafo que, ao se interligar, possibilita a criação de conjuntos de dados ligados (Silva, Santarem Segundo e Freire 2021).

2.1.2. Simple Protocol and RDF Query Language – SPARQL

Da mesma forma que o SQL é utilizado para a realização de consultas em bases de dados comuns, o SPARQL é utilizado para a consulta em base de dados RDF. No caso do SPARQL, as buscas são feitas pelas triplas RDF, tendo vocábulos e sintaxe próprios para a manipulação e tratamento da informação obtida das triplas. O SPARQL possibilita ao utilizador fazer consultas complexas através do uso de filtros, recolhendo apenas as informações relevantes ou específicas da base (Minichiello 2017). Segundo Schiessl (2015), o SPARQL viabiliza a extração de dados (semi) estruturados, a exploração de dados através de consultas por relações desconhecidas, a operação de múltiplas combinações de bases de dados com consultas simples, a transformação de dados RDF de um vocabulário para outro e a construção de novos grafos RDF a partir de buscas em outros vocabulários. Esses vocabulários são listas de termos que podem ser usados nos processos de indexação, assim como de representação temática, com o intuito de padronizar e ajudar na recuperação da informação em sistemas de documentação

(Webmuseu 2022). Embora alguns comandos de busca do SPARQL sejam similares aos do SQL, a sua sintaxe possui comandos exclusivos, assim como uma estrutura específica para o tratamento da informação. Alguns desses comandos são:

PREFIX – Usado como forma de oferecer uma abreviação das URIs para tornar os códigos mais simples;

SELECT – Usado para adquirir informações na base RDF e mostrá-las em modo de tabela;

CONSTRUCT – Usado para tornar o resultado da pesquisa em uma tripla RDF.

DISTINCT – Usado para retirar linhas duplicadas.

WHERE – Usado para restringir as buscas, mostrando apenas os resultados relevantes para o utilizador (Minichiello 2017).

É possível ao utilizador fazer uso do SPARQL de forma a obter resultados, muito bem estruturados e detalhados, apenas fazendo uso dos conceitos básicos da linguagem.

2.1.3. Ontologia

Além da camada RDF, outra camada essencial da web semântica é a da Ontologia. Nela são estabelecidas as relações entre as palavras e os seus conceitos. As ontologias são formadas por uma taxonomia, que estabelece as classes de objetos e os seus respectivos relacionamentos, podendo ser vistas como um modelo de conhecimento. Para a identificação do sentido de estrutura XML nas páginas web passou a ser empregada uma ontologia (Lima e Carvalho 2004). Os programas de computação conseguem, por regras de inferência, utilizar os termos de maneira mais eficiente do que seres humanos (Patrício 2012). Contudo, com o aumento de páginas na web utilizando ontologias, problemas em torno de terminologias começaram a surgir.

Algumas relações de equivalência de termos precisam ser estabelecidas previamente a fim de permitir o alinhamento de conceitos com o mesmo significado. As ontologias têm a capacidade de unificar o conhecimento e gerar consenso sobre determinados conceitos, sendo um elemento estruturador de informação. Um exemplo disso são as aplicações que têm como objetivo aumentar a precisão das buscas na web (Lima e Carvalho 2004).

Para a aplicação da ontologia alguns preceitos devem ser levados em consideração, como: clareza, coerência, extensibilidade, compromissos ontológicos e de codificação mínimos. A ontologia consegue oferecer algumas vantagens importantes para os sistemas apoiados em conhecimento: a partilha de informações e identificação de conceitos; o reuso de conceitos correntes para a formação de novos conceitos ou adequação de outros já existentes,

propiciando um melhor entendimento pelas máquinas, o que permite a interoperabilidade do sistema (Minichiello 2017). De acordo com Moraes e Ambrósio (2007) podemos classificar a ontologia em cinco categorias:

Ontologias Genéricas: representam conceitos amplos, como tempo, espaço, ações, eventos;

Ontologias de Domínio: representam concepções e nomenclaturas de domínios específicos, como nutrição ou engenharia;

Ontologias de Tarefas: representam atividades genéricas, não importando o domínio, tais como procedimentos de vendas ou diagnóstico;

Ontologias de Aplicação: representam conceitos fundamentados tanto em um domínio particular como em atividades específicas;

Ontologias de Representação: esclarecem as ideias que fundamentam os princípios de representação de conhecimento.

Dessa forma, a ontologia desempenha um papel essencial para a criação de sistemas fundamentados em informação. Além de contribuir para a construção do sistema com a reutilização das informações, uma ontologia favorece a interoperabilidade entre os sistemas e o aperfeiçoamento nas buscas semânticas (Minichiello 2017). No futuro, as ontologias serão empregadas para inferir novos conhecimentos, tornando a web ainda mais útil (Lima e Carvalho 2004).

2.2. Bases de Dados em RDF

2.2.1. DBpedia

A DBpedia é um trabalho coletivo de crowdsourcing a fim de retirar conteúdo estruturado das informações geradas em diversos projetos da Wikimedia. Essas informações organizadas são parecidas com um grafo de conhecimento aberto (OKG) que se encontra acessível na web para qualquer pessoa. Esse grafo é uma espécie de banco de dados que guarda conhecimento de uma maneira que é legível por máquina e disponibiliza uma maneira das informações serem recolhidas, estruturadas, distribuídas, pesquisadas e usadas. Esse banco de dados é atualizado periodicamente, de 6 a 18 meses, a partir do repositório da Wikipedia e o procedimento de extração de informação é demorado (Pillai, Soon e Haw 2019).

A DBpedia em Outubro de 2016 já possuía 13 mil milhões de RDF triplas. Os dados da DBpedia são aproveitados como dados ligados (“linked data”), o que vem transformando a maneira como os aplicativos se relacionam com a web; sendo assim possível utilizar navegadores web padrão, rastreadores automatizados e realizar pesquisas com linguagens de consulta parecidas

a SQL, como SPARQL (DBpedia Association 2022). A DBpedia atua como o centro dos dados interligados e possui ligações para a maioria das bases de conhecimento: além da citada Wikipédia, para a Wikidata, Yago, Freebase e CYc, entre outras. Visto que o conteúdo da Wikipédia está a ser atualizado a cada segundo, também existe uma versão dinâmica da DBpedia, a DBpedia Live. Essa versão possibilita a sincronia da DBpedia com a Wikipedia, e possui informação atualizada (Pillai, Soon e Haw 2019).

2.2.2. Wikidata

A Wikidata é um grafo de conhecimento colaborativo, construído através de crowdsourcing, livre e aberto que pode ser formatado e lido por máquinas e seres humanos. A Wikidata é hospedada pela Wikimedia Foundation, e opera como um centro de armazenamento para dados estruturados de projetos diversos, entre eles: Wikipedia, Wikivoyage, Wikcionário, Wikisource. Além disso, ela fornece apoio a diversos outros sites e serviços. O seu conteúdo pode ser acessado gratuitamente, exportado fazendo uso de formatos padrão e ser ligado a demais conjuntos de dados abertos na web (Wikidata.Org 2022).

A grande base de conhecimento colaborativa Freebase, que contava com uma coleção online de dados estruturados colhidos de diversas fontes, incluindo contribuições individuais da wiki submetidas pelos utilizadores, foi incorporada à Wikidata depois do seu término em Junho de 2015. A Wikidata importa factos de diversas fontes e inclui as suas referências. A inclusão de referências é importante para a verificação da validade das informações. Os seus rótulos, pseudónimos e descrições de entidades são fornecidos em mais de 350 línguas. Uma página na Wikidata é criada para cada artigo na Wikipedia. Essas páginas são denominadas “itens”. Esses itens são identificados de forma única por um nome que começa com a letra “Q”, seguidas por um número, por exemplo, Q1532667 é utilizado para Mouzinho da Silveira. As propriedades começam com a letra “P”, seguidas por um número, por exemplo, a propriedade P570 é “data da morte”. A propriedade e o seu valor são conhecidos em conjunto como declarações (Pillai, Soon e Haw 2019).

2.3. Dados Ligados no Patrimônio Cultural

Quando se faz referência a dados ligados, se diz respeito ao grupo de práticas para a divulgação de dados na web, que consiste em acomodar os dados de uma maneira ordenada, estruturada, com a conexão dos dados entre si. Já é possível a propagação do conhecimento com base em informação de várias esferas da sociedade com o uso de programas de dados ligados que possibilitam as relações entre as informações para o recetor (Minichiello 2017). De acordo com Bizer, Heath e Berners-Lee (2009) dados ligados são dados divulgados na web de uma maneira

que os tornam possíveis de serem lidos por máquinas. Eles têm os seus conceitos claramente definidos, estando conectados a outros conjuntos de dados (“datasets”) externos que igualmente podem ser acedidos através de um conjunto de dados externos.

Os dados ligados vêm se tornando uma enorme nuvem de informações, tendo um grande aumento nos últimos anos. Esses conjuntos de dados ligados dispõem de segmentos distintos tais como: redes sociais, linguística, geografia. Na Figura 6, podemos ver a nuvem de dados abertos ligados onde se pode notar as conexões de cada conjunto de dados, assim como as divisões dos segmentos. É possível ver os conjuntos de dados apresentados por cores diferentes de acordo com o seu segmento, e as ligações entre os conjuntos de dados. No meio da nuvem se encontra a DBpedia, que está interligada no nível RDF com diversos conjuntos de dados abertos na web (Minichiello 2017).

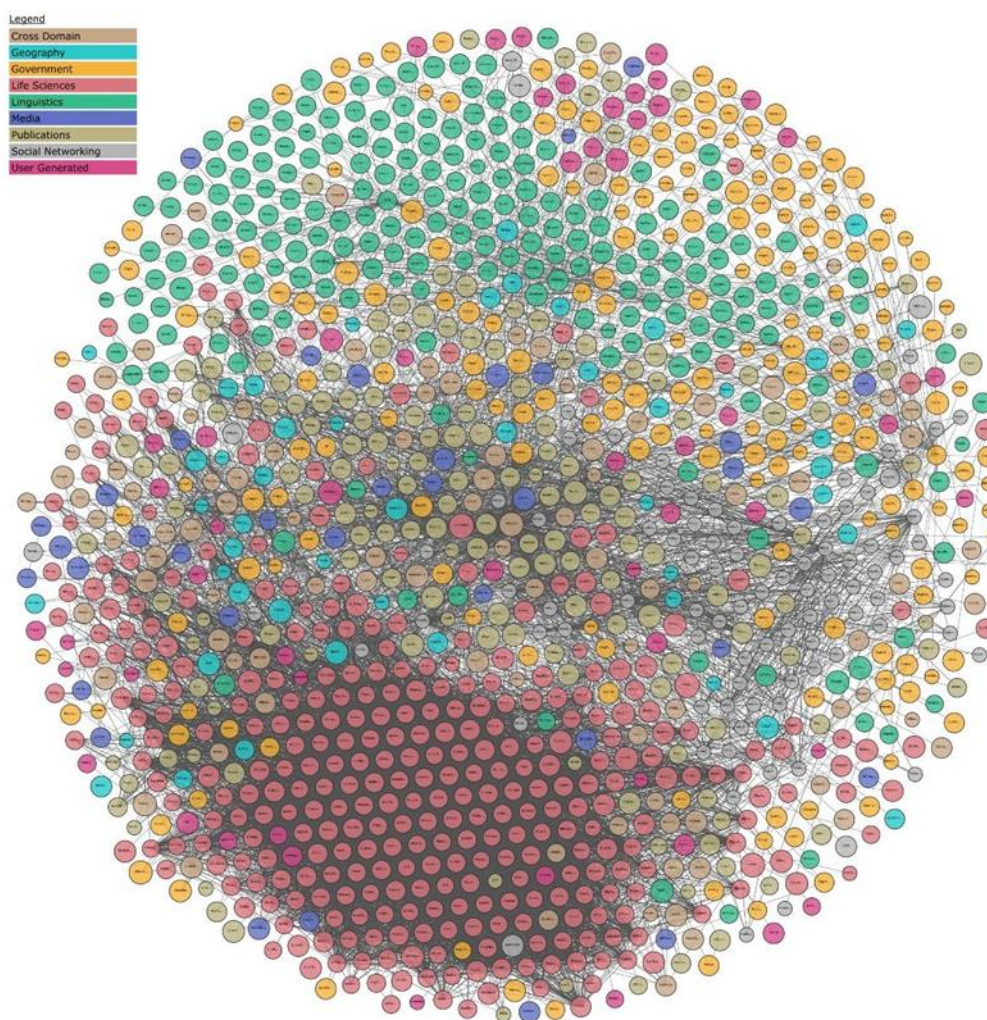


Figura 6 - Nuvem de dados abertos ligados (Lod-Cloud.Net 2022).

Segundo Bizer, Heath e Berners-Lee (2009), os dados ligados apresentam regras para a sua

criação e divulgação, como: utilização de URIs como nomes para entidades; emprego de URIs fazendo uso de http, de maneira que seja possível achar informações por busca de nomes na web; utilização de padrões como o RDF para a associação de informações relevantes às URIs; inserção de ligações com outras URIs, de forma a permitir a descoberta de novas entidades e implementar de facto a ligação de dados.

O uso dos dados ligados é somente uma etapa da implantação e transformação da web atual para a web semântica, buscando tornar mais simples o acesso das máquinas à informação. Dessa forma, se favorece a recolha dos dados e o seu processamento pelas máquinas. Como resultado dessa associação das pessoas com as máquinas, a web se tornará mais potente, com a capacidade de desenvolver aplicações que possam a vir a melhorar a vida das pessoas, como a elaboração de programas para a fiscalização de empreendimentos governamentais. Esse controle pode ser feito através do acesso às bases de dados divulgadas pelo governo. Essas aplicações também poderiam ser úteis para as empresas, que teriam os dados ligados com várias bases governamentais ou institucionais, com o intuito de as ajudar nas suas atividades (Minichiello 2017).

A publicação dos dados ligados não tem, necessariamente, que ser em formato aberto. Uma empresa pode querer fazer a ligação de dados, mas não os tornar em formato aberto, por exemplo, por questões de sigilo empresarial. Por outro lado, organizações como a W3C (World Wide Web Consortium) vêm estimulando a criação de dados abertos ligados como forma de gerar ligações entre recursos de origens diversas e incentivar a reutilização e o enriquecimento de dados na web. (Isotani e Bittencourt 2015). O diretor da W3C, Tim Berners-Lee, apresentou um plano para a publicação de dados abertos ligados que é baseado no seguinte sistema de 5 estrelas:

- 1ª estrela: quando os dados são colocados na internet através de uma licença aberta;
- 2ª estrela: quando os dados são cedidos de uma maneira estruturada, como em uma grelha de Excel;
- 3ª estrela: quando se disponibiliza os dados fazendo uso de formatos não proprietários, exemplo: CSV;
- 4ª estrela: quando se faz uso de URI para a identificação de recursos de maneira a simplificar o seu direcionamento;
- 5ª estrela: quando se liga os dados de outras pessoas com os seus próprios dados de forma a fornecer contexto (Minichiello 2017).

Conforme se aumenta o número de estrelas, maiores vantagens vão sendo oferecidas ao

publicador e ao consumidor. Ademais, a cada acréscimo de estrelas os integrantes se comprometem mais com o objeto de conhecimento. Os dados ligados proporcionam um enorme conjunto de dados dos quais instituições, empresas e pessoas podem obter dados relevantes para os seus propósitos, além de ajudar a expandir esse conjunto de dados com a sua divulgação. Dessa forma, a cada dia cresce o volume de dados, e as tecnologias estão sendo aperfeiçoadas de maneira a possibilitar que as informações sejam mais bem utilizadas (Minichiello 2017). Os dados abertos ligados têm a capacidade de serem usados, reutilizados e compartilhados por qualquer indivíduo, sob a condição apenas da citação das fontes e, se a licença original obrigar, o compartilhamento das mesmas licenças que foram utilizadas para os dados apresentados (Silva, Santarem Segundo e Freire 2021).

2.3.1. Modelo CIDOC CRM

Quando se trata de dados abertos ligados no patrimônio cultural, percebe-se que nos últimos anos, várias organizações culturais transformaram o seu patrimônio cultural em informações que, anteriormente, eram utilizadas somente para propósitos internos, a dados associados com conjuntos de dados abertos ligados. Eles utilizam a infraestrutura da “web” para a publicação e associação de conjuntos de dados que podem ser analisados e executados por aplicativos que os acessam e são constantemente atualizados e associados a novos ativos (Aydin et al. 2013).

Assim, os dados estruturados conseguem favorecer tanto a organização quanto a comunidade, dando a vantagem de produzir novas interfaces para os utilizadores experimentarem, além de ampliar a web semântica e assegurar a credibilidade de uma organização como provedora de dados de excelente qualidade (Marden et al. 2013). Como o patrimônio cultural abrange instituições bastante variadas, os modelos de dados empregados também são bastante diversificados. Os museus, bibliotecas e arquivos têm uma natureza transnacional, o que leva ao uso de múltiplas práticas de descrição de recursos entre eles, sendo algumas estabelecidas dentro de um contexto regional. Uma das maiores dificuldades para a gestão de dados de patrimônio cultural é a falta de uniformidade na aplicação dos modelos, linguagens e padrões de dados (Freire et al. 2019).

Em busca de uma normalização na área do patrimônio cultural, uma série de ontologias para a representação semântica vem sendo recomendadas, uma delas é o CIDOC CRM (Modelo de Referência Conceitual do CIDOC). O modelo é uma ferramenta para a inclusão de informação relacionada à área do patrimônio, especialmente desenvolvido como um modelo entidade-relacionamento para informações de museus. Ele permite a pesquisa de questões múltiplas sobre o passado em conjuntos de dados variados e separados. Isso é possível pelo estabelecimento de uma estrutura formal e de definições para descrever os conceitos

(implícitos e explícitos). As relações utilizadas nos documentos também facilitam as buscas desses dados. Uma outra forma de denominar esses modelos é como ontologias formais, sendo possível às descrições formais integrar dados de fontes diversas (CIDOC CRM 2022).

Um dos objetivos do CIDOC CRM é oferecer um entendimento das informações relacionadas ao patrimônio cultural, através de uma estrutura semântica compartilhada e extensível para a inclusão de informações fundamentadas em evidências. Assim, busca fornecer uma linguagem que sirva tanto para implementadores de domínio desenvolverem requisitos para sistemas de informação quanto para ser utilizado como um manual de boas práticas para a criação de modelos conceituais. O modelo pode ser a ligação entre fontes de informação variadas do patrimônio cultural, como as divulgadas por museus, bibliotecas e arquivos. O CIDOC CRM foi desenvolvido pelo CIDOC Documentation Standards Working Group e, atualmente, pelo CIDOC CRM SIG, sendo fruto de mais de duas décadas de trabalho de elaboração e manutenção. Ele é um padrão ISO oficial desde 2006, sendo encontrado na ISO 21127:2014 (CIDOC CRM 2022).

O modelo permite tanto a recuperação de informação de alto nível quanto a elaboração e documentação de “data point” e questionamentos mais específicos, sendo um padrão CRM que oferece as classes básicas e os relacionamentos criados para o patrimônio cultural. A ontologia base é completada por extensões modulares que são planejadas para dar apoio a vários tipos de perguntas de pesquisa e documentação. As extensões são desenvolvidas de maneira articulada com a ontologia base, de forma a garantir a compatibilidade dos dados com o sistema de base de conceitos e relações. Como consequência, se conquista um nível de integridade e de integração da informação não encontrada em outros sistemas de informação. A ontologia possui classes que reproduzem muitos eventos, atributos e conceitos. A versão 7.1.1., liberada em 2021, possui 81 classes e 160 propriedades (CIDOC CRM 2022).

2.3.2. Modelo de Dados da Europeia – EDM

Um outro modelo que merece destaque quando se trata de representação semântica é o da Europeia, que divulga uma quantidade grande de dados culturais através do padrão EDM, sendo um dos grandes destaques da área das Humanidades Digitais (Coneglian e Santarem Segundo 2017).

A Europeia tem como finalidade tornar acessível na internet o patrimônio cultural europeu, permitindo a preservação e a consulta dos materiais que são digitalizados. O seu começo foi no ano de 2000, quando o Projeto Gateway and Bridge to Europe’s National Libraries (GABRIEL) criou um portal com o material de quarenta e três bibliotecas nacionais europeias. Assim, a Comissão Europeia patrocinou programas de pesquisa e estabeleceu uma maior colaboração

entre estados-membros que demonstraram desejo de tornar o seu patrimônio cultural acessível. Com esse patrocínio, o projeto Gabriel expandiu para a The European Library (TEL), sendo uma associação de dados abertos para coleções de bibliotecas (Jefcoate 2006).

Em 2005, Jacques Chirac, na época presidente da França, juntamente com seis primeiros-ministros europeus, mandou uma carta oficial ao presidente da Comissão Europeia, José Manuel Durão Barroso, expondo o seu interesse na preservação do patrimônio cultural e propondo a formação de uma biblioteca virtual para guardar o registo cultural e científico europeu. A proposta foi aceita e se seguiu uma estratégia para a criação de bibliotecas digitais, cumprindo práticas de digitalização de coleções correlacionadas, acesso online, preservação e armazenamento. O projeto da Biblioteca Digital Europeia foi apresentado pela Comissão Europeia no mesmo ano, e produzido com o suporte das bibliotecas nacionais, algumas entidades culturais dos Estados-Membros e do Parlamento Europeu (Jefcoate 2006).

Nos dias de hoje, as chamadas Coleções Europeanas (ou apenas Europeana) oferecem uma rica e variada coleção com o patrimônio cultural e científico europeu, formada por documentos digitalizados, como: revistas, livros, pinturas, manuscritos, mapas, jornais, músicas, programas de tv, vídeos e documentários. Devido ao êxito do projeto, a TEL foi suspensa em 2016, e os seus dados passaram a ser acessados pela plataforma da Europeana. A direção das Coleções Europeanas é de responsabilidade da Fundação para a Biblioteca Digital Europeia, que engloba as principais instituições de bibliotecas, museus, arquivos e associações culturais europeias. A coleção da Europeana é formada por um grande e crescente número de objetos digitalizados, que são disponibilizados por fornecedores de dados dos setores de entidades culturais de mais de 36 países. Os fornecedores utilizam distintos modelos de metadados para representar os objetos. Uma das tarefas da Europeana é padronizar os dados para que possam ser usados novamente sem grandes custos e impedir a duplicação de trabalho. Com a finalidade de mostrar os dados de uma maneira relevante em um ambiente de múltiplas culturas e idiomas, com o uso de recursos semânticos do patrimônio cultural, foi criado o Europeana Data Model (EDM) (Silva, Santarem Segundo e Freire 2021).

O EDM foi elaborado por especialistas técnicos das áreas de patrimônio cultural envolvidos com a Europeana, e consiste em um modelo para organizar os dados gerenciados e divulgados pela biblioteca digital. Ele foi feito através de padrões e técnicas de domínio público, tendo como fundamento a web semântica. O EDM tem a habilidade de organizar e desenvolver dados de diferentes domínios, tais como: Lightweight Information Describing Objects - LIDO (museus); Encoded Archival Description – EAD (arquivos); Metadata Encoding and Transmission Standard - METS (bibliotecas digitais). O propósito é estruturar os metadados de domínios distintos e oferecer em um mesmo espaço digital a fim de atingir um número

maior de pessoas (EDM PRIMER 2013).

O primeiro modelo de dados da Europeia foi o European Semantic Elements - ESE, baseado no padrão Dublin Core, suportando um grupo mínimo de dados que podem ser empregados em materiais diversos. O modelo gera um registo plano, o que significa que em um registo de metadados estão reunidos campos relacionados a várias entidades, sendo difícil saber se determinado valor se refere ao objeto primário ou à sua representação digital. Alguns fornecedores ainda enviam alguns dados utilizando o modelo ESE, que são posteriormente migrados para o modelo EDM. A representação feita diretamente para o modelo EDM apresenta uma maior flexibilidade e um grupo maior de elementos do que os oferecidos pelo ESE. O EDM possui metadados mais ricos e com uma granularidade mais fina, o que possibilita um grau mínimo de interoperabilidade entre os dados enviados por múltiplos colaboradores (Doerr et al. 2010).

