ArchOnto, a CIDOC-CRM-based Linked Data Model for the Portuguese Archives [⋆]

Inês Koch[1][0000−0002−9363−9713], Cristina Ribeiro[1][0000−0001−6150−0090], and Carla Teixeira Lopes[1][0000−0002−4202−791X]

INESC-TEC, Faculty of Engineering of the University of Porto, Porto, Portugal

Abstract. Archives are faced with great challenges due to the vast amounts of data they have to curate. New data models are required, and work is underway. The International Council on Archives is creating the RiC-CM (Records in Context), and there is a long line of work in museums with the CIDOC-CRM (CIDOC Conceptual Reference Model). Both models are based on ontologies to represent cultural heritage data and link them to other information. The Portuguese national archives hold a collection with over 3.5 million metadata records, described with the ISAD(G) standard. The archives are designing a new linked data model and a technological platform with applications for archive contributors, archivists, and the public. The current work extends CIDOC-CRM into ArchOnto, an ontology-based model for archives. The model defines the relevant archival entities and properties and will be used to migrate existing records. ArchOnto accommodates the existing ISAD(G) information and takes into account its implementation with current technologies. The model is evaluated with records from representative fonds. After the test on these samples, the model is ready to be populated with the semi-automatic transformation of the ISAD records. The evaluation of the model and the population strategies will proceed with experiments involving professional and lay users.

Keywords: EPISA · Linked Data · Archival Description · CIDOC-CRM · ArchOnto

1 Introduction

The Portuguese National Archives, Torre do Tombo (ANTT), one of the oldest institutions in Portugal, curates a unique collection of historical and contemporary objects that it has been accumulating since the 9th century. With a large number of documents in its custody, including large volumes of administrative data, organized in series and covering extended periods and the evolution of the institutions that create them, descriptive metadata is essential to the management of the archives. As the content of the documents is currently not searchable,

---

metadata is the basis for browsing and querying the archives remotely, and remote access now exceeds the direct contact with the archives by several orders of magnitude. Considering the central role of metadata, it is essential that the information in the descriptive records is thoroughly explored and it is recognized that this is not always the case.

In a changing world where open information is supposed to be accessible to all, archives are also redefining their mission and considering the access to documents on a par with the preservation of the information therein. The ANTT is, therefore, concerned with the transformation of the archives, which involves the development of a new data model, a new information system, and new workflows and services to the public. A new data model is central to this transformation, as it stands at the core of the whole design and is instrumental in many aspects: in the expressiveness of the metadata created in the archives, in the ability to integrate existing records into the new model, in the fitness of the model to interoperate with other systems. The archives are no longer isolated, they need to link documents and their descriptions with external data that can provide context and enrich their contents.

The work described here is part of the EPISA project (Entity and Property Inference for Semantic Archives), a research project that brings together a team from Information and Computer Science and the archival experts from ANTT. EPISA intends to design and prototype an open-source knowledge platform representing archival information on a linked data model. Additionally, the project will work on the existing records to extract the relevant entities and their properties and take advantage of the wealth of information built by specialists over the years.

The project will assist the ANTT in moving from the ISAD(G) multilevel description model to a graph data model based on state-of-the-art technologies that can provide data for Artificial Intelligence algorithms to extract resources and infer relationships between those resources, having the current textual descriptions as a starting point.

The ArchOnto model is presented here and evaluated based on a selection of archival records. These records include fonds with a large volume of data, like parochial records, and fonds of unique objects, the so-called treasures.

2   Standards for Cultural Heritage

Considering the initiative of creating a new data model for archives, it is essential to have a knowledge of the standards that are currently used for cultural heritage.

The General International Standard for Archival Description, ISAD(G), and its associated standards issued by the International Council on Archives (ICA) is widely adopted and has been the basis for DigitArq [9], the platform currently used for archival description by the ANTT. The ISAD(G) [6] is an archival description standard that is characterized by a multilevel structure that allows the archival description to be made from general to specific, representing the context and hierarchical structure of the fonds and its components.

It is also necessary to consider data models that include the atomization of cultural heritage records, i.e. the transformation of flat textual fields into structured subgraphs with meaningful entities, and their connection to external information as Linked Open Data. In what follows, we will have a sense of what already exists and how they can contribute to our goals.

