Applying Records in Contexts in a Federal University Record*

Marcos Vinícius Bittencourt de Souza[1,2][0000−0001−5292−7658] and Daniel Flores[3][0000−0001−8888−2834]

[1] Universidade Federal de Santa Maria
[2] Universidad de Salamanca
`marcos.cpd@ufsm.br`
[3] Universidade Federal Fluminense
`df@id.uff.br`

Abstract. The International Council on Archives, through the EGAD, is developing a conceptual model to replace the current archival description standards, bringing a new vision and methodology to the archival description task. The Records in Contexts will have a multidimensional aspect, arranging the information elements like a graph, with the described entities being the nodes linked by the relationships between them. The present study applies the current draft version (v 0.2) of the conceptual model in the description of a record instance of the Federal University of Santa Maria, projecting which entities can be used to represent the information elements related to the record. The aim of this study is to help institutions, archivists and IT professionals to foresee how the description process and description tools will be impacted by the RiC model in a brief future.

Keywords: Archival description · Conceptual model · Record management · Records in Contexts.

1 Introduction

In 2016, the International Council on Archives (ICA) started the work to build a new international archival description standard, through the Experts Group on Archival Description (EGAD), publishing the first draft of the Records in Contexts. The ICA´s idea is to incorporate a new vision of the archival description task at the same time that updates the currents standards, task recognized as a "daunting challenge" [4, p.1] because of the great diversity of cultures, thoughts and methods used by professionals around the world.

As part of the standard elaboration process, a conceptual model was developed and published as a draft version to the archival community express its opinion and to propose changes. In 2019, it was made available the corresponding model ontology with the conceptual model preview, modified to incorporate the received suggestions [5]. The process of establishment the new standard still is in progress and will have a new stage of consultation with the archival community.

The present study, of qualitative nature, has the aim to analyze the application of RiC model in a Federal University of Santa Maria record instance projecting the possible modifications that the migration of the standard will bring. To this study, it was chosen the born digital record named "Diário de Classe" (Class Diary) that is "[...] the instrument that registers the related activities to the teaching-learning process, containing the essential information about the students frequency, performance, taught content as well as the professor frequency registry"[8, p.69].

2   Records in Contexts - RiC

The publication named "Records in Contexts: A Conceptual Model for Archival Description" [4], made public in 2016 under the authorship of the Experts Group on Archival Description (EGAD), presents the first proposal of a new international archival description model, result of the four years work of this group, previously known as Committee on Best Practices and Standards (CBPS).

The RiC presents the archival entities, their respective attributes and the possible relations between each type of entity, exactly as the Comisión de Normas Españolas de Descripción Archivística [1, p.23] defines an archival description conceptual model. In parallel with the conceptual model elaboration, the EGAD made the corresponding ontology, named RiC-O, that will allow the implementation of the records description and the integration with the computerized systems, as it plans to be made available in the OWL format (Web Ontology Language) and represented in the RDF (Resource Description Format).

Before the RiC, the ISAD(G) standard was designed in the 1990 decade and later, in 2000, its second edition was elaborated with the appropriated updates. Based in a one-dimensional and multilevel approach, the information elements are arranged in a hierarchical way, starting from the most general level, the fond, to the most specific element that represents the record, using the methodology named as "part-to-whole" [3, p.12]. The first goal of ICA, with RiC, is to unify the current standards ISAD(G), ISAAR, ISDF and ISDIAH, adding changes that reflects the new concepts observed in the archives. The centralization of the models tends to make easy the understanding and utilization of the standard by the archivists at the same time that allows the standardization of tools that support the archival description.

Considering the current world, this view may not contemplate the reality in which the born digital records are inserted, given that many times it is difficult to identify the provenance of certain records [7, p.527]. The concern about the digital world and the growing production of records can be seen in the RiC model, which says [4, p.5]:

> Both the emergence of collaborative editing in the networked digital environment, which is creating many records that have complex authorship, and the use of remote storage services that are not fully controlled by the users, raise issues of ownership and custody. All of these observations and more lead to the conclusion that provenance is much more

complex than the long established understanding of fonds: a fonds does not exist in isolation, but within layers of interconnected contexts, past, present, and future.