Dessa forma, é possível afirmar que o EDM garante uma maior qualidade aos processos através dos quais a Europeia adquire, produz e divulga os seus dados. A maneira de organizar os dados e os formatos através do EDM permite a sua divulgação como dados abertos ligados. Para isso, a Europeia faz a estruturação dos múltiplos padrões de metadados retirados dos provedores de dados das diversas áreas do patrimônio cultural, sendo posteriormente esses metadados convertidos para o padrão de dados EDM através de folhas de estilos que alteram para EDM com URIs apresentados no padrão RDF/XML. Após essa conversão dos metadados, começa a fase de enriquecimento dos dados através de conexões com fontes exteriores tais quais o DBPedia, VIAF, GeoNames e demais vocabulários (Silva, Santarem Segundo e Freire 2021).

Os componentes do EDM são formados por vocabulários exclusivos, que são um meio de partilhar informação e de desenvolver e documentar definições de dados padrão de forma a reduzir a ambiguidade e melhorar a comunicação. Contudo, alguns vocabulários e modelos de dados legitimados e recomendados pelo W3C são igualmente usados de forma a descrever os dados e suportar as funções da Europeia. Alguns dos mais utilizados pelo EDM são: o RDF; RDF(S); Dublin Core; W3C Data Catalog Vocabulary - DCAT OWL; Open Archives Initiative Object Reuse and Exchange - OAI-ORE; SKOS; e a Creative Commons Zero - CCo (EDM Primer 2013). Como forma de apoiar a incorporação desses modelos distintos na utilização de dados no campo do patrimônio cultural, a identificação de entidades e relacionamentos pertinentes aos metadados, o modelo EDM faz uso do vocabulário OAI-ORE, que estabelece padrões descritivos e de intercâmbio de conjuntos de recursos da web (Patrício 2012).

2.3.3. Modelo Schema.org

O modelo EDM mudou o foco da modelagem de dados de escolhas específicas de formato de metadados para princípios gerais que são compatíveis com vários modelos de dados. Isto permite à Europeia utilizar outros modelos de dados, tais como o Schema.org, que é um novo modelo de dados para o intercâmbio de dados estruturados na web apoiado pelos principais motores de busca da Internet que o desenvolveram (Freire et al. 2020). Atualmente, mais de 10 milhões de sites utilizam o Schema.org, e inúmeros aplicativos fazem uso do seu vocabulário com o intuito de enriquecer e expandir as suas experiências na web (Schema.Org 2022).

O Schema.org faz parte de um movimento colaborativo que tem a tarefa de criar, manter e promover esquemas para dados estruturados na Internet. O seu vocabulário pode ser utilizado com diversas codificações, e permite que entidades, os seus respectivos relacionamentos e ações possam ser expostos através de um modelo bem fundamentado (Schema.Org 2022). Para as bibliotecas digitais, o Schema.org permite a descrição de livros, mapas, arte visual, gravações musicais e muitos outros tipos de recursos culturais (Freire et al. 2020).

Para representar o objeto sobre o qual as instituições de património cultural oferecem uma descrição há classes como o schema:CreativeWorkFootnote27 e várias das suas subclasses tais como schema:VisualArtwork, schema:Book, schema:Painting e schema:Sculpture. Já as representações digitais dos objetos podem ser feitas com: schema:MediaObject, schema:ImageObject, schema:VideoObject, schema:AudioObject. E as classes contextuais podem ser representadas com as classes: schema:Pessoa, schema:Place e schema:Organização. Uma outra classe bastante importante é a Bibliographic ExtensionFootnote28 para a descrição de recursos bibliográficos (Freire et al. 2020).

Como o Schema.org foi delineado para a indexação de páginas web, em sua maior parte encontra-se em formato HTML. Uma forma de agregar dados do Schema.org pode ser iniciado com o “crawling” de páginas web comuns e depender de um processo baseado em seguir hiperligações dentro do HTML. As aplicações somente interessadas nos dados do Schema.org, e não no conteúdo textual, tem a possibilidade de rastrear as páginas web de forma similar aos motores de busca, descartar o conteúdo textual, extrair hiperligações do HTML e prosseguir o rastreamento seguindo estas ligações (Freire et al. 2020).

2.3.4. Modelo Dublin Core

Um outro modelo de metadados amplamente utilizado em razão de sua facilidade e interoperabilidade é o Dublin Core ou Dublin Core Metadata Element Set (DCMES). A iniciativa é uma comunidade aberta que coopera para a inovação na criação de metadados e

para as melhores práticas em todas as etapas de sua organização. A sua semântica foi definida e é mantida por um grupo internacional e interdisciplinar de profissionais da área de ciência da informação e campos relacionados (Dublincore.Org 2022). O Dublin Core Metadata Initiative (DCMI) publicou o seu primeiro conjunto de metadados em 1998, em HTML, e os seus quinze elementos de metadados (propriedades) para descrever recursos se tornaram ISO 15836. Esses elementos são: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title e type. (Evans 2019).

O Dublin Core tornou-se um dos vocabulários mais populares para utilização com o RDF. Os seus metadados descritivos têm em vista a interoperabilidade através da utilização de vocabulários compartilhados mundialmente, baseados nos princípios dos dados ligados (Dublincore.Org 2022). Desde a sua concepção, o conjunto de elementos tem sido aplicado a uma variedade de materiais diferentes, incluindo uma variedade de itens audiovisuais (Evans 2019).

Embora de simples compreensão, o DCMES é insuficiente em termos de capacidade de representação. Assim, em 2008, a especificação Dublin Core Terms (DCTERMS) foi lançada com o intuito de difundir uma semântica de descrição mais refinada e detalhada. Ela possui o conjunto original de quinze descritores como subpropriedades dos descritores originais, de forma a garantir a compatibilidade com os registos já presentes, e adiciona vários outros. A sua representação no RDF possibilita que os seus registos de metadados sejam publicados como dados abertos ligados, além de fornecer um modelo de metadados mais detalhado. Essa ampliação no esquema pode ajudar a enriquecer os registos de metadados, porém também intensifica a dificuldade da tarefa de descrição dos dados já que nem todos os elementos são pertinentes na descrição de resultados de pesquisas diversas, principalmente pelos utilizadores sem conhecimento em gestão de dados (da Silva, Ribeiro e Lopes 2019).

Atualmente, toda modificação feita no modelo é revisada por um Conselho de Utilização seguindo uma Política de “Namespace”. Essa política descreve como os termos são concedidos e determina limites à quantidade de mudanças editoriais permitidas às etiquetas, definições, e comentários de utilização (Dublincore.Org 2022).

2.4. Trabalhos Relacionados

A cada ano surgem mais projetos relacionados com o uso de dados ligados no patrimônio cultural. A fim de exemplificar como eles vêm sendo utilizados por diversas instituições culturais ao redor do mundo, abaixo cito alguns dos principais trabalhos que vêm sendo elaborados na área.

- **Linked Open Data Infrastructure for Digital Humanities in Finland (LODI4DH)** - Infraestrutura de Dados Abertos Vinculada para Humanidades Digitais na Finlândia. É um projeto cooperativo entre a Universidade de Aalto, o Departamento de Ciência da Computação e a Universidade de Helsinque, Centro HELDIG para Humanidades Digitais. O intuito é o desenvolvimento de serviços nacionais de dados ligados para a ciência aberta. Os serviços possibilitam a divulgação e uso de grupos de dados para pesquisa de Humanidades Digitais (HD) através do uso abundante de dados em formatos estruturados e formalizados através de interfaces abertas. O projeto LODI4DH é apoiado pela cadeia de colaboração e programas nacionais em HD desenvolvidos desde 2002, pela UH e Aalto, que concebeu protótipos de infraestrutura como o serviço de ontologia ONKI, o serviço de ontologia Finto na Biblioteca Nacional de Finlândia e a plataforma Linked Data Finland LDF.fi (Semantic Computing Research Group 2022);
- **MuseumFinland - Finnish museums on the semantic web** - Museus finlandeses na web semântica. O projeto é de um portal semântico para a divulgação de coleções diversificadas de museus na web semântica, evidenciando como os museus com as suas coleções semanticamente fartas e conectadas podem gerar um enorme portal de recolha semântica na web através do compartilhamento de um conjunto de ontologias. Assim, se tornou viável a criação de coleções semanticamente interoperáveis e oferecendo serviços de busca e navegação eficientes para os visitantes do museu. O sistema possui uma exposição de mais de 4.000 artefatos culturais, como tecido, mobiliário, utensílios, entre outros. Ademais, foram inseridos no sistema metadados relacionados com aproximadamente 260 locais históricos na Finlândia. Os objetivos para a construção do sistema foram: oferecer uma perspectiva global para coleções distribuídas; recuperar informações fundamentadas em conteúdo; criar conteúdo ligados semanticamente; facilitar a divulgação de conteúdo local. Os serviços gerais de pesquisa e navegação dos recursos e metadados do aplicativo são separados por uma camada de regras lógicas. A estrutura e o programa desenvolvidos para a criação do portal já foram empregados com sucesso em outros tipos de domínios. Em 2004, o MuseumFinland ganhou o Semantic Web Challenge Award (segundo prêmio) (Hyvönen et al. 2005);
- **TOMS (Thailand Open Museum System)** - Sistema de Museu Aberto da Tailândia. O projeto consiste em um aplicativo criado para possibilitar a colaboração e a partilha de uma grande quantidade de informações entre os museus nacionais. Isso é possível através do uso de dados aberto ligados e do modelo de ontologia CIDOC (CIDOC CRM), que é um reconhecido padrão de ontologia utilizado em museus. O sistema do TOMS é constituído pelas três respectivas camadas: de armazenamento de dados, com a função de guardar o banco de dados dos museus envolvidos, possuindo material digital em 3D e descrições de metadados

apresentados em um formato distribuído padrão; de manipulação e tratamento de dados, que manipula e processa os dados e realiza as pesquisas; de interface do sistema, que é criada para cada site com a função de trocar e incorporar dados com outros sites na rede através de um endpoint LOD SPARQL (Chanhom e Anutariya 2019);

- Swiss Art Research Infrastructure (SARI) - Infraestrutura de Pesquisa Artística Suíça. É um projeto que oferece o acesso a dados de pesquisa de domínio particular, dados de coleção, materiais visuais digitalizados e dados de referência concernentes aos setores do patrimônio cultural, assim como das humanidades digitais. O objetivo é juntar a experiência acadêmica de entidades de pesquisa de todo o país, como museus, arquivos, institutos acadêmicos em geral entre outros. O SARI oferece de forma on-line dados essenciais de pesquisa e operadores de coleta, atendendo à procura da comunidade acadêmica por recursos de alta qualidade. A SARI integra o Roteiro de Infraestruturas de Pesquisa de Relevância Nacional da Secretaria de Estado de Educação, Pesquisa e Inovação. E se encontra instalada e alojada na Universidade de Zurique, sendo utilizada em colaboração com o ETH Zurique (GTA Institute) e o Swiss Institute for Art Research, Zurique (Swissartresearch.Net 2022).
- Wisski - Infraestrutura de Comunicação Científica. O projeto criado em colaboração com o Germanisches Nationalmuseum (Nuremberg), o Zoologisches Forschungsmuseum Alexander Koenig (Bonn) e o Lehrstuhl Informatik (Departamento de Ciência da Computação) para inteligência artificial na Universidade Friedrich-Alexander Nürnberg-Erlangen, busca produzir um sistema fundamentado nas tecnologias da internet que ofereça apoio à comunicação e documentação científica do seu patrimônio cultural. Para isso, utiliza o padrão CIDOC-CRM para a gestão e análise dos dados internos. Tem como objetivo mostrar como o tratamento de conhecimento apoiado por ontologias deve ser realizado e as vantagens alcançadas nas trocas e agregação de dados. O WissKI procura criar soluções de aplicações e o desenvolvimento de portais de informação fundamentados na internet, intermediado por grupos de investigação, sendo o seu conteúdo resultado de objetos de museu. Esses portais podem ser usados como instrumentos de ajuda para a busca de informações ou para o desenvolvimento de diferentes projetos. Um exemplo desse tipo de portal de informações cooperativas são as plataformas Wiki (Nationalmuseum 2022).

3. Entidades em Registos de Arquivo e em Bases de Dados Ligados

Quando se trata da preservação da memória nacional portuguesa, o Arquivo Nacional da Torre do Tombo (ANTT) ocupa a posição mais importante. O arquivo é gerido pela Direcção-Geral do Livro, Arquivos e Bibliotecas (DGLAB), e detém a maior coleção do património cultural português, com documentos originais desde o século IX até a atualidade. O ANTT possui uma coleção amplamente digitalizada e com o seu próprio sistema de descrição online estabelecido há 20 anos, em conformidade com as normas do Conselho Internacional de Arquivos (ICA). Os metadados de descrição arquivística consistem, na sua maioria, em descrições textuais relacionadas à história e ao conteúdo dos documentos. Com o aprimoramento das descrições de objetos de arquivo, que passaram a incluir informações do meio digital e a demandar condições de interoperabilidade, se tornou necessário a criação de novos instrumentos de descrição que englobassem os documentos de bibliotecas, arquivos e museus. Para isso, será necessário serem instrumentos mais flexíveis, apurados e fáceis de serem processados por máquinas, sendo essas as principais particularidades dos dados abertos ligados (LOD) nas redes semânticas. Com esses objetivos traçados, o ANTT está a explorar a norma elaborada pela comunidade dos museus, o Modelo de Referência Conceitual (CRM) do CIDOC (INESC TEC 2022).

Essa nova etapa apresenta alguns problemas como: a migração satisfatória entre as normas ICA e as do CIDOC CRM, que utiliza os modelos de tradução vigentes, assim como, as inferências das relações com o uso de métodos semi-automatizados; o apoio à descrição com o uso de instrumentos que tornam automático parte da criação dos registos de metadados com maior complexidade do CIDOC CRM; interface apropriada para utilizadores e máquinas, favorecendo a interoperabilidade tanto com os arquivos quanto com as redes semânticas mundiais e otimizando o uso dos utilizadores aos arquivos (INESC TEC 2022).

3.1. O Projeto DigitArq

Atualmente, há um enorme volume de informação arquivística e de materiais disponibilizados na internet, porém toda essa informação se encontra compartimentada entre instituições. Cada instituição dispõe de suas próprias estruturas e interfaces de buscas, concedendo acesso às informações de uma forma específica e com diferentes graus de qualidade. A fim de reduzir esses problemas, a Direcção-Geral de Arquivos (DGARQ), criou a Rede Portuguesa de Arquivos (2007), que teve como um dos seus objetivos o desenvolvimento de um portal de pesquisa que centralizasse toda a informação de arquivo criada em território nacional. Nascia assim o Portal

Português de Arquivos, que atualmente utiliza o software DigitArq para suportar os seus trabalhos (Ferreiros, L. M., Ferreira, M. e J. C. Ramalho 2010).

O DigitArq é um software composto por 5 módulos operacionais que buscam ajudar um profissional de arquivo em atividades, tais como: descrição arquivística, gestão de projetos de digitalização, divulgações e navegação na web e pesquisa. O programa tem como propósito garantir um serviço ágil, satisfatório e transparente para os utilizadores. Em 2004, o projeto DigitArq recebeu o prêmio Fernandes Costa (Agência para a Sociedade do Conhecimento) de melhor trabalho no quesito inovação e contribuição para o progresso da Sociedade da Informação em Portugal. O software se baseia em três normas internacionais: International Standard Archival Description (ISAD), Encoded Archival Description (EAD) e Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). A norma ISAD e a EAD visam apoiar os sistemas de descrição arquivística. A OAI-PMH pretende suportar a incorporação do DigitArq com portais agregadores de metainformação, como é o caso da Europeana, Driver e do próprio Portal Português de Arquivos.

O módulo do DigitArq utilizado pelo Portal Português de Arquivos é o OAI-PMH, que possibilita a exportação de registos de agregadores de conteúdo. A DGARQ estabeleceu ainda juntar as instruções de uso sob o formato de uma wiki, ao invés da forma tradicional de manual. Isso permite aos utilizadores consultarem as instruções na internet, fazerem sugestões de melhorias e ainda colaborar na edição das instruções (Barbedo, Francisco 2011).

3.2. Introdução ao Projeto EPISA

Iniciado em 1 de janeiro de 2019, o projeto EPISA (Inferência de Entidades e Propriedades para Arquivos Semânticos) concebido no Centro de Sistemas de Informação e de Computação Gráfica (CSIG) do INESC TEC em colaboração com a Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB) e com a Universidade de Évora, tem a previsão de ser finalizado em dezembro de 2022. O programa tem financiamento da Fundação para a Ciência e a Tecnologia (FCT) (INESC TEC 2022).

O EPISA pretende desenvolver um novo protótipo de descrição para os arquivos, e possibilitar a geração semiautomática de metadados. Dessa forma, os arquivos nacionais poderão ser inseridos na rede mundial de dados abertos ligados e semânticos. O projeto é norteado pelos seguintes objetivos:

- Desenvolver um protótipo de descrição para os arquivos fundamentado em modelos recomendados para museus e arquivos;
- Suportar a descrição de arquivo na geração semi-automática de metadados;

- Transferir as descrições presentes de acordo com as normas ISAD (arquivos) para normas segundo o modelo CIDOC-CRM (museus) e RiC (ICA);
- Integrar os arquivos na rede mundial de dados ligados e semânticos;
- Conectar os dados dos arquivos a diferentes fontes de informação;
- Sugerir interfaces que viabilizem a interatividade para públicos diversos;
- Incentivar a participação dos arquivos nacionais nos agregadores internacionais;
- Promover o acesso aos arquivos;
- Viabilizar o acesso à informação;
- Ajudar na associação dos recursos administrativos e de pesquisa no Arquivo Digital da Administração Pública (INESC TEC 2022).

3.3. Expansão de Entidades nos Registos de Arquivo do DigitArq

Uma das finalidades do trabalho foi o de procurar entidades de interesse entre os documentos do Arquivo Nacional da Torre do Tombo (ANTT), disponibilizados no Portal Português de Arquivos através do DigitArq, com o propósito de expandir os registos tendo como foco o utilizador. Para isso, foram analisados os 25 registos que estão sendo utilizados pelo projeto EPISA para a identificação de entidades e relações, assim como para a estruturação de interoperabilidade e relações com outros recursos não-arquivísticos. A Figura 7 apresenta um fragmento da grelha com os 25 registos utilizados pelo projeto EPISA.

Arquivo	Código de referência	Título	Nível
ANTT	PT/TT/JIM	Juizo da Índia e Mina	Fundo
ANTT	PT/TT/JIM/A/0001/00001	Autos civis de petição e justificação para protesto de avaria grossa e rectificação do mesmo em que é autor Acácio José Correia, mestre e piloto do navio "Castor e Pólux" e réus os interessados na carga do navio e seus seguradores.	Documento composto
ANTT	PT/TT/JIM-JJU	Juizo das Justificações Ultramarinas	Subfundo
ANTT	PT/TT/JIM/E	Cartório do escrivão Lino José de Almeida Lobo da Torre do Vale	Secção
ANTT	PT/TT/JIM-JJU/004	Oriente	Série
ANTT	PT/TT/JIM-JJU/003/0005	Maço 5	Unidade de instalação
ANTT	PT/TT/JIM-JJU/003/0005/00008	Autos de sentença de José Borges Leal, Dionísio José, Jacinto Borges Leal, Mariana de Jesus, Joaquim Machado, filhos de Manuel Leal e de Rita Mariana, naturais e moradores no distrito de Vila da Praia, Ilha Terceira (Açores)	Documento composto
ANTT	PT/TT/JIM/JIE/0001/00038	Extrato do jornal a bordo do navio sueco "Joanes de Gotemburgo"	documento simples
ANTT	PT/TT/JIM/JIE/0002/00035	Procuração de Vasco Salter Mascarenhas e Mendonça dando poderes a João Manuel da Costa	documento simples
ANTT	PT/TT/JIM/A/0008/00012	Ação civil de fretes em que é autor António Henriques Leal e réu António Esteves Costa	Documento composto
ANTT	PT/TT/EPJS/SF/008/10253	Venda de louça preta: as Sras...	Documento simples
ADPTG	PT/ADPTG/MON/MSBP	Mosteiro de São Bernardo de Portalegre	Fundo
ADSTB	PT/ADSTB/AC/GAS	Gabinete da Área de Sines	Fundo
ADCTB	PT/ADCTB/PSS/APASC/00001	Carta de Antão Santos da Cunha para o Sub-secretário de Estado	documento simples
ADBGC	PT/ADBGC/AC/GCBGC/PAS/077/01140	Registo de passaportes defrídos	Unidade de instalação
ANTT	PT/TT/AS/A/001/000142	"Um aquário de 2,40x1,00 m"	documento simples
ADAVR	PT/ADAVR/NOT/CNVER2/002/0001/000007	Testamento	documento simples
ANTT	PT/PNA/DGFP/0002/0001/00047/00002	Planta de parte do 3.º pavimento do Convento das Necessidades	Documento simples
ANTT	PT/TT/CRT/198	Carta náutica, de João Teixeira Albernaz I	Documento simples
ANTT	PT/TT/BUL/0003/06	Bula "Romani Pontificis" do Papa Adriano VI concedida ao Infante D. Henrique pela qual o proveu no Mosteiro de São Cristóvão de Laões, e no priorado do Mosteiro de São Jorge como comenda, por óbito de Diogo da Gama	Documento simples
ANTT	PT/TT/TSO-IE/021/00203	Processo de Jerónima da Cruz	documento composto
ANTT	PT/TT/DP	Desembargo do Paço	Fundo
ADPRT	PT/ADPRT/JUD/TCVNG/077/00053	Inventário por óbito de Manuel Joaquim Ferreira Lino	documento composto
ADBJA	PT/ADBJA/PRQ/AJT01/001/0013	Registo de batismo	Unidade de instalação
ADSTR	PT/ADSTR/NOT/07CNABT/001/0023	Notas para escrituras diversas	Unidade de instalação

Figura 7 - Fragmento da grelha dos 25 registos utilizados pelo Projeto EPISA (Projeto EPISA 2022).

Após a análise dos registos na Figura 7, procurou-se inferir quais das entidades poderiam ser promissoras para expandir, ou seja, que entidades contidas nos registos seriam interessantes para serem enriquecidas através da ligação a dados de outras fontes. Na descrição arquivística alguns dos elementos descritivos incluem informação sobre pessoas, lugares, eventos, organizações, que são entidades que os arquivistas consideram importantes para serem apresentadas de uma forma estruturada. Tendo em consideração estes aspectos e a informação textual dos elementos dos registos, as classes de entidades consideradas de interesse para expansão foram: person, place, product, vessel, organization, legal system, activity, project, thing, document, and event. É pertinente esclarecer que os registos do arquivo estão em português, porém como os vocabulários do CIDOC que estão sendo utilizados no projeto estão em inglês, tanto na construção das tabelas nos apêndices quanto nos testes, foi escolhido utilizar os nomes das classes de entidades e propriedades em inglês de forma a padronizar os termos. A seguir, são apresentados três dos vinte e cinco registos analisados para ilustrar o meu processo de escolha.