2.1 CIDOC-CRM

The CIDOC-CRM (Conceptual Reference Model) [4], a formal ontology, was developed by the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM). This model, which aims to exchange, mediate, and integrate heterogeneous sources of information related to cultural heritage, is being actively developed by the CRM-SIG (CRM - Special Interest Group). After several changes over a period of over 10 years, it was considered an ISO standard in 2006 (ISO 21127), and a 2014 version is under review (ISO 21127: 2014). Building on the concept of the event, CIDOC-CRM is quite complete with regard to the representation of people, places, and time periods, concepts that are also central to the archival description.

Due to its origin in the museum community, it is in this domain that examples of institutions that applied the model are found. These include the British Museum[1] and the Museo del Prado[2]. Although CIDOC-CRM is recognized as the base model for the implementation of linked data in these institutions, it is articulated with other models, as in the case of Museo del Prado, where FRBR (Functional Requirements for Bibliographic Records, a model that originated in the library community) [5] is applied, among some more specific vocabularies.

In addition to museums, CIDOC-CRM has already been applied in other areas, namely in Archaeological Heritage, where it has been promoted in the context of the Ariadne Project [1]. Within the archives, work has been carried out mapping the EAD (Encoded Archival Description) standard to CIDOC-CRM. EAD is an XML language that represents the ISAD(G) standard. The mapping of EAD to CIDOC-CRM [2] took into account concepts that are extremely relevant for the transformation of the data model currently used by the Portuguese National Archives. Among these is the level of description, central to ISAD(G). These first experiments were focused on archival requirements but did not evolve to the proposal of a more substantial model.

2.2 RiC-CM

A new data model is currently under development in the area of archives that incorporates the existing archival standards, following their principles in a conceptual data model, the RiC-CM [7]. In addition to the data model, an ontology, called RiC-O (Records in Context Ontology), is being defined by the Expert

---

[1] https://www.britishmuseum.org/
[2] https://www.museodelprado.es/modelo-semantico-digital/modelo-ontologico

Group on Archival Description (EGAD) from the ICA. This model aims to represent all archival concepts, taking into account the main descriptive entities. Thus, properties, classes, and attributes that represent the essential relationships present in the archives are considered. This opens the ground for the cooperation between our project and the RiC-CM initiative, in the same spirit we have adopted with regard to CIDOC-CRM, and more so given the fact that this proposal is aimed specifically at archives.

2.3 Comparison

A comparison of standards is summarized in Table 1 that highlights several commonalities and differences. The models originating in the archives take into account the hierarchical structure intrinsic to the domain. CIDOC-CRM and RiC-CM are both based on semantic web concepts and therefore aim to represent cultural heritage data as linked data. While the ISAD(G) standard has a limited number of elements for which values typically have a rich structure, the more recent models have a number of properties an order of magnitude larger, attesting to their more atomized representation of knowledge. The number of properties presented for RiC is taken from the RiC-O ontology. All models have institutional support in the corresponding working groups under well-established cultural heritage institutions, and while ISAD(G) currently supports the implementation of archival information systems, the other still lack the test of actual deployment.

Table 1. Comparison of ISAD(G), CIDOC-CRM and RiC-CM

	ISAD(G)	CIDOC-CRM	RiC-CM
Hierarchical	✔	✖	✔
LOD	✖	✔	✔
Ontology	✖	✔	✔
Number of properties	26 elements	285 properties	449 properties
Institutional support	ICA (+25 years)	ICOM (+15 years)	ICA (5 years)
Implementation on archives	Custom	(none known)	None
Most recent version	2000	2020	2020
Support	ICA	CIDOC-CRM SIG	ICA (EGAD)

3 ArchOnto, a modular ontology for archives

The Portuguese National Archives were early adopters of the ISAD(G) standard. They created a set of rules and recommendations for consistent use of the standard at national and regional level [3]. Moreover, a custom-designed information system was developed in close collaboration with the archives experts and deployed in all archives, enforcing the aggregation of records at national level. After

almost two decades and several system updates, the system no longer supports the requirements of the archives, in more than one aspect: the data model that embodies the ISAD(G) is limited, the technologies that support the information system are too rigid, expensive and difficult to maintain, the mission of the archives has extended into new processes, and more user profiles have to be considered. In the sequel, we focus on the choice of the data model.

