ICA EGAD Meeting, April 13-15, 2016, Paris

Attendees:

- Florence Clavaud
- Beatriz Franco Espiño
- Padré Lydie Gnessougou Baroan-Dioumency
- Gavan McCarthy
- Vitor Manoel Marques da Fonseca
- Victoria Peters
- Daniel Pitti
- Bogdan-Florin Popovici
- Javier E. Requejo Zalama
- Stéphanie Roussel
- William Stockting
- Kat Timms
- Stefano Vitali

Caveats:

- *Revisions to the entity definitions as discussed and developed during the meetings are not reflected here, rather just a summary or overview of many of the high-level discussion points (not all could be captured, unfortunately).*
- *Apologies up front for any misunderstandings or omissions on my part!*

DISCUSSION NOTES

Relations:

- Comment that it might be helpful to provide comments (definitions or examples) to help explain their meaning.

- For properties shared by relation, FC suggested adding Certainty and Source to explain the relation. DP countered by stating that this information is not about the thing directly, but rather about the information about the thing. BS mentioned that this is about control/administrative metadata, and DP responded that yes and that it was in scope and will have to be addressed in the intro to the draft but hadn't yet been worked on. Description is part of the context of the Record while also being a Record. Certainty is also not easy to determine, but there may be a place for it elsewhere.

Inheritance

- Discussion of tree (hierarchy) vs. directed acyclic graph. One node with children nodes that are siblings, vs. (almost inverse view) with one node up top with three nodes below that can inherit. BP: in

hierarchy, children can share properties with the parent, not just vice versa. Sequence is a relationship as well.

Entities Generally

- Need to clarify intent or label of 'Data Type' column, as there is inconsistency at the moment – some are specific types of entities (e.g. Agent: Group, Person), while others are types of data that could be used (e.g. Controlled Vocabulary, Text).

- JR/BF/SV: Could have higher class for 'entity' with entity types such as Agent, Record, etc.

- DP: Prefer not to take this approach, though it could be reflected through a particular implementation (e.g. table1 for Agent entity separate from table2 for Record entity).

Properties shared by Agent, Record, Record Component, Record Set, Function, Activity, Mandate

- Note that these are also applicable to Relation.

Properties Shared by Relation

- Note that these are in addition to the Properties otherwise shared by Agent, Record, Record Component, Record Set, Function, Activity, Mandate and Relation (i.e. Global Persistent Identifier, Local Identifier, Name, General Note).

Agent Entity

- DP: Archivists are concerned with agents as identities. How do we identify the real-world entity? Although we can perceive and interact with living people, most of them we only know indirectly through information about them, and dead ones are only accessible indirectly. Also, agents can have more than one identity.

- DP: Will want to have a corporate body/person/family (CPF) as an agent with the possibility of having more than one identity associated with them. The certainty of the assertion that the real-world CPF exists/existed cannot be 100%, but rather on a spectrum from absolutely sure to unsure.

- FC: We are able to identify an existence of an Agent through traces of what they did, their impact, and will collect information about them in this way.

- DP: We're modeling one or more identities that point to the real thing.

- SV: There is also attributed identity. Assigning 'anonymous' is also done by someone.

- DP: Yes, this is the case for any identities, including anonymous. No different, just sparser evidence.

- FC: Some aspects of identity are equivalent to occupation - e.g. president.

- DP: In a legal sense, being president is an identity.

- FC: There is a French group revising rules for describing corporate bodies, looking to Resource Description and Access (RDA) to see if it can work.

- BP: President of the USA is distinct from any individual person who is president.

- DP: Example of Bill Clinton's authority records - one for the man, one for him as President, and two more for when he was governor. Four separate identities as per his roles/legally recognized positions.

- FC: This is not an identity but a position/post/office.

- DP: There is the person, the office held by the person, and these kept separate and linked together.

- FC: Do not create a separate authority for the official position. RDA treats the presidency as a corporate body, which doesn't work in the French context.

- DP: Common practice to include executive positions among identities.

- FC: This is fine but there should be flexibility to accommodate other approaches - e.g. for executive roles to be handled as occupations. Planning to use ISDF to describe official positions such as Prime Minister.

- Could there be authorities for individual persons (ISAAR-CPF), an abstract authority for the office (ISDF) and other authorities for each instance of a particular person who holds that office (ISAAR-CPF).

- BS: Grey area for records related to Ministers (those not strictly personal, not strictly official, but falling somewhere in between).

- DP: Assumed identity should not be confused with the identity of the official position a person may hold (e.g. presidency). Referring to Obama is a shortcut for pointing to the Office of the President.

Agent Properties

- Note: 'Assumed Identity' has accidentally been repeated in the 'Type' row – should be removed.

- Note: No decision was stated whether 'Individual Assumed' and 'Joint Assumed' are sub-types of 'Assumed Identity' in row 'Identity Type'.

