

Results of Tests of Digitarq Archive Record Entities and Their Respective Properties in Wikidata and in Dbpedia, Performed in the Wikidata Query Service and in the Dbpedia SPARQL Explorer.

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 was to provide additional information to the archival interfaces. After selecting the entities in the records, Wikidata and DBpedia 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 and DBpedia SPARQL Explorer. 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.

In archival description some of the descriptive elements include information about people, places, events, organizations, which are entities that archivists consider important to be presented in a structured way. Taking into consideration these aspects and the textual information of the record elements, the classes of entities considered of interest for expansion were: person, place, product, vessel, organization, legal system, activity, project, thing, document, and event.

The tables below show the entities and properties selected for testing in Wikidata Query Service and DBpedia SPARQL Explorer. The goal was to verify if the tests confirmed the relevance of the chosen properties and retrieved meaningful results from them, enabling us to conclude that these properties carry important information for understanding and explaining

the respective entities. Moreover, the fact that information about a particular entity exists on the web makes it easier to fit these entities into the user interfaces that are being designed.

Class Table and Ranking of the Top Properties in Wikidata After Tests in the Wikidata Query Service.

Classes	Associated Properties in Wikidata	
person	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	label also known as coordinate location location country capital shares border with image	code population area koppen climatic classification address continent
natural product/product	alternativeName equivalent class manufacturer image raw material processed	source of material fabrication method
primary raw material /product	alternativeName equivalent class image made from material source of material	has use
ship	equivalent class vessel class homeport number of masts image	
religious military order/religious order	alternativeName image inception location of formation member of military unit	foundational text founded by member count dissolved, abolished or demolished date
payment system	alternativeName	instance of (type of system)
handicraft	alternativeName image practiced by	country of origin made from material fabrication method

incentive program	subclass of (subsidy) inception dissolved, abolished or demolished date	objective of project or action wikimedia import URL country of origin
nautical chart	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	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	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	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	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	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	alternativeName subclass of (ceremony; religious concept) point in time diocese	location participant religion image street address

Class Table and Ranking of the Top Properties in DBpedia After Tests in the DBpedia SPARQL Explorer.

Classes	Associated Properties in DBpedia	
foaf:Person/ person	foaf:gender foaf:name;birthname foaf:familyName alternativeName birthDate deathDate	description profession birthplace deathplace picture/foaf:img spouse
place	dc: title alternativeName dct: description kindOfCoordinate country city foundedBy foundingDate	code agglomerationPopulationTotal area climate address picture/foaf:imgpouse
product	dc:title alternativeName owl:sameAs dct: description manufacturer	picture/foaf:img ingredient origin
product/ plant	dc:title alternativeName owl:sameAs dct: description picture/foaf:img	material origin originallyUsedFor cultivatedVariety
ship	dc: title owl:sameAs rdf:type dct: description homeport	picture/foaf:img

religious organisation	rdfs:label alternativeName picture/foaf:img formationDate	foundationPlace organisationMember comment
system of law	rdfs:label/name alternativeName	domain comment
activity	rdfs:label/name owl:sameAs picture/foaf:img dct:description	country material colour comment
project	rdfs:label rdfs:subClassOf projectStartDate	projectEndDate projectObjective rdfs:comment
thing	rdfs:label/name alternativeName rdfs:subClassOf picture/foaf:img creator	material date dct:description comment
document	originalTitle alternative title author dc:description	originalLanguage completionDate picture/foaf:img comment
historical event	rdfs:label/name alternativeName domain	picture/foaf:img rdfs:comment
historical event	rdfs:label/name alternativeName domain	picture/foaf:img rdfs:comment
document	originalTitle alternative title rdfs:subClassOf author picture/foaf:img	originalLanguage completionDate dc:description comment
natural event	rdfs:label/name alternativeName rdfs:subClassOf picture/foaf:img damage	endDate startDate followingEvent comment
event	rdfs:label/name alternativeName rdfs:subClassOf abstract date diocese	location participant religion religious head picture/foaf:img comment