



# Records in Context (RiC)

An international standard for archival description  
Progress report

International Council on Archives  
Expert Group on Archival Description (EGAD)

Cleveland  
2015

---

# Presentation Overview

---

## **Introduction: background and context**

Daniel Pitti

Institute for Advanced Technology in the Humanities, University of Virginia

## **RiC Conceptual Model**

Bill Stocking

British Library

## **RiC Ontology**

Aaron Rubinstein

University of Massachusetts Amherst



# Introduction

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

---

## ICA and 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 <sup>st</sup>	1990-1993	1994
ISAAR	1 <sup>st</sup>	1993-1995	1996
ISAD	2 <sup>nd</sup>	1996-2000	1999
ISAAR	2 <sup>nd</sup>	2000-2004	2004
ISDF	1 <sup>st</sup>	2005-2007	2007
ISDIAH	1 <sup>st</sup>	2005-2008	2008

- Principle of Provenance
  - All records created or accumulated and used by the same agent kept together
- Respect for the Original Order
  - Preserve the interrelation of the records established in the context of their creation and use
- General International consensus on principles
  - But historical, cultural differences in understandings
- Records in Context
  - Embodies both principles
  - Embodies the international consensus

Since at least mid-19<sup>th</sup> 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

- 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



- Conceptual Model for archival description
  - The Conceptual Model will resemble the current ICA standards
  - Document the key components of archival description and the properties of each
  - With diagrams illustrating how the components are interrelated to form complete archival description
  - Draft for public review fall 2015
- An Ontology based on the Conceptual Model
  - 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
  - Draft for public review winter 2016

---

## Some Highlights

---

- 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 and ...
  - Enables more flexible description (relational and graph) that is more expressive of the complex realities than possible in a single hierarchical description



# RiC Conceptual Model

Today I will talk about:

- What do we mean by a ‘Conceptual Model’ as a document – Records in Context (RiC)
- Update on progress:
  - Provide insight to debates
  - Very much work in progress!
- Next steps...



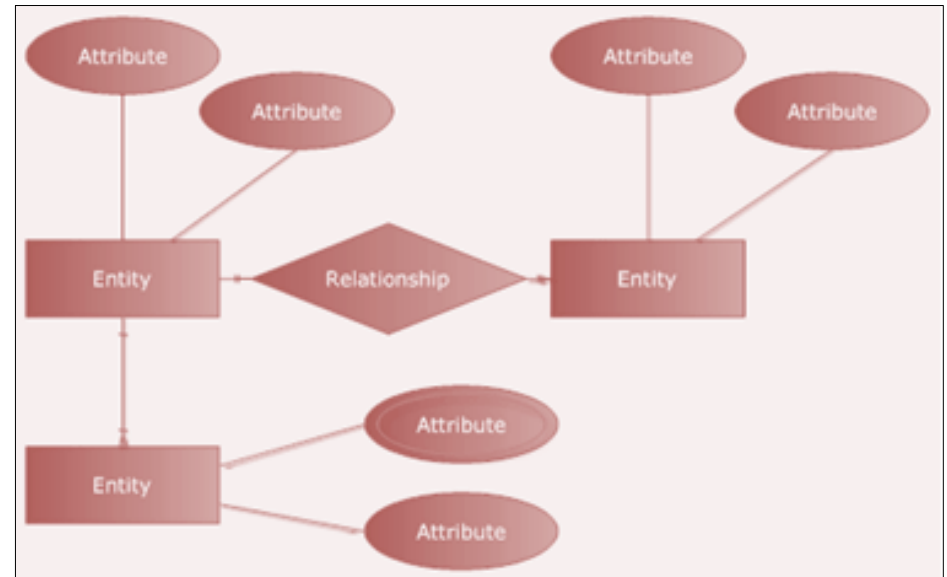
Experts Group on Archival Description  
Group d'Experts sur la Description Archivistique

**Records in Context:  
A Conceptual Model for Archival  
Description**

## A 'CONCEPTUAL' MODEL?

The conceptual model will be a document that:

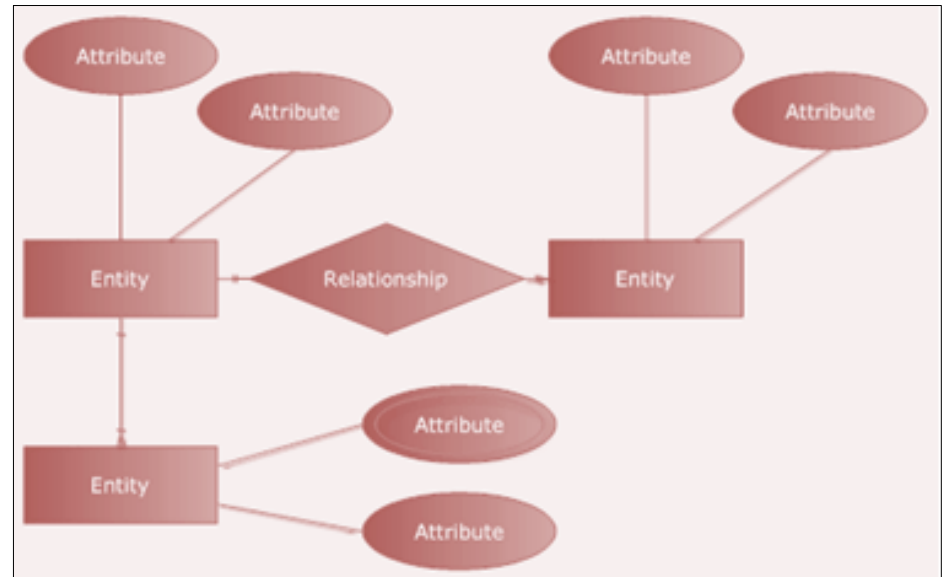
- describes key high level concepts for archival description:
  - Entities (Record, Agent etc)
  - Their properties (id, title, description)
  - The relations between them (Record 'is created by' Agent)
- A generalised view of archival description that does not replace:
  - detailed logical and physical data models
  - cataloguing guidelines (e.g. DACS, RDA)
  - XML schema for data transmission (e.g. EAD3 or MARC)



## A 'CONCEPTUAL' MODEL?

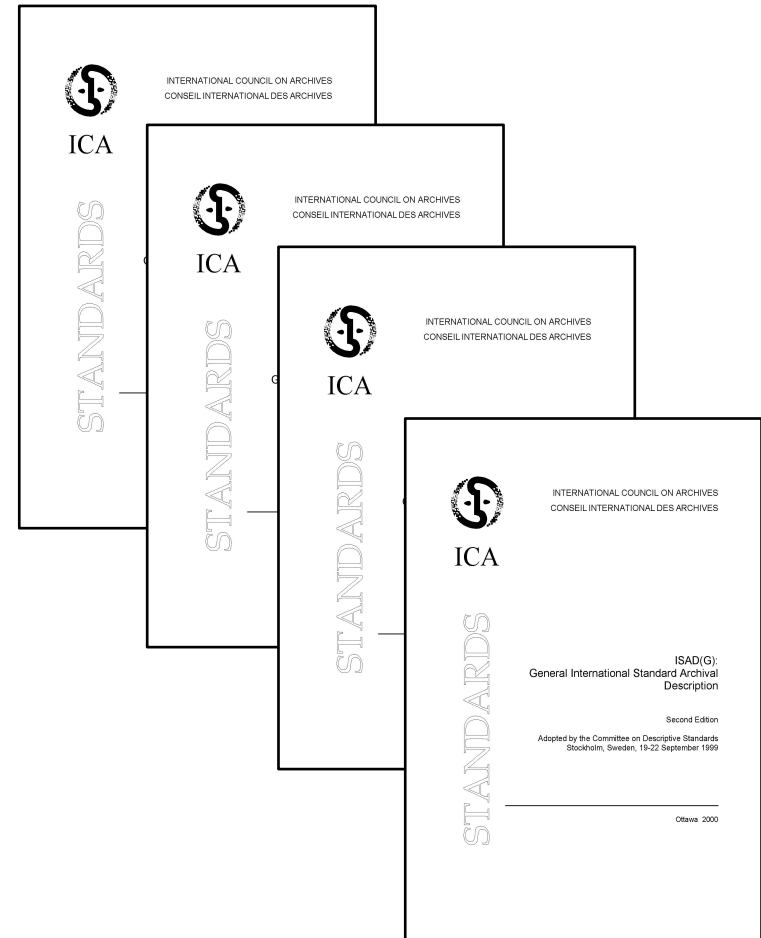
The conceptual model will be a document that:

- frames and respects current practice, systems and process yet provides a basis for their ongoing development
- informs ongoing professional discussions, education and training
- enables us to collaborate with other information professionals



