As I found through many resources on the web (the most of which, in the “useful links” file), I’ll add in this draft the main conceptual models and web ontologies for archival description, with a brief summary, in order to better focusing the attention on the most appropriate.

**\*\*\***

**ARCHIVAL - An OWL/RDF vocabulary for describing archival collections**

Archival is a very small vocabulary, with only 8 classes (Collection and the subclass Manuscript for the archive, foaf:Agent with subclasses foaf:Group [which has subclass “Family”], foaf:Organization [which has subclass “Archive”] and foaf:Person), 15 object properties and no datatype property: it seems not to be very used, probably -I suppose – because of its kind of “general purpose” approach (eg: Dublin Core), but addressed at a very specific and specialized professional field. The only form of hierarchy in this vocabulary is between “Collection” and “Manuscript”, but it could be useful only in a very small number of situations, not considering the complexity of the structure of an archive.

\*\*\*[I don’t know, but maybe are there some tools to check whether an ontology is used (and in how many projects), or not?]\*\*\*

**RELOAD Project - Repository for Linked Open Archival Data**

* The conceptual background of the ontology is the standard for Archival Description ISAD(G), [the SKOS version of ISAD(G) is no more available on the project’s website] in order to achieve the interoperability with the EAD-coded data already improved in official mappings with ISAD(G); from ISAD(G) come OAD’s classes and properties;
* The ontology itself (OAD) starts from ISAD(G), but also adds some informative elements for a better result on the web of data (information about creators and holdings). [41 Classes; 42 ObjectProperties; 31 DatatypeProperties]
* The OCSA ontology was created to link more non-archival-but-useful-to-know-(and-in-some-way-strictly-related) data about archival holdings (e.g.: opening hours of the structures, members of the staff, services to the public, …) [ 8 Classes; 7 ObjectProperties; 3 DatatypeProperties]
* The EAC-CPF ontology is the “semantic version” of the EAC-CPF XML Standard, to describe creators of archival material. [ 37 Classes; 32 ObjectProperties; 44 DatatypeProperties]

I’ve also talked to Marilena Daquino (UniBO) and she said to me that the ontologies from ReLoad project are still used, but only from IBC (Istituto per i Beni Culturali – Emilia Romagna), because from when SAN released its own ontology (which have elements from EAD, OAD, EAC-CPF and OCSA) the archives have used that ontology instead of the ReLoad’s ones. (But SAN’s triples were last updated in December 2014) [the SAN ontology counts 40 Classes; 34 ObjectProperties; 50 DatatypeProperties]

[I’ve checked the SAN’s SPARQL endpoint today (16/11/2016) and it sends to me a 500 server error: I don’t know if it’s only a temporary or local problem or maybe the project has stopped. \*\*\*(same error: 17/11/2016) \*\*\*(18/11/2016: the endpoint works properly)]

\*\*\* As a personal observation, I’ve noticed that – in general – there are no or very few restrictions on the characteristics of the object properties (functional, transitive, symmetrical, inverse functional, ...) \*\*\*

**APEx – the Archives Portal Europe network of excellence**

About APEx:  
They have structured the data in three main layers:

* *The first layer, the Archival Landscape, consists of one EAD document and has the function of a sort of umbrella for the whole providing future users of the Archives Portal Europe with a starting point for navigational research by*
  + *naming all participating archival institutions in a structured way (f.i. sorted by country) with links to their more detailed descriptive information on the next lower level and*
  + *granting access to detailed information on the institutions themselves like contact details or opening hours given in EAG files linked to the Archival Landscape.*
* *The second layer, the Holdings Guides, consists of a set of EAD documents containing each a structured list   
  of fonds or record groups of every archival institution. It aims at*
  + *giving an overview over the fonds and collections of the single archival institutions, intended to include links to more detailed information on records creators in EAC-CPF files,* (→ **~~the same as EAC-CPF from ReLoad~~** → I’ve checked this point: they’re not using the same: in fact EAC-CPF from ReLoad is a “semantic” translation of the standard, while they’re using an XML schema application profile [apeEAC-CPF] frome the EAC-CPF XML held by the Berlin State Library of which the ReLoad’s one is the semantic version; → apeEAC-CPF according to the “State of the art report on EAC-CPF and recommendations for implementation in Archives Portal Europe” is an application profile derived from the comparison between the profiles of partner institutions [f.i. some elements which were not used by partners were discussed to be excluded from the apeEAC-CPF profile])
  + *providing short information on fonds and collections incl. f.i. conditions concerning their access and use and*
  + *linking to their detailed description on the next lower level.*
* *The third layer, the Finding Aids, consists of EAD documents containing detailed description for the single fonds or record groups structured according to series and arrangement groups. Finding Aids will*
  + *provide detailed information on the archival material from collection level down to units’ level indicating the reference numbers necessary for ordering or communication with the archival repositories and*
  + *grant access to digitisations of the archival material, if applicable*

