

Records in Contexts (RiC)

An Archival Description Draft Standard



INTERNATIONAL COUNCIL ON ARCHIVES
EXPERTS GROUP ON ARCHIVAL DESCRIPTION

AVEC 2017

Visby, Sweden

2017

Presentation Overview

Background, Overview of Process and Products

Observations on Comments, Concerns, Design Considerations, Challenges

Next Steps

Background

Overview of Process and Products

Experts Group on Archival Description (EGAD)

- Formed by the ICA Programme Commission in late 2012
- Partial successor to the Committee on Best Practices and Standards (CBPS)
- Term 2012-2016; 2016-
- Charged with developing a Conceptual Model for Archival Description
 - Based on four current ICA descriptive standards
 - Employing formal information modeling techniques

ICA Standards for Archival Description 1988-2008

<u>Standard</u>	<u>Edition</u>	<u>Development Dates</u>	<u>Publication Date</u>
Principles		(1988) 1989-1992	1992
ISAD	1 st	1990-1993	1994
ISAAR	1 st	1993-1995	1996
ISAD	2 nd	1996-2000	1999
ISAAR	2 nd	2000-2004	2004
ISDF	1 st	2005-2007	2007
ISDIAH	1 st	2005-2008	2008

Archival Principles: Records in Contexts

- Principle of Provenance
 - Respect des fonds
 - The Records created, accumulated, and used by a person or group in the course of life and work are to be kept together and not intermixed with records from other sources
 - Respect for Original Order
 - The intellectual grouping of and sequencing imposed on the records in the context of accumulation and use is essential to understanding the interrelations among them as well as being evidence of how they were used
- General international consensus on the principle
- But historical, cultural differences in understandings
- Records in Contexts
 - Embodies both facets of the principle
 - Though more expansive understanding of Provenance
 - Based on intellectual and practical critique of archival description
 - Records and the people that create, manage, and use them do not exist in isolation but in complex layers of interrelated, interdependent contexts

Historical Context

- Since at least mid-19th century, cultural heritage communities
 - Reimagine description in relation to emerging and new communication technologies
 - Trend
 - Separate the components of description
 - To efficiently and more effectively create prevailing access tool (e.g., book catalog, finding aids)
 - At the same time, enable new tools, new perspectives, new paths, based on recombining the components
- Four ICA standards reflect this trend
- Though the separation and new perspectives not realized

Current and Emerging Technology Landscape

- Network, of course, and Markup (XML), and Database (SQL) ...
- XML and SQL have dominated but ...
- Emergence of Graph technologies: RDF, Semantic technologies and Linked Open Data
 - More expressive, but also more challenging: complexity, quality ...
- Opportunities: separation, recombining, interrelating, opening domain borders, new perspectives, new paths ...
- Reposition community to take advantage of the opportunities

The RiC Products

- Conceptual Model for archival description (RiC-CM)
 - The Conceptual Model resembles the current ICA standards
 - Documents the key entities of archival description and the properties of each
 - With diagrams illustrating how the components are interrelated to form complete archival description
- An Ontology for archival description (RiC-O)
 - Based on RiC-CM
 - Expressed using the W3C OWL language
 - Will map archival description concepts to similar concepts employed by allied communities: integrated access to cultural heritage
 - Will enable archival community to participate on its own terms, so-to-speak
 - RiC-O draft available late 2017
- Implementation Guidelines

From ISAD(G)

- Predominant form of archival description today
 - Hierarchical (top-down) description of a single fonds
 - Description of the whole, the parts of the whole, parts of the parts
 - Largely if not exclusively self-contained, inward “looking”
 - That is, not connected to the broader context
 - All contained in a single apparatus
 - ISAD(G) a model for this approach; EAD a method for communicating it

To RiC: Pivotal Changes

- Records and aggregations of records treated as two distinct entities
 - Records
 - Record Sets
 - Over the course of its existence, a record may be a member of more than one record set, and at the same time
- Multilevel description
 - Multilevel or hierarchical description one among other possible methods of description
 - Multilevel description predominates, and will do so for the foreseeable future: well understood and economic
- Multidimensional description
 - Encompasses multilevel description
 - Within a network of interrelated records, fonds, people ...: context within context
 - Enables more flexible description (relational and graph) that is more expressive of the complex realities of records than possible in a single hierarchical description
- Description of a vast social-document network

