

Bogdan Florin Popovici¹
(Romania)

"YET ANOTHER ONE?!" SOME CONSIDERATIONS ON INSTANTIATION ENTITY IN 'RECORDS IN CONTEXTS' MODEL

ABSTRACT

This paper seeks to compare the use of different, but apparently synonym terms, like Instantiation, Manifestation or Representation, terms used in different models and projects in library, archives and digital preservation. For the comparison, we reviewed the definitions and examples in several authoritative documents, like PREMIS, FRBR, Records in Contexts etc. This exercise revealed that similar words may have different usage in different context, which make it hard to use one term to communicate the same meaning for everybody. The conclusion is that, at least for the time being, Records In Context model of archival description needed to use a new term and to define its scope of use within the framework of the descriptive model, since it present specific features and aim to be general, carrier-neutral.

Keywords: Records in contexts, PREMIS, digital preservation, archival description, FRBR, PAAST

1 INTRODUCTION

During working meetings of the Experts Group for Archival Description (EGAD), under the auspices of the International Council on Archives, in the process of identifying the relevant entities, several existing models have been analysed. For one specific case, it was agreed to use a "new" term (i.e. "not used in other archival-related standard"), as it was easier to model a new entity that fits to the needs of the model than to create potential confusions with other models' entities. However, I was inclined ever since to have a closer review of such other entities, similar with the one used in "Records in Contexts" model (hereafter, RiC).

Despite the fact a certain amount of time was devoted to the issue of Instantiation, a question that bothered me was if it was really necessary to employ a new term to label this entity, while some other terms, apparently similar, already existed. *Manifestation* (in FRBR or PAAST) or *representation* (in PREMIS) seem quite similar with Instantiation. Hence, the purpose of this paper is to comparatively examine these terms and check their suitability.

¹ Bogdan-Florin Popovici graduated the Faculty of Archival Science In Bucharest (Romania) (1999). He holds a Ph.D. in History (2004). He attended several training programs in Europe (France, Italy). He works with the National Archives of Romania, Brasov County Division (since 2009), as archivist and, since 2013, as head of Division. He has been actively involved with ICA/Committee on Best Practices and Standard, respectively Experts Group for Archival Description (2008-). He participated at several international professional meetings (Slovenia, Italy, Oman, Russia, Hungary, Mexico, South Korea, Spain), both as a speaker and as a professional trainer. He is also involved in translating professional standards into Romanian. Contact: bogdan.popovici@arhivelenationale.ro.

2 THE CASE

Despite having as a condition of its existence the attribute of "fixity", a record passes through a lot of changes in its being, mainly of context, but sometime also of form and appearance. It is, as the Australian archival scholars called it, the Records Continuum. Sometime, these transformations do not change significantly the record, one observer being still able to call it "the same" record². In other situation though, such a change may lead to a new record, related with the initial one. From the perspective of an archivist, describing records that encompass transformations may mean repetition of bits of description, that are shared by each instance.

Here are some examples. We may have:

- a charter from 15th century, which is described by archivist. We also have a microfilm copy of this charter, that needs to be described. No matter if we consider that it is "the same" record or not, some descriptive values are identical (for example informational content), while others are not (e.g., carrier).
- A historian's personal paper. Among the documents, there is a transcription of a 18th century record. When describing it, we may re-use parts of the description of the original document for describing the focus of the historian's record.
- a 14th century record which embeds an extended *transumptum* of an older charter. When compiling a document chronological edition, that *transumptum* will be used as a placeholder for the original document, which is missing today; the *transumptum* will be the document, in its abridged form.
- a letter, that was in the same time produced as paper record and as a DOCX file, which, for preservation reasons, was then converted to PDF. The informational content (the message) is the same, but each instance has its own characteristics.
- a paper original, which is digitized, then the paper original is destroyed. Within that organisation, it is considered "a copy in the form of original"; and so, we may have a nice digital record from 1950...

