Modeling Archives by Means of OAI-ORE

Nicola Ferro and Gianmaria Silvello

Department of Information Engineering, University of Padua, Italy
{ferro,silvello}@dei.unipd.it

Abstract. Currently, archival practice is moving towards the definition of complex relationships between the resources of interest as well as the constitution of compound digital objects. To this end archives can take advantage of using the *Open Archives Initiative - Object Reuse and Exchange (OAI-ORE)* providing additional and flexible visualizations of archival resources.

In this paper we define a formal basis that provides a means for defining OAI-ORE instances which are consistent with the fundamental archival principles.

1 Motivation

Archives are composed of aggregations of interrelated material and their significance lies in their aggregate, or collective nature. Archivists work to preserve the *original order* of the documents within an archive – i.e. principle of provenance – because the context and the physical order in which the documents are held are as valuable as their content [3]. The principle of provenance leads archivists to evaluate records on the basis of the importance of the creator's mandate and functions, and fosters the use of a hierarchical method for describing the archives. Although this practice is still vitally important for the archives, the archivists also need more powerful tools to capture the complexity of the reality of interest. Indeed, the reality of modern records creation is that documents may exist in "multiple contexts and have multiple and complex relationships that describe their significance and value" [9]. Furthermore, new archival trends encourage the adoption of "plural, provisional and interpretative perspective" [12] in the description of the archives.

The archival practice is thus experiencing a transformation process which promotes the definition of complex relationships between the resources of interest and the constitution of compound digital objects [9]. For similar reasons in the wider context of digital libraries we are experiencing a wide-ranging diffusion of the *Open Archives Initiative - Object Reuse and Exchange (OAI-ORE)*[1].

Archives as a meaningful part of the DL can take advantage of using the OAI-ORE [9]; indeed, a methodology for representing the archives in OAI-ORE would allow richer methods for modeling archival descriptions and can also provide additional and flexible visualizations of the documents that would not be restricted

[1] http://www.openarchives.org/ore/

M. Agosti et al. (Eds.): IRCDL 2012, CCIS 354, pp. 216–227, 2012.

to the "old linear view inspired by the paper tradition" [9]. At the same time, it is commonly agreed [12,18,9] that new approaches, such as the adoption of OAI-ORE model, should add to, but not undermine, the fundamental archival theory.

We can see an archive as a compound object composed by atoms of information which have to be identifiable and we need to define the granularity of this atoms. In this paper we adopt the NESTOR Model [1] to provide an alternative way to model archives allowing us to manipulate archival resources as atoms of information without loosing their multileveled relationships. Therefore, in this work we lever on the *NEsted SeTs for Object hieRarchies (NESTOR)* Model [5] to:

- define a formal basis that allows us to model an archive as an OAI-ORE instance while retaining its hierarchical structure;
- propose a methodology to map archival descriptions into OAI-ORE showing how it enables both the preservation of their original order and the definition of new types of relationships.

This paper is organized as follows: Section 2 presents a brief overview on archives and archival metadata, the NESTOR Model, and OAI-ORE. Section 3 describes the formal basis for modeling the archives as OAI-ORE instances while respecting the fundamental archival principles. Section 4 introduces a methodology which shows how we can represent a sample archive as an OAI-ORE instance. Lastly, Section 5 draws some final remarks.

2 Background

Archives: Metadata and Digital Objects. An archive is the trace of the activities of a physical or juridical person in the course of their business which is preserved because of their continued value. Archives have to keep the context in which their records have been created and the network of relationships between them in order to preserve their informative content and provide understandable and useful information over time [6]. The context and the relationships between the documents are preserved thanks to the hierarchical organization of the documents inside the archive. Indeed, an archive is divided by fonds and then by sub-fonds and then by series and then by sub-series and so on – see Figure 2a for an example; at every level we can find documents belonging to a particular division of the archive or documents describing the nature of the considered level of the archive. The union of all these documents, the relationships and the context information enables the full informational power of the archival documents to be maintained. The archival documents are analyzed, organized, and recorded by means of *archival descriptions* [15] that have to reflect the peculiarities of the archive [3].