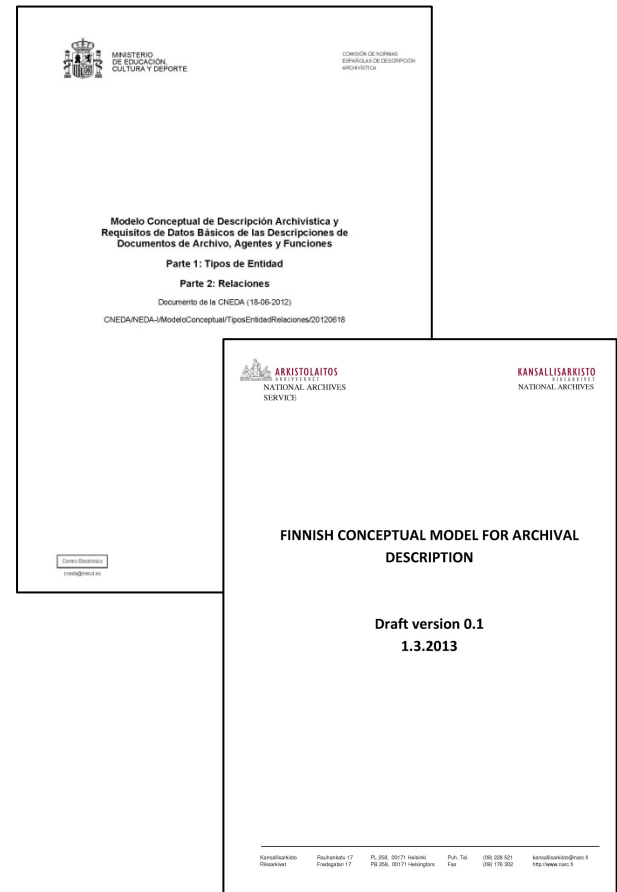
## ICA Framework for Archival Description:

- ISAD(G) and related standards for archive creators, functions and archive institutions
- Previous ICA Committee on Best Practice Standards' reports on harmonisation and relationships



## Archive description models:

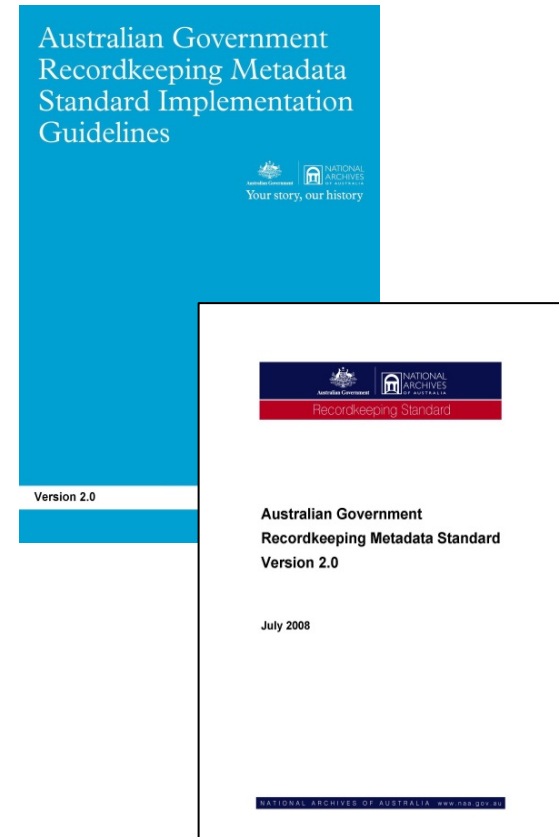
- CNEDA, *Modelo conceptual de descripción archivística y requisitos de datos básicos de las descripciones de documentos de archivo, agentes y funciones*, 2012
- Kilkki, Jaana, Outi Hupaniittu, Pekka Henttonen, *Towards the new era of archival description - the Finnish Approach*, 2012 and updates as the work has progressed





## Record Keeping Model:

- National Archives of Australia, *Australian Government Recordkeeping Metadata Standard Version 2.0*, 2008 (AGRkMS);
- and *Implementation Guidelines*, 2011