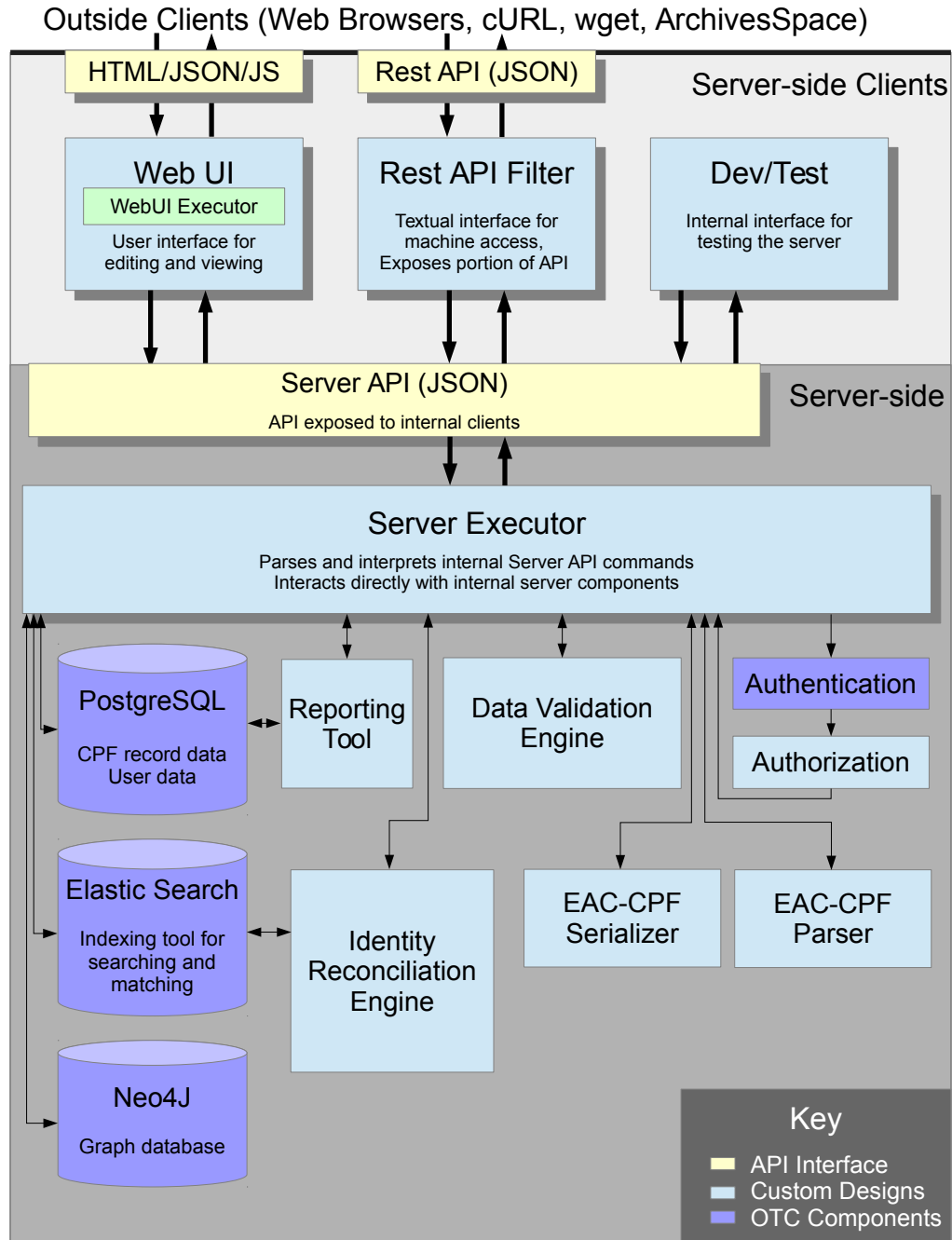
Observations on Comments, Concerns, Design Considerations, Challenges

Comments

- Comment Period from September 2016 through January 2017
- Sixty-two sets of comments from individuals and groups
- From 19 countries
- Overall positive
 - Some very enthused
 - Some pretty enthused
 - Some supportive but worried and anxious about what it all means
- Or put another way, the comments were “Yes, very good, but ...”
- What comes after the “but,” when compiled, is over 200 pages! Some high level suggestions, some requests for clarification, some very specific observations ...
- To generate such interest is a very good thing, we think!
- But it also will require a lot of work to sift through and digest all of the comment, giving each due consideration, discussing, making decisions

Observations: Development Strategy

- Develop a high-level model of the essential components of archival description, the identifying characteristics of each, and the relations among them
- Focus on a model that supports the creation, maintenance, and publication of archival description
 - Current ICA standards focus on modeling the publication (end-product) without consideration of the creation and maintenance, of the system used to produce it
 - RiC focuses on modeling data to support the design and creation of maintenance *and* publications systems
 - In information technology terminology, the way the data is created and maintained is optimized for those functions, is configured in such a way, that multiple products, multiple ways of presenting the data are possible
 - Including the familiar finding aid, the fonds described hierarchically in a single apparatus
- Is it feasible: it is possible to build such a system, right now, using all open source software
- Social Networks and Archival Context Cooperative (SNAC) is an example