In a digital environment an archive and its components are described by using the metadata that have to be able to express and maintain such structure

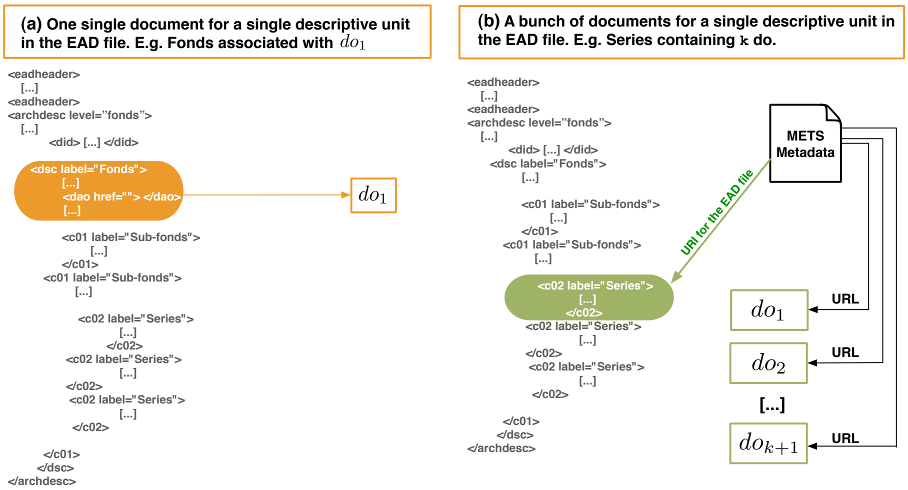

Fig. 1. A solution to link the EAD file with the described digital objects

and relationships. The standard metadata format for representing the hierarchical structure of the archive is the *Encoded Archival Description (EAD)*[2], which reflects the archival structure and holds relations between documents in the archive [17]; an EAD file is an *eXtensible Markup Language (XML)* file with a deep hierarchical internal structure. In an EAD file the information about fonds, sub-fonds and series are mapped into several nested elements and the archival structure is maintained by a collection of nested `<cN>` tags (e.g. `<c02 label="Series">` in Figure 1). EAD describes an archive as a unique monolithic resource; indeed, it is not an aggregation of metadata each describing a single part of the archive, but a monolithic metadata where every subcomponent describes a different division or document of the archive. In order to access a specific division of an archive described by EAD we may need to navigate the whole XML hierarchy; otherwise, it is also possible to define an ad-hoc solution, for instance using XPointers[3], to provide direct access to frequently requested archival divisions encoded by EAD subcomponents.

Each `<cN>` tag of the EAD may contain a description of a digital object or a bunch of digital objects. These objects are usually reachable by means of an *Uniform Resource Identifier (URI)*; the link from EAD to a digital object or group of objects can be made at any level, but "*it should be made at the level where the object(s) is described or implied in EAD*" [13]. To this end EAD provides a `<dao>` tag which allows us to specify a URI to an external digital object which is part of the described material (see Figure 1a); furthermore, EAD also provides an `<extptr>` element to point to a digital object that is not part of the described materials [13]. By means of these tags we can link one external

[2] `http://www.loc.gov/ead/`
[3] `http://www.w3.org/XML/Linking/`

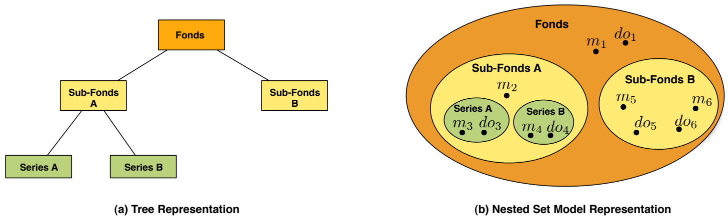

Fig. 2. The structure of a sample archive represented by: (a) a tree; (b) an Eulero-Venn diagram

digital object to each archival division; if we need to link more than one digital object to a specific division we have to exploit third-party components – i.e. the so-called "digital wrappers"[4]; a relevant example is the *Metadata Encoding and Transmission Standard (METS)* metadata that is used as an in-between component for relating a bunch of digital objects to an EAD component [19,14] – see Figure 1b.