For their purposes they have made specific schemas of EAD → apeEAD; EAC-CPF → apeEAC-CPF; METS → apeMETS; EAG → EAG 2012. I think that this “personalization” of the various profiles was pursued in order to improve the quality, with specific constraints and limitations on input, of data. [EAG 2012 «*Encoded Archival Guide (EAG) in its version of 2012, which has been developed within the APEx project in the context of expanding and enhancing the Archives Portal Europe. EAG 2012, forthwith referred to as EAG, is the result of reviewing the existing version EAG 0.2, created initially in the context of the Censo-Guía de los Archivos de España e Iberoamérica, and relating this to the International Standard for Describing Institutions with Archival Holdings (ISDIAH) as published by the International Council on Archives (ICA)*»].

Anyway, it doesn’t seem to me to be a “real” semantic ontology, but a structured metadata system, compliant with the international standards.

**LOCAH (Linked Open Copac and Archives Hub?)**  
About LOCAH:  
The LOCAH project was a JISC funded project. The main aims of the project were to make available data from the Archives Hub and from Copac in the form of Linked Data.They had used common vocabularies such as DublinCore, FOAF, SKOS, OAI-ORE (object reuse and exchange), LinkedEvents, to build (? I don't really well understand right now if I can say that they have built their own ontology <http://data.archiveshub.ac.uk/def/> starting from concepts and definition taken from other ontologies and vocabularies or if I have to say that they made a sort of mix of the existing ones but without really "making" their own ontology)...Anyway every class in the ontology is marked with a custom label (in the "Annotations" window in protégé) "term\_status: unstable", maybe meaning that the ontology it's functional to the purpose of the project, but not really totally compliant with international standards.

The unique and main class (no subclasses) for describing the structure of the archive is “Archival Resource” with the objectProperty “Level”(of description, I suppose), but not in a structured hierarchical way. [The “only” because in my opinion the “EAD Document” class is only a way to exchange interoperable data with institutions.]   
[ 24 Classes; 17 ObjectProperties; 43 DatatypeProperties]

**SNAC - Social Network and Archival Context (EAC-CPF)**About SNAC:   
They have mapped the EAC-CPF directly from the standard, not reusing the existing ontology from IBC. However, the project, born in 2011, has evolved through time (they have dismissed the SPARQL endpoint in favour of "Prototype", a navigation interface). They have a well documented section on the gitlab portal <http://gitlab.iath.virginia.edu/snac/Documentation>, and in one section they claim that:  
"*Currently there is no existing ontology for archival description, and thus the classes and properties used in exposing graph data expressed in RDF are based on classes and attributes selected from existing, well-known and widely used ontologies and vocabularies: Friend of a Friend, OWL, SKOS, Europeana Data Model (EDM), RDA Group 2 Element Vocabulary, Schema.org, and Dublin Core elements and terms* [ → they are now using BIBFRAME from LOC probably - I suppose - because of the use of MARC in the US for the archival description]*. In the long term, it should be noted that the International Council on Archives' Expert Group on Archival Description is developing an ontology (Records in Contexts (RiC)) for archival entities and the description thereof. The SNAC Cooperative will transition to the ICA RiC semantics when it becomes available*."  
Data, probably for better results in browser visualization, are coded in JSON format. The purpose of describing relations between the instances is well pursued, but maybe the constraints on the identification (they have a high threshold for automatic matching of the "sameOf: instance"; they prefer the "hypotetical" matching "maybeSameOf: instance") and on the description of the archives are a little weak. The project, in this evolution from the past, is a pilot funded by the Mellon foundation; it could be possibly become more compliant to the standard with the release of RiC-O.