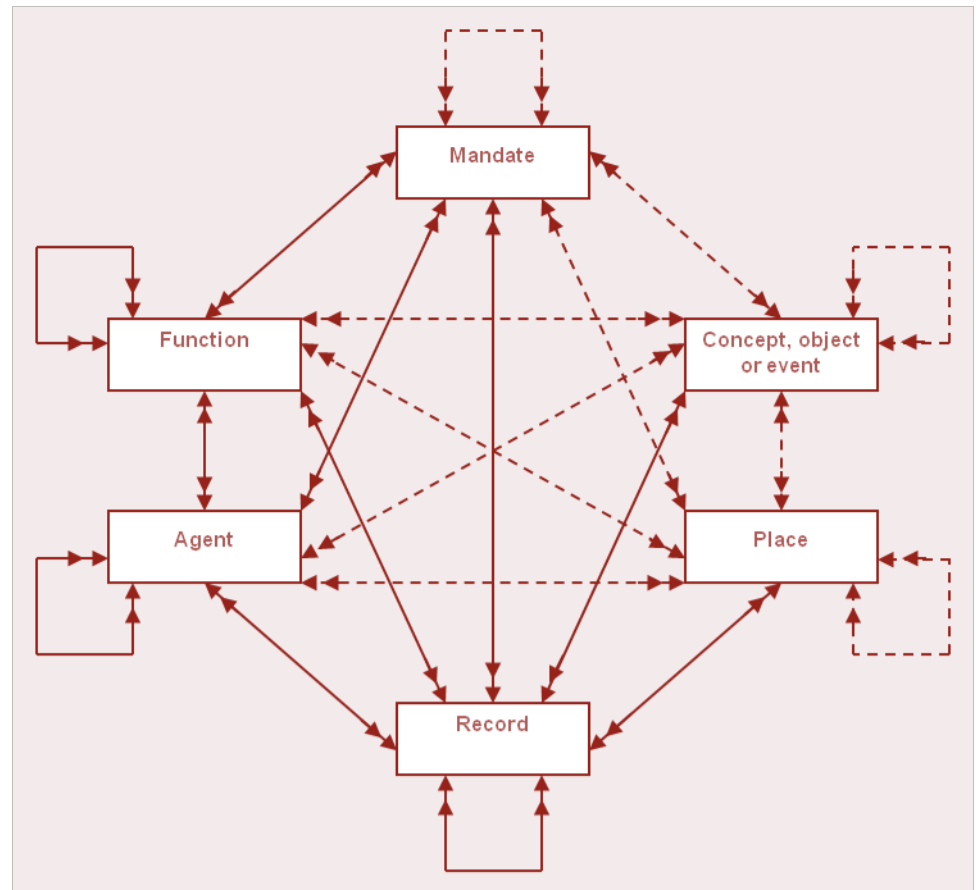
Also looking at current versions of key models from other cultural heritage domains:

- Libraries: Functional Requirements for Bibliographic Records (FRBR) and related standards for authority data – now being rationalised in to a single reference model (FRBR-LRM)
- Museums, galleries etc: The CIDOC Conceptual Reference Model (CRM) and as harmonised bibliographically in FRBRoo



We found:

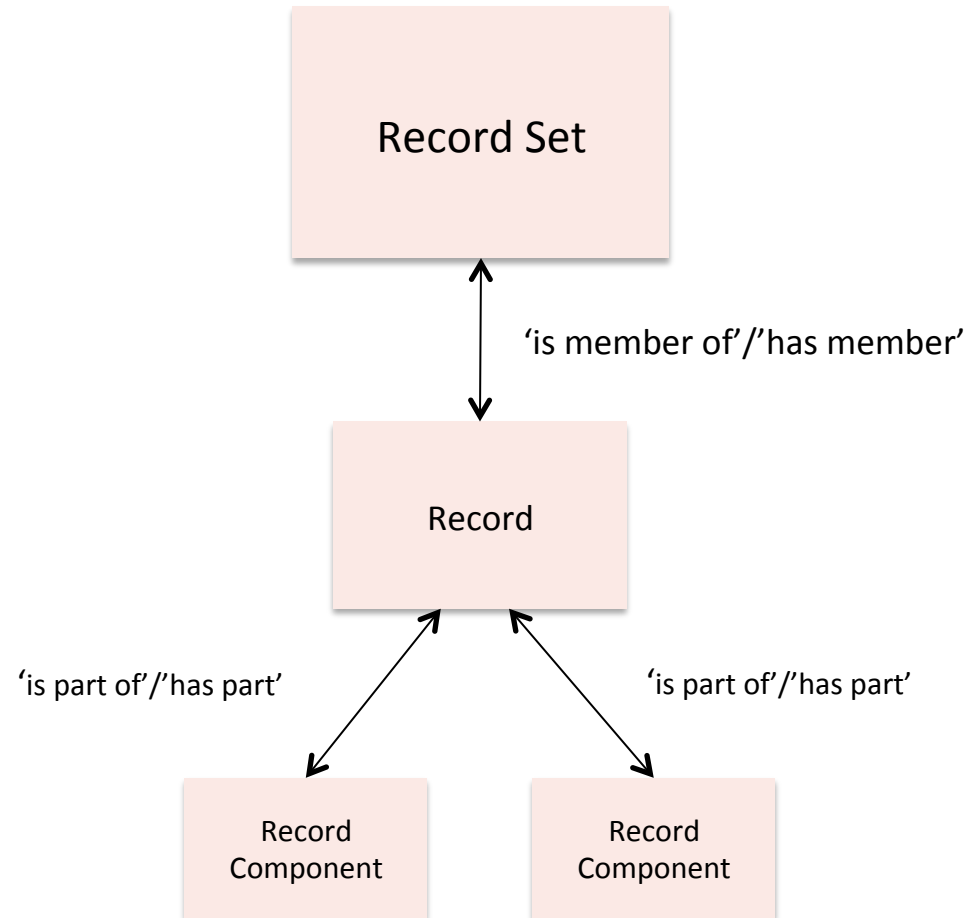
- Differences in perspective due to context of modelling
- A great deal of agreement:
  - Key entities: Record, Agent, Function or Business and Mandate
  - Stress contextual entities but model differently: Topics, Dates, Events and Places
  - significant properties and relations between entities for archival description