It was clear from the requirements of the archives that the new model has to be more fine-grained and able to identify documents but also events, people, and their roles and connections. This is in line with recent work in knowledge graphs and linked open data, where bits and pieces of information from various sources are linked using properties defined in many different contexts. The information in the archives is no longer regarded as isolated, but rather able to connect to information created by other instances.

Looking for existing models in this line, we considered the RiC-CM. However, as this model was still in a preliminary phase in 2018 when this work started, it was necessary to find a more mature model, and the CIDOC-CRM stood as a strong candidate.

Although CIDOC-CRM has evolved over more than a decade and is now a stable model, it is in steady development, and the EPISA project team has been following its evolution. The ArchOnto model proposed here takes into account the current version of CIDOC-CRM, version 6.2.7 [4].

3.1   The process of adapting CIDOC-CRM for archival use

With CIDOC-CRM as the foundation, we began to structure the ArchOnto model, considering the information present in the ISAD(G) descriptors and how it might be mapped to CIDOC-CRM. Due to the expressiveness of CIDOC-CRM, we concluded that most of the information present in the ISAD(G) elements would be easily represented there. However, we also found that not all information from ISAD(G) would turn into CIDOC-CRM.

A data model for archives began to be composed based on the CIDOC-CRM, following the general principle that the elements of the model should be implemented as a knowledge graph via a graph database. The model is represented in the ArchOnto ontology, which aims to include information from all the records in the Portuguese National Archives and is being embedded in ArchGraph [8], a knowledge graph that will support the archival information system. We will focus here on ArchOnto and its evolution, but the operational concerns raised by ArchGraph have been essential in the design of ArchOnto.

The first approach to ArchOnto [8] aimed at using just CIDOC-CRM, including its recommendations for the representation of non-binary relationships, and some extensions already validated by the CIDOC-CRM. As CIDOC-CRM was created for museums, there are core concepts for the archives that are not present in this model, such as the level of description. The process of adapting CIDOC-CRM for archival use took several steps, detailed as follows.

CIDOC-CRM extensions Our first approach was to create Data Property extensions to cope with the limited number of data properties in CIDOC-CRM, using them to capture the semantics of the elements from the descriptions we observed in the various archival objects [8]. Note that most of CIDOC-CRM properties are Object Property, used to relate individuals.

Most of these data properties were created to accomodate information from the text fields of the ISAD(G) elements.

From CIDOC-CRM extensions to separate ontologies The CIDOC-CRM extension approach required a large number of data properties, so we evolved to create the ISAD Ontology to put them together, rather than including them as CIDOC-CRM extensions. This separate ontology is then imported into the ArchOnto model. Besides these properties, subproperties of *P3 has note*, the ISAD Ontology contains all elements of ISAD(G), which will be atomized with CIDOC-CRM in order to have finer-grained descriptions. Table 2 shows some examples of data properties that were captured first as CIDOC-CRM extensions and then as part of the ISAD Ontology.

Table 2. CIDOC-CRM Data Properties extensions to ISAD Ontology

CIDOC-CRM Extension	ISAD Ontology
ARP1 has administrative history	ISAD7 has administrative history
ARP2 has archival history	ISAD8 has archival history
ARP3 has scope	ISAD9 has scope

Classes and object properties that existed as CIDOC-CRM extensions in the preliminary ontology were also moved to the N-ary ontology (presented below). This organization in separate ontologies is more flexible in the sense that, if CIDOC-CRM changes and these ontologies are no longer necessary, they can be dropped.