The multidimensional aspect of the RiC becomes its main difference from previous standards, intending to express in a more real and trustworthy way the environment in which the records are inserted. In contrast to the previous models, which are represented in an tree format that reflects the hierarchy of the archives, the RiC is represented through a non-directional graph, with the nodes being the entities with their properties and the edges being the relationships between the entities, also with their properties. Technologically, one of the solutions to represent the model's graph is the RDF format, which describes nodes and edges as subject-predicate-object tuples reaching the non-hierarchical nature of the model [4, p.9]. Even though the RiC represents a paradigm shift, it allows the description of all archive's entities, but their main difference is the possibility of a reorganization of these entities to adapt the archive to the constant changes of the digital world. According to Llanes-Padrón and Moro-Cabero [7, p.529]:

> The conception of level of arrangement specifying the description levels it is not seen in this new model where the hierarchy is not decisive, but unique representations based on concise relationships between independent entities that are able to capture the differences and complex contexts of creation.

The way the RiC was elaborated indicates the concern with the description of the context in which the records and their components are inserted, reflected by some dedicated entities. Entities like "RiC-E18 Date", "RiC-E22 Place", for example, allows to describe the context through the properties of each one and relating them with the records and other entities already described. Thinking about the dissemination of records, it will be possible to explore the records navigating since some date, for example, improving multiple ways to view the institutional archive.

3   The "Class Diary" Record

The Federal University of Santa Maria (UFSM) has in its classification scheme the record named "Diário de Classe" (Class Diary) that registers the life cycle of the courses offered to the students. The record has information related to the students grades, the content taught in each class, and the students presences and absences of all classes, at the same time that registers the professors activities [8]. All information is filled by the professors in the institution's system named "Portal do Professor" (Professor Portal).

After registering the grades equivalent to students evaluations, the system does the appropriate calculations that informs which students were approved or not and persist the situations of the students, signaling that the professor's job was done successfully [8]. The system generates a record that documents all the

class activity, identifying and gathering all the necessary metadata to ensure its authenticity. The generated record is stored digitally with no need of manual signature and printing [10].

The UFSM, as well as all the other brazilian federal universities, has its classification scheme defined by the government, elaborated by the National Archive (Arquivo Nacional). The universities have the option to adapt the general classification scheme adding all needed records to reflect their realities but respecting the retention schedule and its rules of preservation or disposition of the records. The records of the type "Diário de Classe" are stored in the business system for six months, corresponding to the active phase of the document, after that, the records are transferred to the institutional digital repository to be preserved forever [11].

4 RiC mapping to the "Class Diary"

In order to study the application of concepts, it was chosen an instance of a "Diary Class' record to do the mapping of its metadata using the conceptual model RiC. The metadata set used in this study was defined by Luz [8] that uses the brazilian metadata standard named E-Arq [2].

To map the archival units corresponding to the UFSM classification scheme, it was chosen to use the "RiC-E03 Record Set" entity that is defined as "one or more records that are associated by categorization and/or physical aggregation"[5, p.8] according some predefined criteria being required that all records of the same Record Set needs to have at least one common characteristic between them. All instances of the "Diário de Classe" records were mapped to the "RiC-E04 Record" entity according its definition "information inscribed at least once by any method on any physical carrier in any persistent, recoverable form[...]"[5, p.9]. The PDF file generated by the system with all information about the class is represented by the "RiC-E06 Instantiation" entity because it has the data in a bit stream format and it is the focus of the preservation [5, p.10].

The professor that fills the class information and generates the record is represented by the "RiC-E08 Person" entity that is a specification of the "RiC-E07 Agent" entity, defined as the "person, or group, or an entity created by a person or group (Delegate Agent), or a Position, that acts in the world."[5, p.11]. If the class has more than one professor, all of them will be agents related to the generated class record. The system itself could be considered an agent and could have its properties mapped to a instance of the "RiC-E13 Mechanism" entity. The specification of the professor's contract with the institution can be mapped using the "RiC-E12 Position" as it represents "the functional role of a Person within a Group"[5, p.14].