Currently we have three Record entities in

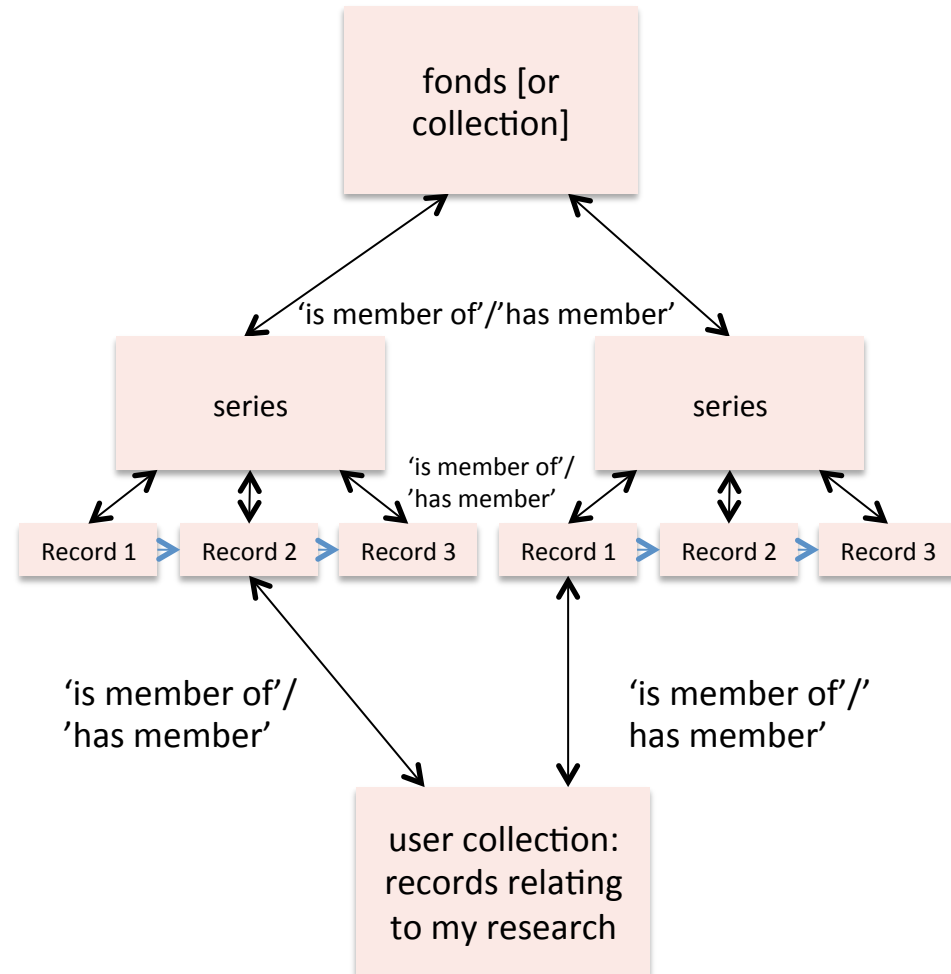
RiC:

- Record: 'information created, received and maintained in the conduct of business or affairs' - an 'item'
- Record Component: the parts that make up a record
- Record Set: the groups or aggregations of which a Record may be part