Remaining CIDOC-CRM extensions Despite the move of properties and classes to other ontologies, there were cases of Classes and Object Properties that remained as extensions of the CIDOC-CRM model. These include the ones considered essential to archives, such as the level of description, which is represented by the *ARE1 Level of Description* class and the *ARP12 has level of description*, *ARP8 upper level* and *ARP9 lower level* properties (see Fig. 1). The basic principles of the archival organization require that each *Unit of Description* be assigned a description level and that levels be organized hierarchically. This takes into account organization principles that are well established in the archives, but can also be considered an enduring principle for large collections, in that description can be performed for more or less vast collections of documents and then inherited if their organization is maintained.
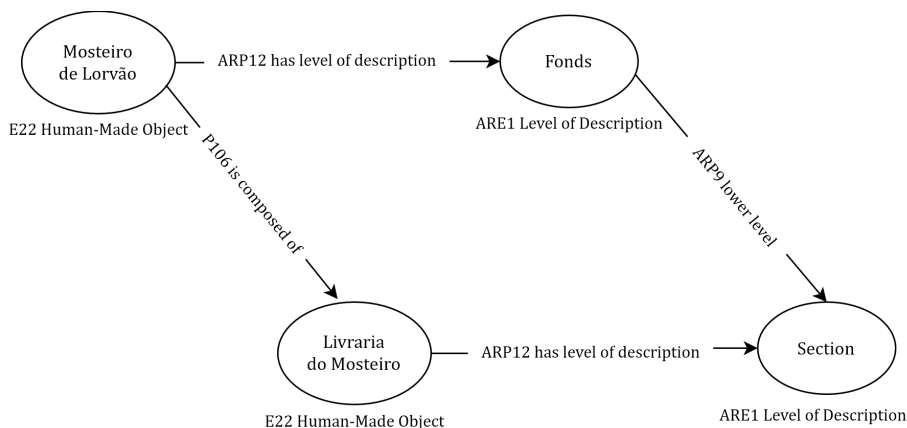
Fig. 1. Levels of description in CIDOC-CRM. Original record at https://digitarq.arquivos.pt/details?id=4381091.

3.2   The ontologies in ArchOnto

With a better knowledge of CIDOC-CRM and archival standards, some decisions made the model for the archives more complete and, therefore, more able to represent the universe of archives. As such, ArchOnto went from an ontology where CIDOC-CRM was imported with some extensions to an ontology where more ontologies are imported to represent more accurately the existing archival records. This is quite in line with the semantic web principles of reusing existing ontologies whenever they are available and working on the additional concepts for our domain.

ArchOnto[3] currently has five ontologies at its base, which complement each other. Besides the CIDOC-CRM (base ontology), we will briefly describe N-ary, the ISAD Ontology, DataObject, and an ontology for the connection between CIDOC-CRM and DataObject.

The N-ary ontology was created taking into account the CIDOC-CRM recommendations[4] for the representation of tuples with an arity higher than two. With this ontology, it is possible to represent all instances in which it is necessary to build associations that are not binary, i.e., that connect more than two individuals.

The ISAD Ontology is in place to represent the elements of the ISAD(G) standard without atomization. It has as its base data properties that capture each of the ISAD(G) elements, thus maintaining the information from the original records, making sure that what was previously described with ISAD(G) is not lost when atomized for ArchOnto. It also allows, whenever necessary, the

---

[3] OWL version available on GitHub - ArchOnto2020: https://github.com/feup-infolab/archontology

[4] http://www.cidoc-crm.org/sites/default/files/Roles.pdf

validation of the contents that have been atomized, checking if they comply with the information present in the ISAD(G) description. With this ontology, it will be possible to maintain the interoperability of all data represented in ISAD(G). Moreover, as information extraction algorithms will be applied to the ISAD(G) fields, it is easier to expose all information in a single system to be able to compare legacy descriptions with the corresponding atomized subgraphs.

The DataObject ontology is present to validate the literal values used in the properties of the new data model. It has as its base classes and data properties that are used to validate simple types in the ontology, ensuring that validation for each object is performed based on the corresponding class.

Finally, to link the ontology of CIDOC-CRM with the DataObject ontology, we created an ontology with a hasValue property that connects both ontologies.

3.3   Issues faced in the adaptation of CIDOC-CRM for archival use

In the adaptation of CIDOC-CRM to the archives, we had to make sure that ISAD(G) descriptions were mapped to the new model. In this mapping process, we tried, as much as possible, to use CIDOC-CRM features. However, when this model was unable to satisfy our requirements, we created classes and properties to represent the ISAD(G) attributes, making their semantics explicit in the new model.