Registo 1 - Juízo da Índia e Mina

JM Juízo da Índia e Mina - AA Juízo das Justificações Ultramarinas 1700/1833 - A Cartório do escrivão Claudio da Costa Pereira 1772/1820 - B Cartório do escrivão Henrique José Saravia da Guerra 1781/ - C Cartório do escrivão Antonio Joaquim Peixoto da Silva 1759/ - D Cartório do escrivão José Duarte de Sousa 1779/1802 - E Cartório do escrivão Lino José de Almeida Lobo da Torre do 1 - F Cartório do escrivão José Antonio Xavier de Lemos 1785/1811 - G Cartório do escrivão Antonio Xavier de Lemos 1806/1822 - H Cartório do escrivão Manuel Francisco de Araújo Monteiro 18 - I Cartório do escrivão Martinho Bartolomeu Rodrigues 1818/181 - J Cartório do escrivão José Gomes 1762/1780 - JA Cartório do escrivão João Pereira da Costa 1806/1818 - JB Cartório do escrivão Manuel José da Silva 1761/1783 - JC Cartório do escrivão José Tomás Delfim 1756/1777 - JD Cartório do escrivão José Noronha Lobo 1771/1781 - JE Cartório do escrivão Domingos Duarte Muradas Renna 1802 - JF Cartório do escrivão Manuel Francisco de Araújo Monteiro 1 - JG Cartório do escrivão Luis José Moreira de Carvalho 1793/1 - JH Cartório do escrivão Francisco José do Avelar e Silva 1825/ - JI Cartório do escrivão Bartolomeu Rodrigues Couto 1749/1781 - JJ Cartório do escrivão Toraz José Soares 1810/1810 - JJA Cartório do escrivão João Teodoro Baradas 1778/1778 - JJB Cartório do escrivão José Antonio da Campos 1810/1810 - JJC Cartório do escrivão José Batista Moreira 1772/1772 - JJD Informação não disponível - JJE Informação não disponível - JJF Informação não disponível - JJG Informação não disponível - JJH Informação não disponível - JJI Informação não disponível - JJJ Informação não disponível - JJK Informação não disponível - JJL Informação não disponível - JJM Informação não disponível - JJN Informação não disponível - JJO Informação não disponível - JJP Informação não disponível - JJQ Informação não disponível - JJR Informação não disponível - JJS Informação não disponível - JJT Informação não disponível - JJU Informação não disponível - JJV Informação não disponível - JJW Informação não disponível - JJX Informação não disponível - JJY Informação não disponível - JJZ Informação não disponível	JUÍZO DA ÍNDIA E MINA NÍVEL DE DESCRIÇÃO Fundo CÓDIGO DE REFERÊNCIA PT/TTUJIM TIPO DE TÍTULO Atribuído DATAS DESCRITIVAS [1700-1833] DIMENSÃO E SUPORTE 2040 u.l.; papel HISTÓRIA ADMINISTRATIVA/BIOGRÁFICA/FAMILIAR Ao Juízo da Índia e Mina, criado na sequência dos descobrimentos e do comércio com o Ultramar, competia conhecer dos processos cíveis e crime concernentes ao comércio e às cargas e descargas dos navios. As decisões tomadas fora dele eram consideradas nulas. Era constituído por juiz, escrivães, distribuidor, inquiridor, contador, depositário, meirinho, escrivão do meirinho e porteiro. Das sentenças proferidas neste Juízo cabia recurso para o Juízo das Apelações e Agravos da Casa da Suplicação. Na dependência do Juízo da Índia e Mina existia o Juízo das Justificações Ultramarinas. ÂMBITO E CONTEÚDO Processos referentes a avarias de navios, pagamento de soldadas, cobrança de fretes, fretes e descarga, habilitação de herdeiros (de marinheiros que morreram nas embarcações ou nos lugares do ultramar), processos de indivíduos que queriam provar ser portugueses e que os seus navios tinham sido feitos em estaleiros portugueses, não havendo estrangeiros interessados nas suas cargas. Processos sobre corsários e pirataria. A maioria dos processos refere os produtos comercializados: açúcar branco, açúcar mascavado, aguardente, algodão, arroz, atados, cacau, milho, feijão, carne, favas, manteiga, bolacha, biscoito, queijo, salsaparrilha, café, chá, canela, couro, couro em cabelo, goma, farinha, escravos vindos de Benguela (Angola), mel, melão, pau de tatajuba, sola, tabuado, vinhático, aças de lenha, toros de jacarandá, vaquetas, vinho, arroz, pano de café, caçoeiras de jacarandá, seda, louça esmaltada fina, toucador de charão, madeira de carvalho, tabaco, móveis (baú, piano), pimenta, cravo, tecidos, leques, tabuado, canela, tabuado tapainhã, chifres, farinha, quina e anil, entre outros. Referem, também, as rotas comerciais: Açores, Angola, Ilha do Príncipe, Baía, Benguela, Maranhão, Bengala, Bombaim, Cabo Verde, Ceilão, China, Costa da Mina, Costa de Coromandel, Damão, Goa, Guiné Bissau, Índia, Macau, Madeira, Mangalor, Pará, Pernambuco, Rio de Janeiro, Surrate, etc. Referem ainda o tipo de embarcações: navio, corveta, bergantim, galera, escuna, brigue, iate, caïque, nau, sumaca, barco, corsário, polaca, e o respectivo nome de que se apresentam alguns exemplos: "Nomes das embarcações", "Mãe de Deus e Santa Ana", "Nossa Senhora do Carmo, Leão", "Nossa Senhora da Esperança e Santa Rita, Lusitana", "A flor do Porto, Bom Jesus de Vila Nova", "Águia Dourada", "Anna" (navio inglês), "Anibal", "Marta", "Belisário", "Campeões", "Flor do Porto, Bom Jesus de Vila Nova", "Gratidão", "Indiano", "Mãe de Deus e Santa Ana", "Ódio Maranhão", "Negrinha", "Nossa Senhora da Conceição e Bom Jesus dos Navegantes", "Nossa Senhora da Conceição e Santa Ana", "Nossa Senhora da Conceição e São Castano", "Nossa Senhora da Conceição e São Francisco", "Nossa Senhora da Conceição e São Francisco de Paula", "Nossa Senhora da Conceição e São Francisco Xavier", "Nossa Senhora da Conceição e São José", "Nossa Senhora da Conceição e Princesa de Portugal", "Nossa Senhora da Conceição, Rei David", "Nossa Senhora da Conceição, Santo António e Almas", "Nossa Senhora da Esperança, Neptuno", "Nossa Senhora da Graça, Maranhão", "Nossa Senhora da Lampadosa e São João Baptista", "Nossa Senhora da Nazaré e Bomfim", "Nossa Senhora da Nazaré e São Miguel", "Nossa Senhora da Nazaré, Negrinha", "Nossa Senhora da Nazaré, Santa Ana, Santo António e Almas", "Nossa Senhora da Oliveira e São José", "Nossa Senhora da Penha de França, Rainha de Nantes", "Nossa Senhora da Piedade e Almas", "Nossa Senhora da Piedade e São Boaventura", "Nossa Senhora da Piedade e São Francisco de Paula", "Nossa Senhora da Piedade e São José", "Nossa Senhora da Piedade, Destino", "Nossa Senhora da Vitória, Princesa Divina", "Nossa
--	--

Figura 8 - Registo 1 - Juízo da Índia e Mina (Arquivo Nacional Torre do Tombo 2022).

Entidades para Expansão: atados; salsaparrilha; pau de tatajuba; sola; tabuado; vinhático; vaquetas; pano de café; caçoeiras de jacarandá; tabuado tapainhã; quina; corveta; bergantim; galera; escuna; brigue; iate; caïque; nau; sumaca; corsário; polaca; Açores; Angola; Ilha do Príncipe; Baía; Benguela; Maranhão; Bengala; Bombaim; Cabo Verde; Ceilão; China; Costa da Mina; Costa de Coromandel; Damão; Goa; Guiné Bissau; Índia; Macau; Madeira; Mangalor; Pará; Pernambuco; Rio de Janeiro; Surrate; Galiza; Brasil; Félix José da Costa; D. Francisca Xavier; Manuel de Almeida.

Essas foram algumas das entidades que se mostraram interessantes de serem enriquecidas com dados de outras fontes. Dentre elas estão: produtos, embarcações, lugares e pessoas. Os produtos, por não terem a mesma designação de antigamente ou simplesmente não serem mais comercializados, não são de total conhecimento popular. Logo, mais informações sobre a sua origem e uso seriam interessantes. As embarcações seriam entidades interessantes de ter alguma informação extra associada que esclarecesse a diferença entre elas, por exemplo, uma imagem. Os lugares e pessoas são sempre entidades interessantes de terem alguma informação extra associada sempre que possível, pois poderia esclarecer e enriquecer muitas das pesquisas realizadas em Arquivos. Muitos dos lugares, como cidades e países, mudaram de nome e território ao longo dos anos, sendo relevantes maiores esclarecimentos.

Registo 11 - "Venda de Louça Preta: as Sras. D. Maria Luísa de Brito Rocha e Melo, D. Helena de Brito Rocha e Melo, D. Ana da Cunha e Meneses Pinto Cardoso

(Lumiares), D. Maria Luísa Barroso da Câmara, o Sr. Francisco Queirós de Andrade Pinto, e as Sras. D. Francisca Seabra de Lacerda, D. Maria de Brito Rocha e Melo e D. Maria Emília Cabral da Silva"

Figura 9 - Registo 11 – Venda de Louça Preta (Arquivo Nacional Torre do Tombo 2022).

Entidades para Expansão: Louça preta; D. Maria Luísa de Brito Rocha e Melo; D. Helena de Brito Rocha e Melo; D. Ana da Cunha e Meneses Pinto Cardoso; D. Maria Luísa Barroso da Câmara; Sr. Francisco Queirós de Andrade Pinto; D. Francisca Seabra de Lacerda; D. Maria de Brito Rocha e Melo; D. Maria Emília Cabral da Silva.

As entidades escolhidas neste registo são: atividade e pessoas. A atividade em questão é o artesanato, sendo a fabricação da louça preta considerada um patrimônio cultural intangível. Por isso, seria muito interessante associar e ligar novas informações a seu respeito. As outras entidades são pessoas, que como já mencionado, são sempre relevantes para expansão com novos dados, sobretudo em um arquivo.

Registo 12 – Mosteiro de São Bernardo de Portalegre

MON15201 Mosteiro de São Bernardo de Portalegre 1443-11-12 A Comunidade 103844-2015201-04-20 Mosteiro 1443-11-1215201-04-20	MOSTEIRO DE SÃO BERNARDO DE PORTALEGRE NÍVEL DE DESCRIÇÃO Fundo CÓDIGO DE REFERÊNCIA PTIADPTGIMONIMSSP TIPO DE TÍTULO Atribuído DATAS DE PRODUÇÃO 1443-11-12 ✓ a 1882-04-20 ✓ DIMENSÃO E SUPORTE 22 ex. (1899 doc.) HISTÓRIA ADMINISTRATIVA/BIOGRÁFICA/FAMILIAR O Convento de São Bernardo, da Ordem de Cister, igualmente conhecido pelo nome de Mosteiro de Nossa Senhora da Conceição, foi fundado pelo bispo da Guarda, D. Jorge de Melo, em 1518, numa das encostas da Serra de S. Mamede, no sítio da Fontedeira, junto à fonte do mesmo nome, para esse efeito cedido pelo conde da vila de Portalegre por doação que viria a ser confirmada por alvará de D. João III, lavrado em Santarém aos 14 de Junho de 1526. Anos mais tarde, em apostilha ao citado alvará, com data de 27 de Novembro de 1549, o mesmo soberano concedeu à abadesa do mosteiro a licença necessária para a construção da cerca do convento, tendo a igreja sido consagrada, em 10 de Março de 1572, pelo então bispo de Portalegre, D. André de Noronha. As obras de edificação ficaram, no essencial, concluídas em 1533 e foram complementadas por outras campanhas de construção localizáveis nos períodos de 1533-1550, 1587-1607/8 e, finalmente, durante o séc. XVIII. No edifício merece especial relevo, pela excelência do trabalho, o faustoso túmulo do fundador, atribuído a Nicolau Chanterene. A congregação cujos estatutos foram assinados em 19 de Agosto de 1531, directamente filiada à abadia de Claraval, foi extinta pelo decreto de 30 de Maio de 1834 e o mosteiro encerrado em 23 de Abril de 1878, por falecimento da última religiosa residente. De então para cá, o antigo mosteiro de São Bernardo albergou o Seminário Diocesano (1879), uma secção do Liceu de Portalegre (1880-1887), o Museu Municipal (1932-1991). As instalações são, presentemente, utilizadas pela Escola Prática do Agrupamento de Instrução da GNR de Portalegre. Imóvel classificado (Dec. 16 de Junho de 1910; DG 136 de 23 de Junho de 1910; Dec. 30.782, DG 225 de 28 de Setembro de 1940; Dec. 32.973, DG 175 de 18 de Agosto de 1943 e Dec. 30.838, DG 254 de 1 de Novembro de 1948). HISTÓRIA CUSTODIAL E ARQUIVÍSTICA Documentação transferida da Direcção de Finanças do Distrito de Portalegre para o Arquivo Distrital de Portalegre, em cumprimento do disposto na alínea a) do § 1º do art.º 28 do decreto com força de lei n.º 19952, de 27 de Junho de 1931. A documentação deu entrada no Arquivo Distrital em Março de 1938. FONTE IMEDIATA DE AQUISIÇÃO OU TRANSFERÊNCIA Incorporação directa proveniente da Direcção de Finanças do Distrito de Portalegre. ÂMBITO E CONTEÚDO Inclui alvarás, cartas e provisões régias, bulas, inventários, tombo, escrituras, sentenças, petições, inquirições, receitas e despesas, dotes, pareceres jurídicos, e testamentos INGRESSOS ADICIONAIS Não estão previstos quaisquer ingressos adicionais SISTEMA DE ORGANIZAÇÃO Organizado em séries cronologicamente ordenadas
--	---

Figura 10 - Registo 12 – Mosteiro de São Bernardo de Portalegre (Arquivo Nacional Torre do Tombo 2022).

Entidades para Expansão: Convento de São Bernardo (=Mosteiro de Nossa Senhora da Conceição); Ordem de Cister; D. Jorge de Melo; Serra de S. Mamede; Portalegre; D. João III; Santarém; D. André de Noronha; Nicolau Chanterene; Abadia de Claraval.

As entidades seleccionadas neste registo são: lugares, pessoas e organização religiosa. Anteriormente, já foi mencionado a importância do enriquecimento de dados sobre lugares e pessoas em Arquivos. Da mesma forma, organizações religiosas são ricas em informações relacionadas, logo, interessantes para expansão.

Após a análise e seleção das entidades consideradas relevantes nos 25 registos do Arquivo Nacional da Torre do Tombo (ANTT) trabalhados pelo projeto EPISA, o próximo passo foi fazer uma investigação das propriedades associadas a essas entidades no CIDOC CRM, Schema.org e Dublin Core. O CIDOC CRM foi escolhido por já ser o modelo conceitual utilizado no projeto EPISA, e pela sua capacidade de fazer a ligação entre fontes de informação diversas do património cultural, possibilitando tanto a recuperação de informações de alto nível quanto a elaboração de questionamentos mais específicos (CIDOC CRM 2022). O Schema.org e o Dublin Core foram os vocabulários escolhidos para fazer a comparação de propriedades durante a etapa de reuniões do projeto. O Schema.org se mostrou importante de ser analisado, dada a abrangência das suas propriedades e variedade de usos, não se limitando apenas à área

do patrimônio cultural. E o Dublin Core por conta do seu amplo emprego e popularidade internacional.

A Tabela de Entidades e Propriedades Associadas aos 25 Registos no CIDOC CRM, Schema.org e Dublin Core (APÊNDICE A) mostra como as respectivas entidades podem ser trabalhadas em cada um deles e possibilita uma comparação entre eles. Ao examinar a Tabela 1, que consiste em um fragmento da tabela do APÊNDICE A, podemos perceber que o Schema.org faz uso de algumas propriedades do Dublin Core para compor o seu vocabulário.

Os registos 6 e 8 foram os únicos onde não foram encontradas entidades relevantes para a expansão. O registo 6 intitulado Maço 5 não contém nenhuma entidade que poderia se enriquecer com mais informações. Da mesma forma, o registo 8 denominado Extrato do Jornal a Bordo do Navio Sueco “Joanes de Gotemburgo” não apresentou entidades consideradas interessantes para expansão em seu documento.

Nº de Registo	Entidades	CIDOC CRM	Schema.org	Dublin Core
Registo 2	Acácio José Correia	P14 carried out by (performed); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P152 has parent (is parent of)	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee	abstract; bibliographicCitation; coverage; description; memberOf

Tabela 1 - Fragmento da Tabela de Entidades e Propriedades Associadas aos 25 Registos no CIDOC CRM, Schema.org e Dublin Core (APÊNDICE A).

3.4. Fontes Exteriores de Enriquecimento e Ligação de Dados

Após a etapa de buscas por formas de identificar e expandir as entidades encontradas nos registos, o próximo passo foi o de buscas por fontes exteriores para o enriquecimento e ligação dos dados. Desse modo, foi escolhido explorar as bases de dados utilizadas pelo projeto EPISA: Wikidata, DBpedia e Europeana.

Essa etapa começou com a elaboração de uma lista de classes de entidades e de propriedades encontradas associadas ao registo 1 na Wikidata, Europeana e DBpedia. Essa lista (APÊNDICE

B), mostra a enorme quantidade de propriedades relacionadas com as classes pertencentes às entidades desse primeiro registo. Devido ao extenso número de propriedades, principalmente no que se refere à DBpedia, foi necessário fazer uma seleção das propriedades que fossem pertinentes às entidades dos registos. Com esse intuito, foi elaborada uma tabela de classes de entidades e propriedades encontradas associadas ao registo 1 na Wikidata, Europeana e DBpedia, após passar por uma primeira seleção (APÊNDICE C).

Nº de Registo	Classes das Entidades	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeana	Propriedades Associadas na DBpedia
Registo 1	product (wikidata); edm:Physical Thing (europeana); product (DBpedia)	product: subclass of (goods and services; product; good); image; has quality (function; product naming; equivalent class; different from (tangible good; goods and services); said to be the same as (product; goods and services); commons category (products); topic's main category (category:products); identifiers.	edm:PhysicalThing (as contextual entities); dc:coverage; dc:creator; dc:description; dc:format; dc:identifier; dc:relation; dc:rights; dc:source; dc:subject; dc:title; dc:type; dcterms:alternative; dcterms:created; dcterms:extent; dcterms:hasFormat; dcterms:hasPart; dcterms:isFormatOf; dcterms:isPartOf; dcterms:isReferenced By; dcterms:isRequiredBy; dcterms:medium; dcterms:provenance; dcterms:references; dcterms:temporal; edm:currentLocation; edm:hasType; edm:incorporates; edm:isDerivativeOf; edm:isSimilarTo;	product: abbreviation; abstract; alternativeName; cost; dc: subject; dc: title; dc: type; dct: description; ingredient; ingredientName; manufactory; manufacturer; material; origin; originalName; originallyUsedFor; owl: differentFrom; owl: sameAs; picture; price; produces;productShape; sourceName; sourcePlace; sourceRegion; foaf:img

Tabela 2 – Fragmento da Tabela de Classes das Entidades e Propriedades Encontradas Associadas ao Registo 1 na Wikidata, Europeana e DBpedia Após Primeira Seleção (APÊNDICE C).

Na tabela no APÊNDICE C, de que se mostra um fragmento na Tabela 2, foram reunidas apenas as propriedades que pudessem descrever algum aspecto significativo sobre as várias classes das entidades selecionadas do registo 1. Foi necessário avaliar uma lista grande de propriedades e ponderar sobre a pertinência de cada uma delas, ou seja, se ofereciam informações relevantes para o esclarecimento dessas entidades. A lista no APÊNDICE B foi importante para ilustrar a imensa quantidade de propriedades relacionadas com algumas

entidades na Wikidata, Europeana e DBpedia apenas no registo 1. Como o objetivo é fazer testes com essas propriedades, não seria viável fazer listas tão extensas de propriedades em cada base de dados, sendo o intuito apenas escolher as mais utilizadas nas pesquisas pelos utilizadores. Por esse motivo, um refinamento maior das propriedades se mostrou necessário. A tabela no APÊNDICE C mostra as propriedades que foram selecionadas nas três bases de dados associadas somente ao registo 1, e embora bem mais sucinta que a lista no APÊNDICE B, ainda se mostrou bastante ampla. Como o objetivo seguinte era mostrar as principais propriedades associadas às classes das entidades dos 25 registos, foi necessário passar as propriedades por mais um filtro. Desse modo, foi criada a Tabela Resumida de Classes das Entidades e Propriedades Associadas aos 25 Registos na Wikidata, Europeana e DBpedia (APÊNDICE D), de que se mostra um fragmento na Tabela 3. Nessa tabela podemos ver de forma bem mais resumida as propriedades que foram consideradas relevantes para serem utilizadas em um posterior processo de enriquecimento dessas classes de entidades.

Nº de Registo	Classes das Entidades	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeana	Propriedades Associadas na DBpedia
Registo 1	product (wikidata); edm:Physical Thing (europeana); product (DBpedia)	product: subclass of image; has quality; function; product naming; equivalent class; different from; image; identifiers	edm:PhysicalThing; dc:creator; dc:description; dc:format; dc:title; dc:type; dcterms:isRequiredBy; edm:currentLocation; edm:isDerivativeOf; edm:isSimilarTo; dcterms:alternative; edm:hasType	product: alternativeName; dc: title; dc: type; dct: description; ingredient; manufacturer; material; origin; originallyUsedFor; picture; sourcePlace; sourceRegion; owl:sameAs; foaf:img

Tabela 3 - Fragmento da Tabela Resumida de Classes das Entidades e Propriedades Associadas aos 25 Registos na Wikidata, Europeana e DBpedia (APÊNDICE D).

O próximo passo foi criar uma Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia (APÊNDICE E). Nessa etapa, o foco passou a ser as classes das entidades e as principais propriedades utilizadas para as descrever. Como resultado, as propriedades passaram por mais um crivo. A intenção foi a de fazer um comparativo das propriedades equivalentes ou semelhantes nas três bases de dados. Contudo, muitas das propriedades existentes em uma base de dados não existem nas outras, como se pode ver pela Tabela 4. Por exemplo: na classe geographic location da Wikidata existe a propriedade *country* enquanto na classe equivalente edm:Place da Europeana essa propriedade não existe.

Classes	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeana	Propriedades Associadas na DBpedia
geographic location (wikidata); edm:Place (europeana); place (DBpedia)	label	skos:prefLabel	dc: title
	also known as	skos:altLabel/ owl:sameAs	alternativeName
	description	skos:note	dct: description
	coordinate location	wgs84_pos: lat;wgs84_pos:long	kindOfCoordinate/ location
	country	x	country
	x	x	city
	x	x	foundedBy
	x	x	foundingDate
	x	x	code
	x	x	agglomerationPopulation Total
	x	x	area
	x	x	climate
	x	x	address

Tabela 4 – Fragmento da Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia (APÊNDICE E).

A base de dados com o maior número de propriedades é a DBpedia, em segundo lugar está a Wikidata. A base de dados da Europeana possui não apenas a menor variedade de propriedades, como também a menor variedade de classes. As chamadas classes contextuais se restringem às seguintes: edm:Agent, edm:Place, edm:TimeSpan, skos:Concept e cc:License. Assim, as entidades que não se adequavam a pessoas, lugares, intervalo de tempo e licença/permissão foram inseridas como concept (conceito/noção/ ideia). Nessa Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia (APÊNDICE E), foram incluídas apenas as propriedades que pudessem trazer maiores informações acerca das classes em um futuro processo de expansão.

De modo a clarificar com alguns exemplos dos 25 registos, foi elaborada a Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia (APÊNDICE F). Nessa tabela, além das classes, estão alguns exemplos das entidades previamente seleccionadas dos 25 registos do Arquivo da Torre do Tombo. Como se pode notar no fragmento da Tabela 5, algumas classes possuem vários exemplos de entidades, como em pessoas (person; edm:Agent; foaf:Person/ person), enquanto outras

apenas um exemplo, como em handicraft; skos:Concept; activity.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeiaana	Propriedades Associadas na DBpedia
person (wikidata); edm:Agent (europeana); foaf:Person/person (DBpedia)	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	gender given name family name also known as date of birth x biographic information x x x x	rdaGr2: gender foaf:name x owl:sameAs rdaGr2:dateOfBirth rdaGr2:dateOfDeath rdaGr2:biographicalInformation rdaGr2:professionOrOccupation rdaGr2:placeOfBirth rdaGr2:placeOfDeath x	foaf:gender foaf:name/birthname foaf:familyName alternativeName birthdate deathDate description profession birthplace deathplace spouse
handicraft (wikidata); skos:Concept (europeana); activity (DBpedia)	louça preta	X said to be the same as image practiced by x x x x	skos:prefLabel skos:altLabel x x skos:note x x x	name x x x det:description country material colour

Tabela 5 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeiaana e DBpedia (APÊNDICE F).