### Record Set:

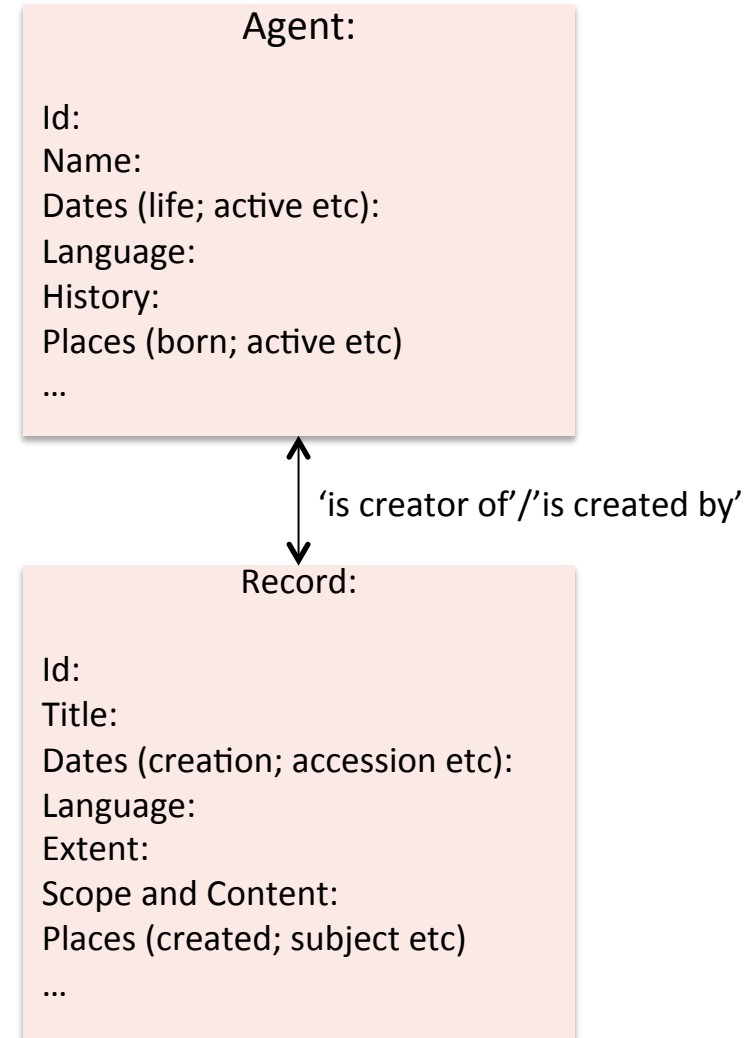
- Records grouped together by an Agent based on shared properties in order to serve their own ends: provenance, topic, place etc
- May hold other record sets in a hierarchy so caters for traditional multi-level provenance based description at fonds or series 'level' and other types of 'collections'
- Also allows records to be grouped in other ways by creators, archivists and users, simultaneously or over time



Good progress made setting out the key properties for our entities:

Two types of property:

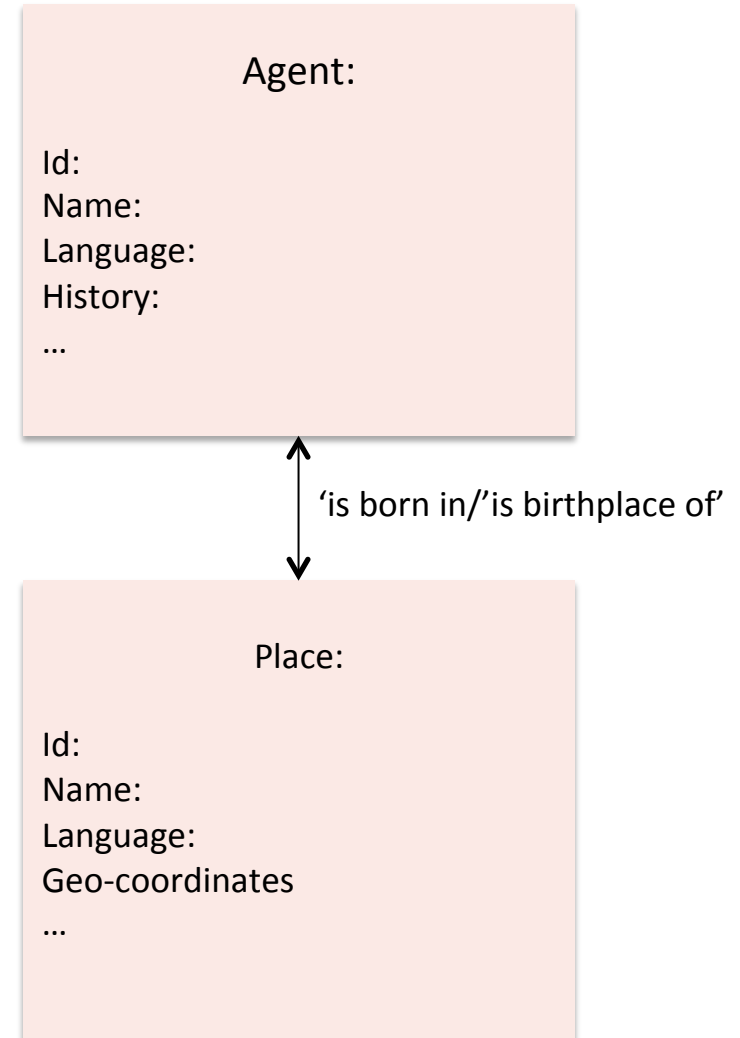
- Those that identify the entity that are common to all:
  - globally unique id
  - other ids
  - name or title
- Those that describe the entity which may differ. Record:
  - will include 'language', 'access conditions' and 'extent' etc
  - but not 'creator' as that modelled as a relation between a record and an agent