The NESTOR Model. The NESTOR Model relies on two set data models called *Nested Set Model* (NS-M) and *Inverse Nested Set Model* (INS-M) [1]. Both these set data models, formally defined in the context of axiomatic set theory [8], can be used to model an archive [5]. Indeed, we can represent the archival structure by means of a collection of nested sets where each set represents an archival division and contains the metadata describing the resources belonging to that division [4]. An extensive analysis of the NESTOR Model and its applications in the context of DL and archives can be found in [1]; in this paper we exploit the NS-M and thus we focus our presentation on this model.

The most intuitive way of understanding how the NS-M works is to see how a sample tree is mapped into an organization of nested sets based on the NS-M. An organization of sets in the NS-M is a collection of sets in which any pair of sets is either disjoint or one contains the other. In Figure 2 we can see how a sample tree representing an archive is mapped into an organization of nested sets based on the NS-M – for the moment please ignore the elements belonging to the sets. We can see that each node of the tree is mapped into a set, where child nodes become *proper subsets* of the set created from the parent node. Every set is subset of at least one set; the set corresponding to the tree root is the only set without any supersets and every set in the hierarchy is subset of the root set. The external nodes are sets with no subsets. The tree structure is maintained thanks to the nested organization and the relationships between the sets are expressed by the set inclusion order. Even the disjunction between two sets brings information; indeed, the disjunction of two sets means that these belong to two different branches of the same tree.

[4] Digital wrappers "*are pieces of software for binding digital content files and their metadata together and for specifying the logical relationships among the content files*" [14].

Resource Map 1

rm1	ore:describes	a1
a1	ore:aggregates	ar1
a1	ore:aggregates	ar2
a1	ore:aggregates	a2
p1	ore:isProxyFor	ar1
p1	ore:isProxyIn	a1
p2	ore:isProxyFor	ar2
p2	ore:isProxyIn	a1
p1	xyz:hasNext	p2

Resource Map 2

rm2	ore:describes	a2
a2	ore:isDescribedBy	rm2
a2	ore:aggregates	ar3

Entity to URI association

rm1	urm1
a1	ua1
p1	up1
ar1	uar1
ar2	uar2
a2	ua2
rm2	urm2
ar3	uar3
p2	up2
ore:describes	http://www.../ore/terms/describes
[...]	[...]

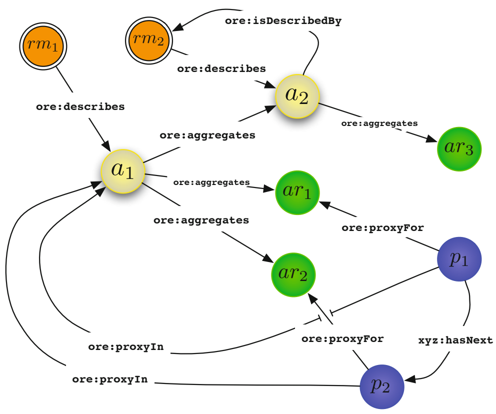

Fig. 3. An instance of the OAI-ORE data model represented by an RDF graph

In [4] a methodology is described for mapping an EAD file into the NESTOR Model which preserves the full informative power of the metadata. [4] shows that the EAD is mapped into a NS-C which retains the EAD structure and a collection of lightweight metadata – e.g. Dublin Core Application Profile[5] – which contains the content of archival descriptions. In this way, the NESTOR Model can be used as a model to describe an archive from scratch as well as a mapping component that allows us to manipulate and transform the EAD files while respecting archival principles [5].

OAI-ORE. The OAI-ORE defines a machine-readable and standard mechanism for defining aggregations of resources on the Web. By means of OAI-ORE we can identify a bunch of resources related to each other as a single entity enabling the access and exchange of them at an aggregation level of granularity. The OAI refers these aggregations as "*compound objects*". Compound units are aggregations of distinct information units that, when combined, form a logical whole. Some examples [20] of these are a digitized book that is an aggregation of chapters, where each chapter is an aggregation of scanned pages, and a scholarly publication that is an aggregation of text and supporting materials such as datasets, software tools, and video recordings of an experiment; also the archives can be seen as aggregations of archival metadata describing archival objects which in turn can have a digital form.