All the interaction made between the people and the record can be interpreted as an "RiC-E15 Activity", a kind of "RiC-E14 Event". The record generation did by the class professor creates an instance of the "RiC-E15 Activity" because represents the action of a human activity, as well as the access of the class students. If the system performs some type of operation in the records in some

point of its life cycle, this operation can be represented by "RiC-E14 Event" entity, containing the all details of that operation.

The first draft version of RiC (v0.1) [4] declared 792 types of possible relations between the entities, which was the target of much criticism by the archival community. The second draft version (v0.2) [5] has 78 types of relations, making the concept model simpler and easier to understand [9, p.59].
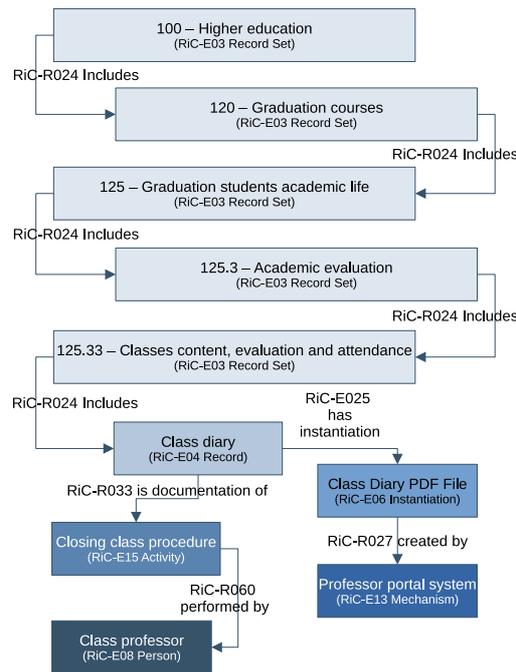

Fig. 1. RiC application.

The Fig. 1 shows the mapped entities with their respective relation. As the UFSM's classification scheme has a hierarchy between the Record Sets, it was chosen to use the "RiC-R024 includes" relation to express it, so the most generic Record Set includes its child Record Set and so on. Additionally, the last Record Set also has the "RiC-R024 includes" relation with the Record of the "Diário de Classe" that has "RiC-R025 has instantiation" with the entity that represents its physical file "RiC-E06 Instantiation". The professor is represented by "RiC-E08 Person" entity that has a relation "RiC-R060 performed by" to represent the process of the record generation represented by "RiC-E15 Activity" entity.

The corresponding mapping of entities using RDF/XML is shown in the Listing 1.1 based on the RiC-O ontology [6]. All the entities are assigned to an identifier that is used to do the references in their relationships. The first archival

unit, named "100 - Higher education", has the identifier #100 and has a relation to the entity #120 (120 - Graduation courses) declared as includesOrIncluded, equivalent to the "RiC-R024 includes" relation. The identifiers used in the RiC representation use URI form, allowing to reference entities described by external archives, making them linked.

Listing 1.1. RDF/XML entities representation

```
<owl:NamedIndividual rdf:about="#100">
 <name>100 - Higher education</name>
 <rdf:type rdf:resource="#RecordSet"/>
 <includesOrIncluded rdf:resource="#120"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#120">
 <name>120 - Graduation courses</name>
 <rdf:type rdf:resource="#RecordSet"/>
 <includesOrIncluded rdf:resource="#125"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#125">
 <name>125 - Graduation students academic life</name>
 <rdf:type rdf:resource="#RecordSet"/>
 <includesOrIncluded rdf:resource="#125.3"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#125.3">
 <name>125.3 - Academic evaluation</name>
 <rdf:type rdf:resource="#RecordSet"/>
 <includesOrIncluded rdf:resource="#125.33"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#125.33">
 <name>125.33 - Classes content, evaluation and attendance
 </name>
 <rdf:type rdf:resource="#RecordSet"/>
 <includesOrIncluded rdf:resource="#Class_diary"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#Class_diary">
 <name>Class diary</name>
 <rdf:type rdf:resource="#Record"/>
 <hasInstantiation rdf:resource="#Class_diary_PDF_file"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#Class_diary_PDF_file">
 <name>Class diary PDF file</name>
 <rdf:type rdf:resource="#Instantiation"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#Class_professor">
 <name>Class professor</name>
 <rdf:type rdf:resource="#Person"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#Closing_class_procedure">
 <name>Closing class procedure</name>
 <rdf:type rdf:resource="#Activity"/>
```

```
<documentedBy rdf:resource="#Class_diary"/>
<performsOrPerformed rdf:resource="#Class_professor"/>
</owl:NamedIndividual>
<owl:NamedIndividual rdf:about="#Professor_portal_system">
<name>Professor portal system</name>
<rdf:type rdf:resource="#Mechanism"/>
<isCreatorOf rdf:resource="#Class_diary_PDF_file"/>
</owl:NamedIndividual>
```