Many similar cases are met in archival practice. In all the examples, there is "something" which is not the original document, but a different entity, that may be called a "copy", a "reproduction", a "variant", a "transformation" etc. Describing each one of these implies inevitably to use some shared elements (form, appearance, content), in addition to specific elements of each of that "something". In some cases, the two related documents may be considered "the same", while for others they are regarded as being different; the odd situation is that all sides are right, depending upon the context.

These cases are not new and in practice they were found, and descriptive method of referencing was used for long time. While referencing may be used within digital tools for archival description, the question is whether that "something" does not reveal the existence of a new entity, tight connected with record, but with its own attributes.

The Records in Context Conceptual Model calls this "something" an *instantiation* and considers it as one entity of archival description.

² For another contribution to the topic see our paper presented in Kazan (Russia) and published in Romania (Popovici, 2019).

3 THE "INSTANTIATION"

"Records in Contexts", as a conceptual model for archival description, has not yet been released formally as a "definitive" version (EGAD, 2019). Despite that, for the purposes of present paper, we shall analyse the concepts considering the present status of development.

Traditionally, "record" or similar terms were used to define "archival materials", but RiC splits it into several different entities. A first separation is within the archival material types, where RiC individualizes record, component of a record and aggregation of records as Record, Record Component and Records Set entities. Archival material as such is labeled Records Resource. A second differentiation is made between Records Resource, as intellectual entity, and its physical presence, as Instantiation.

The current definition for Instantiation in RiC is *"The inscription of information made by an Agent on a physical carrier in any persistent, recoverable form as a means of communicating information through time and space"*, while record is *"Information inscribed at least once by any method on any physical carrier in any persistent, recoverable form by an Agent in the course of life or work Activity."*

As resulting from the definition, an instantiation is a *sine qua non* condition for the birth of record; if no instantiation exists, we cannot talk of a record (e.g., a radio broadcast or a "signature" on water surface). However, the original instantiation may disappear in time, and record may only be preserved either a copy or a notice (e.g., *transumptum*). It means that, although the record does not exist in its original instantiation, the message in context may be considered to survive through its new, derived, instantiations, which is a new record, but functionally equivalent.

A record may have multiple instantiations. They can be simultaneous (e.g., a scanned record, generated in one step as TIFF and JPG), or successive in time (a paper record, microfilmed, then digitized). The latter re-instantiated record, depending on the context, can be considered the same (functionally equivalent of the original instantiation) or a new record. For example, the microfilm in the Archives is considered as a surrogate for the original, and many users find this is quite "the same" record (while it is obvious it is a new record, with different date and carrier). However, a police officer would hardly consider a photocopy of the driver licence as being the original.

In case of Records Set (e.g., file, series or even fonds), the original instantiation is a sum of each component record's instantiation. But in the case of re-instantiations, a reshape may occur. For instance, a file is a collection of records on paper; despite being treated as a whole, each of them keeps also individuality of carrier, form etc. After a digitisation, all the records in the physical file may be assembled in one PDF file, in which case their individuality as items can be disputable. Reciprocally, a ledger (one physical item) can be re-instantiated digitally as a JPG file for each page. The new instantiation may be, therefore, a fragment or a sum of previous instantiations.

By their nature, Instantiations are the main target of preservation. Either analogue or digital, Instantiations implies a carrier and/or a form of encoding that may require an intermediation device (from paper, vinyl disk and magnetic tape to digital file format). This aspect is important within the comparison with other models and explains why digital preservation focuses on instantiation related issues.

Existence of an intellectual entity and its instantiation were contemplated in other models too. The labels used were different. In the following we shall examine some of them.

4 THE "REPRESENTATION"

PREMIS (*Preservation Metadata: Implementation Strategies*) is an international standard seeking to offer support for practical implementation of metadata in digital preservation processes. Started in 2003, the project has become one of the basic intellectual tools for digital preservation worldwide.