4. Testes de Entidades e de Propriedades dos Registos

Após as fases de descrição das entidades encontradas nos registos e da pesquisa por fontes exteriores para o enriquecimento e ligação dos dados, o passo seguinte consistiu em testar as entidades e propriedades selecionadas. O objetivo era verificar se os testes confirmavam a relevância das propriedades escolhidas e recuperavam resultados significativos das mesmas, permitindo-nos concluir que estas propriedades trazem informação importante para a compreensão e esclarecimento das respectivas entidades. Além disso, o facto de existir informação sobre uma determinada entidade na web facilita o enquadramento destas entidades nas interfaces de utilizador que estão a ser concebidas. Contudo, para a recuperação desta informação, os termos utilizados devem ser interoperáveis, ou seja, compatíveis com os termos utilizados noutros repositórios digitais.

Essa etapa obteve a colaboração de outro participante do projeto EPISA, Ricardo Pinto, do mestrado em Engenharia Informática e Computação. O seu trabalho centra-se na ligação das informações dos arquivos com outras informações presentes em repositórios internacionais, tais como Wikidata, DBpedia e Europeana. A sua tarefa consiste em expandir sistematicamente o conjunto de arquivos de confiança nos repositórios a serem analisados para que mais informação possa ser acrescentada aos arquivos. Como os nossos objetivos se mostraram complementares, essa fase do projeto contou com a contribuição e auxílio em ambos os trabalhos a fim de aprimorar as nossas atividades.

4.1. Testes com Wikidata Query Service

Os primeiros testes foram feitos com as propriedades da Wikidata, sendo utilizadas as propriedades da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia (APÊNDICE F). Contudo, novas propriedades foram adicionadas após sugestões dos serviços de consulta durante os testes, mesmo não sendo tão utilizadas com essas classes de entidades. Assim, foi necessário elaborar uma nova tabela, de que se mostra um fragmento na Tabela 6, com a inclusão dessas propriedades mais genéricas, com o intuito de aprimorar a futura recuperação de informação. A Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata Após Testes no Wikidata Query Service (APÊNDICE G) mostra essas novas propriedades.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na Wikidata	
person	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	gender birth name given name family name also known as date of birth	date of death occupation place of birth place of death spouse image

Tabela 6 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata Após Testes no Wikidata Query Service (APÊNDICE G).

Para garantir que pudessem ser reproduzidos, os exemplos a seguir foram realizados utilizando o Wikidata Query Service com as entidades pertencentes à classe pessoa e com as propriedades que constam do APÊNDICE G. A Figura 9 mostra o teste realizado utilizando a entidade Mouzinho da Silveira e as propriedades *data de nascimento* e *data da morte*.

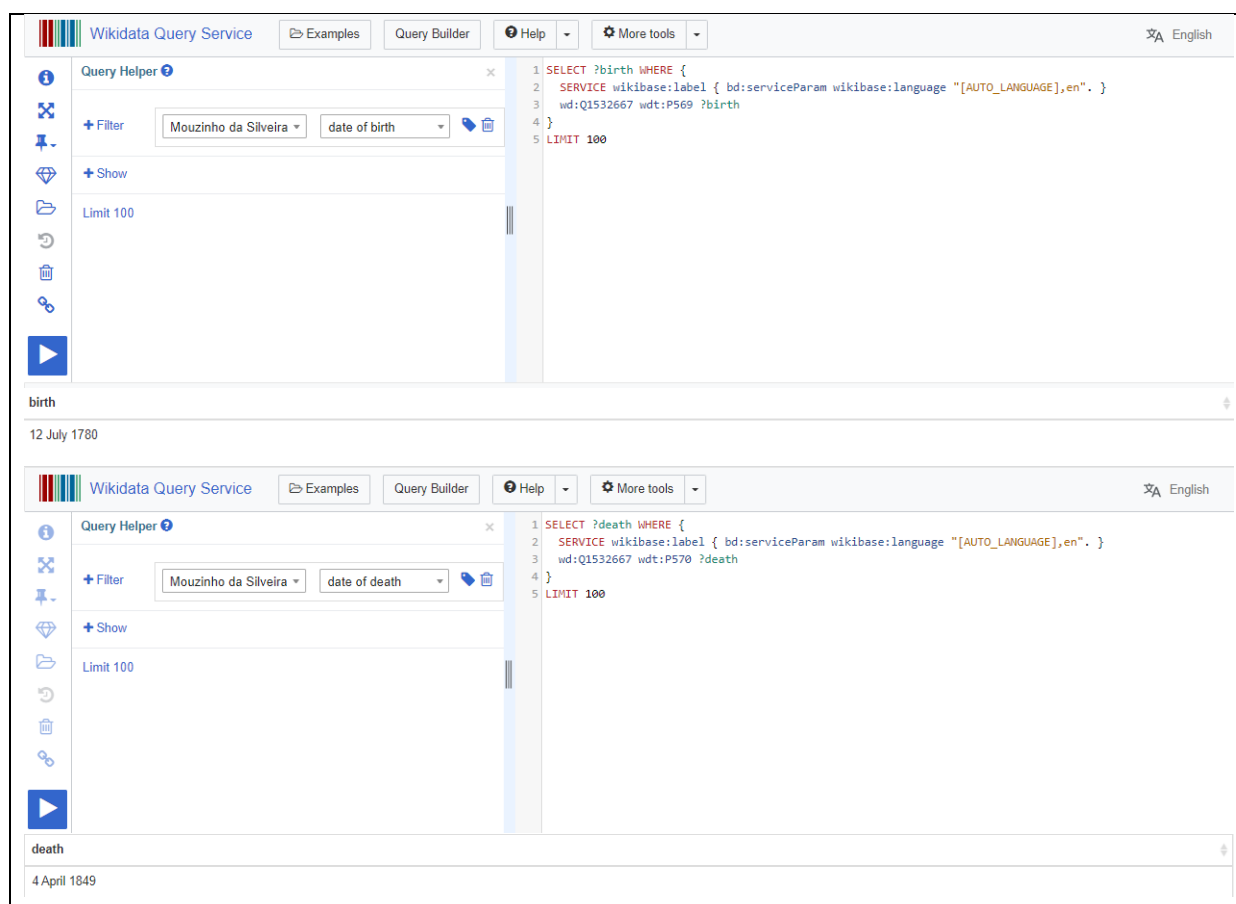


Figura 11 - Teste com a entidade Mouzinho da Silveira na Wikidata

Estas buscas recuperaram as informações 12 de Julho de 1780 e 4 de Abril de 1849, respectivamente. A Figura 10 apresenta um outro teste realizado com a entidade Adriano VI e as propriedades *nome de nascimento* e *local da morte*.

The figure consists of two screenshots of the Wikidata Query Service interface. The top screenshot shows a query for the birth name of 'Adrian VI'. The query is: `SELECT ?birthName WHERE { SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". } wd:Q57087 wdt:P1477 ?birthName. } LIMIT 100`. The result is 'Adriaan Florenszoon Boeyens d'Edel'. The bottom screenshot shows a query for the place of death of 'Adrian VI'. The query is: `SELECT ?placeOfDeath ?placeOfDeathLabel WHERE { SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". } wd:Q57087 wdt:P20 ?placeOfDeath. }`. The results are 'placeOfDeath: wd:Q220' and 'placeOfDeathLabel: Rome'.

Figura 12 - Teste com a entidade Adriano VI na Wikidata

A pesquisa obteve os seguintes dados: Adriaan Floreszoon Boeyens d'Edel, por nome de nascimento e Roma, por local da morte. De acordo com Cheney (2022), o nome seria Adriaan Florenszoon Dedel.

A entidade Acácio José Correia também foi pesquisada, mas o Wikidata Query Service não encontrou quaisquer resultados. É visível deste exemplo com isso que algumas entidades dos arquivos podem não estar presentes em outros repositórios.

4.2. Testes com DBpedia SPARQL Explorer

O teste seguinte foi realizado com as propriedades da DBpedia, sendo também utilizadas as propriedades da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeia e DBpedia (APÊNDICE F). Assim como com a Wikidata, propriedades mais genéricas foram adicionadas após sugestões dos serviços de consulta durante os testes.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na DBpedia	
foaf:Person/ person	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	foaf:gender foaf:name;birthname foaf:familyName alternativeName birthDate deathDate	description profession birthplace deathplace picture/foaf:imgpouse

Tabela 7 - Fragmento da Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na DBpedia Após Testes no DBpedia SPARQL Explorer (APÊNDICE H).

Logo, a Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na DBpedia Após Testes no DBpedia SPARQL Explorer (APÊNDICE H) foi elaborada para a incorporação dessas outras propriedades. Os testes foram feitos com o DBpedia SPARQL Explorer, a fim de facilitar a sua reprodução. Com a finalidade de comparação, as entidades e propriedades testadas foram as mesmas utilizadas no teste anterior da Wikidata, sendo todas as entidades pertencentes à classe pessoa e as propriedades encontradas no APÊNDICE H, de que se mostra um fragmento na Tabela 7. A Figura 11 mostra o teste realizado utilizando a entidade Mouzinho da Silveira e as propriedades *data de nascimento* e *data da morte*.

SPARQL Explorer for https://dbpedia.org/sparql

SPARQL:

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX dbpedia: <http://dbpedia.org/resource/>
PREFIX dbpedia2: <http://dbpedia.org/property/>
PREFIX dbpedia: <http://dbpedia.org/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?birth WHERE {
  ?person dbo:birthDate ?birth .
  ?person foaf:name "Mouzinho da Silveira"@en .
}
```

Results: Browse Go! Reset

SPARQL results:

birth

"1780-07-12"^^xsd:date

Powered by [OpenLib Vireos](#) and [dbpedia](#)

SELECT ?birth WHERE {
 ?person dbo:birthDate ?birth .
 ?person foaf:name "Mouzinho da Silveira"@en .
}

Results: Browse Go! Reset

SPARQL results:

birth

"1780-07-12"^^xsd:date

SELECT ?death WHERE {
 ?person dbo:deathDate ?death .
 ?person foaf:name "Mouzinho da Silveira"@en .
}

Results: Browse Go! Reset

SPARQL results:

death

"1849-04-04"^^xsd:date

Figura 13 - Teste com a entidade Mouzinho da Silveira na DBpedia

A pesquisa também recuperou as informações 12 de Julho de 1780 e 4 de Abril de 1849, respectivamente. A Figura 12 mostra o teste realizado com a entidade Adriano VI com as propriedades *nome de nascimento* e *local da morte*.

SPARQL Explorer for https://dbpedia.org/sparql

SPARQL:

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdfs: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbpedia: <http://dbpedia.org/property/>
PREFIX dbpedia: <http://dbpedia.org/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?birthName WHERE {
  ?person dbo:birthName ?birthName .
  ?person foaf:name "Adrian VI"@en .
}
```

Results:

SPARQL results:

birthName
"Adriaan Floriszoon Boeyens"@en

Powered by [OpenLink Virtuoso](#) and [dbpedia](#)

```
SELECT ?birthName WHERE {
  ?person dbo:birthName ?birthName .
  ?person foaf:name "Adrian VI"@en .
}
```

```
SELECT ?placeOfDeath WHERE {
  ?person dbo:deathPlace ?placeOfDeath .
  ?person foaf:name "Adrian VI"@en .
}
```

Results:

SPARQL results:

birthName	placeOfDeath
"Adriaan Floriszoon Boeyens"@en	:Papal_States ↗
	:Rome ↗

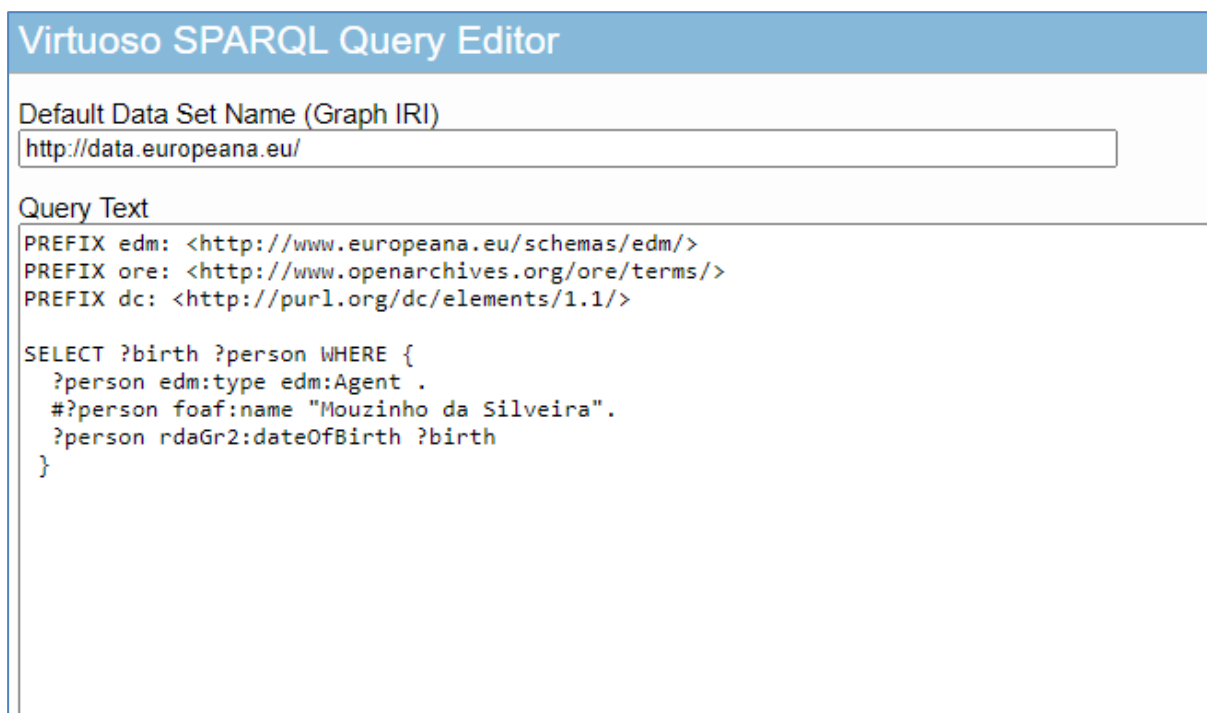
Figura 14 - Teste com a entidade Adriano VI na DBpedia

A pesquisa obteve as seguintes informações: Adriaan Floriszoon Boeyens, por nome de nascimento; Roma e Estados papais, por local da morte. Os Estados papais foram um grupo de territórios no centro da Itália governado pelos papas entre 754 e 1870. No seu apogeu, os Estados papais cobriram a maioria das regiões italianas modernas do Lácio (que incluía Roma), Marcas, Úmbria e Romagna, e partes da Emília (En.Wikipedia.Org. 2022). A propriedade nome de nascimento recuperou um resultado ligeiramente diferente do obtido pelo Wikidata Query Service, mostrando que a entidade pode ser encontrada com o seu nome de nascimento escrito de maneiras distintas.

A entidade Acácio José Correia também foi pesquisada, mas assim como no Wikidata Query Service, não encontrou quaisquer resultados. Dessa forma, fica evidente que nem todas as entidades dos arquivos possuem informações em outros repositórios.

4.3. Testes com Europeana SPARQL API

Diferente das bases de dados da Wikidata e DBpedia que são nativas de dados ligados, a Europeana evoluiu de um esquema simples de metadados para um modelo de integração e enriquecimento de dados agregados em redes semânticas e a sua efetivação como dados abertos ligados. O seu modelo de agregação de metadados para o setor do património cultural já possui informações sobre mais de 3000 instituições ao redor da Europa (Patrício, 2012). Contudo, nem todos esses metadados se encontram em formato de dados abertos ligados. Por esse motivo, não foi possível recuperar informações sobre propriedades de entidades como realizado nas bases de dados anteriores. A Europeana possui um sistema de buscas que faz uso de um SPARQL API, assim como a Wikidata e a DBpedia, sendo alimentado por uma base de dados separada (Virtuoso). A busca foi feita com entidades e propriedades diversas, contudo, nenhum dos testes obteve resultados. A figura 13 mostra um dos testes feitos na plataforma do Virtuoso SPARQL Query Editor.



Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
PREFIX edm: <http://www.europeana.eu/schemas/edm/>
PREFIX ore: <http://www.openarchives.org/ore/terms/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>