Types As we explored the mapping between ISAD(G) and CIDOC-CRM, we became aware that the CIDOC-CRM *E55 Type* class was extensively used. This class is very versatile and is used with many of the concepts that are present in ISAD(G) and not in CIDOC-CRM. The broad use of this class did not contribute to separate the specific semantics of each concept that was being represented. Naturally, if used in a large number of concepts, they would no longer be differentiated.

To face this challenge, we decided to create subclasses of *E55 Type*, to have specific types to distinguish identifiers from personal names, date from language or legal status, while considering the concepts of ISAD(G). Many of these concepts already correspond to controlled vocabularies in the archives, and some are listed in Table 3, as well as the subclasses that represent them.

Table 3. Proposal of CIDOC-CRM Extensions - Some subclasses of *E55 Type*

Proposal of CIDOC-CRM Extension	Examples
ARE5 Identifier Type	Reference Code, Physical Location
ARE6 Date Type	Predominant Date
ARE7 Name Type	First name, Surname, Nickname
ARE8 Role Type	Material Author, Bride, Godfather

Conceptual Object vs Physical Object While applying CIDOC-CRM, we noticed the existence of concepts that are not central to ISAD(G) and archival practice. The distinction between the physical object and the conceptual object is an example of this. The two concepts emerged with the need to identify the language of a document.

Initially, all the ISAD(G) elements were being mapped as being related to the physical object. However, when mapping the language of a document, we found that in CIDOC-CRM it should not be associated with a physical object, but rather to a conceptual object. The language should be related to the expression of a work, an abstract concept distinct from the material object that embodies such an abstraction. The language should, therefore, appear as a property of a conceptual object.

According to CIDOC-CRM, objects that do not have a physical dimension, but transmit information about the physical world, are considered conceptual objects. These objects cannot be destroyed, they exist as long as an individual can conceive them through their memory.

In the mapping process, a substantial effort went into exploring the possibilities of turning extensive textual elements in ISAD(G) into their atomized versions, namely in the association of entities mentioned in the ISAD(G) records with conceptual objects. Among these elements are, in addition to the language, the scope and content, the notes, the publication notes and the access conditions.

As for the physical object, CIDOC-CRM considers it an item of a material nature with clear boundaries that can be independently documented. The physical object, unlike the conceptual object, has a physical dimension and, therefore, can be moved, if its weight allows it. As with the Conceptual Object, there are also ISAD(G) elements that are sources of information to atomize and associate with a physical object. Among these attributes are the titles, the support, the dimension and the location of the documents (in the sense of physical location).

Considering that in the ISAD(G) standard the distinction between the physical and the conceptual object is not explicit, in the ISAD Ontology it was decided to relate all attributes to physical objects.

Validation of data types Data types have a careful treatment in ArchOnto as they stand at the interface between the higher-level concepts of the domain and the implementation and validation that applications are supposed to perform in order to enforce the validity of the knowledge graph. A set of basic data types have therefore been represented in the DataObject ontology and articulated with CIDOC-CRM for the properties whose ranges involve objects such as strings, dates, or identifiers.

We can take dates as an example, as they are omnipresent in archival records, appear under different formats, and require strong validation. CIDOC-CRM provides classes and properties to deal with dates that go as far as considering them as individuals. DataObject handles the transition to the actual representation as values with validated formats. Like the dates, the titles also need validation, going through a similar validation process.

In the case of dates, the format presented is intended to be uniform. For that, all mapped dates will have a defined format to specify the date and time when a given event happens, according to the dateTime data type – xsd:dateTime, which has the format "YYYY-MM-DDThh:mm:ss".

To validate dates, it is necessary to use the CIDOC-CRM, the DataObject, and the ontology that links these two ontologies, and it is the DataObject that supports validation of the date format. All dates, according to CIDOC-CRM, must be related to an event, which is the starting point for its validation. Fig. 2 illustrates the validation of a date using CIDOC-CRM classes and properties complemented with those from DataObject.
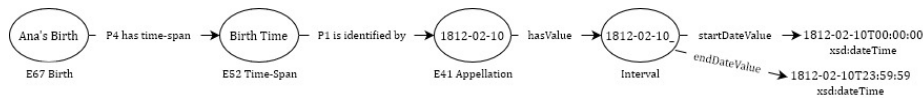

Fig. 2. Validation of a date in CIDOC-CRM. Original record at https://pesquisa.adporto.arquivos.pt/details?id=1374655.