The PREMIS Data Model defines 4 entities: Object, Right Statement, Agent and Event. For the purposes of the present paper, the entity of interest is Object, which is defined as "a discrete unit of information, subject to digital preservation". Object has 4 subcategories: Intellectual Entity, Representation, File or Bitstream. The Intellectual Entity is defined as "*distinct intellectual or artistic creation that is considered relevant to a designated community in the context of digital preservation*". A Representation is "*the set of files, including structural metadata, needed for a complete rendition of an Intellectual Entity*". One Intellectual Entity can have one or multiple Representations; in fact, it will always have at least one Representation. A File is considered a "*named and ordered sequence of bytes that is known to an operating system*" and it designates what is commonly understood as "file" in a computer systems, while Bitstream goes deeper in the structure of a File, indicating "*the contiguous or non-contiguous data within a file that has meaningful common properties for preservation purposes*" (PREMIS Data Dictionary for Preservation Metadata, 2015, pp. 7-8). The conceptual view is represented in the quoted resource like this:

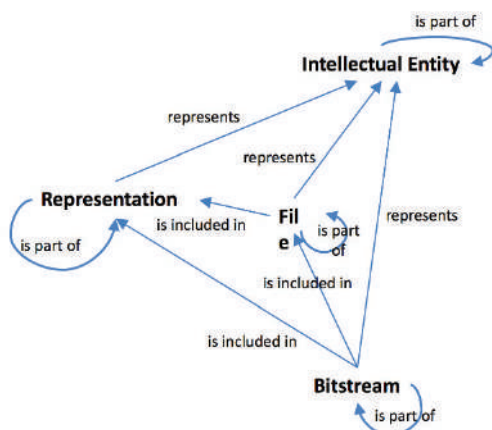


Figure 1 Subcategories of Object and their relations in PREMIS (apud PREMIS)

In a practical example, the image of a mountain would be the Intellectual Entity, as an intellectual creation and unit of information. The image as a digital photograph is saved on the card in 2 files, in format JPG and RAW, both of them being Representations of the picture.

Beside digital Representation, in the latest version of PREMIS (v.3), it is accepted that an Intellectual Entity may also have non-digital Representations. In the examples above, a charter, a transcription, a *transumptum*, the letter printed are all Representations of their original record (which are the Intellectual Entities), either physical or digital. Moreover, as indicated in PREMIS Data Dictionary, the Intellectual Entity may not be only regarded at a granular level, but also encompassing large components, like archival fonds or series. While the inclusion of non-digital Representation is rather non-consistent just yet (for instance, it is clear that non-digital Representation may not include Files, as depicted in the Figure 1 above), it seems that Representation, in its core characteristics, is similar to Instantiation.

5 THE "MANIFESTATION"

One term heavily used in cases of multiplication of the content in various instances is *manifestation*. Despite the common term, the meaning seems inconsistent among various documents.

A). FRBR

Functional Requirements for Bibliographic Resources is a user-centred conceptual model for describing objects of interest for librarians/users of bibliographic information (IFLA, 1998, p. 2). Unlike PREMIS, that embraces physical entities later in its development, FRMR was addressing physical and digital as a core option from the beginning.

The model comprises several entities, from which four of them in the first group (Products of Intellectual & Artistic Endeavor) are relevant for the present paper. **Work** is defined as "a distinct intellectual or artistic creation". **Expression** is "the realisation of a work, wither intellectual or artistic", which may take a variety of forms (textual, sound, musical notation, images or a combination of them). These two entities are purely abstract. **Manifestation** represents the physical embodiment of an expression of a work (that is, information put in a container or on a carrier). It is normally associated with a particular medium, such as a printed volume, photographic print or film reel. **Item** is the copy (exemplar) of a manifestation, that is, a physical instance. A manifestation may have more than one item, but one item can only have one manifestation. (IFLA, 1998, pp. 17-24). *"A manifestation represents all the physical objects that bear the same characteristics of intellectual content and physical form. In actuality, a manifestation is itself an abstract entity, but describes and represents physical entities, that is all the items that have the same content and carrier."* (Tillett, 2003, p. 10)

B). DIGITAL PRESERVATION – BROWN

Reflecting the findings of several projects and of the practical activity in the field, Adrian Brown approaches, in an influential book, some practical issues in digital preservation. Due to the methodologies used in this process, it is inevitable an analysis of the concepts associated with various "embodiment" of a record.