SELECT ?birth ?person WHERE {
  ?person edm:type edm:Agent .
  #?person foaf:name "Mouzinho da Silveira".
  ?person rdaGr2:dateOfBirth ?birth
}
```

Figura 15 - Teste com a entidade Mouzinho da Silveira no Virtuoso SPARQL Query Editor

No espaço reservado para buscas no sítio web da Europeana, é possível encontrar documentos digitalizados, imagens e vídeos sobre as entidades Mouzinho da Silveira e Adriano VI, porém não há como filtrar as buscas por dados específicos. Um dos motivos dos testes no Virtuoso SPARQL Query Editor não ter recuperado resultados sobre as propriedades desejadas pode ser a forma de agregação de metadados pela Europeana. Quando uma instituição cultural portuguesa, por exemplo, deseja fornecer conteúdo para ser inserido na base de dados da

Europeana, ela faz isso através do RNOD (Registo Nacional de Objetos Digitais). O RNOD é o agregador de conteúdos bibliográficos digitais e digitalizados, viabilizados em rede por instituições portuguesas, coordenado pela Biblioteca Nacional de Portugal. Ele assegura a integração desses conteúdos na Europeana e em outros portais internacionais. Como agregador nacional para a Europeana, o RNOD garante a normalização dos formatos e metadados de cada fornecedor, conduzindo-os para a Europeana conforme as diretrizes e procedimentos definidos. O RNOD incorpora todos os metadados importados de outros sistemas, especificamente os dados inscritos em registos UNIMARC. Os metadados mínimos para a descrição dos objetos são os seguintes: tipo de documento descrito, data, língua do documento, título e dados de acesso ao objeto digital (RNOD 2022). Por isso, mesmo encontrando as entidades almejadas, não é garantido que haja metadados inseridos sobre determinadas propriedades. Além disso, os metadados nas descrições são para objetos, e não para as entidades referidas nesses objetos, o que leva a uma limitação na interrogação, mesmo que estes metadados estejam representados em dados ligados. A entidade Mouzinho da Silveira só é possível de ser encontrada porque está mencionada em algum metadado de descrição, por exemplo, indexado em texto integral. Nesse caso, como as entidades não são identificadas, isso não transforma essa informação em dados ligados.

Conforme Freire (2021), um piloto de dados ligados desenvolvido como parte do projeto Europeana Common Culture buscou inovar a agregação de metadados na Europeana através de dados ligados. Esse trabalho realizou-se entre Maio de 2019 e Junho de 2020, sendo coordenado pelo Netherlands Institute for Sound and Vision (NISV) em parceria com a Dutch Digital Heritage Network (NDE). O projeto abrangeu fornecedores de dados, agregadores e a Fundação Europeia. Com o intuito de oferecer um conjunto de dados ligados para a Europeana, o trabalho utilizou uma abordagem de agregação de dados ligados baseada em duas especificações: a primeira é que os metadados a nível de conjunto de dados devem ser oferecidos usando vocabulários amplamente utilizados, com os tipos de distribuições de conjunto de dados que os fornecedores podem usar e os metadados primordiais para cada um deles; a segunda especificação refere-se ao uso de dados ligados usando o modelo Schema.org para descrever objectos do património cultural, segundo as exigências da Europeana e do Europeana Data Model (EDM).

No momento, a Europeana apenas suporta a ingestão de metadados na EDM, porém as tentativas de emprego do Schema.org à descrição de metadados de objetos do patrimônio cultural demonstraram que ele pode proporcionar dados de boa qualidade, ou seja, é possível verificar a sua origem, como foram produzidos, armazenados e manipulados. Esta segunda

especificação orienta o uso de metadados Schema.org, que depois de serem convertidos para EDM, terão metadados capazes de serem agregados pela Europeana. Da mesma forma, este piloto desenvolveu ferramentas para a agregação de dados ligados com o objetivo de serem usados por agregadores da Europeana e de redes similares. As ferramentas criadas são fundamentadas em "container"Docker que protegem a independência técnica das ferramentas, podendo ser utilizadas em diferentes ambientes, em pequenas ou grandes coleções. As ferramentas desenvolvidas e o seu código se encontram acessíveis em Github (Freire 2021).

Contudo, os desafios técnicos durante o projeto foram maiores do que o esperado. Nem todos os fornecedores conseguiram oferecer um conjunto de dados como dados ligados, alguns forneceram conjuntos com dados insuficientes para o processo de agregação na Europeana. Desta experiência é possível constatar que, mesmo com os fornecedores demonstrando interesse na implementação de dados ligados, é necessário um nível de recursos considerável quando não há experiência anterior nesse processo por parte das organizações fornecedoras de dados. Em 2021, um novo trabalho começou a ser desenvolvido, coordenado pelo The Netherlands Institute for Sound and Vision, focando-se nos componentes para a interoperabilidade e integração do conjunto de ferramentas nos sistemas dos agregadores (Freire 2021).

5. Conclusões e perspectivas de desenvolvimento

Com a evolução dos objetos de arquivo, a demanda pela inclusão de metadados digitais e por requisitos de interoperabilidade cresceu nos repositórios de patrimônio, havendo assim a necessidade de uma nova ferramenta de descrição de metadados, mais flexível e processável por máquinas. O intuito é tornar os arquivos mais inclusivos, acessíveis, e ligados ao resto do mundo. Este trabalho tinha como objetivo expandir os registos arquivísticos com abordagens de questionamentos mais amplas e variadas, tendo como foco a recuperação da informação pelo utilizador, o apoio à descrição pelos profissionais de arquivo, além da divulgação de informação sobre o património cultural.

Para isso, foram analisados 25 registos dos Arquivos Nacionais Portugueses, escolhidos pelos arquivistas para entidades. Assim, considerando os elementos mais descritivos na descrição arquivística e a informação textual destes elementos, foram escolhidas as seguintes classes de entidades para serem expandidas: *person, place, product, vessel, organization, legal system, activity, project, thing, document*, e *event*.

Após a selecção das entidades dos registos, foram pesquisadas fontes externas para o enriquecimento de dados e a interligação, tendo sido escolhidas as bases de dados da Wikidata, DBpedia e Europeana. Com base nas entidades dos registos, foram investigadas centenas de propriedades diretamente relacionadas com as classes nas respectivas bases de dados. Com o propósito de filtrar estas propriedades, foram realizadas buscas adicionais para as propriedades que tinham mais dados introduzidos nas bases de dados de acordo com cada classe. Em seguida, essas entidades e propriedades foram submetidas a testes utilizando o Wikidata Query Service, o DBpedia SPARQL Explorer e o Virtuoso SPARQL Query Editor.

As pesquisas realizadas com as entidades no Wikidata Query Service e no DBpedia SPARQL Explorer apenas forneceram resultados para determinadas propriedades, uma vez que nem todas as entidades possuem dados na base de dados. Por outro lado, as propriedades testadas provaram possibilitar dados relevantes sobre as entidades, fornecendo informação adicional e outras entidades com que se relacionam. Contudo, com a Europeana as pesquisas com as entidades não se mostraram produtivas, não tendo sido possível recuperar informações sobre propriedades de entidades como realizado nas bases de dados anteriores. Mesmo possuindo um sistema de buscas que faz uso de um SPARQL API, assim como a Wikidata e a DBpedia, nenhum dos testes obteve resultados. Um dos motivos pode ser a forma de agregação de metadados pela Europeana, já que ela utiliza metadados das informações originais. Por isso, mesmo encontrando as entidades almejadas, não é garantido que haja metadados inseridos sobre determinadas propriedades. A Europeana é uma grande agregadora de metadados,

porém esses metadados ainda se encontram em processo de conversão em dados abertos ligados.

Uma das preocupações na etapa da procura por fontes externas para o enriquecimento dos metadados, foi com a credibilidade das bases de dados, principalmente com as informações da Wikidata e DBpedia. A informação textual proveniente da Wikipedia é transformada em informação estruturada na Wikidata, embora essa também possua outras fontes de contribuição. Por isso, primeiramente, é importante a validação dessas informações. Por causa de alegações a respeito dos conteúdos postados, a Wikipédia precisou reconsiderar os seus procedimentos, exigindo cadastro dos utilizadores na página antes que eles pudessem fazer edições nos verbetes. Além de também utilizar critérios mais rígidos e seletivos, que vieram a aumentar o controle de qualidade dos conteúdos publicados. A colaboração na Wikipédia é controlada por procedimentos técnicos com parâmetros de classificação dos artigos, o que facilita a identificação no que diz respeito à credibilidade do artigo, além de aspectos como cumprimento a direitos autorais, impropriedades intelectuais, inconsistência e inconclusividade dos temas publicados (Vieira e Christofolletti 2017). Já que as informações da Europeana provêm diretamente de instituições culturais, a questão da credibilidade não foi uma preocupação.

Um outro aspecto que mereceu a atenção ao longo dos testes com as propriedades da Wikidata e DBpedia foi o do controle de autoridade. O controle de autoridade é fundamental para identificar pessoas ou outros conceitos de forma inequívoca e, dessa forma, facilitar a pesquisa. Contudo, alguns valores de propriedades podem gerar ambiguidade, dando margem a erros. Um exemplo é a propriedade GivenName (que serve para especificar o primeiro nome ou outro nome próprio de uma pessoa) e a propriedade BirthName (que especifica o nome completo de uma pessoa à nascença, se diferente do seu nome atual geralmente utilizado). Logo, é necessário estar atento aos metadados procedentes de fontes externas.

Em conclusão, o trabalho se mostrou bastante desafiador e, por razões de limitação de tempo e de coordenação com as tarefas do projeto EPISA, não prosseguiu para o desenho de uma interface capaz de realizar testes com os arquivistas e utilizadores comuns, por isso, esse objetivo inicialmente previsto não foi prosseguido. Por conseguinte, esse trabalho requer continuação para uma posterior validação com arquivistas, assim como avaliação da interface pelos utilizadores.

Este trabalho abre promissoras perspectivas de trabalhos futuros, tanto no que diz respeito ao enriquecimento dos registos de arquivo tendo em mente a recuperação da informação pelo utilizador, como na integração dos arquivos na rede mundial de dados ligados e semânticos e

na sugestão de interfaces que possibilitem uma melhor experiência com utilizadores diversos. Além disso, os testes feitos com a Europeana SPARQL API mostraram a necessidade de mais estudos com o intuito de proporcionar a agregação de metadados na Europeana através de dados ligados, promovendo a interoperabilidade e a integração de ferramentas já existentes nos sistemas dos agregadores.

6. Referências bibliográficas

- Aliprantis, John, Eirini Kalatha, Markos Konstantakis, Kostas Michalakis, e George Caridakis. 2018. "Linked Open Data as Universal Markers for Mobile Augmented Reality Applications in Cultural Heritage." Em *Digital Cultural Heritage*. Ioannides, M. (eds). Lecture Notes in Computer Science (), vol 10605. Springer, Cham. https://doi.org/10.1007/978-3-319-75826-8_7.
- Aydin, B., et al. 2013. "Extending Augmented Reality Mobile Application with Structured Knowledge from the LOD Cloud." Conference: Proceedings of the 3rd International Workshop on Information Management for Mobile Applications. IMMoA. Riva del Garda, Italia. <https://dbis.rwth-aachen.de/IMMoA2013/>.
- Arquivo Nacional Torre Do Tombo. 2022. "Reprodução Arquivo Nacional Torre Do Tombo". <https://antt.dglab.gov.pt/servicos/reprografia/>.
- Barbedo, Francisco. 2011. "Nova Versão DigitArq". *Boletim DGARQ Direcção-Geral de Arquivos*, 17 de Abril- Junho de 2011. <https://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2014/01/DGARqBolt-17.pdf>.
- Berners-lee, T., Ora Lassila e James Hendler. 2001. "The Semantic Web: A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities." *Scientific American*, 1 de Maio de 2001. <https://www.scientificamerican.com/article/the-semantic-web/>.
- Bizer, C., T. Heath e T. Berners-Lee. 2009. 2009. "Linked data – The Story So Far". *International Journal on Semantic Web & Information Systems*, p. 9. <https://doi.org/10.4018/jswis.2009081901/>.
- Chanhom, Weeraphan e Chutiporn Anutariya. 2019. "TOMS: A Linked Open Data System for Collaboration and Distribution of Cultural Heritage Artifact Collections of National Museums in Thailand". *New Generation Computing* 37 (4): 479-498. <https://doi.org/10.1007/s00354-019-00063-1/>.
- Cheney, David. 2022. "Pope Hadrian VI (Adriaan Florenszoon Dedel)". *Catholic-Hierarchy.Org*. <http://www.catholic-hierarchy.org/bishop/bflo.html>.
- CIDOC CRM. 2022. "What is the CIDOC CRM?" <https://www.cidoc-crm.org/>.
- Coneglian, C. S. e J. E. Santarem Segundo. 2017. "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*, 22, n. 48 (janeiro/abril) p. 88-99, Florianópolis. ISSN 1518-2924. <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n48p88>.
- da Silva, J.R., C. Ribeiro e J.C. Lopes. 2019. "Ranking Dublin Core descriptor lists from user interactions: a case study with Dublin Core Terms using the Dendro platform". *Int J Digit Libr* 20, 185–204. <https://doi.org/10.1007/s00799-018-0238-x>.
- DBpedia Association. 2022. "About DBpedia - DBpedia Association". <https://www.dbpedia.org/about/>.

- Doerr, M., S. Gradmann, S. Henricke, A. Isaac, C. Meghini e H. Sompel. 2010. "The Europeana Data Model (EDM)." World Library and Information Congress: 76th IFLA General Conference and Assembly, 10-15, Gothenburg, Suécia, 10 de agosto de 2010. https://www.researchgate.net/publication/303058300_The_Europeana_Data_Model_EDM.
- Dublincore.Org. 2022. "Dublin Core Metadata Innovation". <https://www.dublincore.org/>.
- EDM PRIMER. 2013. "Europeana Data Model Primer." Editado por Antoine Isaac. <http://pro.europeana.eu/edm-documentation>.
- Evans, Emma. 2019. Semantic Interoperability of Moving Image Records in Aggregating Digital Repositories: A Comparative Study of Dublin Core and Pbc Core Descriptive Metadata Schemas. <https://doi.org/10.17615/qfn5-qr47>.
- Ferros, Luís Miguel, Susana Ramalho Filipe e J. C. Ramalho. 2010. "DigitArq e o Novo Módulo de Interoperabilidade OAI-PMH." Em *Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas*, 10, Guimarães, Portugal, 2010 – "Políticas de informação na sociedade em rede : actas" [CD-ROM]. [S.l.]: APBAD, 2010. ISBN 978-972-9067-39-6. <https://hdl.handle.net/1822/10532>.
- En.Wikipedia.Org. 2022. "Papal States - Wikipedia". https://en.wikipedia.org/wiki/Papal_States.
- Ferreira, Miguel João. 2020. "Descrição da Informação: 2. Metadados, Meta-Informação e Meta-Confusões." <https://www.linkedin.com/pulse/descr%C3%A7%C3%A3o-da-informa%C3%A7%C3%A3o-2-metadados-meta-informa%C3%A7%C3%A3o-e-ferreira/?originalSubdomain=pt>.
- Freire, Nuno. 2021. "Innovating Metadata Aggregation in Europeana Via Linked Data". <https://pro.europeana.eu/post/innovating-metadata-aggregation-in-europeana-via-linked-data>.
- Freire, Nuno, René Voorburg, Roland Cornelissen, Sjors de Valk, Enno Meijers, e Antoine Isaac. 2019. "Aggregation of Linked Data in the Cultural Heritage Domain: A Case Study in the Europeana Network". *Information* 10, (agosto): 252. <https://doi.org/10.3390/info10080252/>.
- Freire, Nuno, G. Robson, J.B. Howard et al. 2020. "Cultural heritage metadata aggregation using web technologies: IIIF, Sitemaps and Schema.org". *Int J Digit Libr* 21, 19–30. <https://doi.org/10.1007/s00799-018-0259-5>.
- Germanisches Nationalmuseum. 2022. "Wisski Germanisches Nationalmuseum". <https://www.gnm.de/en/research/archive-research-projects/wisski/>.
- Gil, A. C. 2008. *Métodos e Técnicas de Pesquisa Social*. 6º ed. São Paulo: Editoras Atlas S A, p. 27.
- Hyvönen, Eero, Eetu Mäkelä, Mirva Salminen, Arttu Valo, Kim Viljanen, Samppa Saarela, Miikka Junnila, e Suvi Kettula. 2005. "Museumfinland—Finnish Museums on the Semantic Web". *Journal Of Web Semantics* 3 (2-3): 224-241. <https://doi.org/10.1016/j.websem.2005.05.008/>.

- INESC TEC. 2022. "Entity and Property Inference for Semantic Archives - INESC TEC". <https://www.inesctec.pt/en/projects/episa#about>.
- Isotani, Seiji e Ig Ibert Bittencourt. 2015. *Dados Abertos Conectados: em Busca da Web do Conhecimento*. São Paulo: Novatec.
- Jefcoate, Graham. 2006. "Gabriel: Gateway to Europe's National Libraries." *Program: electronic library and information systems* 40, n.4 (outubro): 325-333. <https://doi.org/10.1108/00330330610707908/>.
- Koivune, Marja-Riitta e Eric Miller. 2001. "W3C Semantic Web Activity". Semantic Web Kick-off Seminar, Finlândia, 2 de novembro de 2001. <http://www.w3.org/2001/12/semweb-fin/w3csw>.
- Laney, D. 2001. "3D Data Management: Controlling Data Volume, Velocity and Variety." *META Group Research Note*. <http://gtnr.it/1bKfLKH>.
- Lima, Júnio César e Cedric L. Carvalho. 2004. *Uma Visão da Web Semântica*. [S.l.]: Technical Report INF_001/94, Instituto de Informática-Universidade Federal de Goiás. https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-04.pdf.
- Lod-Cloud.Net. 2022. "The Linked Open Data Cloud". <https://lod-cloud.net/>.
- Lopes, Yayiane dos Santos de Freitas. 2017. "Ontologia Como Interface de Apresentação de Resultados de Busca: Uma Proposta Baseada no Modelo Espaço Vetorial". Dissertação de Pós-Graduação, Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP. https://repositorio.unesp.br/bitstream/handle/11449/151715/lopes_tsf_me_mar.pdf?sequence=3.
- Marden, J., et al. 2013. "Linked Open Data for Cultural Heritage: Evolution of an Information Technology." SIGDOC '13: Proceedings of the 31st ACM International Conference on Design of Communication, Greenville Carolina do Norte, EUA, setembro de 2013, pages 107–112. <https://doi.org/10.1145/2507065.2507103>.
- Minichiello, Alexandre. 2017. "Protótipo De Aplicação Web Para Dados Abertos Conectados". Trabalho de Conclusão de Curso (TCC). Tecnólogo em Gestão da Tecnologia da Informação, Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina. https://repositorio.ifsc.edu.br/bitstream/handle/123456789/670/TCCAlexandre_Andrade_Minihiello-GTI-2017.pdf?sequence=1&isAllowed=y.
- Morais, Edison Andrade Martins e Ana Paula L. Ambrósio. 2007. *Ontologias: Conceitos, Usos, Tipos, Metodologias, Ferramentas e Linguagens*. Universidade Federal de Goiás. https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-07.pdf.
- NISO - National Information Standard Organization. 2004. "Understanding Metadata". <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.
- Patrício, H. S. 2012. "A Europeana e a Agregação de Metadados na Web: Análise dos Esquemas ESE/EDM e da Aplicação de Standards da Web Semântica a Dados de Bibliotecas." Actas do Congresso Nacional de Bibliotecários, Arquivistas e

- Documentalistas, 11., Lisboa. Anais... Lisboa, outubro de 2012.
<https://doi.org/10.5195/biblios.2020.759>.
- Pillai, S.G., LK. Soon e SC. Haw. 2019. "Comparing DBpedia, Wikidata, and YAGO for Web Information Retrieval." Em *Intelligent and Interactive Computing*. Piuri, V., Balas, V., Borah, S., Syed Ahmad, S. (eds). Lecture Notes in Networks and Systems. vol 67. Springer, Singapore. https://doi.org/10.1007/978-981-13-6031-2_40.
- RNOD - Registo Nacional de Objectos Digitais. 2022. "Contribuições para Europeana".
<https://rnod.bnportugal.gov.pt/rnod/winlib.aspx?skey=50FD87AD135F4826AA62F4F1228A88&option=generic2>.
- Sayão, Luís Fernando. 2010. "Uma Outra Face dos Metadados: Informações para a Gestão da Preservação Digital." *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação* 15, no. 30:1-31. <https://www.redalyc.org/articulo.oa?id=147169190002>.
- Schema.Org. 2022. "Organization of Schemas". <https://schema.org/>.
- Schiessl, Marcelo. 2015. "Lexicalização de Ontologias: O Relacionamento Entre Conteúdo e Significado no Contexto da Recuperação da Informação." Tese de Doutorado – Doutorado em Ciência da Informação, Universidade de Brasília.
<https://repositorio.unb.br/handle/10482/18663>.
- Semantic Computing Research Group (SeCo). 2022. "Linked Open Data Infrastructure for Digital Humanities in Finland (LODI4DH)". <https://seco.cs.aalto.fi/projects/lodi4dh/>.
- Silva, Luciana Candida, José Eduardo Santarem Segundo e Nuno Freire. 2021. "Web Semântica na Europeana: Uma Abordagem do Padrão EDM". *Biblios: Journal of Librarianship and Information Science*, no. 80: 15-26.
<https://doi.org/10.5195/biblios.2020.759>.
- Swiss Art Research Infrastructure. 2022. "SARI – Swiss Art Research Infrastructure".
<https://swissartresearch.net/>.
- Vieira, Marli Vick e Rogério Christofolletti. 2017. "Confiabilidade no Uso da Wikipédia Como Fonte de Pesquisa Escolar". *Revista On Line de Política e Gestão Educacional*, nº 15 (fevereiro). Araraquara. <https://doi.org/10.22633/rpge.voi15.9351>.
- Webmuseu. 2022. "Recursos Web para GLAM". <https://webmuseu.org/recursos/conteudo-digital-codigo-exemplo-de-rdf-da-dbpedia/>.
- Wikidata.Org. 2022. "Wikidata". https://www.wikidata.org/wiki/Wikidata:Main_Page.
- Yuji T. e P. Jung-ran. 2013. "RDA: Resource Description and Access – A Survey of the Current State of the Art." *Journal of the American Society for Information Science and Technology* 64(4): 651–662. 20 (fevereiro). <https://doi.org/10.1002/asi.22825>.

Apêndices

APÊNDICE A – Tabela de Entidades e Propriedades Associadas aos 25 Registos no CIDOC CRM, Schema.org e Dublin Core

Nº de Registo	Entidades	CIDOC CRM	Schema.org	Dublin Core
Registo 1	atanados; salsaparrilha; pau de tatajuba; sola; tabuado; vinhático; vaquetas; pano de cafre; caçoeiras de jacarandá; tabuado tapainhã;qui na; corveta; bergantim; galera; escuna; brigue; iate; caíque; nau; sumaca; corsário; polaca; Açores; Angola; Ilha do Príncipe; Baía; Benguela; Maranhão; Bengala; Bombaim; Cabo Verde; Ceilão; China; Costa da Mina; Costa de Coromandel; Damão; Goa; Guiné Bissau; Índia; Macau; Madeira; Mangalor; Pará;	P14 carried out by (performed); P16 used specific object (was used for); P19 was intended use of (was made for); P21 had general purpose (was purpose of); P45 consists of (is incorporated in); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P101 had as general use (was use of); P103 was intended for (was intention of); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; name; category (physicalActivity or product); countryOfLastPro cessing; countryOfOrigin; material(product) ; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNu mber; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; alternative; bibliographicCitation; coverage; description; title; type; physicalObject; spatialCoverage; location; memberOf.

	Pernambuco; Rio de Janeiro; Surrate; Galiza; Brasil; Félix José da Costa; D. Francisca Xavier; Manuel de Almeida.			
Registo 2	Acácio José Correia.	P14 carried out by (performed); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P152 has parent (is parent of).	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; memberOf.
Registo 3	Mouzinho da Silveira; Francisco da Silva Braga; João Caetano da Silva Pereira; José Gomes da Cruz; Joaquim Manuel da Silva; Bento Gualdino da Silva Valadares; José Miguel Guilherme.	P14 carried out by (performed); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P152 has parent (is parent of).	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; memberOf.
Registo 4	Lino José de Almeida Lobo da Torre do Vale; Lisboa.	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought	description; sameAs; containedInPlace; containsPlace; geo; hasMap; latitude; longitude; alternateName;	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.

		into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	
Registo 5	Oriente.	P89 falls within (contains); P121 overlaps with; P122 borders with.	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location.
Registo 6	X	X	X	X
Registo 7	José Borges Leal; Dionísio José; Jacinto Borges Leal; Mariana de Jesus; Joaquim Machado; Manuel Leal; Rita Mariana; Inácio Mateus Leal; Bento Gualdino da Silva Valadares; Vila da Praia - Ilha Terceira (Açores); Estado da Índia (Índia Portuguesa).	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.
Registo 8	X	X	X	X

Registo 9	Vasco Salter Mascarenhas e Mendonça; João Manuel da Costa; Joaquim José da Silva Santos; ordem de cristo; precatório.	P11 had participant (participated in); P14 carried out by (performed); P20 had specific purpose (was purpose of); P21 had general purpose (was purpose of); P43 has dimension (is dimension of); P44 has condition (is condition of); P45 consists of (is incorporated in); P95 has formed (was formed by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P101 had as general use (was use of); P102 has title (is title of); P104 is subject to (applies to); P105 right held by (has right on); P107 has current or former member (is current or former member of); P151 was formed from; P152 has parent (is parent of).	description; name; potentialAction; about; category (Thing); sameAs; alternateName; member; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; coverage; date created; creator; description; subject; title; licenseDocument; Jurisdiction; PhysicalObject; bibliographicCitation; alternative; jurisdiction; locationPeriodOrJurisdiction; PeriodOfTime; memberOf.
Registo 10	Pernambuco (Brasil); António Henriques Leal; António Esteves Costa.	P14 carried out by (performed); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P89 falls within (contains); P100 was death of (died in); P121	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation;	abstract; bibliographicCitation; coverage; description; memberOf; spatialCoverage; location.

		overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	memberOf; nationality; parent; relatedTo; sibling; spouse; employee; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber.	
Registo 11	Louça preta; D. Maria Luísa de Brito Rocha e Melo; D. Helena de Brito Rocha e Melo; D. Ana da Cunha e Meneses Pinto Cardoso; D. Maria Luísa Barroso da Câmara; Sr. Francisco Queirós de Andrade Pinto; D. Francisca Seabra de Lacerda; D. Maria de Brito Rocha e Melo; D. Maria Emília Cabral da Silva.	P14 carried out by (performed); P19 was intended use of (was made for); P32 used general technique (was technique of); P33 used specific technique (was used by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P102 has title (is title of); P125 used object of type (was type of object used in); P152 has parent (is parent of).	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee; category (PhysicalActivity or Product); countryOfLastProcessing; countryOfOrigin; material(product) .	abstract; bibliographicCitation; coverage; description; memberOf; alternative; title; type; PhysicalObject.