The OAI-ORE data model is based on three main kinds of resources: *Aggregation*, *Aggregated Resources* and *Resource Map*. An Aggregation is defined as a resource representing a logical collection of other resources. An Aggregation is a logical construct and thus it has no representation; it is described by a Resource Map which can be seen as a materialization of the Aggregation. A Resource

[5] http://www.dublincore.org/

Map must describe a single Aggregation and must enumerate the constituent Aggregated Resources; a resource is an "Aggregated Resource" in an Aggregation only if it is asserted in a Resource Map. Each resource in the OAI-ORE data model is identified by a URI. The OAI-ORE data model is expressed by the *Resource Description Framework (RDF)*[6], so its instances are expressed as RDF graphs as we can see in Figure 3. An RDF graph is defined by a set of triples (s, p, o) expressing the relationship defined by a predicate p between a subject s and an object o; s and o may be a URI with an optional fragment identifier, a literal or a blank (having no separate form of identification). Properties p are URI references[7]. In Figure 3, we can see a set of subject-property-object triples represented as an RDF graph.

Although OAI-ORE is a relatively young specification, it has been becoming a standard reference in the context of digital libraries and its use is widespread in many systems and applications that deal with aggregations of digital objects. The use of OAI-ORE was adopted firstly for the management, access, and curation of scholarly publications and now it is spreading into the management and representation of scientific data [16] and of complex cultural objects [2].

3 A Formal Basis for Modeling Archives by Means of OAI-ORE

The aim of this work is to define a way to model archives by means of the OAI-ORE data model and the formal basis we propose provides a means to produce OAI-ORE instances which are consistent with the fundamental archival principles.

In order to explain how an archive can be properly modeled as an instance of the OAI-ORE data model we need to introduce several formal definitions. First-of-all, we present the definition of the NS-M which is based on the basic set-theoretical concept of "collection of subsets" [8].

Definition 1. *Let B be a set and let $\mathcal{C}$ be a collection of subsets of B. Then $\mathcal{C}$ is a **Nested Set Collection** if:*

$$B \in \mathcal{C}, \tag{3.1}$$

$$\forall H, K \in \mathcal{C}, \mid H \cap K \neq \emptyset \Rightarrow H \subseteq K \vee K \subseteq H. \tag{3.2}$$

Thus, we define a *Nested Set Collection* (NS-C) as a collection of subsets where two conditions must hold. The first condition (3.1) states that set B which contains all the subsets of the collection must belong to the NS-C. The second condition states the intersection of every couple of sets in the NS-C is not the empty-set only if one set is a subset of the other one. This formulation of the NS-C follows the original definition of "nested sets representation" of a tree given by [10] and that we informally explained in the background section.

[6] `http://www.w3.org/RDF/`
[7] `http://www.w3.org/TR/rdf-concepts/`

Now, we can introduce a compact representation of the OAI-ORE data model in order to clarify the relationships between the entities and to manipulate them in a formal environment. We express OAI-ORE in terms of sets and functions in order to establish a direct connection with the NS-M by using the same mathematical formalism. We define with R the set of all the resources[8] we take into account, with U the sets of all possible URIs identifying the resources and with $\eta : U \rightarrow R$ the bijective function[9] which associates a URI in U with one resource in R.

We indicate with $UA \subset U = \{ua_1, \ldots, ua_k, \ldots, ua_n\}$ the set of URI identifying the Aggregations and with $\eta_A : UA \rightarrow R$ the restriction of η ($\eta|_A$) to UA; the image of η_A is the set of Aggregations $A \subset R = \{a_1, \ldots, a_k, \ldots, a_n\}$. In the same way, we indicate with $URM \subset U$ the set of URI identifying the Resource Maps and we define $\eta_{RM} : URM \rightarrow R$ to be the restriction $\eta|_{RM}$ where $RM \subset R$ is the set of Resource Maps. Finally, we indicate with $UAR \subset U$ the set of URI identifying the Aggregated Resources[10]. We define $\eta_{AR} : UAR \rightarrow R$ to be the restriction $\eta|_{AR}$ where $AR \subset R$ is the set of Aggregated Resources. Every $rm_i \in RM$ must describe one and only one $a_j \in A$, but a_j may be described by more than one Resource Map; thus, we indicate with $\varphi_{RMA} : RM \rightarrow A$ a function which maps a Resource Map to the Aggregation it materializes. Every $ar_i \in AR$ may be aggregated by more than one $a_j \in A$. An example of the use of these URIs is shown in the tables in Figure 3.