Now I’m looking at the NEDA (Spain) and the Finnish Conceptual Model for Archival Description (but they don’t seem to me to be already in “an ontological form”)

**Finnish Conceptual Model for Archival Description**

The draft of the project (01/03/2013) outlines the needs for a common shared ontology among the various memory institutions in Finland and gives a brief description of the conceptual model. In my humble opinion it’s a small set, probably because of the “influence” of BIBFRAME/RDA in the conceptualization; the granularity of information and the perspective of a “broad-range” use give to each class and their properties (developed with a punctual “relationship approach”, in which each property is peculiar not to its function but rather to the classes it refers to) a sort of “lack of specificity”. From these slides (IFLA WLIC 2016 <http://www.oclc.org/content/dam/oclc/events/2016/IFLA2016/presentations/RDA-based-Data-Model-of-the-Finnish-Memory-Organizations.pdf>) it appears more clear the influence of RDA on the model, even for future evolutions. Because on my ignorance about Finnish language I can’t understand what precisely are the vocabularies of the project (<https://github.com/NatLibFi/Finto-data/tree/master/vocabularies>) but I can’t recognize any familiar name, such as EAC-CPF, EAD, OAD,...(and not even RDA). It seems to me to be a project that aims to share the same model and the same resources for LAM on national basis, so, probably, libraries seem to be more advanced in this field and push ahead Archives and Museums, and that’s (could be) why RDA is used as main model. [In the conceptual model draft one class and two properties are missing, maybe for a numbering error.]

**SAN Ontology**

SAN’s ontology uses the same models from Reload Project (OAD, EAC-CPF, OCSA), with a great importance given to Finding Aids, Holdings and context description, but with no depth in the description of the structure of the archive itself: they have only one defined class for “complesso archivistico”. “Complesso archivistico” is an equivalent class to “Unity of Description”(UOD) in the OAD ontology, but there is a main difference between them: while OAD’s “UOD” has a “level of description” property which allows at least to “label” the resource at the right level of description (and it also has a “related UOD” property, but maybe them both aren’t yet enough to create a proper hierarchy), SAN’s equivalent class to UOD (“complesso archivistico”) has only one and “broader” property: hasUpperLevel (“ha livello superiore”), that is probably even less than the OAD ontology because those weak constraints are useful to represent the basic superficial structure of an archive, but not functional to describe properly its hierarchical structure.

**Record in Context (RiC)**

I’m working now on RiC: i’m trying to make an OWL ontology out from the conceptual model.

I have some doubt about the right way to do it: in the “translation” from “conceptual-model-language” to “ontological-language” I think that “Entity” should match “Class”, “Relationship” should match “Object Property”, and so “Property” should match “Datatype property”, is it right?

It could be a quite long work, cause RiC-CM (if I’ve matched correctly the “translation”) has [ 14 Classes; ?(not 792) ObjectProperties; 67 DatatypeProperties]

About the compliance with the standards and the capability of in-depth descriptions, they claim that:

«*RiC-CM models what may be described as “multidimensional description.” Rather than a hierarchy, the description may take the form of a graph or network. Modelling description as a graph accommodates the single, fonds-based, multilevel description modelled in ISAD(G), but also enables addressing the more expansive understanding of provenance described above. The multidimensional model thus enables the description of the fonds, but also sees the fonds existing in a broader context, in relation to other fonds. [...] While RiC-CM is a model that makes it possible to describe records and the environments in which they are created, accumulated, used, and managed [...] it does not repudiate hierarchical description as such. In fact, the model assumes that Record Sets, in addition to the possibility of having individual member Records, may also have member Record Sets, hierarchically   
arranged, such as a hierarchy that represents a series that contains subseries that in turn contain files.*»

\*\*\*

I have the impression that the scarce or “not-so-effectively-improved” description of hierarchical structure in the ontologies could be owed to the fact that archival LOD are mainly used (not in theory but in the daily practice) to convert already existing archival descriptions (which infrequently cover the full hierarchical structure from “fond” to “item”) in LOD ( → therefore ontologies are useful for the upper level/levels) OR to describe digitized/digital collection at the item level ( → so ontologies are useful for the lower level). This is only my very personal opinion, but, I think, this could be one of the reasons: because of the extent of documentation (and the description cannot account of all the complexity item by item) or because they have to deal only with items (and maybe not even considering the item as part of its original fond, but as part of a collection by a digitization project).

\*\*\*

[\*\*\*

I’ve add a folder (“Ontologies”) with the main ontologies in OWL or RDF; inside the subfolder “Reload” there are the OAD, EAC-CPF and OCSA ontologies, and in the “Demo” subfolder there are the “Finnish model”, and a first attempt to put RiC-CM in the form of an ontology, for now only with classes, later I will try to add also all the properties. You can obviously open all the files with Protégé.