Observations: It Will Take Time, No Need to Rush

- RiC is an aspiration, pointing to where we want to go
- Developing widely available, affordable systems that will enable archives, large and small, to describe in a way that takes advantage of existing and emerging technologies is a challenge, it will take time to unfold
- Existing systems, many though not all based on ISAD(G)
 - Relational databases (records management tools built "over" them
 - EAD encoded finding aids
 - Word processing files
 - Spreadsheets
- The EGAD assumes that all of these will continue to exist for some time while a few pioneers begin to implement RiC and demonstrate its feasibility
- This is all perfectly understandable and acceptable, there is no need to rush
- But we do hope that we are laying the foundation for the next generation of description

Observations: Entities and Properties

- Terminology is a challenge:
 - Entity-relation models
 - Ontologies: class, properties
 - XML: elements and attributes
 - Compounded by all in different languages
- RiC-CM uses “entities,” “properties,” “relations”
- When is something an entity, when a property, when a relation?
- The strictly archival entities do not present too much of a challenge: Record, Record Set, Agent, Mandate, Function but what about the other entities? Documentary Form, Date, ...
- Here are the criteria employed
 - An entity is defined/identified by its characteristics
 - When the characteristics are exclusively related to the entity, one-to-one, or many-to-one, they are modeled as properties
 - When the characteristic is shared by many different entities *and* itself has characteristics that we want to define, Documentary Form, for example, then the it is “elevated” to being an entity

Observations: Date

- Date is in a category of its own
- Date do have their own characteristics: calendar system being a major one, but also that we typically want to express dates in human-readable form and in a standardized way
- Date is traditionally treated as a property in a one-to-one, or likely, many-to-one relations with an entity. A record may have a creation date as well as additional dates associated with curation events of one kind or another.
- As so while date is, on one hand, a many-to-one with the various entities, it is treated as an entity rather than a property because each date has characteristics that we want to represent.

Observations: Relations

- Relations are in a category all of their own
- Two kinds of relations:
 - Property relation or “Hold of” relation, for example, the thing is black
 - One side of the relation is not in itself a distinct entity
 - Between relation or “Hold between” relation
 - Each side of the relation is a distinct entity with its own properties
- RiC-CM uses the term Relation only for the latter, the “Hold between”
- Though not in the model because we could not figure out how to present it, RiC also intends to treat the Relation as an entity, that is, a thing that has properties
- It joins two entities but at the same time it has characteristics:
 - When the relation holds
 - Where the relation holds
- Every relation of both types is also an assertion of some observed fact, to assert that an entity has a characteristic, or to assert two entities are related to one another some how
- While I will not go into greater detail at this time, the last statement is essential in modeling the role that records managers and archivists play in description

Observations: How Much Detail is Enough?

- The EGAD is endeavoring to be comprehensive, to cover all aspects of record description as we understand them
- While we know that not all or even most of the detail will be needed or used regularly, there are very important reasons for trying to be comprehensive
 - Focusing on each and every thing we can think of; struggling to push our understanding, ensures that we understand the whole of what is possible and ideal, what the parts are, and how they related to one another
 - Such an approach provides us with the best opportunity to model the best description, whether that description is minimal, or full, and that the minimal may, over time be augmented if need and resources permit
- The risk, though, in such an approach, is that the detail is overwhelming, confusing, that the essential is lost in the detail
- Thus the EGAD needs to find a way to present RiC in a manner where the essential is understood and in the foreground, but without sacrificing being comprehensive and as detailed as necessary to be comprehensive

Next Steps

- We are in the process of organizing and analyzing the large number of comments, classifying them
- Over the course of the summer and early fall, we will hold a series of teleconferences to begin to discuss and build consensus on revisions
- The discussions will not only focus on particular entities, properties, and relations, but the way in which RiC-CM is presented
- Among other things, more examples, both traditional format records and electronic records
- Meet in Rome the October 2017
- We hope to have RiC-CM second draft completed by the end of the year

Next Steps

- RiC-O, the ontology, depends on a stable draft of RiC-CM
- Nevertheless, provisional drafting of RiC-O is underway, as RiC-O is also a way of testing RiC-CM concepts and structures
- We hope to make a draft of RiC-O available later this year

Thank you

Draft RIC-CM available at:
<http://www.ica.org/egad-ric-conceptual-model>