In Figure 3 we can see the set of triples constituting two Resource Maps (rm_1 and rm_2) materializing two Aggregations (a_1 and a_2). This triple states that the Resource Map rm_i identified by urm_i describes the Aggregation a_i identified by ua_i.

OAI-ORE comes with another two important features: *Proxy* and *Nested Aggregations*. A Proxy is a resource that indicates an Aggregated Resource in the context of a specific Aggregation; a Proxy is associated with an Aggregated Resource via an assertion in a Resource Map describing the Aggregation that is the context of the Proxy [11]. We indicate with $UP \subset U = \{up_1, \ldots, up_k, \ldots, up_z\}$ the set of URI identifying the Proxies. We define $\eta_P : UP \rightarrow R$ to be the restriction $\eta|_P$ where $P \subset R$ is the set of Proxies. Proxies allow us to define relationships between Aggregated Resources; in Figure 3 we can see two Proxies p_1 and p_2 defining an order of precedence between the Aggregated Resources ar_1 and ar_2 in the context of Aggregation A_1. We indicate with $\varphi_{PAR} : P \rightarrow AR$ a function which maps a Proxy to the Aggregated Resource *for which* it is a Proxy and with $\varphi_{PA} : P \rightarrow A$ a function which maps a Proxy to the Aggregation *in which* it is a Proxy.

---

8 In this context a *resource* can be a metadata or a digital object.

9 We choose to define η as bijective function to keep the problem as straightforward as possible; in a different context, a resource could be identified by more than one URI.

10 Please note that the definition of the sets UA, URM, UAR is a mere convention to indicate URIs pointing to different kind of resources in OAI-ORE and they do not stand for different kind of URIs [20].

The *Nested Aggregations* feature enables the definition of Aggregations of Aggregations; this is consistent in the OAI-ORE data model because an Aggregation is a Resource which can also be seen as an Aggregated Resource of another Aggregation. Thanks to this feature, an order exists between Aggregations, call it $\prec_a$; more formally: for all $a_i, a_j \in A$ we say that $a_i \prec_a a_j$ if and only if the Aggregation a_i is aggregated by a_j; in Figure 3 we show two nested Aggregations $a_1, a_2 \in A$ where $a_2 \prec_a a_1$. It is important to notice that $\prec_a$ cannot define any orders between any OAI-ORE entities other than Aggregations; in fact, to define an order between Aggregated Resources we must use Proxies. Now, we can summarize the concept of *OAI-ORE Data Model* thanks to the next definition.

Definition 2. *Let* $\mathcal{E} = \{A, R, AR, P, UA, UR, UAR, UP\}$ *be the collection of OAI-ORE entity sets and* $\Phi = \{\eta_A, \eta_{RM}, \eta_{AR}, \eta_P, \varphi_{RMA}, \varphi_{PAR}, \varphi_{PA}\}$ *be the set of OAI-ORE functions. We define* $\mathcal{O} = \langle \mathcal{E}, \Phi \rangle$ *to be an OAI-ORE Data Model.*

In order to model an archive by means of OAI-ORE we need a methodology to identify the archival resources and to express the relationships between them. We have seen that we can represent a tree by means of the NS-M and that an archive can be modeled by means of a tree as well as by a NS-C. Therefore, we can model an archive throughout OAI-ORE by starting from its representation in the NS-M. We need to define a mapping between a NS-C $\mathcal{C}$ and an OAI-ORE model $\mathcal{O} = \langle \mathcal{E}, \Phi \rangle$; in order to do this we have to take into account the two main entities of the NESTOR Model which are: the sets and the resources belonging to them.

The intuitive idea is that every set $H \in \mathcal{C}$ becomes an Aggregation $a_h \in A$ and consequently, every resource $r_t \in R$ belonging to H becomes an aggregated resource $ar_t \in AR$ aggregated by a_h. Furthermore, for every pair of sets $\{H, K\} \in \mathcal{C} \mid H \subseteq K$ it is possible to create a pair of aggregations $\{a_h, a_k\} \in A$ such that $a_h \prec_a a_k$ where $\prec_a$ is a binary relation between aggregations.