4   Evaluation of CIDOC-CRM for archives

Following the design of the ArchOnto model, the team that is developing the model tested it with several pilot cases extracted from the DigitArq database of archival records. As DigitArq has a great diversity of records, we used records of different kinds of fonds present in this database. Three examples are used in this evaluation - one from a parish fonds, one from judicial records and a unique object, classified as a treasure.

Among the parish fonds, a series of documents related to baptism records[5], where homogeneity of information was observed, were selected. In the judicial records, a document related to an orphan record[6] was chosen, which proved to be quite rich in information. Finally, a treasure[7] was selected, which has a wide variety of information present, making it a very extensive record.

Throughout the ArchOnto mapping, some of the concepts used in the archives, the ones also present in CIDOC-CRM, were mapped directly to the new model. These were the first elements to be mapped, and therefore, to be evaluated, and include the reference code, title, dates, dimension and support, language, and physical location.

From the ISAD(G) elements available, the reference code and the physical location can be considered identifiers of the document, and are therefore mapped through the class *E42 Identifier*. On the other hand, the dimension is mapped

---

[5] Sample baptism record: https://pesquisa.adporto.arquivos.pt/details?id=1374655

[6] https://digitarq.adevr.arquivos.pt/details?id=1174365

[7] https://digitarq.arquivos.pt/details?id=4381091

through the *E54 Dimension* class, support by *E57 Material*, titles by *E35 Title*, dates by *E52 Time-Span*, and the language by *E56 Language*. In Fig. 3, the different ISAD(G) elements can be observed, with the title of the document being portrayed having a formal type, and, therefore, a subclass of *E35 Title* is used to indicate its type, in this case *ARE2 Formal Title*. The dates, as they need validation, are not represented in Fig. 3. Dates give rise to more detailed mini graphs, such as the one illustrated in Fig. 2.
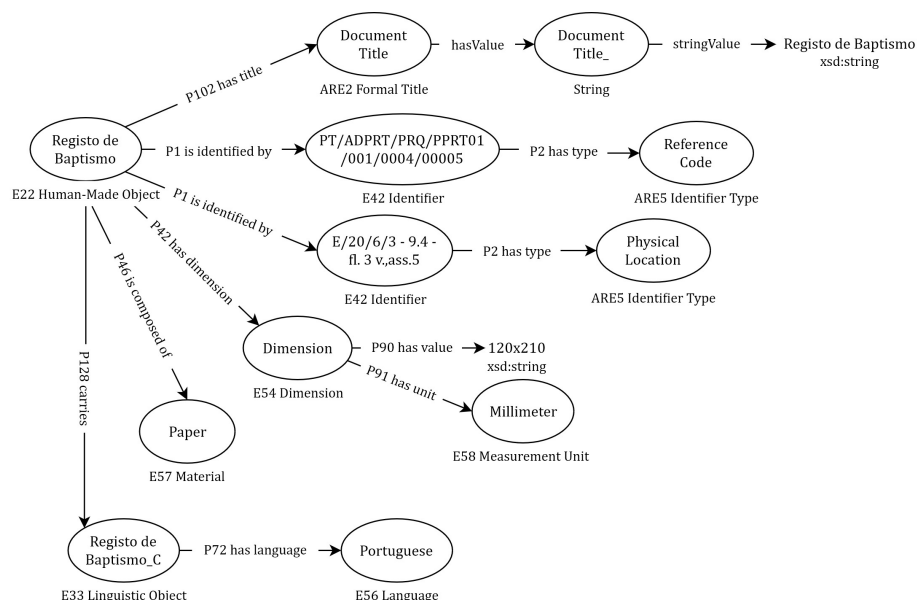

Fig. 3. Evaluation of ISAD(G) concepts in CIDOC-CRM. Original record at https://pesquisa.adporto.arquivos.pt/details?id=1374655.

As the representation of events is central to the CIDOC-CRM, it is essential to evaluate if these can capture some of the concepts that are present in the ISAD(G) standard. The elements where events appear frequently are those that capture the Archival History, Biographical History, and Scope and content.