Registo 12	Convento de São Bernardo (=Mosteiro de Nossa Senhora da Conceição); Ordem de Cister; D. Jorge de Melo; Serra de S.	P11 had participant (participated in); P14 carried out by (performed); P21 had general purpose (was purpose of); P89 falls within (contains); P95 has formed (was	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; potentialAction;	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; date created; creator; subject; title; alternative; jurisdiction; locationPeriodOrJurisdiction;

	Mamede; Portalegre; D. João III; Santarém; D. André de Noronha; Nicolau Chanterene; Abadia de Claraval.	formed by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P102 has title (is title of); P107 has current or former member (is current or former member of); P121 overlaps with; P122 borders with; P151 was formed from; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	about; member; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	periodOfTime; memberOf.
Registo 13	Sines; Santiago do Cacém; Vila Nova de Santo André (Santo André).	P89 falls within (contains); P121 overlaps with; P122 borders with; P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap; latitude; longitude; alternateName; globalLocationNumber.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative.
Registo 14	Antão Santos da Cunha; Pedro Santos da Cunha.	P14 carried out by (performed); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P152 has parent (is parent of).	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation;	abstract; bibliographicCitation; coverage; description; memberOf.

			memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	
Registo 15	programa ADAI.	P3 has note; P107 has current or former member (is current or former member of); P11 had participant (participated in); P67 refers to (is referred to by); P101 had as general use (was use of); P104 is subject to (applies to); P105 right held by (has right on); P106 is composed of (forms part of); P129 is about (is subject of); P148 has component (is component of); P149 is identified by (identifies); P151 was formed from.	description; name; subjectOf; about; founder; foundingDate; foundingLocation ; member; sponsor; suborganization; creator; legislationPassed By; legislationRespon sible; parentOrganizati on; potentialAction.	abstract, contributor, coverage, created, creator, date, description; subject; jurisdiction; periodOfTime; policy.
Registo 16	aquário Vasco da Gama; Dafundo (Oeiras).	P89 falls within (contains); P121 overlaps with; P122 borders with; P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNu mber.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative.
Registo 17	José Cândido Marques de Azevedo; Vila da Feira; José Alves Domingues; Silvalde.	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude;	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.

		for); P98 brought into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	
Registo 18	Convento das Necessidades ; António dos Santos Lucas; Joaquim Freire dos Santos Callado; Ludgero Maria de Lima e Quina; João d' Oliveira e Sousa.	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.
Registo 19	João Teixeira Albernaz I; carta náutica.	P14 carried out by (performed); P43 has dimension (is dimension of); P44 has condition (is condition of); P45 consists of (is incorporated in); P65 shows visual item (is shown	description; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation;	abstract; bibliographicCitation; coverage; description; memberOf; created; creator; date; format; hasFormat; temporal; subject; title; physicalResource; physicalObject.

		by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P101 had as general use (was use of); P152 has parent (is parent of).	homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee; image; subjectOf; about; object; itemCondition; material; size; weight.	
Registo 20	bula (pontifícia); Papa Adriano VI; Infante D. Henrique; Mosteiro de São Cristóvão de Lafões; Mosteiro de São Jorge; Diogo da Gama.	P14 carried out by (performed); P43 has dimension (is dimension of); P44 has condition (is condition of); P45 consists of (is incorporated in); P65 shows visual item (is shown by); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P101 had as general use (was use of); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap; latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee; size; weight; image; subjectOf; about; object; itemCondition; material.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf; created; creator; date; format; hasFormat; temporal; subject; title; physicalResource; physicalObject.
Registo 21	Jerónima da Cruz; Sevilha;	P7 took place at (witnessed); P8 took place on or	description; sameAs; containedInPlace;	abstract; bibliographicCitation; coverage; description;

	Évora; auto-de-fé; tribunal do santo ofício.	within (witnessed); P9 consists of (forms part of); P11 had participant (participated in); P12 occurred in the presence of (was present at); P14 carried out by (performed); P20 had specific purpose (was purpose of); P21 had general purpose (was purpose of); P89 falls within (contains); P92 brought into existence (was brought into existence by); P93 took out of existence (was taken out of existence by); P95 has formed (was formed by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P102 has title (is title of); P107 has current or former member (is current or former member of); P121 overlaps with; P122 borders with; P151 was formed from; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172	containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee; potentialAction; about; member; actor; audience; contributor; location; organizer.	spatialCoverage; location; alternative; memberOf; date created; created; creator; subject; title; jurisdiction; locationPeriodOrJurisdiction; PeriodOfTime; audience; contributor; spatial; relation; temporal; event.
--	--	--	--	---

		contains; P189 approximates (is approximated by).		
Registo 22	D. João II; D. Manuel I; ordenações manuelinas; D. João III; D. Sebastião; D. João de Melo; rei D. Filipe I; ordenações filipinas; Filipe II; D. João IV; Corte, Estremadura e Ilhas; Minho e Trás-os-Montes; Beira; Alentejo; Algarve; José Anastácio Guerreiro; terramoto (1 de Novembro de 1755); Belém; Aveiro; D. José de Mascarenhas de Lencastre; Ferragial; Senhor Jesus da Boa Morte; Lisboa; Terreiro do Paço.	P7 took place at (witnessed); P8 took place on or within (witnessed); P12 occurred in the presence of (was present at); P13 destroyed (was destroyed by); P14 carried out by (performed); P43 has dimension (is dimension of); P44 has condition (is condition of); P45 consists of (is incorporated in); P46 is composed of (forms part of); P70 documents (is documented in); P72 has language (is language of); P73 has translation (is translation of); P89 falls within (contains); P92 brought into existence (was brought into existence by); P93 took out of existence (was taken out of existence by); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P101 had as general use (was use of); P102 has title (is title of); P103 was	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee; image; about; object; itemCondition; material; size; weight; audience; endDate; location; startDate.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf; contributor; creator; created; date; format; hasFormat; language; subject; title; locationPeriodOrJurisdiction; physicalResource; image; physicalObject; text; audience; temporal; spatial; periodOfTime; SizeOrDuration; event.

		intended for (was intention of); P104 is subject to (applies to); P105 right held by (has right on); P121 overlaps with; P122 borders with; P128 carries (is carried by); P149 is identified by (identifies); P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).		
Registo 23	Manuel Joaquim Ferreira Lino; Sta Isabel, Olival (Vila Nova de Gaia).	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.
Registo 24	batismos.	P11 had participant (participated in); P12 occurred in the presence of	description; name; about; actor; attendee; audience; duration;	abstract; audience; date; description; spatial; temporal; subject; title; event.

		(was present at); P86 falls within (contains); P92 brought into existence (was brought into existence by); P98 brought into life (was born).	endDate; location; startdate.	
Registo 25	arquivo da Universidade de Coimbra; João Monteiro.	P14 carried out by (performed); P89 falls within (contains); P96 by mother (gave birth); P97 from father (was father for); P98 brought into life (was born); P100 was death of (died in); P121 overlaps with; P122 borders with; P152 has parent (is parent of); P168 place is defined by (defines place); P171 at some place within; P172 contains; P189 approximates (is approximated by).	description; sameAs; containedInPlace; containsPlace; geo; hasMap;latitude; longitude; alternateName; globalLocationNumber; name; affiliation; birthDate; birthPlace; deathDate; deathPlace; familyName; gender; givenName; hasOccupation; homeLocation; memberOf; nationality; parent; relatedTo; sibling; spouse; employee.	abstract; bibliographicCitation; coverage; description; spatialCoverage; location; alternative; memberOf.

APÊNDICE B - Lista de Classes das Entidades e Propriedades Encontradas Associadas Apenas ao Registo 1 na Wikidata, Europeia e DBpedia.

Registo 1

Wikidata:

person: Language; label; description; also known as; subclass of (person or organization; individual; entity whose item has the given name property; legal person); image; pronunciation audio (language of work or name); has quality (personal life; pseudonym; date of birth; marital status; person property; age of a person; personhood; given name; gender; family name; personality; personal identity; personal infrastructure; mental state; adaptive capacity; biographic information); opposite of (social consciousness; society); on focus list of Wikimedia project (WikiProject Human rights); category for eponymous categories (category:Wikipedia categories named after people); female form of label; male form of label; NCI Thesaurus ID; equivalent class; different from (personhood; legal personhood; human; fictional person; fulbe people); said to be the same as (natural person); commons category (people); topic's main category (category:People); wikidata property (significant person); wikidata usage instructions; identifiers.

geographic location: Language; label; description; also known as; instance of (property); subclass of (location); part of (Earth's surface); named after (locus; relative location); studied by (geography); has quality (geographic entity); OpenStreetMap tag or key (Key:location); equivalent class; properties for this type (coordinate location; Wi-Fi access; associated hazard; country; located in the administrative territorial entity; elevation above sea level); different from (geographic region; geographical feature; location); commons category (Places); topic's main template; topic's main category (Category:Places); Wikidata property; identifiers.

natural product: instance of (group or class of chemical substances); subclass of (chemical substance; product; primary raw material); different from (natural material; biomolecule); commons category (natural products); topic's main category (Category:Natural products); Wikidata property (this taxon is source of; natural product of taxon; found in taxon); identifiers.

primary raw material: subclass of (raw material); part of (natural environment); image; opposite of (manufactured good); described by source; commons category (materials); topic's main category; identifiers.

product: subclass of (goods and services; product; good); image; has quality (function; requirement; product life-cycle; quality; design; specification; product naming; product innovation; product differentiation; functionality); OpenStreetMap tag or key; equivalent class; properties for this type (brand; inception; transmitted signal type; received signal type; consumption rate per capita; consumption rate; carbon footprint); different from (tangible good; goods and services); said to be the same as (product; goods and services); commons category (Products); topic's main category (Category:Products); Wikidata property (GS1 Global Product Classification brick code; product or material produced); identifiers.

ship: instance of (vessel class); subclass of (watercraft; vessel); image; pronunciation audio; has use (transport); has quality (ship type); described by source; on focus list of Wikimedia project (Wikipedia:Vital articles); maintained by WikiProject (WikiProject Ships); category for the view of the item (Category:Views of ships); small logo or icon; OpenStreetMap tag or key;

Unicode character; equivalent class; has part or parts (hull; ship element; bridge); properties for this type (gross tonnage; operator; IMO ship number; MMSI; image; beam; home port; port of registry; vessel class; manufacturer; location of creation; yard number; pennant number; length; significant event; conflict; draft; width; net tonnage; has part or parts; number of masts; owned by; service entry; service retirement); different from (boat); commons gallery (ships); topic's main template; topic's main category (Category:Ships); identifiers.

Europeana:

edm:Agent: skos:prefLabel; skos:altLabel skos:hiddenLabel; skos:note (for e.g., biographical notes); dc:date; dcterms:hasPart; dc:identifier; dcterms:isPartOf; edm:begin; edm:end; edm:hasMet; edm:isRelatedTo (for generic relations to other agents, especially); edm:wasPresentAt; foaf:name; rdaGr2:biographicalInformation; rdaGr2:dateOfBirth; rdaGr2:dateOfDeath; rdaGr2:placeOfBirth; rdaGr2:placeOfDeath; rdaGr2:dateOfEstablishment; rdaGr2:dateOfTermination; rdaGr2:gender; rdaGr2:professionOrOccupation; owl:sameAs.

edm:Place: wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:hiddenLabel; skos:note; dcterms:hasPart; dcterms:isPartOf; edm:isNextInSequence; owl:sameAs.

edm:PhysicalThing (as contextual entities): dc:contributor; dc:coverage; dc:creator; dc:date; dc:description; dc:format; dc:identifier; dc:language; dc:publisher; dc:relation; dc:rights; dc:source; dc:subject; dc:title; dc:type; dcterms:alternative; dcterms:conformsTo; dcterms:created; dcterms:extent; dcterms:hasFormat; dcterms:hasPart; dcterms:hasVersion; dcterms:isFormatOf; dcterms:isPartOf; dcterms:isReferencedBy; dcterms:isReplacedBy; dcterms:isRequiredBy; dcterms:issued; dcterms:isVersionOf; dcterms:medium; dcterms:provenance; dcterms:references; dcterms:replaces; dcterms:requires; dcterms:spatial; dcterms:tableOfContents; dcterms:temporal; edm:currentLocation; edm:hasMet; edm:hasType; edm:incorporates; edm:isDerivativeOf; edm:isNextInSequence; edm:isRelatedTo; edm:isRepresentationOf; edm:isSimilarTo; edm:isSuccessorOf; edm:realizes; edm:wasPresentAt.

DBpedia:

foaf:Person: abbreviation; abstract; access; accessDate; activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation; ageRange; agglomerationPopulationTotal; aitaCode; album; alias; alongside; alternativeName; animal; anthem; apcPresident; appearance; appointer; area; areaMetro; asWikiText; ascent; associate; associateStar; associatedBand; associatedMusicalArtist; astrazenca; astrazencaCumul; atcCode; atcPrefix; atcSuffix; authority; autonomy; average; averageSpeed; award; awayColourHexCode; background; battery; battle; bibsysId; biome; blazon; blazonLink; blazonRatio; bnfId; book; border; bpmId; budget; builder; buildingEndDate; buildingStartDate; calculationNeeds; callSign; capacity; casNumber; casSupplemental; category; chain; chairLabel; chairman; chancellor; channel; chiefPlace; city; class; classification; closingDate; closingYear; coach; coachedTeam; coatOfArms; code; collectionSize; colour; colourHexCode; colourName; commandant; commander; comment; committee; commonName; company; conflict; connotation; constituencyDistrict; cosparId; cost; country; county; courseArea; cpu; creationYear; creator; crest; criteria; cuisine; currency; currentCity; currentStatus; dailyVaccinationsPerMillion; dailyVaccinationsRaw; date; dateBudget; dateOfAbandonment; dc:creator; dc:identifier; dc:language; dc:rights; dc:subject; dc:title; dc:type; dct:created; dct:dateSubmitted; dct:description; dct:extent; dct:format; dct:medium; dct:modified; dct:spatial; dct:subject; dct:temporal; dct:type;

deFactoLanguage; deaths; deliveryDate; demonym; density; depictionDescription; deputy;
 description; designCompany; designer; destructionDate; developer; diameter; diocese;
 discoverer; disease; distance; distributor; division; documentDesignation; dose; dosesFirst;
 dosesSecond; draft; draftYear; drug; drugbank; duration; dynasty; editor; einecsNumber;
 electionMajority; elementAbove; endOccupation; endYear; engineer; era;
 ethnicGroupsInYear; event; executiveProducer; exhibition; existence; extinctionDate;
 failedLaunches; fdaUniiCode; feastDay; fighter; filmPolskiId; firstLeader; firstOwner;
 firstPlace; firstPopularVote; flag; flagBorder; flagLink; floodingDate; foaf:based_near;
 foaf:depiction; foaf:familyName; foaf:gender; foaf:givenName; foaf:homepage; foaf:img;
 foaf:isPrimaryTopicOf; foaf:logo; foaf:mbox; foaf:name; foaf:nick; foaf:page; foaf:phone;
 foaf:surname; foaf:thumbnail; foaf:topic; focus; followedBy; follows; forces; format;
 formerName; foundedBy; founder; foundingDate; foundingYear; frequency; fuel; fuelType;
 gameModus; gender; genre; geology; governmentType; governor; governorGeneral;
 greekName; hasChannel; hasInput; hasSurfaceForm; hasVariant; height; highestMountain;
 highestPlace; highestPosition; highschool; homeArena; homeColourHexCode; homeStadium;
 honours; imageSize; imdbId; incumbent; individualisedGnd; influenced; influencedBy;
 inscription; institution; intercommunality; introduced; isPartOf; isPartOfName; isnild;
 jureLanguage; jurisdiction; keyPerson; kinOfLanguage; knownFor; language;
 languageRegulator; latinName; lccnId; leader; league; length; licensee; lieutenant; listItemOf;
 localAuthority; location; locationCountry; locationName; logo; longName; mainInterest;
 majorityFloorLeader; majorityLeader; management; managementElevation;
 managementMountain; managementPlace; managementPosition; managerYearsEndYear;
 managerYearsStartYear; mandate; manufactory; manufacturer; mascot; mass; material; max;
 maximumDischarge; mayor; mayorCouncillor; mbaId; meaning; mediaItem; medlinePlus;
 meltingPoint; memberOfParliament; militaryGovernment; militaryUnitSize; min;
 minimumDischarge; minister; minorityFloorLeader; minorityLeader; mission; moderna;
 modernaCumul; monarch; month; motto; mouthElevation; municipality;
 musicBrainzArtistId; name; nameAsOf; nameInCantoneseChinese; nameInHangulKorean;
 nameInHanjaKorean; nameInJapanese; nameInMindongyuChinese;
 nameInMinnanyuChinese; nameInPinyinChinese; nameInSimplifiedChinese;
 nameInTraditionalChinese; nameInWadeGilesChinese; namedAfter; ndId; newspaper; nlaId;
 nominee; notableWork; notes; number; numberOfClassess; numberOfClassessWithResource;
 numberofDisambiguates; numberofIndegree; numberofOutdegree; numberofPixels;
 numberofPredicates; numberofProperties; numberofPropertiesUsed;
 numberofRedirectedResource; numberofResource; numberofResourceOfClass;
 numberofResourceOfType; numberofResourceWithType; numberofSports;
 numberofSportsEvents; numberofTriples; numberofUniqueResources;
 numberofUseOfProperty; numberofVisitors; occupation; odor; office; openingDate;
 openingYear; operator; orbits; orcidId; orderInOffice; organisation; origin; originalName;
 originallyUsedFor; orogeny; owl:differentFrom; owl:sameAs; owner; owningCompany;
 owningOrganisation; pandemic; parentCompany; parliament; part; partialFailedLaunches;
 party; partyNumber; pastMember; patronSaint; pccSecretary; peopleFullyVaccinated;
 peopleVaccinated; peopleVaccinatedPerHundred; percentage; percentageOfAreaWater;
 period; personsFirstDosesCumul; personsFullDosesCumul; pfizer; pfizerCumul; picture;
 pictureDescription; picturesCommonsCategory; place; playRole; playingTime; pole;
 populationTotal; portfolio; position; postalCode; predecessor; prefix; president; price;
 primeMinister; produces; productionEndDate; productionEndYear; productionStartDate;
 productionStartYear; prominence; pronunciation; provides; publicationDate;
 publiclyAccessible; purpose; quotation; race; raceHorse; rank; rating; rdf:type; rdfs:comment;
 rdfs:label; rdfs:seeAlso; rdfs:subClassOf; rebuildDate; recordLabel; region; registry; related;
 releaseDate; religion; religiousHead; reopened; requirement; restoreDate; restriction; retired;
 ridId; rkdArtistsId; role; royalAnthem; runningMate; sales; scaleFactor; score; secondLeader;
 secondPlace; secondPopularVote; selibrId; senator; seniority; series; sessionNumber;

setupTime; show; similar; siren; sizeBlazon; sizeLogo; sizeMap; sizeThumbnail; skills;
 skos:broader; skos:prefLabel; skos:related; slogan; soundRecording; source;
 sourceConfluenceElevation; sourceDistrict; sourceElevation; sourceMountain; sourceName;
 sourcePlace; sourcePosition; sourceRegion; sourceState; sourceText; sourceWebsite; speaker;
 specialist; specialization; sport; sportGoverningBody; stadium; startOccupation; startYear;
 statName; statValue; state; stateDelegate; status; subdivision; subdivisions; sublimationPoint;
 subtitle; successfulLaunches; successor; sudocId; surfaceFormOccurrenceOffset; surfaceType;
 synonym; systemRequirements; tag; taoiseach; team; teamSize; televisionSeries; temperature;
 tennisSurfaceType; termPeriod; thirdPlace; thumbnail; thumbnailCaption; time;
 timeInSpace; title; topSpeed; totalLaunches; totalVaccinations;
 totalVaccinationsPerHundred; tradeMark; translatedMotto; type; ulanId; university; updated;
 uses; usk; vaccine; value; variantOf; vehicle; viafId; viceChancellor; vicePresident;
 vicePrimeMinister; victim; victims; voice; volume; webcast; websiteLabel; weight; width;
 wikiPageCharacterSize; wikiPageDisambiguates; wikiPageEditLink; wikiPageExternalLink;
 wikiPageExtracted; wikiPageHistoryLink; wikiPageID; wikiPageInDegree;
 wikiPageInterLanguageLink; wikiPageLength; wikiPageModified; wikiPageOutDegree;
 wikiPageRedirects; wikiPageRevisionID; wikiPageRevisionLink; wikiPageUsesTemplate;
 wikiPageWikiLink; wikiPageWikiLinkText; wikidataSplitIri; wins; year; zodiacSign; کتاب

person: achievement; activeYears; activeYearsEndDateMgr; activeYearsEndYearMgr; activeYe
 arsStartDateMgr; activeYearsStartYearMgr; activity; affair; agency; allegiance; almaMater; anno
 uncedFrom; approach; arrestDate; artisticFunction; astrologicalSign; awardName; badGuy; best
 Lap; bigPoolRecord; birthDate; birthName; birthPlace; birthSign; birthYear; bloodGroup; bloodT
 ype; board; bodyDiscovered; breeder; britishWins; bustSize; bustWaistHipSize; careerStation; ca
 se; catch; causeOfDeath; child; circumcised; citizenship; clothSize; clothingSize; coalition; collabo
 ration; colleague; college; competitionTitle; complexion; contest; continentalTournament; contin
 entalTournamentBronze; continentalTournamentGold; continentalTournamentSilver; convicti
 onDate; copilote; cousurper; created; currentTeamManager; dam; damsire; danseCompetition; d
 anseScore; dateOfBurial; deadInFightDate; deadInFightPlace; deathAge; deathCause; deathDate
 ; deathPlace; deathYear; debutWork; decoration; derivedWord; detractor; diploma; dubber; educa
 tion; educationPlace; electionDate; employer; endCareer; endReign; escalafon; ethnicity; exp; ion;
 externalOrnament; eyeColor; eyeColour; feat; federation; filmNumber; foalDate; foot; freeDanseS
 core; freeProgCompetition; freeProgScore; freeScoreCompetition; friend; fullCompetition; fullSc
 ore; grandsire; grave; gymApparatus; hairColor; hairColour; hand; handedness; handisport; hasN
 aturalBust; hipSize; homage; idNumber; iihfHof; imposedDanseCompetition; imposedDanseSco
 re; individualisedPnd; juniorYearsEndYear; juniorYearsStartYear; jutsu; killedBy; lahHof; leader
 ship; leagueManager; linguisticsTradition; littlePoolRecord; livingPlace; mainDomain; masters
 Wins; matchPoint; measurements; media; member; militaryService; mood; mount; movie; nation
 alTeamMatchPoint; nationalTeamYear; nationalTournament; nationalTournamentBronze; nati
 onalTournamentGold; nationalTournamentSilver; nationality; networth; nnadbId; nonProfessio
 nalCareer; notableIdea; noteOnRestingPlace; numberOfRun; olympicGames; olympicGamesBr
 onze; olympicGamesGold; olympicGamesSilver; olympicGamesWins; opponent; orientation; ori
 ginalDanseCompetition; originalDanseScore; otherActivity; otherFunction; otherMedia; otherO
 ccupation; parent; parliamentaryGroup; particularSign; partner; patent; penisLength; personFu
 nction; pgaWins; philosophicalSchool; piercing; placeOfBurial; playerStatus; podium; polePositi
 on; politicalFunction; profession; project; pseudonym; publication; raceWins; radio; reign; reignN
 ame; relatedFunctions; relation; relative; residence; restingDate; restingPlace; restingPlacePositi
 on; retirementDate; salary; scene; school; seasonManager; second; seiyu; sentence; sex; sexualOrie
 ntation; shoeNumber; shoeSize; shoot; shortProgCompetition; shortProgScore; sibling; signature
 ; significantProject; sire; skinColor; specialTrial; speciality; sportDiscipline; sportsFunction; spou
 se; spouseName; startCareer; startReign; startWct; startWqs; stateOfOrigin; statusManager; stud
 ent; tattoo; teamManager; temple; templeYear; tessitura; third; titleDate; tvShow; usopenWins; us

urper;victory;victoryAsMgr;victoryPercentageAsMgr;waistSize;weddingParentsDate;world;worldTournament;worldTournamentBronze;worldTournamentGold;worldTournamentSilver.

ship: acquirementDate; assembly; captureDate; christeningDate; commissioningDate; decommissioningDate; dischargeAverage; enginePower; engineType; homeport; introductionDate; layingDown; maidenVoyage; modelEndDate; modelEndYear; modelLineVehicle; modelStartDate; modelStartYear; numberOfCrew; numberOfLaunches; numberOfPassengers; numberOfSeats; orderDate; powerType; range; rebuilder; recommissioningDate; relatedMeanOfTransportation; shipBeam; shipCrew; shipDisplacement; shipDraft; shipLaunch; version.