Every set in a collection of subsets can be mapped into an Aggregation in the OAI-ORE model; the inclusion order between the sets is maintained by the binary relation defined between the nested Aggregations of OAI-ORE. Then, by the means of the function φ_{RMA} a Resource Map is associated with each Aggregation. Every resource belonging to a set H in the NS-C is mapped into an Aggregated Resources belonging to the Aggregation mapped from H. Thus, we can map a NS-C into a correspondent OAI-ORE model being sure that the hierarchical dependencies are properly retained. This means that if we model an archive through a NS-C then we define an OAI-ORE instance of the archive which retains the original hierarchical structure of the archive.

4 How to Model an Archive as an OAI-ORE Instance

The presented formal basis guarantees that an archive modeled by means of the NS-M can be mapped into an instance of the OAI-ORE Data Model retaining the

Sets	Aggregations
fonds	a_1
subFondsA	a_2
subFondsB	a_3
seriesA	a_4
seriesB	a_5

Table A
Mapping of sets into aggregations

Nested Sets	Nested Aggregations
subFondsA ⊂ fonds	$a_2 \prec_a a_1$
subFondsB ⊂ fonds	$a_3 \prec_a a_1$
seriesA ⊂ subFondsA	$a_4 \prec_a a_2$
seriesB ⊂ subFondsA	$a_5 \prec_a a_2$

Table B
Mapping of nested sets into nested aggregations

Elements	Aggregated Resources
m_1	ar_a
do_1	ar_b
m_2	ar_c
m_3	ar_d
[...]	[...]
do_6	ar_m

Table C
Mapping of elements into aggregated resources

Elements and Sets	Aggregations and Aggregated Resources
$m_1 \in$ fonds	a_1 aggregates ar_a
$do_1 \in$ fonds	a_1 aggregates ar_b
$m_2 \in$ subFondsA	a_2 aggregates ar_c
$m_3 \in$ seriesA	a_4 aggregates ar_d
[...]	[...]
$do_6 \in$ subFondsB	a_3 aggregates ar_m

Table D
Mapping of the elements belonging to sets into aggregated resources belonging to aggregations

Aggregated Resources	Proxies
ar_a	p_a
ar_b	p_b
ar_d	p_d
ar_e	p_e
[...]	[...]
ar_m	p_m

Table E
Proxies for the aggregated resources

p_a isMetadataOf p_b
p_d isMetadataOf p_e
[...]
p_l isMetadataOf p_m

Table F
The use of property "isMetadataOf"

fundamental archival hierarchy. In this section we show how we can define different kinds of relationships between the resources; furthermore, we show how a proper use of Proxies can preserve the order between the resources within the same archival division. It is worthwhile to provide a concrete example of how this formal basis can be applied to a sample archive modeled by the NS-M; we describe the mapping methodology step-by-step with the help of some mapping tables.

Let us take into account the sample archive represented in Figure 2b; this archive is composed by five archival divisions – i.e. one fonds, two sub-fonds and two series – each containing metadata and digital objects. In NS-M these divisions are represented by means of five sets and the hierarchical relationships are retained by means of the inclusion dependencies between the sets. In "Table A" we can see the mapping of the sets into the OAI-ORE Aggregations and in "Table B" we can see how the inclusion dependencies are mapped into Nested Aggregations. These two mappings show us how to represent the structure of a sample archive into an instance of the OAI-ORE data model.

Each set in the NS-C contains several elements which are metadata or digital objects. For instance, the set "fonds" contains two elements: a metadata (i.e m_1) and an associated digital object (i.e. do_1). The set "sub-fondsA" contains only a metadata (i.e. m_2), the set "seriesA" contains a metadata (i.e m_3) and an associated digital object (i.e. do_3), and so on and so forth. In "Table C" we can see how the elements are mapped into Aggregated Resources and in "Table D" we can see how the Aggregated Resources are associated with the correct Aggregations. We can see that an element belonging to a set – e.g. $m_i \in H$ – is mapped into an Aggregated Resource – e.g. ar_i – aggregated by the Aggregation a_h which corresponds to the set H. "Table E" and "Table F" show how we can use Proxies to associate the metadata with the digital objects they describe. OAI-ORE allows us to define different kinds of relationships between the Aggregated Resources using the Proxies. For instance, in Table F we can see that two Proxies p_a and p_b associated to ar_a and ar_b respectively are related by the relationship "`isMetadataOf`"; thus, throughout p_a and p_b we can say that