Describing processes and entities, A. Brown uses several terms, as "manifestation", "version" or "representation". Manifestation is defined as "specific data object that instantiates an information object. Multiple manifestations can exist for any given information object" (Brown, 2013, p. XIV). Later, the author emphasizes a very important feature: *"it is critical to distinguish between manifestations, being technical representation of the same information object and different editorial version of an information object"* (p. 216). The example given covers different ways that the same information object³ (a book) can be physically stored:

3 Defined, in the book, as in OAIS, as "a conceptual object of preservation... realized as meaningful information by interpreting a data object through its associated representation information".

Table 1.1 Alternative representations of a book	
Version	Technical representation
Physical	1 printed volume (comprising 12 chapters and 700 pages)
Word 2000	12 DOC files
Word 2007	12 DOCX files (each containing various XML files)
Digitized masters	700 TIFF files
Digitized access copy	1 PDF file
e-Book	1 EPUB container file (containing various XML, XHTML and image files)
Web	12 HTML files, 1 CSS file and 15 GIF images

Figure 2 Manifestations (apud Brown)

Adrian Brown also scrutinizes the origins of term "manifestation" within the realm of digital preservation. His conclusion is that term "almost certainly evolved independently within a number of different initiatives" (Brown, 2013, p. 215) and he gives no indication that the term might have been borrowed from library conceptual model.

C). DIGITAL PRESERVATION – INTERPARES PAAST

Another relevant document for using the term "manifestation" is PAAST, a product resulted from last InterPARES Trust project. Preservation as a Service for Trust (PaaST) presents functional and data requirements for digital preservation (Interpares_Trust, 2017). In this project usage, Manifestation is *"A concrete, physical implementation of an Intellectual Entity capable of communicating information as originally intended"*, while an Intellectual Entity is *"an artifact that is intended to communicate information"*, a record being one type of Intellectual Entity.

In PAAST model, an Intellectual Entity is digitally encoding its specification in an entity called Binary Encoding. The Binary Encoding has 2 type of parts: Digital Component (that is, *"an ordered string of bits that encodes numbers, characters, signs, symbols, sounds, images or other graphics in a digital form that is both suitable for storage and appropriate to instantiate an Intellectual Entity"*) and Manifestation. This perspective is consistent with previous InterPARES findings, which discussed about the stored and manifested records: *"Stored record is constituted of the digital component(s) used in re-producing it, which comprise the data to be processed in order to manifest the record (content data and form data) and the rules for processing the data, including those enabling variations (composition data)"* and *"manifested record is the visualization or instantiation of the record in a form suitable for presentation to a person or a system"* (Duranti & Thibodeau, 2006, p. 39), (Duranti, 2016, p. 8).

Manifestation, from PAAST view, enables Intellectual Entity to be instantiated. For example, a record (Intellectual Entity) is digitally encoded as a PDF file (record stored); manifestation is what would be rendered on/readable from a screen by a human reader (that is, the message conveyed reached to interested party)⁴. Therefore, a PAAST Manifestation is a physical materialization of an Intellectual Entity, which is not stored and preserved, but exists on an output device (Interpares_Trust, 2017, p. 23)

4 This is an intentional simplification for our paper; PAAST specifies that Manifestation may be under the form of a Runtime version (when the manifestation is intended for a machine) or of a rendering (when the IE is output in a form accessible to humans).