Several documents that have events present in their contents were analyzed. In these documents, there are events such as birth, death, and marriage. The CIDOC-CRM provides an explicit representation for the first two events through specific classes and properties, but not to the third (marriage).

In the documents used here, the events come from the textual content of the ISAD(G) elements, since they are not identified separately in this standard. It is, therefore, essential to extract these contents, so that they can be represented through CIDOC-CRM. For this preliminary evaluation, the events were manually

identified with the analysis of the records, and their mapping into CIDOC-CRM used the ArchOnto ontology.

As ISAD(G) has descriptive attributes at its base, it is necessary to bear in mind that the events represented may not be the main point of the description, but rather a reference to an event related to the document being described. For example, the *Registo de Baptismo*[8], where the goal is to describe Ana's baptism registry, also mentions the event of her birth, as a secondary event.
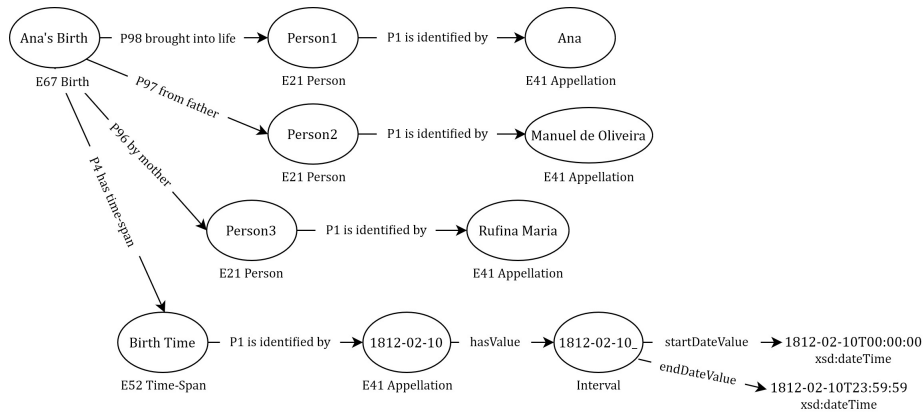

Fig. 4. Birth event at CIDOC-CRM.

Fig. 4 shows part of a graph with the event of Ana's birth, in which the three people who witnessed the event (the baby, the mother, and the father) and the date on which it happened are mentioned. As there is no reference to the time at which the birth occurred, we take into account the time interval in which it may have occurred on the indicated day.

Looking at the description of the document "Processo de inventário orfanológico por óbito de Maria Henriqueta Fragoso Barahona Carvalho e Mira", a record describing the assets related to orphaned children, the events of death and marriage are present. Unlike the death event, which bears the date on which it occurred, the wedding event has to be inferred through the description, since this only refers to Maria Henriqueta as being married to José Paulo.

In Fig. 5, the marriage event is represented, but the fact that the event is marriage has to come from the event object itself, as CIDOC-CRM provides no specific class for this kind of event, as is the case in the birth and death events. With this in mind, the wedding event was mapped based on a ternary relationship, where two people had specific roles in the event. In this case, Person 1 in the role of bride and Person 2 in the role of fiancé. This graph excerpt makes use of the n-ary pattern twice, linking people with events and the roles they play therein.
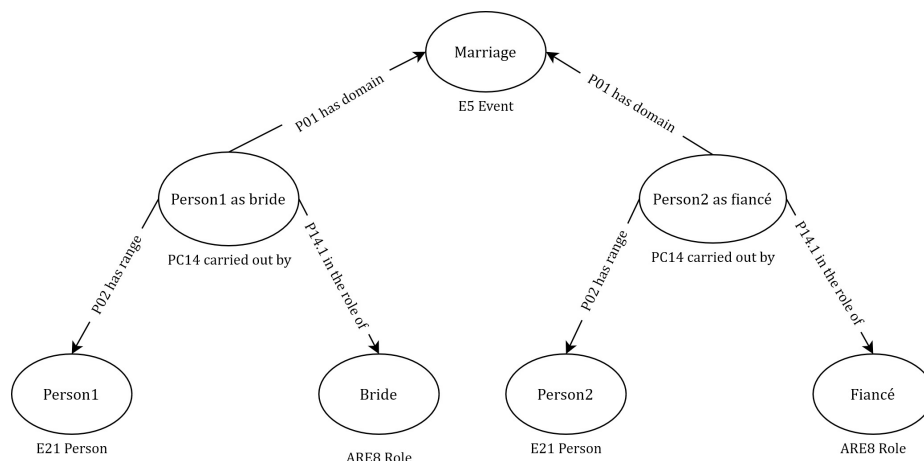
---