5 Conclusion

A new archival description standard is under development by ICA and promises to replace the four previous standards ISADG(G), ISAAR(CPF), ISDF and ISDIAH. The process is not easy and needs to cover the most distinct realities of archives around the world. The description process will need to be adapted to reflect the graph way of thinking, allowing to archivists and systems to add relations between the elements that were not possible before. As consequence, these new relations will improve the records contexts of the archives.

The use case made with the selected record at the UFSM is a pilot study and will need to be done to all records in order to migrate to the RiC description model. The systems that currently generate records will need to be changed to store the constant growing relations that records will have, as well as the diffusion tools that will need to show the network of entities related to the records. Additionally, the description tools will need to have adaptions to provide to archivists functionalities that make the description task more user-friendly.

The road to establish RiC as the new standard is too long and will require efforts from all the archival community, involving different professionals to make RiC a facto-standard. While the model still is in the elaboration process, other use cases studies will need to be done to support EGAD.

References

1. Comisión de Normas Españolas de Descripción Archivística: Vocabulario para la descripción archivistica. Terminología y glosario técnico. Ministerio de Educación, Cultura e Deporte (2017), `http://www.mcu.es/ccbae/es/catalogo_imagenes/grupo.do?path=41406`
2. Conselho Nacional de Arquivos: E-Arq Brasil: Modelo de requisitos para sistemas informatizados de gestão arquivística de documentos. Arquivo Nacional, Rio de Janeiro (2011)
3. International Council on Archives: ISAD(G): General International Standard Archival Description, second edn. (2000)
4. International Council on Archives: Records in Contexts: A Conceptual Model for Archival Description. v. 0.1. consultation draft edn. (2016)
5. International Council on Archives: Records in Contexts: A Conceptual Model for Archival Description. v. 0.2. consultation draft edn. (2019)

6. International Council on Archives: Records in Contexts Ontology (2021), `https://www.ica.org/standards/RiC/RiC-O_v0-2.html`

7. Llanes-Padrón, D., Moro-Cabero, M.: Records in contexts: Un nuevo modelo para la representación de la información archivística en el entorno de la web semántica **26**(3), 525–533 (2017), `https://doi.org/10.3145/epi.2017.may.19`

8. Luz, D.E.S.: Metadados para preservação e segurança do diário de classe eletrônico da UFSM. Master's thesis, Universidade Federal de Santa Maria (2011)

9. Souza, M.V.B., Flores, D.: Possíveis impactos do modelo records in contexts para os usuários de arquivos. Acervo **33**(3), 49–67 (2020)

10. Universidade Federal de Santa Maria: Instrução Normativa N. 03/2016 - PROGRAD/UFSM (2016), `https://www.ufsm.br/app/uploads/sites/342/2018/05/instrucao_normativa_03_2016_prograd_diario_de_classe_nato_digital.pdf`

11. Universidade Federal de Santa Maria: Código de Classificação e Tabela de Temporalidade de Documentos da UFSM (2019), `https://www.ufsm.br/orgaos-suplementares/dag/wp-content/uploads/sites/400/2019/03/TTD-UFSM.pdf`