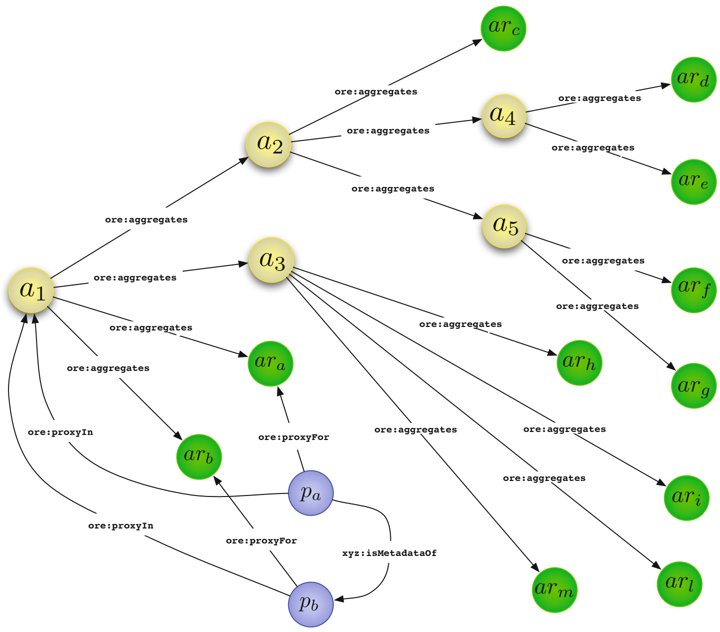

Fig. 4. An instance of OAI-ORE which models a sample archive

the Aggregated Resource ar_a is a metadata describing the digital object ar_b. In the same way we can define a linear order between the Aggregated Resources as we have shown in Figure 3 where we defined a "`hasNext`" relationship between the Proxies "p_1" and "p_2". The relationships between the Aggregated Resources can reflect the order between the archival descriptions within a common archival division; in this way, we are sure that the OAI-ORE representation of the archive respects the original order principle. We can see that within this methodology it is quite simple to extend the range of the relationships connecting the Aggregated Resources and to define in this way new semantic associations between the archival resources.

In Figure 4 we can see the RDF graph representing the OAI-ORE instance of the sample archive in Figure 2b. In this figure we represent the Aggregations, the Aggregated Resources and the Proxies associated to a_1; for space reasons we have omitted showing the other Proxies and the Resource Maps. This methodology makes it possible to model and describe the archives from scratch by means of OAI-ORE while allowing archivists to easily express relationships between archival metadata and digital objects. Archival principles are preserved and still have primary importance for understanding archival resources; at the same time, OAI-ORE offers the possibility of defining new relationships between the resources enabling the definition of new services over the archives. Moreover, this methodology provides a means to define archival compound objects that can be shared with the systems which already employ OAI-ORE and related technologies.

On the other hand, this methodology and the described formal basis guarantee the backward compatibility with other archival descriptive standards; for instance, a methodology to map the archival descriptions modeled by OAI-ORE into EAD can be easily defined. Indeed, we know how to map EAD into a NS-C and a NS-C into an instance of the OAI-ORE data model. In the same way, we can map the archival descriptions modeled by OAI-ORE into an EAD file by reversing the presented methodology[11]. In this context, the NESTOR Model can act as an interoperability layer between EAD and OAI-ORE and guarantee the possibility of going from one model to the other.

5 Final Remarks

In this paper we present a formal basis and a methodology to model an archive by means of the OAI-ORE data model consistent with the fundamental archival principles. OAI-ORE is widely-employed in the context of Digital Libraries but is still not completely exploited within archives; the formal basis reported in this paper can settle the ground for further investigations about the adoption of OAI-ORE in the archival context. This research direction can bring into archival practice the expressive power of OAI-ORE to allow for a multitude of non-linear relationships, providing richer and more powerful access and descriptions.

Furthermore, the use of OAI-ORE is increasing in several systems and digital library federations such as Europeana[12] the aim of which is to collect and make available resources from a wide spectrum of cultural institutions including the archives. A further step toward this direction will be to investigate how the NESTOR Model may allow different ways of modeling archival resources easing the integration of these resources with the Europeana Data Model (EDM). It will be interesting to consider the proposed methodology under the lens of other approaches trying to map archival resources into EDM [7].