[8] Original record at: https://pesquisa.adporto.arquivos.pt/details?id=1374655

Fig. 5.   Marriage   event   at   CIDOC-CRM.   Original   record   at
https://digitarq.adevr.arquivos.pt/details?id=1174365.

5   Conclusions

This work has shown that the CIDOC Conceptual Reference Model can be extended and used as a model for archives. We presented a data model, ArchOnto, represented as an ontology and based on the CIDOC-CRM. Special attention was given to entities and properties that are essential for archives and for the applications that run on their information.

Two principles were followed in this first approach to an archival model. The first is to accommodate existing information. The records created by archival specialists are a wealth of information with high standards. It is essential to turn this information into its Linked Data counterpart without losing its integrity. Moreover, archivists will continue to generate such records, and therefore the new model should be intelligible to them. The second principle is to favour implementation. The model has been developed alongside the selection of the technology stack that is expected to support the new archival information system. Choices took into account the ease of implementation and the extent to which validation of information added to the records can be done at various points: when migrating information from existing records, when archivists create new records, when records are imported into the archives from the public administration sources.

Given the large number and diversity of available records, the first step to validate the model has used samples of records from different fonds, having in common the fact that they are frequently accessed. The first is a set of parish records, for which there is a large number in series that span centuries. This illustrates records with a common format and well-known relations for the individuals involved. The second is a judicial record and the third a record from a Portuguese National Treasure, for which detailed metadata was created by

archivists, linking it to historical sources. At the moment, the records' analysis is ongoing and will be followed by an evaluation of the model performance.

This work illustrates the use of CIDOC-CRM in the archival domain and will be pursued in two main lines. The first is incorporating a large set of documents, in a process that will use a mix of automatic migration and revision by archivists. The second is user testing, and user interfaces are under development considering the professional users but also the growing interest of the public for the archives. This work will continue in close collaboration with the implementation of the knowledge graph.

References

1. ARIADNE: The Way Forward to Digital Archaeology in Europe (2014), https://arachne.dainst.org/project/ariadne
2. Bountouri, L., Gergatsoulis, M.: The semantic mapping of archival metadata to the CIDOC CRM ontology. Journal of Archival Organization **9**(3-4), 174–207 (2011)
3. Direcção Geral de Arquivos, Grupo de Trabalho de Normalização da Descrição em Arquivo: Orientações para a Descrição Arquivística. Direção Geral de Arquivos, Lisboa, 2nd edn. (2007)
4. ICOM/CIDOC CRM Special Interest Group: Definition of the CIDOC Conceptual Reference Model. ICOM, 6.2.7 edn. (2019)
5. IFLA Study Group: Functional Requirements for Bibliographic Records. K.G. Saur Verlag, Munich (2009)
6. International Council on Archives: ISAD(G) Second Edition. International Council on Archives (2000)
7. International Council on Archives: Records in Context: a conceptual model for archival description. International Council on Archives (2016)
8. Koch, I., Freitas, N., Ribeiro, C., Lopes, C.T., da Silva, J.R.: Knowledge Graph Implementation of Archival Descriptions Through CIDOC-CRM. In: Doucet, A., Isaac, A., Golub, K., Aalberg, T., Jatowt, A. (eds.) Digital Libraries for Open Knowledge. pp. 99–106. Springer International Publishing, Cham (2019)
9. Ramalho, J.C., Ferreira, J.C.: DigitArq: creating and managing a digital archive. In: Engelen, J., de Souza Costa, S.M., Moreira, A.C.S. (eds.) Building Digital Bridges: Linking Cultures, Commerce and Science: 8th ICCC/IFIP International Conference on Electronic Publishing (2004)