place: abbreviation; abstract; access; accessDate; activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; address; administrativeStatus; affiliation; ageRange; agglomerationPopulationTotal; aitaCode; album; alias; alongside; alternativeName; altitude; animal; annualTemperature; anthem; apcPresident; appearance; appointer; area; areaCode; areaDate; areaLand; areaMetro; areaOfCatchmentQuote; areaQuote; areaRank; areaTotal; areaWater; artificialSnowArea; asWikiText; ascent; associate; associateStar; associatedBand; associatedMusicalArtist; astrazenca; astrazencaCumul; atcCode; atcPrefix; atcSuffix; authority; authorityMandate; autonomy; average; averageDepth; averageDepthQuote; averageSpeed; avifaunaPopulation; award; awayColourHexCode; background; battery; battle; bibsysId; biggestCity; bioclimate; biome; bird; blackLongDistancePisteNumber; blackSkiPisteNumber; blazon; blazonLink; blazonRatio; blueLongDistancePisteNumber; blueSkiPisteNumber; bnfId; book; border; bpnId; budget; budgetYear; builder; buildingEndDate; buildingStartDate; cableCar; calculationNeeds; callSign; cannonNumber; capacity; capitalCoordinates; casNumber; casSupplemental; category; chain; chairLabel; chairman; chancellor; channel; chiefPlace; city; cityLink; class; classification; climate; climbUpNumber; closeTo; closingDate; closingYear; coach; coachedTeam; coastLength; coatOfArms; code; codeMunicipalMonument; codeNationalMonument; codeProvincialMonument; collectionSize; colour; colourHexCode; colourName; commandant; commander; comment; committee; commonName; company; conflict; connotation; constituencyDistrict; cosparId; cost; country; countryCode; county; courseArea; cpu; creationYear; creator; crest; criteria; cuisine; currency; currencyCode; currentCity; currentStatus; dailyVaccinationsPerMillion; dailyVaccinationsRaw; date; dateAgreement; dateBudget; dateOfAbandonment; daylightSavingTimeZone; dc:creator; dc:identifier; dc:language; dc:rights; dc:subject; dc:title; dc:type; dct:created; dct:dateSubmitted; dct:description; dct:extent; dct:format; dct:medium; dct:modified; dct:spatial; dct:subject; dct:temporal; dct:type; deFactoLanguage; deaths; deliveryDate; demonym; density; depictionDescription; depth; depthQuote; depths; deputy; description; designCompany; designer; destructionDate; developer; diameter; diocese; discoverer; disease; distance; distributor; district; division; documentDesignation; dose; dosesFirst; dosesSecond; draft; draftYear; drug; drugbank; duration; dynasty; eastPlace; or; electionsNumber; electionMajority; elementAbove; elevation; elevationQuote; endOccupation; endPoint; endYear; endangeredSince; engineer; era; ethnicGroupsInYear; event; eventDate; executiveProducer; exhibition; existence; extinctionDate; failedLaunches; fauna; fdaUniiCode; feastDay; fighter; filmPolskiId; firstLeader; firstOwner; firstPlace; firstPopularVote; flag; flagBorder; flagLink; floodingDate; flora; flower; foaf:based_near; foaf:depiction; foaf:familyName; foaf:gender; foaf:givenName; foaf:homepage; foaf:img; foaf:isPrimaryTopicOf; foaf:logo; foaf:mbox; foaf:name; foaf:nick; foaf:page; foaf:phone; foaf:surname; foaf:thumbnail; foaf:topic; focus; followedBy; follows; forces; foresterDistrict; format; formerName; foundedBy; founder; foundingDate; foundingYear; frequency; frontierLength; fuel; fuelType; gameModus; gender; genre; geolocDual; geologicPeriod; geology; governingBody; governmentType; governor; governorGeneral; greekName; greenLongDistancePisteNumber; greenSkiPisteNumber; gridReference; grossDomesticProductPerPeople; grossDomesticProductRank; hasChannel; hasInput; hasInsidePlace; hasOutsidePlace; hasSurfaceForm; hasVariant; height; heritageRegister; highestAltitude; highestMountain; highestPlace; highestPosition; highschool; historicalMap; historicalName; homeArena; homeColourHexCode; homeStadium; honours; humanDevelopmentIndexRank; imageSize; imdbId; incumbent; in

individualisedGnd;influenced;influencedBy;information;informationName;inscription;institution;intercommunal;introduced;isPartOf;isPartOfName;isRouteStop;isniId;iso31661Code;isoCode;jureLanguage;jurisdiction;keyPerson;kinOfLanguage;kindOfCoordinate;kindOfRock;knownFor;lake;land;landArea;landPercentage;language;languageRegulator;latinName;lawCountry;lccnId;leader;league;length;lengthQuote;licensee;lieutenant;limit;linkedSpace;listItemOf;localAuthority;locatedInArea;location;locationCountry;locationName;logo;longDistancePisteKilometre;longDistancePisteNumber;longName;lowest;lowestAltitude;lowestMountain;lowestPlace;lowestPosition;mainInterest;mainIsland;majorityFloorLeader;majorityLeader;management;managementElevation;managementMountain;managementPlace;managementPosition;managerYearsEndYear;managerYearsStartYear;mandate;manufactory;manufacturer;map;mapCaption;mapDescription;mascot;mass;material;max;maximumArea;maximumAreaQuote;maximumDepth;maximumDepthQuote;maximumDischarge;maximumElevation;mayor;mayorArticle;mayorCouncillor;mbaId;meaning;mediaItem;medlinePlus;meltingPoint;memberOfParliament;mergerDate;militaryGovernment;militaryUnitSize;min;minimumArea;minimumAreaQuote;minimumDischarge;minimumElevation;minister;minorityFloorLeader;minorityLeader;mission;moderna;modernaCumul;monarch;month;motto;mouthElevation;municipality;municipalityCode;musicBrainzArtistId;name;nameAsOf;nameInCantoneseChinese;nameInHangulKorean;nameInHanjaKorean;nameInJapanese;nameInMindongyuChinese;nameInMinnanyuChinese;nameInPinyinChinese;nameInSimplifiedChinese;nameInTraditionalChinese;nameInWadeGilesChinese;namedAfter;namedByLanguage;nationalTopographicSystemMapNumber;ndId;nearestCity;neighbourRegion;newspaper;nextEntity;nlaId;nominee;northEastPlace;northPlace;northWestPlace;notableWork;notes;number;numberOfClasses;numberOfClassesWithResource;numberOfDisambiguates;numberOfIndegree;numberOfIsland;numberOfOutdegree;numberOfPersonBornInPlace;numberOfPixels;numberOfPredicates;numberOfProperties;numberOfPropertiesUsed;numberOfRedirectedResource;numberOfResource;numberOfResourceOfClass;numberOfResourceOfType;numberOfResourceWithType;numberOfSoccerPlayersBornInPlace;numberOfSports;numberOfSportsEvents;numberOfTriples;numberOfUniqueResources;numberOfUseOfProperty;numberOfVisitors;nutsCode;occupation;odor;office;openingDate;openingYear;operator;orbits;orcidId;orderInOffice;organisation;origin;originalName;originallyUsedFor;orogeny;output;owl:differentFrom;owl:sameAs;owner;owningCompany;owningOrganisation;pandemic;parentCompany;parliament;parliamentType;part;partialFailedLaunches;party;partyNumber;pastMember;patronSaint;pccSecretary;peopleFullyVaccinated;peopleVaccinated;peopleVaccinatedPerHundred;percentage;percentageOfAreaWater;perimeter;period;personsFirstDosesCumul;personsFullDosesCumul;pfizer;pfizerCumul;picture;pictureDescription;picturesCommonsCategory;pisciculturalPopulation;place;plant;playRole;playingTime;pole;politicalLeader;populationQuote;populationTotal;portfolio;position;postalCode;predecessor;prefix;president;presidentGeneralCouncilMandate;presidentRegionalCouncilMandate;previousEntity;price;primeMinister;produces;productionEndDate;productionEndYear;productionStartDate;productionStartYear;prominence;pronunciation;protectionStatus;provides;province;provinceLink;publicationDate;publiclyAccessible;purpose;quotation;quote;race;raceHorse;rank;rankAgreement;rating;rdf:type;rdfs:comment;rdfs:label;rdfs:seeAlso;rdfs:subClassOf;rebuildDate;recordLabel;redLongDistancePisteNumber;redSkiPisteNumber;refcul;refgen;refgeo;refpol;regency;region;regionLink;regionType;regionalPrefecture;registry;related;relatedPlaces;releaseDate;relief;religion;religiousHead;reopened;representative;requirement;restoreDate;restriction;retentionTime;retired;ridId;river;rkdArtistsId;role;route;royalAnthem;runningMate;sales;scale;scaleFactor;score;sea;secondLeader;secondPlace;secondPopularVote;selibrId;senator;seniority;series;sessionNumber;settlement;setupTime;show;similar;siren;sizeBlazon;sizeLogo;sizeMap;sizeThumbnail;skiLift;skiPisteKilometre;skiPisteNumber;skiTow;skills;skos:broader;skos:prefLabel;skos:related;slogan;snowParkNumber;soundRecording;source;sourceConfluenceElevation;sourceDistrict;sourceElevation;sourceMountain;sourceName;sourcePlace;sourcePosition;sourceRegion;sourceState;sourceText;sourceWebsite;southEastPlace;southPlace;southWestPlace;sovereignCountry;speaker;specialist;specialization;sport;sportGoverningBody;stadium;startOccupation;start

Point;startYear;statName;statValue;state;stateDelegate;status;statusYear;subPrefecture;sub
 division;subdivisionLink;subdivisions;sublimationPoint;subregion;subsystem;subsystemLin
 k;subtitle;successfulLaunches;successor;sudocId;supply;surfaceFormOccurrenceOffset;sufa
 ceType;synonym;systemRequirements;tag;taoiseach;team;teamSize;televisionSeries;tempera
 ture;tennisSurfaceType;termPeriod;thirdPlace;thumbnail;thumbnailCaption;time;timeInSpa
 ce;timeZone;title;topSpeed;totalLaunches;totalVaccinations;totalVaccinationsPerHundred;tr
 adeMark;translatedMotto;tree;type;typeCoordinate;ulanId;unesco;university;unloCode;upda
 ted;uses;usk;utcOffset;vaccine;value;variantOf;vehicle;vehicleCode;viafId;viceChancellor;vic
 ePresident;vicePrimeMinister;victim;victims;voice;volume;volumeQuote;water;waterArea;w
 aterPercentage;webcast;websiteLabel;weight;westPlace;wholeArea;width;widthQuote;wikiPa
 geCharacterSize;wikiPageDisambiguates;wikiPage;Link;wikiPageExternalLink;wikiPageExtr
 acted;wikiPageHistoryLink;wikiPageID;wikiPageInDegree;wikiPageInterLanguageLink;wiki
 PageLength;wikiPageModified;wikiPageOutDegree;wikiPageRedirects;wikiPageRevisionID;
 wikiPageRevisionLink;wikiPageUsesTemplate;wikiPageWikiLink;wikiPageWikiLinkText;wik
 idataSplitIri;wins;year;yearOfConstruction;zodiacSign;کتاب.

city: adjacentSettlement; administrativeCollectivity; administrativeDistrict;
 agglomerationPopulation; agglomerationPopulationYear; associationOfLocalGovernment;
 bourgmestre; canton; coastline; codeSettlement; collectivityMinority; commune;
 crownDependency; daira; day; distanceToBelfast; distanceToCapital; distanceToCardiff;
 distanceToCharingCross; distanceToDouglas; distanceToDublin; distanceToEdinburgh;
 distanceToLondon; distanceToNearestCity; employersCelebration; federalState; frazioni;
 geolocDepartment; highestPoint; inseeCode; isCityState; isoCodeRegion; jointCommunity;
 landRegistryCode; largestMetro; locality; lowestPoint; mergedSettlement; minority;
 officialName; ofsCode; ons; otherInformation; otherLanguage; phonePrefixLabel;
 politicalMajority; politicalSeats; populationTotalReference; provCode; reffBourgmestre;
 sameName; settlementAttached; summerTemperature; twinTown; urbanArea; wilaya.

country: administrativeHeadCity; agglomeration; agglomeration;
 agglomerationDemographics; apskritis; areaRural; areaTotalRanking; areaUrban;
 arrondissement; barangays; borough; capital; catholicPercentage; ccaState; censusYear;
 ceremonialCounty; circle; citySince; cityType; colonialName; commissionerDate;
 communityIsoCode; continent; councilArea; countySeat; delegateMayor; delegation; deme;
 demographics; demographicsAsOf; department; departmentPosition; disappearanceDate;
 ekatteCode; emblem; ethnicGroup; europeanUnionEntranceDate; fipsCode; firstMention;
 foundation; gdpPerCapita; giniCoefficient; giniCoefficientAsOf; giniCoefficientCategory;
 giniCoefficientRanking; gnisCode; gnl; governorate; grossDomesticProduct;
 grossDomesticProductAsOf; grossDomesticProductNominalPerCapita;
 grossDomesticProductPurchasingPowerParityPerCapita; historicalRegion; hraState;
 humanDevelopmentIndex; humanDevelopmentIndexAsOf;
 humanDevelopmentIndexRankingCategory; illiteracy; income; infantMortality;
 internationalPhonePrefix; internationalPhonePrefixLabel; ist; istat; largestCity;
 largestSettlement; latestElection; leaderName; leaderParty; leaderTitle; legislature;
 licenceNumber; licenceNumberLabel; lieutenantancy; lieutenantancyArea; lifeExpectancy;
 localPhonePrefix; localizationThumbnail; localizationThumbnailCaption;
 metropolitanBorough; nation; neighboringMunicipality; nisCode; numberOfCantons;
 numberOfCapitalDeputies; numberOfCounties; numberOfFederalDeputies;
 numberOfMinistries; numberOfMunicipalities; numberOfMusicalArtistEntities;
 numberOfPersonEntities; numberOfSettlementsInCountry; numberOfStateDeputies;
 officialLanguage; okatoCode; oldDistrict; oldName; oldProvince; otherName; outskirts;
 parish; peopleName; perCapitaIncome; perCapitaIncomeAsOf; perCapitaIncomeRank;
 percentageLiteracyMen; percentageLiteracyWomen; percentageLiterate; phonePrefix;
 pluviometry; policeName; politicGovernmentDepartment; population; populationAsOf;

populationDensity; populationMetro; populationMetroDensity; populationPctChildren; populationPctMen; populationPctWomen; populationRural; populationRuralDensity; populationTotalRanking; populationUrban; populationUrbanDensity; populationYear; previousDemographics; previousName; previousPopulation; previousPopulationTotal; principalArea; protestantPercentage; provinceIsoCode; purchasingPowerParity; purchasingPowerParityRank; purchasingPowerParityYear; rankArea; rankPopulation; regionalLanguage; saint; senior; seniunija; sharingOutPopulation; sharingOutPopulationYear; sheading; simcCode; termOfOffice; terytCode; topLevelDomain; topic; touristicSite; twinCountry; unitaryAuthority; viceLeader; viceLeaderParty; watercourse; wordBefore; zipCode.

product: abbreviation; abstract; access; accessDate; activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation; ageRange; agglomerationPopulationTotal; aitaCode; album; alias; alongside; alternativeName; animal; anthem; apc President; appearance; appointer; approximateCalories; area; areaMetro; asWikiText; ascent; associate; associateStar; associatedBand; associatedMusicalArtist; astrazenca; astrazencaCumul; atcCode; atcPrefix; atcSuffix; authority; autonomy; average; averageSpeed; award; awayColourHexCode; background; battery; battle; bibsysId; biome; blazon; blazonLink; blazonRatio; bnfId; book; border; bpnId; budget; builder; buildingEndDate; buildingStartDate; calculationNeeds; callSign; capacity; carbohydrate; casNumber; casSupplemental; category; chain; chairLabel; chairman; chancellor; channel; chiefPlace; city; class; classification; closingDate; closingYear; coach; coachedTeam; coatOfArms; code; collectionSize; colour; colourHexCode; colourName; commandant; commander; comment; committee; commonName; company; complexity; conflict; connotation; constituencyDistrict; cosparId; cost; country; county; courseArea; cpu; creationYear; creator; creatorOfDish; crest; criteria; cuisine; currency; currentCity; currentStatus; dailyVaccinationsPerMillion; dailyVaccinationsRaw; date; dateBudget; dateOfAbandonment; dc:creator; dc:identifier; dc:language; dc:rights; dc:subject; dc:title; dc:type; dct:created; dct:dateSubmitted; dct:description; dct:extent; dct:format; dct:medium; dct:modified; dct:spatial; dct:subject; dct:temporal; dct:type; deFactoLanguage; deaths; deliveryDate; demonym; density; depictionDescription; deputy; description; designCompany; designer; destructionDate; developer; diameter; diocese; discoverer; disease; distance; distributor; division; documentDesignation; dose; dosesFirst; dosesSecond; draft; draftYear; drug; drugbank; duration; dynasty; or; einecsNumber; electionMajority; elementAbove; endOccupation; endYear; engineer; era; ethnicGroupsInYear; event; executiveProducer; exhibition; existence; extinctionDate; failedLaunches; fat; fdaUniiCode; feastDay; fighter; filmPolskiId; firstLeader; firstOwner; firstPlace; firstPopularVote; flag; flagBorder; flagLink; floodingDate; foaf:based_near; foaf:depiction; foaf:familyName; foaf:gender; foaf:givenName; foaf:homepage; foaf:img; foaf:isPrimaryTopicOf; foaf:logo; foaf:mbox; foaf:name; foaf:nick; foaf:page; foaf:phone; foaf:surname; foaf:thumbnail; foaf:topic; focus; followedBy; follows; forces; format; formerName; foundedBy; founder; foundingDate; foundingYear; frequency; fuel; fuelType; gameModus; gender; genre; geology; glycemicIndex; governmentType; governor; governorGeneral; greekName; hasChannel; hasInput; hasSurfaceForm; hasVariant; height; highestMountain; highestPlace; highestPosition; high school; homeArena; homeColourHexCode; homeStadium; honours; imageSize; imdbId; incumbent; individualisedGnd; influenced; influencedBy; ingredient; ingredientName; inscription; institution; intercommunality; introduced; isPartOf; isPartOfName; isniId; jureLanguage; jurisdiction; keyPerson; kinOfLanguage; knownFor; language; languageRegulator; latinName; lccnId; leader; league; length; licensee; lieutenant; listItemOf; localAuthority; location; locationCountry; locationName; logo; longName; mainInterest; majorityFloorLeader; majorityLeader; management; managementElevation; managementMountain; managementPlace; managementPosition; managerYearsEndYear; managerYearsStartYear; mandate; manufactory; manufacturer; mascot; mass; material; max; maxTime; maximumDischarge; mayor; mayorCouncillor; mbaId; meaning; mediaItem; medlinePlus; meltingPoint; memberOfParliament; militaryGovernment; militaryUnitSize; min; minTime; minimumDischarge; minister; minorityFloorLeader; minorityLeader; mission; moderna; modernaCumul; monarch; month; motto; mouthElevation; municipality; musicBrainzAr

tistId;name;nameAsOf;nameInCantoneseChinese;nameInHangulKorean;nameInHanjaKorean;nameInJapanese;nameInMindongyuChinese;nameInMinnanyuChinese;nameInPinyinChinese;nameInSimplifiedChinese;nameInTraditionalChinese;nameInWadeGilesChinese;nameIdAfter;ndId;newspaper;nlaId;nominee;notableWork;notes;number;numberOfClasses;numberOfClassesWithResource;numberOfDisambiguates;numberOfIndegree;numberOfOutdegree;numberOfPixels;numberOfPredicates;numberOfProperties;numberOfPropertiesUsed;numberOfRedirectedResource;numberOfResource;numberOfResourceOfClass;numberOfResourceOfType;numberOfResourceWithType;numberOfSports;numberOfSportsEvents;numberOfTriples;numberOfUniqueResources;numberOfUseOfProperty;numberOfVisitors;occupation;odor;office;openingDate;openingYear;operator;orbits;orcidId;orderInOffice;organisation;origin;originalName;originallyUsedFor;orogeny;owl:differentFrom;owl:sameAs;owner;owningCompany;owningOrganisation;pandemic;parentCompany;parliament;part;partialFailedLaunches;party;partyNumber;pastMember;patronSaint;pccSecretary;peopleFullyVaccinated;peopleVaccinated;peopleVaccinatedPerHundred;percentage;percentageFat;percentageOfAreaWater;period;personsFirstDosesCumul;personsFullDosesCumul;pfiger;pfigerCumul;picture;pictureDescription;picturesCommonsCategory;place;playRole;playingTime;pole;populationTotal;portfolio;position;postalCode;predecessor;prefix;president;price;primeMinister;produces;productShape;productionEndDate;productionEndYear;productionStartDate;productionStartYear;prominence;pronunciation;protein;provides;publicationDate;publiclyAccessible;purpose;quotation;race;raceHorse;rank;rating;rdf:type;rdfs:comment;rdfs:label;rdfs:seeAlso;rdfs:subClassOf;rebuildDate;recordLabel;region;registry;related;releaseDate;religion;religiousHead;reopened;requirement;restoreDate;restriction;retired;ridId;rkdArtistsId;role;royalAnthem;runningMate;sales;scaleFactor;score;secondLeader;secondPlace;secondPopularVote;selibrId;senator;seniority;series;servingSize;servingTemperature;sessionNumber;setupTime;show;similar;siren;sizeBlazon;sizeLogo;sizeMap;sizeThumbnail;skills;skos:broader;skos:prefLabel;skos:related;slogan;soundRecording;source;sourceConfluenceElevation;sourceDistrict;sourceElevation;sourceMountain;sourceName;sourcePlace;sourcePosition;sourceRegion;sourceState;sourceText;sourceWebsite;speaker;specialist;specialization;sport;sportGoverningBody;stadium;startOccupation;startYear;statName;statValue;state;stateDelegate;status;subdivision;subdivisions;sublimationPoint;subtitle;successfulLaunches;successor;sudocId;surfaceFormOccurrenceOffset;surfaceType;synonym;systemRequirements>tag;taoiseach;taste;team;teamSize;televisionSeries;temperature;tennisSurfaceType;termPeriod;thirdPlace;thumbnail;thumbnailCaption;time;timeInSpace;title;topSpeed;totalLaunches;totalVaccinations;totalVaccinationsPerHundred;tradeMark;translatedMotto;type;typeOfGrain;typeOfStorage;typeOfYeast;ulanId;university;updated;uses;usk;vaccine;value;variantOf;vehicle;viafId;viceChancellor;vicePresident;vicePrimeMinister;victim;victims;voice;volume;webcast;websiteLabel;weight;width;wikiPageCharacterSize;wikiPageDisambiguates;wikiPage;Link;wikiPageExternalLink;wikiPageExtracted;wikiPageHistoryLink;wikiPageID;wikiPageInDegree;wikiPageInterLanguageLink;wikiPageLength;wikiPageModified;wikiPageOutDegree;wikiPageRedirects;wikiPageRevisionID;wikiPageRevisionLink;wikiPageUsesTemplate;wikiPageWikiLink;wikiPageWikiLinkText;wikidataSplitIri;wins;year;zodiacSign;کتاب.

plant: cultivatedVariety; hybrid.

APÊNDICE C - Tabela de Classes das Entidades e Propriedades Encontradas Associadas ao Registo 1 na Wikidata, Europeia e DBpedia Após Primeira Seleção.

Nº de Registo	Classes das Entidades	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeia	Propriedades Associadas na DBpedia
Registo 1	person; geographic location; natural product; primary raw material; product; ship; edm:Agent; edm:Place; edm:PhysicalThing (as contextual entities); foaf:Person; place; city; country; plant.	<u>person:</u> description; also known as; image; has quality (personal life; pseudonym; date of birth; marital status; given name; gender; family name; biographic information); commons category (people); topic's main category (category:People); wikidata property (significant person); identifiers. <u>geographic location:</u> label; description; also known as; named after (locus; relative location); has quality (geographic entity); properties for this type (coordinate location; country; elevation above sea level); commons category (Places); topic's main	<u>edm:Agent:</u> skos:prefLabel; skos:altLabel; skos:hiddenLabel; skos:note (for e.g., biographical notes); dc:date; dcterms:hasPart; dc:identifier; dcterms:isPartOf; edm:isRelatedTo (for generic relations to other agents, especially); edm:wasPresentAt; foaf:name; rdaGr2:biographicalInformation; rdaGr2:dateOfBirth; rdaGr2:dateOfDeath; rdaGr2:placeOfBirth; rdaGr2:placeOfDeath; rdaGr2:gender; rdaGr2:professionOrOccupation; owl:sameAs. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel; skos:hiddenLabel; skos:note; dcterms:hasPart; dcterms:isPartOf;	<u>foaf:Person:</u> abbreviation; abstract; activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; ; activeYearsStartYear; ; administrativeStatus; ; affiliation; ageRange; alternativeName; dc:rights; dc:subject; dc:title; description; foaf:familyName; foaf:gender; foaf:name; jurisdiction; knownFor. <u>person:</u> achievement; activeYears; activeYearsEndDate; Mgr; activeYearsEndYear; Mgr; activeYearsStartDate; Mgr; activeYearsStartYear; Mgr; activity; birthDate; birthName; birthPlace; birthYear; deathAge; deathCause; deathDate; deathPlace; deathYear; profession; spouse; spouseName.

		category (Category:Places) ; Wikidata property; identifiers. <u>natural product:</u> instance of (group or class of chemical substances); subclass of (chemical substance; product; primary raw material); different from (natural material; biomolecule); commons category (natural products); topic's main category (Category:Natural products); Wikidata property (this taxon is source of; natural product of taxon; found in taxon); identifiers. <u>primary raw material:</u> subclass of (raw material); part of (natural environment); image; opposite of (manufactured good); described by source; commons category (materials); topic's main category; identifiers. <u>product:</u>	edm:isNextInSequence; owl:sameAs. <u>edm:PhysicalThing</u> (as contextual entities): dc:coverage; dc:creator; dc:description; dc:format; dc:identifier; dc:relation; dc:rights; dc:source; dc:subject; dc:title; dc:type; dcterms:alternative; dcterms:created; dcterms:extent; dcterms:hasFormat; dcterms:hasPart; dcterms:isFormatOf; dcterms:isPartOf; dcterms:isReferencedBy; dcterms:isRequiredBy; dcterms:medium; dcterms:provenance; dcterms:references; dcterms:temporal; edm:currentLocation; edm:hasType; edm:incorporates; edm:isDerivativeOf; edm:isSimilarTo.	<u>ship:</u> enginePower; engineType; homeport; introductionDate; layingDown; maidenVoyage; modelEndDate; modelEndYear; modelLineVehicle; modelStartDate; modelStartYear; numberOfCrew; numberOfLaunches; numberOfPassengers; numberOfSeats; orderDate; powerType; range; rebuilder; recommissioningDate; relatedMeanOfTransportation; shipBeam; shipCrew; shipDisplacement; shipDraft; shipLaunch; version. <u>place:</u> abbreviation; abstract; address; administrativeStatus; agglomerationPopulationTotal; alternativeName; altitude; annualTemperature; anthem; area; areaCode; areaLand; areaMetro; city; climate; closeTo; code; country; countryCode; county; dc:subject; dc:title; dct:description; dct:extent; dct:spatial; dct:subject; formerName; foundedBy; founder; foundingDate; foundingYear; frontierLength; historicalMap; historicalName; landArea
--	--	---	---	---

		subclass of (goods and services; product; good); image; has quality (function; product naming; equivalent class; different from (tangible good; goods and services); said to be the same as (product; goods and services); commons category (Products); topic's main category (Category:Products); identifiers. <u>ship:</u> instance of (vessel class); subclass of (watercraft; vessel); image; has use (transport); has quality (ship type); small logo or icon; equivalent class; has part or parts (hull; ship element; bridge); properties for this type (image; home port; port of registry; vessel class; manufacturer; location of creation; yard number; pennant number; length; significant event; conflict; draft; width; net tonnage; has part or parts; number of masts; owned	a;length;locatedInArea;location;locationCountry;locationName;municipality;municipalityCode;name;picture;place;populationTotal;temperature;kindOfCoordinate. <u>city:</u> administrativeDistrict; agglomerationPopulation; agglomerationPopulationYear; distanceToNearestCity; federalState; isoCodeRegion; landRegistryCode; locality; officialName; populationTotalReference; summerTemperature; urbanArea. <u>country:</u> administrativeHeadCity; agglomeration; agglomerationDemographics; areaRural; areaTotalRanking; capital; colonialName; continent; foundation; historicalRegion; largestCity; otherName; population; populationDensity; populationTotalRanking; populationUrbanDensity; regionalLanguage; zipCode. <u>product:</u>
--	--	---	---

		by; service entry; service retirement); different from (boat); commons gallery (ships);topic's main category (Category:Ships); identifiers.		abbreviation; abstract; alternativeName; cost; dc: subject; dc: title; dc: type; dct: description; ingredient; ingredientName; manufactory; manufacturer; material; origin; originalName; originallyUsedFor; owl: differentFrom; owl: sameAs; picture; price; produces;productSh ape;sourceName;sou rcePlace;sourceRegi on; foaf:img. <u>plant:</u> cultivatedVariety; hybrid.
--	--	--	--	--

APÊNDICE D - Tabela Resumida de Classes das Entidades e Propriedades Associadas aos 25 Registos na Wikidata, Europeia e DBpedia.

Nº de Registo	Classes das Entidades	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeia	Propriedades Associadas na DBpedia
Registo 1	person; geographic location; natural product; primary raw material; product; ship; edm:Agent; edm:Place; edm:PhysicalThing (as contextual entities); foaf:Person; place; plant.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level); label. <u>natural product:</u> instance of; subclass of; different from; commons category (natural products); Wikidata property (this taxon is source of; natural product of taxon; found in taxon); said to be the same as; identifiers.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. owl:sameAs. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note; owl:sameAs. <u>edm:PhysicalThing:</u> dc:creator; dc:description; dc:format; dc:title; dc:type; dcterms:isRequiredBy; edm:currentLocation; edm:isDerivativeOf; edm:isSimilarTo; dcterms:alternative; edm:hasType.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation; alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace; deathDate; deathPlace; profession; spouse. <u>ship:</u> homeport; modelEndDate; modelStartDate; numberOfCrew; shipLaunch; version. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent; dct: spatial; for