Good progress made setting out the key properties for our entities:

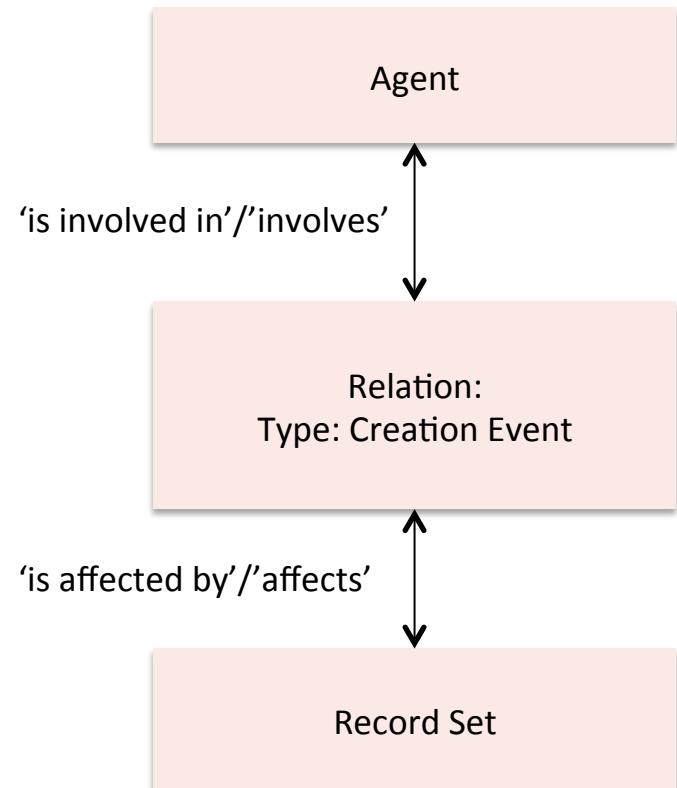
Dates and Places:

- Currently all entities have date and place properties
- As an entity may have many associated dates and places and we might want to represent their properties as well (calendar or geo-coordinates) so discussing having a Date and Place entity



Discussions about how relations between entities are also continuing:

- Lists of simple binary relationships developed for each entity:
  - ‘is creator of’/‘is created by’
  - ‘is part of’/‘has part’
- We may though what to say more about the act of creation for example than this allows so we are discussing inclusion of a Relation or Event entity





So a good start made by leveraging existing work, next:

- Continue to monitor related work
- Finalise upper level entities and their parts including how relationships between them should be represented
- Produce a first draft for feedback to ensure that we have a consensus
- Validate the model against the developing ontology...



**Thank you!**



# RiC Ontology

### **What is the RiC Ontology?**

a machine-readable version of the conceptual model

a tool for wider integration and sharing of archival description

an opportunity for testing and community feedback

### **machine-readable...**

EAD/XML is machine readable!

RiC is an OWL ontology: W3C standard in RDF for defining entities and their relationships

RiC will provide a semantic framework for connecting distributed components and entities and descriptions