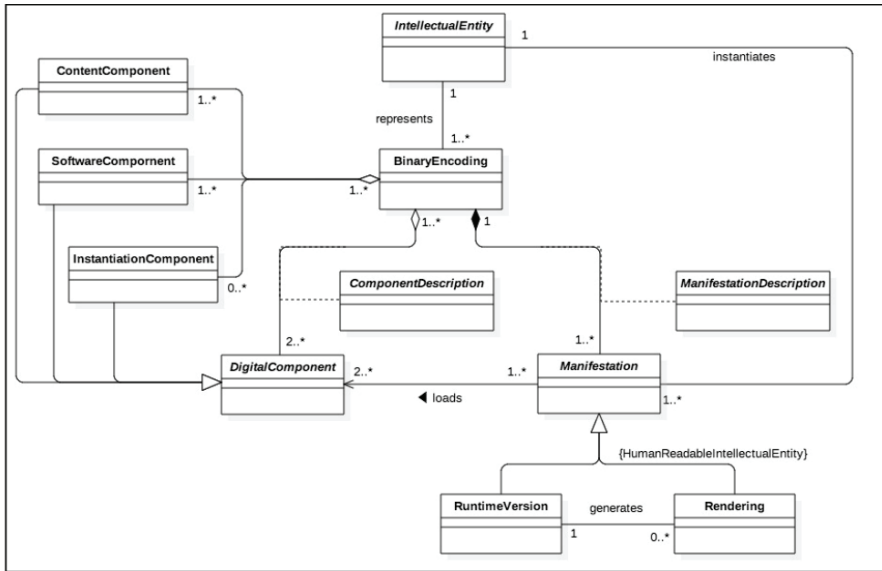


Figure 3 Composition of an Intellectual Entity (apud PAAST)

6 SOME CONCLUSIONS

Reviewing some of the meaning of terms describing physical embodiment of an intellectual entity, a first remark can be the semantic proximity. "Manifestation" is used in several documents, and it seems it has similar, but not identical meanings. On the other hand, visible mainly at A. Brown, manifestation can both "represents" or "instantiates" a piece of information, employing the same wording as in RiC (Instantiation) or PREMIS (Representation). This emphasizes again that the meanings for Manifestation, Representation and Instantiation are close, though not necessarily fully overlapping.

As a scope, FRBR is the closest to RiC, since it aims to encompass all "information objects" without a dedicated view on physical or digital objects. PREMIS is (tentatively) close but, in our opinion, the presence of non-digital Representation is not yet very elaborated, and it is rather peripheral than central to the model (which focus on digital preservation). On the other hand, FRBR has certain features that makes it less suitable for use in archival description and so, in using its entities to describe archival ones. Usually, archivists' interest towards abstract entities, like Work or Expression may be considered rather low, since record is rarely regarded as an intellectual or artistic product, but rather an instrument or a by-product of a practical activity (as definitions cited above indicate). As about Manifestation, as it was noticed, "*As entity, the manifestation is the set of all physical objects that shares...*" (Tillett, 2003, p. 10). But creating many copies of a record is barely the norm. Where many copies exist, they may fulfil different functions in different contexts, and this makes them unique, even if the information content and physical form may be identical. A similar distinction was made in PREMIS, where the term "Representation" was used specifically to avoid the term "manifestation" as it is used in the FRBR. "*In the PREMIS model, a Representation is a single instance of an Intellectual Entity held in a preservation repository*". (PREMIS Data Dictionary for Preservation Metadata, 2015, p. 12), unlike FRBR. This brings PREMIS_Representation very close to RiC_Instantiation.

The term Manifestation used in digital preservation seems to have different meanings for different authors. For A. Brown, according to the definition, Manifestation is data object and it is synonym with Representation. The same word, used in PAAST refers to the actual communication of information, while data object is the Binary Encodings. Using the term Manifestation in RiC would implicitly lead to misunderstandings, since there is no apparent consensus of using the term.