		<u>primary raw material:</u> subclass of; part of; image; opposite of; described by source; identifiers. <u>product:</u> subclass of image; has quality (function; product naming; equivalent class; different from; image; identifiers. <u>ship:</u> instance of; subclass of; has use; has quality; small logo or icon; equivalent class; properties for this type (image; home port; port of registry; vessel class; manufacturer; location of creation; parts; number of masts; owned by; service entry; service retirement); identifiers.		merName;foundedBy; founder;foundingDate ;kindOfCoordinate; location. <u>product:</u> alternativeName; dc: title; dc: type; dct: description; ingredient; manufacturer; material; origin; originallyUsedFor; picture;sourcePlace;sourceRegion; owl:sameAs; foaf:img. <u>plant:</u> cultivatedVariety; hybrid.
Registo 2	person; edm:Agent; foaf:Person.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name;	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative

		biographic information.	placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation.	Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse.
Registo 3	person; edm:Agent; foaf:Person.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse.
Registo 4	person; geographic location; edm:Agent; edm:Place; foaf:Person; place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u>	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name.

		description; also known as; properties for this type (coordinate location; country; elevation above sea level).	professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>Person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession; spouse. <u>Place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 5	geographic location; edm:Place; place.	<u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 6	X	X	X	X
Registo 7	person; geographic location; edm:Agent; edm:Place; foaf:Person; place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName;

		<u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession; spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 8	X	X	X	X
Registo 9	person; edm:Agent; foaf:Person; religious military order; payment system; skos:Concept; religious organisation; system of law.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>religious military order:</u> subclass of; image; religion or world view; has part; different from; commons category; identifiers.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>skos:Concept:</u> _skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch;	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternativeName; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate;

		<u>payment system:</u> instance of (type of system); subclass of (system); different from; commons category; identifiers.	skos:note; skos:related; skos:notation.	deathPlace;profession ;spouse. <u>religious organisation:</u> formationDate; foundationPlace; leaderFunction; organisationMember; numberOfLocations. <u>system of law:</u> name; label; domain; range; comment.
Registo 10	person; geographic location; edm:Agent; edm:Place; foaf:Person; place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;for

				merName;foundedBy; founder;foundingDate .
Registo 11	person; edm:Agent; foaf:Person; handicraft; skos:Concept; activity.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>handicraft:</u> instance of (industry); subclass of (activity; economic concept; craft); image; practiced by; said to be the same as; commons category; identifiers.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInforma tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateM gr; activeYearsStartDate Mgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>activity:</u> abstract; colour; country; dct:format; dct:description; material; name.
Registo 12	person; geographic location; edm:Agent; edm:Place; foaf:Person; place; religious order; religious organisation.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInforma tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u>	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateM gr; activeYearsStartDate

		(coordinate location; country; elevation above sea level). <u>religious order:</u> instance of (group of humans; community); subclass of (religious community; religious organization); described by source; commons category; identifiers.	wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	Mgr; birthDate; birthName; birthPlace; deathDate; deathPlace; profession ;spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate . <u>religious organisation:</u> formationDate; foundationPlace; leaderFunction; organisationMember; numberOfLocations.
Registo 13	geographic location; edm:Place; place.	<u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 14	person; edm:Agent; foaf:Person.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath;	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus;

		name; biographic information.	rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation.	affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse.
Registo 15	incentive program; skos:Concept; project.	<u>incentive program:</u> subclass of (subsidy); Wikimedia import URL; identifiers.	<u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	<u>project:</u> projectEndDate; projectKeyword; projectObjective; projectStartDate.
Registo 16	geographic location; edm:Place; place.	<u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 17	person; geographic location; edm:Agent; edm:Place; foaf:Person; Place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName;

		<u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession; spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 18	person; geographic location; edm:Agent; edm:Place; foaf:Person; place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel;	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternativeName; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession; spouse.

			skos:altLabel skos:note.	<u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate .
Registo 19	person; edm:Agent; foaf:Person; nautical chart; skos:Concept; thing.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>nautical chart:</u> instance of (type of map); subclass of (map); described by source; commons category; identifiers.	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDate Mgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>thing:</u> abstract; dct:format; dct:description; material; name; alternativeName; creator; date.
Registo 20	person; geographic location; edm:Agent;	<u>person:</u> also known as; date of birth;	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInforma	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear;

	edm:Place; foaf:Person; place; papal bull; skos:Concept; document.	marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level). <u>papal bull:</u> subclass of (deed; literary work); image; described by source; has part; different from; commons category; identifiers.	tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note. <u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate . <u>document:</u> alternative title; author; dc:description; originalLanguage; completionDate; originalTitle.
Registo 21	person; geographic location; edm:Agent; edm:Place; foaf:Person; place; auto- da-fé;	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name;	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternativen

	tribunal do santo ofício; skos:Concept; historical event.	biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level). <u>auto-da-fé:</u> subclass of (death by burning); described by source; facet of; commons category; identifiers. <u>tribunal do santo ofício:</u> instance of (fonds); part of (portuguese inquisition); inception; named after; country; location; collection; inventory number; dissolved, abolished or demolished date.	placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note. <u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession; spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;formerName;foundedBy; founder;foundingDate . <u>historical event:</u> name; label; domain; range; comment.
Registo 22	person; geographic location; edm:Agent; edm:Place; skos:Concept; foaf:Person;	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name;	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInformation; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2:	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative

	place; code of law; earthquake; document; natural event.	biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level). <u>code of law:</u> subclass of (legislation; book; compilation); described by source; topic's main category; identifiers. <u>earthquake:</u> subclass of (natural disaster; geologic activity); image; has effect; measurement scale; described by source; commons category; identifiers.	placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel; skos:note. <u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateMgr; activeYearsStartDateMgr; birthDate; birthName; birthPlace; deathDate; deathPlace; profession; spouse. <u>place:</u> address; agglomerationPopulationTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent; dct: spatial; formerName; foundedBy; founder; foundingDate. <u>document:</u> alternative title; author; dc: description; originalLanguage; completionDate; originalTitle. <u>naturalEvent:</u> causedBy; damage; endDate; endDateTime; followingEvent; startDate; startDateTime.
Registo 23	person; geographic location;	<u>person:</u>	<u>edm:Agent:</u>	<u>foaf:Person:</u>

	edm:Agent; edm:Place; foaf:Person; place.	also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	foaf:name; rdaGr2: biographicalInforma tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateM gr; activeYearsStartDate Mgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>place:</u> address; agglomerationPopulat ionTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;for merName;foundedBy; founder;foundingDate .
Registo 24	religious ceremony; edm:Agent; skos:Concept; event.	<u>religious ceremony:</u> subclass of (ceremony; religious concept).	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInforma tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation.	<u>event:</u> abstract; date; diocese; location; participant; religion; religious head.

			<u>skos:Concept:</u> skos:prefLabel; skos:relatedMatch; skos:altLabel; skos:exactMatch; skos:note; skos:related; skos:notation.	
Registo 25	person; geographic location; edm:Agent; edm:Place; foaf:Person; place.	<u>person:</u> also known as; date of birth; marital status; given name; gender; family name; biographic information. <u>geographic location:</u> description; also known as; properties for this type (coordinate location; country; elevation above sea level).	<u>edm:Agent:</u> foaf:name; rdaGr2: biographicalInforma tion; rdaGr2: dateOfBirth; rdaGr2: dateOfDeath; rdaGr2: placeOfBirth; rdaGr2: placeOfDeath; rdaGr2: gender; rdaGr2: professionOrOccupation. <u>edm:Place:</u> wgs84_pos:lat; wgs84_pos:long; wgs84_pos:alt; wgs84_pos:lat; skos:prefLabel; skos:altLabel skos:note.	<u>foaf:Person:</u> activeYearsEndDate; activeYearsEndYear; activeYearsStartDate; activeYearsStartYear; administrativeStatus; affiliation;alternative Name; description; foaf:familyName; foaf:gender; foaf:name. <u>person:</u> activeYearsEndDateM gr; activeYearsStartDate Mgr; birthDate; birthName; birthPlace;deathDate; deathPlace;profession ;spouse. <u>place:</u> address; agglomerationPopulat ionTotal; alternativeName; altitude; area; city; climate; closeTo; code; country; dc: title; dct: description; dct: extent;dct:spatial;for merName;foundedBy; founder;foundingDate .

APÊNDICE E - Tabela de Classes e Ranking das Principais Propriedades na Wikidata, Europeana e DBpedia.

Classes	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeana	Propriedades Associadas na DBpedia
person (wikidata); edm:Agent (europeana); foaf:Person/ person (DBpedia)	gender given name family name also known as date of birth x biographic information x x x	rdaGr2: gender foaf:name x owl:sameAs rdaGr2: dateOfBirth rdaGr2: dateOfDeath rdaGr2:biographicalInfor mation rdaGr2:professionOrOccu pation rdaGr2: placeOfBirth rdaGr2: placeOfDeath x	foaf:gender foaf:name; birthname foaf:familyName alternativeName birthdate deathDate description profession birthplace deathplace spouse
geographic location (wikidata); edm:Place (europeana) place (DBpedia)	label also known as description coordinate location country x x x x x x x x	skos:prefLabel skos:altLabel; owl:sameAs skos:note wgs84_pos: lat;wgs84_pos:long x x x x x x x x	dc: title alternativeName dct: description kindOfCoordinate/ location country city foundedBy foundingDate code agglomerationPopulati onTotal area climate address
natural product/prod uct (wikidata); edm:Physical Thing (europeana); product (DBpedia)	product naming said to be the same as equivalent class has quality x image has quality x	dc:title dcterms:alternative edm:isSimilarTo dc:description dc:creator x x x	dc:title alternativeName owl:sameAs dct: description manufacturer picture/foaf:img ingredient origin
primary raw material /product (wikidata); edm:Physical	product naming said to be the same as equivalent class x image	dc:title dcterms:alternative edm:isSimilarTo dc:description x	dc:title alternativeName owl:sameAs dct: description picture/foaf:img

thing (europeana); product/ plant (DBpedia)	has quality x has quality x	x x x x	material origin originallyUsedFor cultivatedVariety
ship (wikidata); edm:Physical thing (europeana); ship (DBpedia)	x equivalent class has quality (ship type) x homeport number of masts image	dc:title edm:isSimilarTo edm:hasType dc:description x x x	x x x x homeport x x
religious military order/religio us order (wikidata); skos:Concept (europeana); religious organisation (DBpedia)	x x image x x x x	skos:prefLabel skos:altLabel x x x x skos:note	x x x formationDate foundationPlace organisationMember x
payment system (wikidata); skos:Concept (europeana); system of law (DBpedia)	x x instance of (type of system) x	skos:prefLabel skos:altLabel x skos:note	name x domain comment
handicraft (wikidata); skos:Concept (europeana); activity (DBpedia)	x said to be the same as image practiced by x x x x	skos:prefLabel skos:altLabel x x skos:note x x x x	name x x x x dct:description country material colour
incentive program (wikidata); skos:Concept (europeana); project (DBpedia)	x subclass of (subsidy) x x x x wikimedia import URL	skos:prefLabel x x x x skos:note x	x x projectStartDate projectEndDate projectObjective x x
nautical chart (wikidata);	x x subclass of (map)	skos:prefLabel skos:altLabel x	name alternativeName x

skos:Concept (europeana); thing (DBpedia)	x x x x x	x x x x skos:note	creator material date dct:description x
papal bull (wikidata); skos:Concept (europeana); document (DBpedia)	x x x x x image has part x	skos:prefLabel skos:altLabel x x x x x x skos:note	originalTitle alternative title author dc:description originalLanguage completionDate x x x
auto-da-fé (wikidata); skos:Concept (europeana); historical event (DBpedia)	x x x x described by source	skos:prefLabel skos:altLabel x skos:note x	name x domain comment x
tribunal do santo oficio (wikidata); skos:Concept (europeana); historical event (DBpedia)	x x part of (portuguese inquisition) country location dissolved, abolished or demolished date x	skos:prefLabel skos:altLabel skos:related x x x skos:note	name x domain x x x comment
code of law (wikidata); skos:Concept (europeana); document (DBpedia)	x x subclass of (legislation; book; compilation) x x x x	skos:prefLabel skos:altLabel skos:related x x x skos:note	originalTitle alternative title x author originalLanguage completionDate dc:description
earthquake (wikidata); skos:Concept (europeana); natural event (DBpedia)	x x subclass of (natural disaster; geologic activity) image	skos:prefLabel skos:altLabel skos:related x	x x x x

	has effect measurement scale x x x x	x x x x skos:note	damage x endDate startDate followingEvent x
religious ceremony (wikidata); skos:Concept (europeana); event (DBpedia)	x x subclass of (ceremony; religious concept) x x x x x x x x	skos:prefLabel skos:altLabel skos:related x x x x x x x skos:note	x x x abstract date diocese location participant religion religious head x

APÊNDICE F - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata, Europeia e DBpedia.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na Wikidata	Propriedades Associadas na Europeia	Propriedades Associadas na DBpedia
person (wikidata); edm:Agent (europeana); foaf:Person / person (DBpedia)	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	gender given name family name also known as date of birth x biographic information x x x x	rdaGr2: gender foaf:name x owl:sameAs rdaGr2:dateOfBirth h rdaGr2:dateOfDeath h rdaGr2:biographicalInformation rdaGr2:professionOrOccupation rdaGr2: placeOfBirth rdaGr2: placeOfDeath x	foaf:gender foaf:name;birthname foaf:familyName alternativeName birthdate deathDate description profession birthplace deathplace spouse
geographic location (wikidata); edm:Place (europeana); place (DBpedia)	Rio de Janeiro; Convento de São Bernardo (=Mosteiro de Nossa Senhora da Conceição); China	label also known as description coordinate location country x x x x x x x x	skos:prefLabel skos:altLabel/owl:sameAs skos:note wgs84_pos: lat; wgs84_pos: long x x x x x x x x x	dc: title alternativeName dct: description kindOfCoordinate country city foundedBy foundingDate code agglomerationPopulationTotal area climate address
natural product/product (wikidata);	atanados; vaquetas;	product naming said to be the same as equivalent class	dc:title dcterms:alternative edm:isSimilarTo	dc:title alternativeName owl:sameAs

edm:PhysicalThing (europeana) ; product (DBpedia)	pano de cafre	has quality x image has quality x	dc:description dc:creator x x x	dct:description manufacturer picture/foaf:img ingredient origin
primary raw material /product (wikidata); edm:PhysicalThing (europeana) ; product/ plant (DBpedia)	salsaparrilha; vinhático; pau de tatajuba	product naming said to be the same as equivalent class x image has quality x has quality x	dc:title dcterms:alternative edm:isSimilarTo dc:description x x x x	dc:title alternativeName owl:sameAs dct:description picture/foaf:img material origin originallyUsedFor cultivatedVariety
ship (wikidata); edm:PhysicalThing (europeana) ; ship (DBpedia)	corveta; bergantim; galera	x equivalent class has quality (ship type) x homeport number of masts image	dc:title edm:isSimilarTo edm:hasType dc:description x x x	x x x x homeport x x
religious military order/religious order (wikidata); skos:Concept (europeana) ; religious organisation (DBpedia)	ordem de Cister; ordem de Cristo	x x image x x x x	skos:prefLabel skos:altLabel x x x x skos:note	x x x formationDate foundationPlace organisationMember x
payment system (wikidata); skos:Concept (europeana) ; system of law (DBpedia)	precatório	x x instance of (type of system) x	skos:prefLabel skos:altLabel x skos:note	name x domain comment
handicraft(wikidata); skos:Concept (europeana)	louça preta	x said to be the same as image practiced by x	skos:prefLabel skos:altLabel x x skos:note	name x x x dct:description

; activity (DBpedia)		x x x	x x x	country material colour
incentive program (wikidata); skos:Concept (europeana) ; project (DBpedia)	programa ADAI	x subclass of (subsidy) x x x x wikimedia importURL	skos:prefLabel x x x x skos:note x	x x projectStartDate projectEndDate projectObjective x x
nautical chart (wikidata); skos:Concept (europeana) ; thing (DBpedia)	carta náutica	x x subclass of (map) x x x x x x	skos:prefLabel skos:altLabel x x x x x skos:note	name alternativeName x creator material date dct:description x
papal bull (wikidata); skos:Concept (europeana) ; document(DBpedia)	bula (pontificia)	x x x x x image has part x	skos:prefLabel skos:altLabel x x x x x x skos:note	originalTitle alternative title author dc:description originalLanguage completionDate x x x
auto-da-fé (wikidata); skos:Concept (europeana) ; historical event (DBpedia)	auto-de-fé	x x x x described by source	skos:prefLabel skos:altLabel x skos:note x	name x domain comment x
tribunal do santo ofício (wikidata); skos:Concept (europeana) ; historical event (DBpedia)	tribunal do santo ofício	x x part of (portuguese inquisition) country location dissolved, abolished or demolished date x	skos:prefLabel skos:altLabel skos:related x x x skos:note	name x domain x x x comment

code of law (wikidata); skos:Concept (europeana) ; document (DBpedia)	ordenações manuelinas; ordenações filipinas;	x x subclass of (legislation; book; compilation) x x x x	skos:prefLabel skos:altLabel skos:related x x x skos:note	originalTitle alternative title x author originalLanguage completionDate dc:description
earthquake (wikidata); skos:Concept (europeana) ; natural event (DBpedia)	terramoto (1 de Novembro de 1755)	x x subclass of (natural disaster; geologic activity) image has effect measurement scale x x x x	skos:prefLabel skos:altLabel skos:related x x x x x x skos:note	X x x x damage x endDate startDate followingEvent x
religious ceremony (wikidata); skos:Concept (europeana) ; event (DBpedia)	batismos	x x subclass of (ceremony; religious concept) x x x x x x x x	skos:prefLabel skos:altLabel skos:related x x x x x x x skos:note	x x x abstract date diocese location participant religion religious head x

APÊNDICE G - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na Wikidata Após Testes no Wikidata Query Service.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na Wikidata	
person	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	gender birth name given name family name also known as date of birth	date of death occupation place of birth place of death spouse image
geographic location	Rio de Janeiro; Convento de São Bernardo (=Mosteiro de Nossa Senhora da Conceição); China	label also known as coordinate location location country capital shares border with image	code population area koppen climatic classification address continent
natural product/product	atanados; vaquetas; pano de cafre	alternativeName equivalent class manufacturer image raw material processed	source of material fabrication method
primary raw material/product	salsaparrilha; vinhático; pau de tatajuba	alternativeName equivalent class image made from material source of material	has use
ship	corveta; bergantim; galera	equivalent class vessel class homeport number of masts image	
religious military order/religious order	ordem de Cister; ordem de Cristo	alternativeName image inception location of formation member of military unit	foundational text founded by member count dissolved, abolished or demolished date
payment system	precatório	alternativeName	instance of (type of system)
handicraft	louça preta	alternativeName image practiced by	country of origin made from material fabrication method

incentive program	programa ADAI	subclass of (subsidy) inception dissolved, abolished or demolished date	objective of project or action wikimedia import URL country of origin
nautical chart	carta náutica	title alternativeName subclass of (map) creator made from material inception depicts inventory number	collection location image copyright status width length made from material
papal bull	bula (pontifícia)	title alternativeName author language of work or name inception image copyright status publication date main subject	full work available at URL place of publication location length width collection country of origin made from material
auto-da-fé	auto-de-fé	alternativeName described by source country of origin part of inception location point in time	organizer participant number of deaths image dissolved, abolished or demolished date
tribunal do santo ofício	Tribunal do Santo Ofício	alternativeName part of (Portuguese inquisition) country of origin location inception dissolved, abolished or demolished date	point in time organizer participant number of deaths image location of formation member of
code of law	Ordenações Manuelinas; Ordenações Filipinas	title alternativeName subclass of (legislation; book; compilation) author language of work or name applies to jurisdiction country has part or parts image full work available at URL	inception copyright status publication date main subject place of publication effective date length width collection made from material

earthquake	terramoto (1 de Novembro de 1755)	alternativeName subclass of (natural disaster; geologic activity) image damaged has effect measurement scale inception immediate cause of	described by source country location point in time number of deaths cost of damage has cause duration
religious ceremony	batismos	alternativeName subclass of (ceremony; religious concept) point in time diocese	location participant religion image street address

APÊNDICE H - Tabela de Classes, Exemplos de Entidades dos Registos e Ranking das Principais Propriedades na DBpedia Após Testes no DBpedia SPARQL Explorer.

Classes	Exemplos de Entidades dos Registos	Propriedades Associadas na DBpedia	
foaf:Person / person	Acácio José Correia; Mouzinho da Silveira; Papa Adriano VI	foaf:gender foaf:name;birthname foaf:familyName alternativeName birthDate deathDate	description profession birthplace deathplace picture/foaf:imgpouse
place	Rio de Janeiro; Convento de São Bernardo (=Mosteiro de Nossa Senhora da Conceição); China	dc: title alternativeName dct: description kindOfCoordinate country city foundedBy foundingDate	code agglomerationPopulationTotal area climate address picture/foaf:imgpouse
product	atanados; vaquetas; pano de cafre	dc:title alternativeName owl:sameAs dct: description manufacturer	picture/foaf:img ingredient origin
product/ plant	salsaparrilha; vinhático; pau de tatajuba	dc:title alternativeName owl:sameAs dct: description picture/foaf:img	material origin originallyUsedFor cultivatedVariety
ship	corveta; bergantim; galera	dc: title owl:sameAs rdf:type dct: description homeport	picture/foaf:img
religious organisation	ordem de Cister; ordem de Cristo	rdfs:label alternativeName picture/foaf:img formationDate	foundationPlace organisationMember comment
system of law	precatório	rdfs:label/name alternativeName	domain comment
activity	louça preta	rdfs:label/name owl:sameAs picture/foaf:img dct:description	country material colour comment
project	programa ADAI	rdfs:label rdfs:subClassOf projectStartDate	projectEndDate projectObjective rdfs:comment

thing	carta náutica	rdfs:label/name alternativeName rdfs:subClassOf picture/foaf:img creator	material date dct:description comment
document	bula (pontifícia)	originalTitle alternative title author dc:description	originalLanguage completionDate picture/foaf:img comment
historical event	auto-de-fé	rdfs:label/name alternativeName domain	picture/foaf:img rdfs:comment
historical event	tribunal do santo ofício	rdfs:label/name alternativeName domain	picture/foaf:img rdfs:comment
document	ordenações manuelinas; ordenações filipinas	originalTitle alternative title rdfs:subClassOf author picture/foaf:img	originalLanguage completionDate dc:description comment
natural event	terramoto (1 de Novembro de 1755)	rdfs:label/name alternativeName rdfs:subClassOf picture/foaf:img damage	endDate startDate followingEvent comment
event	batismos	rdfs:label/name alternativeName rdfs:subClassOf abstract date diocese	location participant religion religious head picture/foaf:img comment