Acknowledgments. CULTURA[13] (Grant agreement no. 269973) and the PROMISE network of excellence[14] (Contract n. 258191) projects, as part of the 7th Framework Program of the European Commission, have partially supported the reported work.

References

1. Agosti, A., Ferro, N., Silvello, G.: The NESTOR Framework: Manage, Access and Exchange Hierarchical Data Structures. In: Proceedings of the 18th Italian Symposium on Advanced Database Systems, pp. 242–253. Società Editrice Esculapio, Bologna (2010)

[11] Please note that the backward compatibility can be limited by the fact that the EAD expressive power is inferior to that of OAI-ORE.

[12] http://www.europeana.eu/

[13] http://www.cultura-strep.eu/

[14] http://www.promise-noe.eu/

2. Doerr, M., Gradmann, S., Hennicke, S., Isaac, A., Meghini, C., Van de Sompel, H.: The Europeana Data Model (EDM). In: IFLA 2011: World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden (2010)
3. Duranti, L.: Diplomatics: New Uses for an Old Science. In: American Archivists and Association of Canadian Archivists. Association with Scarecrow Press, Lanham (1998)
4. Ferro, N., Silvello, G.: A Methodology for Sharing Archival Descriptive Metadata in a Distributed Environment. In: Christensen-Dalsgaard, B., Castelli, D., Ammitzbøll Jurik, B., Lippincott, J. (eds.) ECDL 2008. LNCS, vol. 5173, pp. 268–279. Springer, Heidelberg (2008)
5. Ferro, N., Silvello, G.: The NESTOR Framework: How to Handle Hierarchical Data Structures. In: Agosti, M., Borbinha, J., Kapidakis, S., Papatheodorou, C., Tsakonas, G. (eds.) ECDL 2009. LNCS, vol. 5714, pp. 215–226. Springer, Heidelberg (2009)
6. Gilliland-Swetland, A.J.: Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment. Council on Library and Information Resources, Washington, DC, USA (2000)
7. Hennicke, S., de Boer, V., Isaac, A., Olensky, M., Wielemaker, J.: Conversion of EAD into EDM Linked Data. In: Proceedings of the First International Workshop on Semantic Digital Archives, SDA 2011 (2011)
8. Jech, T.: Set Theory. Springer, Berlin (2003)
9. Kaplan, D., Sauer, A., Wilczek, E.: Archival Description in OAI-ORE. In: OR 2010: The 5th Int. Conf. on Open Repositories (2010)
10. Knuth, D.E.: The Art of Computer Programming, 3rd edn., vol. 1. Addison Wesley, Reading (1997)
11. Lagoze, C., Van de Sompel, H., Johnston, P., Nelson, M., Sanderson, R., Warner, S.: ORE Specification-Abstract Data Model. Technical report, OAI (2008)
12. Light, M., Hyry, T.: Colophons and Annotations: New Directions for the Finding Aids. The American Archivists 65, 216–230 (2002)
13. Metadata Standards Subcommittee OAC Working Group. OAC Best Practice Guidelines for EAD. Version 2.0. Technical report, 61 pages (April 2005)
14. University of California. California Digital Library. CDL Guidelines for Digital Objects. Version 2.0. Technical report, 34 pages (January 2011)
15. Pearce-Moses, R.: Glossary of Archival and Records Terminology. Society of American Archivists (2005)
16. Pepe, A., Mayernik, M., Borgman, C., Van de Sompel, H.: From Artifacts to Aggregations: Modeling Scientific Life Cycles on the Semantic Web. JASIST (J. of the American Society for Inf. Science and Technology) 61, 567–582 (2010)
17. Pitti, D.V.: Encoded Archival Description. An Introduction and Overview. D-Lib Magazine 5(11) (1999)
18. Ross, S.: Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries. In: Keynote Address at the 11th European Conf. on Digital Libraries (ECDL), Budapest (2007)
19. Sugimoto, G., van Dongen, W.: Archival Digital Object Ingestion into Europeana (ESE-EAD Harmonization). Technical report, Europeana v1.0 (August 2009)
20. Van de Sompel, H., Lagoze, C.: Interoperability for the Discovery, Use, and Re-Use of Units of Scholarly Communication. CT Watch Quarterly 3(3) (August 2007)