PREMIS_Representation, and Manifestation as it is used by A. Brown seem to be very close to RiC_Instantiation. It is about the mean used to inscribe information for conveying it, not about the information conveyed to the user (which is abstract before inscribed). Any Intellectual Entity can have multiple representation/manifestation/instantiations. Brown underline the fact that manifestations are "technical representation of the same information object" and they are different than editorial versions of the same information object (Brown, 2013, p. 215), with the example of a cropped photo, which is regarded as a new conceptual object). In similar way, a record instantiated cannot be changed or modified, so a re-instantiation is a new entity, which can (or not) be considered as functional equivalent with the original.

It needs to be emphasized that RiC_Instantiation aims to be carrier/medium independent; it needs to describe both analogic and digital inscription of records, while PREMIS_Representation/Manifestation were shaped for digital realm. Expanding the understanding of PREMIS_Representation open to path to synonymity to RiC_Instantiation, but in this moment it is not clear if there are some fine differences between RiC/Manifestation or Representation or not. So, for now, it may be a good option to use new term, in order to satisfy the description needs. But very likely, in the future, a convergence of meaning will lead to a unification of concepts.

REFERENCE LIST

- (EGAD), E. G. (2019). Records in Contexts (v.0.2). Preluat de pe <https://www.ica.org/en/public-release-of-records-in-contexts-ontology-v01-and-records-in-contexts-conceptual-model-v02>
- Bogdan-Florin, P. (2019). „Cât de bine poți cunoaște un document?” Despre dimensiunile fizice si intelectuale ale documentelor arhivistice . În *Szekély Századok. Emlékkönyv Pál-Antal Sándor 80 születésnapjára* (pg. 410-424). Marosvásárhelyi, .
- Brown, A. (2013). *Practical Digital Preservation. A how-to guide for organisations of any size*. London: Facet Publishing.
- Duranti, L. (2016). The Concept of Record in the Digital Environment (presentation).
- Duranti, L., & Thibodeau, K. (2006). The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES. *Archival Science*, 6(1), 13–68.
- IFLA. (1998). Functional Requirements for Bibliographic Records. Munchen. Preluat de pe <https://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf>
- Interpares_Trust. (2017). Preservation as a Service for Trust (PaaST). Functional and Data Requirements for Digital Preservation. Preluat de pe https://interparestrust.org/assets/public/dissemination/PreservationasaServiceforTrust1_0.pdf
- PREMIS Data Dictionary for Preservation Metadata. (2015). Preluat de pe <http://www.loc.gov/standards/premis>
- Tillett, B. B. (2003). The FRBR Model (Functional Requirements for Bibliographic Records). Preluat de pe <https://www.loc.gov/catdir/cpsd/frbreng.pdf>

FURTHER REFERENCES:

- Gill, T. (2004). Building semantic bridges between museums, libraries and archives: The CIDOC Conceptual Reference Model. *First Monday*, 9(5–3 May).
- Henttonen, P., & Kilkki, J. (2017). "Records in contexts" and the Finnish conceptual model for archival description. *Letonica*, 36, 60–71.
- Howarth, L. C. (2012). FRBR and Linked Data: Connecting FRBR and Linked Data. *Cataloging & Classification Quarterly*. <https://doi.org/10.1080/01639374.2012.680835>
- Le Bœuf, P. (2012). Modeling Rare and Unique Documents: Using FRBRoo/CIDOC CRM. *Journal of Archival Organization*, 10(2), 96–106. <https://doi.org/10.1080/15332748.2012.709164>
- Llanes-Padrón, D., & Pastor-Sánchez, J.-A. (2017). Records in contexts: The road of archives to semantic interoperability. *Program*, 51(4), 387–405. <https://doi.org/10.1108/PROG-03-2017-0021>
- Riva, P., Doerr, M., & Žumer, M. (2009). FRBRoo: Enabling a common view of information from memory institutions. *ICBC*, 38(2), 30–34.
- Thurman, A. C. (2007). FRBR and archival materials. In A. G. Taylor (Eds.), *Understanding FRBR: what it is and how it will affect our retrieval tools* (ss. 97–102). Libraries Unlimited.