\*\*\*]

Table with unique properties in RiC-CM (and their inverse) [they should be 176 unique properties; 792 is the possible number of relations with these unique properties, but they differ in domain and range (and a couple of times also for their respective inverse properties), so - just because it’s a draft - in the ontology I have put all the 792 relations as unique properties]

|  |  |
| --- | --- |
| HasCopy → IsCopyOf  HasDraft → IsDraftOf  HasOriginal → IsOriginalOf  HasSubject → IsSubjectOf  IsAssociatedWith → IsAssociatedWith  IsPredecessorOf → IsSuccessorOf  HasPart → IsPartOf  IsMemberOf → HasMember  HasRightsHeldBy → IsRightsHolderOf  IsHeldBy → IsHolderOf  IsOwnedBy → Owns  IsEvidenceOf → IsEvidencedBy  HasHoldingLocation → IsHoldingLocationOf  HasCoverageDate → IsCoverageDateOf  IsOwnerOf → Owns  Manages → IsManagedBy  Controls → IsControlledBy  HasAssumedIdentity → IsAssumedIdentityOf  HasFunctionalRelationWith à  IsChildOf → IsParentOf  IsDirectedBy → IsDirectorOf  IsSiblingOf → IsSiblingOf  IsSpouseOf → IsSpouseOf  IsSubordinateOf → IsSuperiorOf  Uses → IsUsedBy  Pursues → IsPursuedBy  Occupies → IsOccupiedBy  Fulfils → IsFulfilledBy  Performs → IsPerformedBy  HasActiveDate → IsActiveDateOf  HasEndDate → IsEndDateOf  IsCitizenOf → HasCitizen  IsLocatedAt → IsLocationOf  RequiresCompetency → IsRequiredCompetencyOf  IsLocatedIn → IsLocationOf  IsExampleOf → HasExample  Defines → IsDefinedBy  HasJurisdiction → IsJurisdictionOf  IsHoldingPlaceOf → HasHoldingLocation | WasAssociatedWith → WasAssociatedWith  HadPart → WasPartOf  WasMemberOf → HadMember  HadRightsHeldBy → WasRightsHolderOf  WasAddressedTo → WasAddresseeOf  WasAuthoredBy → Authored  WasCollectedBy → Collected  WasCreatedBy → Created  WasHeldBy → WasHolderOf  Owned → WasOwnerOf  WasSentBy → Sent  WasWrittenBy → Wrote  ResultedFrom → ResultedIn  HasDocumentaryForm → IsDocumentaryFormOf  HadCreationDate → WasCreationDateOf  HadHoldingLocation → WasHoldingLocationOf  WasCreatedAt → WasCreationLocationOf  WasAccumulatedBy → Accumulated  WasArrangedBy → Arranged  Owned → WasOwnedBy  Assembled → WasAssembledBy  Managed → WasManagedBy  Controlled → WasControlledBy  HadFunctionalRelationWith à  Used → WasUsedBy  WasChildOf → WasParentOf  WasDirectedBy → WasDirectorOf  WasSiblingOf → WasSiblingOf  WasSpouseOf → WasSpouseOf  WasSubordinateOf → WasSuperiorOf  Pursued → WasPursuedBy  Established → WasEstablishedBy  Occupied → WasOccupiedBy  Fulfilled → WasFulfilledBy  Performed → WasPerformedBy  WasAuthorizedBy → Authorized  HadActiveDate → WasActiveDateOf  HadBirthDate → WasBirthDateOf  HadDeathDate → WasDeathDateOf  HadEndDate → WasEndDateOf  HadStartDate → WasStartDateOf  WasCitizenOf → HadCitizen  WasLocatedAt → WasLocationOf  RequiredCompetency → WasRequiredCompetencyOf  WasLocatedIn → WasLocationOf  IsFulfilledByPerformanceOf → IsPerformedToFulfil  WasFulfilledByPerformanceOf → WasPerformedToFulfill  Defined → WasDefinedBy  Revised → WasRevisedBy  HadJurisdiction → WasJurisdictionOf  WasRevisionDateOf → HadRevisionDate  WasHoldingPlaceOf → HadHoldingLocation  WasBirthPlaceOf → HadBirthPlace  WasDeathPlaceOf → HadDeathPlace |