## machine-readable...

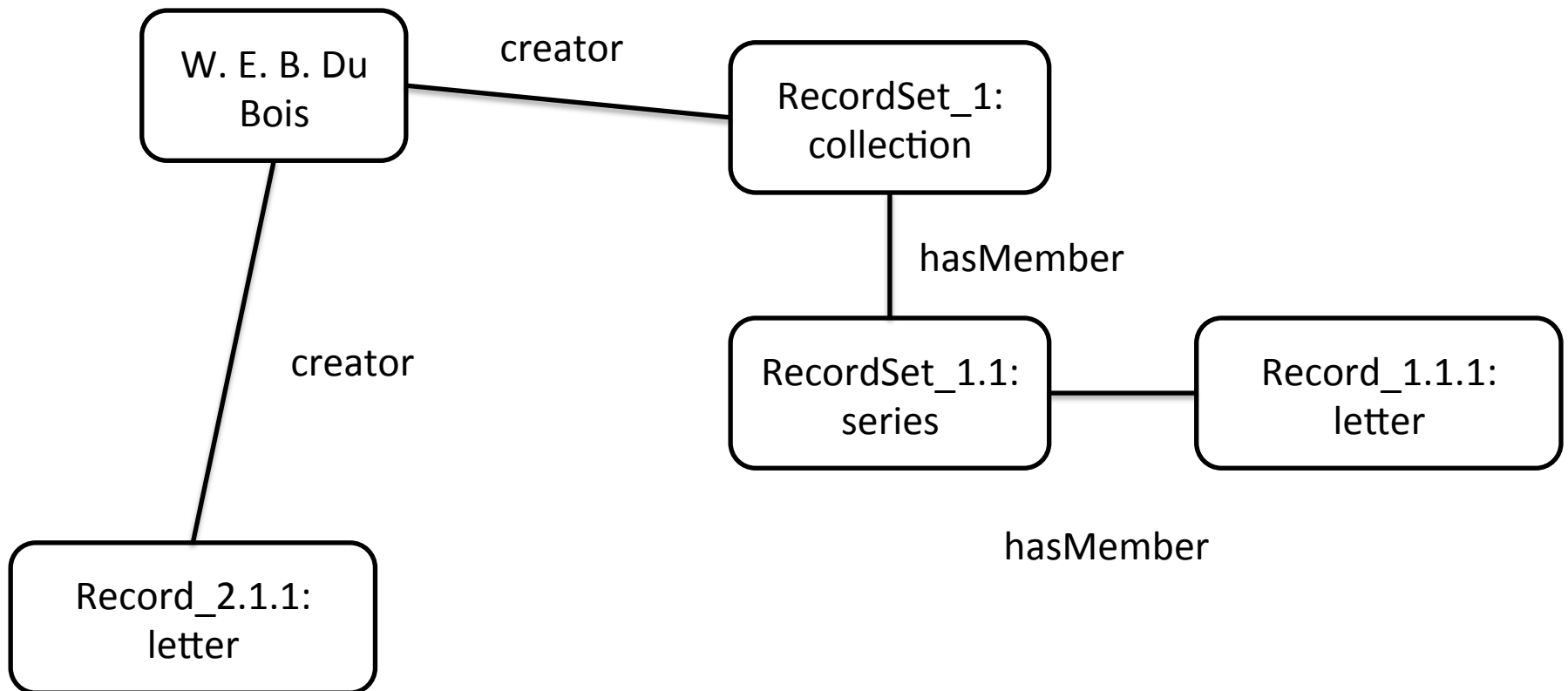
W. E. B. Du  
Bois

RecordSet\_1:  
collection

RecordSet\_1.1:  
series

Record\_1.1.1:  
letter

## machine-readable...



## machine-readable...

### RDF triple statements:

umass:W._E._B._Du_Bois	<i>ric:creator</i>	umass:Du_Bois_Papers
umass:Du_Bois_Papers	<i>rdf:type</i>	ric:RecordSet
umass:Du_Bois_Series1	<i>ric:isMemberOf</i>	umass:Du_Bois_Papers
umass:Aronson_Letter1	<i>ric:isCreatedBy</i>	umass:W._E._B._Du_Bois



## machine-readable...

### RDF triple statements:

umass:W._E._B._Du_Bois	<i>ric:creator</i>	umass:Du_Bois_Papers
umass:W._E._B._Du_Bois	<i>ric:subjectOf</i>	umass:Du_Bois_Exhibit
umass:Du_Bois_Exhibit	<i>rdf:type</i>	ric:RecordSet
umass:Du_Bois_Exhibit	<i>ric:hasMember</i>	umass:Aronson_Letter1

**machine-readable...**

*RiC puts our current practices (EAD) into  
a wider conceptual framework*

## integration and sharing

umass:W.\_E.\_Du\_Bois    *owl:sameAs*    viaf:W.\_E.\_B.\_Du\_Bois  
umass:W.\_E.\_Du\_Bois    *owl:sameAs*    snac:W.\_E.\_B.\_Du\_Bois

**merging multiple datasets**

**open world vs. closed world**

## integration and sharing

mappings and integrations:

*CIDOC-CRM*

*FRBRoo*

*FOAF*

*dcterms*

*PROV-O*

*schema.org*

## **testing and feedback**

feedback loop: developing and testing the ontology will help us build and refine conceptual model

begin to build a practical framework for implementation

## testing and feedback

*milestones:*

**draft: winter 2015**

**version 1: fall, 2016**

all versions will be released on GitHub:

<http://github.com/ICA-EGAD>

## **ontology team**

**Florence Clavaud**, *Archives National de France*

**Pete Johnston**, *ArchivesHub*

**Daniel Pitti**, *University of Virginia*

**Aaron Rubinstein**, *University of Massachusetts Amherst*

**Salvatore Vassallo**, *Archivum Romanum Societatis Iesu*