Record Entity

- DP: Whether or not a Record is evidence/a trace is an absolute (yes or no); whether it is reliable evidence or not is a separate issue. Need to be clear that this quality assessment is separate (and out of scope). The intent of the RiC definition is inclusivity, to accommodate all sorts of information things that archivists deal with. Official Records and everything else held by an archives represents evidence (in the broad sense of the word) of something. Traces constitute evidence (of existence) without any value judgment or interpretation by anyone re: its reliability, quality, etc.

- DP: Need to make intent of this definition clear: the inclusivity, and the fact that 'evidence' is meant in a generic sense. Evidence is a particular attribute of the content of the Record.

Record Properties

- (Note: I missed part of this discussion as it overlapped with a break).

- Extent: Discussion of 'Content Extent' and 'Physical Extent', to differentiate between the intellectual and physical aspects of extent.

Record Component Entity

- There was discussion on the intended meaning of the term 'evidence' (i.e. 'trace').

- Regarding the modifications to the definition, GMcC noted that removing a Record Component from a Record compromises the Record Component's evidential capacity (evidence in a broad sense, or one with specific legal connotation) and/or removing the Record Component from a Record also compromises the Record's evidential capacity (same qualification).

- FC: Components could be separated but still integrally linked. This is true for digital records.

- Distinction between intellectual separation and physical separation while maintaining intellectual linkage.

- Compromised integrity: doesn't state to what extent integrity is affected. For some uses, may still be acceptable. Compromised is a matter of degree, not an absolute.

- DP: Factors that affect evidential quality/value may be intrinsic or extrinsic to the record.

- Compound Record: its parts are Records, versus Record Components, which are one step lower (pieces of records).

- Compound Record: one Record joins with others and together they comprise a Compound Record, but not a Record Set, which itself is not a Record.

- SV: A Compound Record's components (the Records) could also be separated and therefore the overall Record's integrity could be compromised. DP: But we are distinguishing between a Compound Record's components (Records) and Record Components.

- Practicality of Record Component: Use case - isolate particular components across Records to study them as a set.

- Compound Record and a Record with its components: although there is some overlap intellectually, they are not the same thing. Haven't yet finalized a way to present this difference in the model.

- DP: Original intent of Record Component was to provide a hook for the diplomatics perspective (e.g. seal, signatures of witnesses, etc.). Would 'Record Facet' be a better label?

- FC: The diplomatics community hasn't dealt with Compound Record. InterPARES Glossary: 'Record Element' - a constituent part of a record's documentary form; an element is a formal expression visible on the form of the record (e.g. signature).

- DP: Could state in the draft for public review that we're aware of some logical tensions in the model but are not presenting solutions because we're not completely sure how to resolve them.

- ORE Proxies: proxy contains information that relates a record to a specific context. Sequence: article1 appears at time/place X, before article1 appears at time/place Y. Record_ABC could be encompassed into another Record, from the point of view of the creator, as it is a component of that Record. DP: in some cases, just give coordinates or link to the other Record. Functionally equivalent to attaching the other Record as part of that Record. Record exists out there somewhere, with its own context of creation. Later, it gets re-used and becomes part of another Record, or in a certain sense, it is a component of that Record, but not in a diplomatic sense.

Record Set Entity

- "A set of one or more Records…" BP: Cannot group only one Record. Semantic problem? DP: Semantically correct, what about empty set. FC: If Record Set is made by a query and query results only returns 1.

- Discussion largely between DP and SV regarding whether users can group Records or not. Were users just grouping representations of Records or descriptions of Records but not the Records directly? DP noted that in the digital context, only representations of Records existed, that there are no static Records. He also noted a standard by which someone could identify Records across institutions and group and organize them. SV felt that in this case the person was creating a new archive made up of representation of other Records created elsewhere, while DP said that it could be seen as creating an intellectual assemblage of Records without acquiring them. A Record is only a conceptual thing in the digital domain. This was part of the larger discussion of whether or not end users could be conceived as forming Record Sets or not. There was division within the group on this point.

- Discussion of 'miscellaneous': "For convenience and by exception…" – issue of mixing in practical application concerns alongside principles. However, something should be said about miscellaneous sets as they are common practice and it should be acknowledged somehow. FC: Include a note to explain it, but leave it outside of the definition? DP: Prefer to leave it as part of the definition.

- DP: One can physically aggregate Records and create a Set, or intellectually aggregate Records and create a Set. We are not talking about physical aggregation here, rather the intellectual/conceptual groupings. Record Set is an intellectual construct we're super-imposing on the Records.

- SV: It's an intellectual grouping of the metadata about the Records or representation of the Records, not the Records themselves.

- DP: Sees the point, but wondering if making this clarification would be pedantic (overkill) here. Set is intellectual construct comprising representations of Records. Saying intellectual construct here is sufficient.

- BP: One difference for file is that it has a physical reality, while series and fonds don't.

- DP: The fact that they're physically put together isn't a requirement for them to be a file. Physical management reflect, benefits and supports intellectual management. A file in the digital environment is not physically brought together because doesn't need to be for administrative purposes. But it's still a file.

- BP: Remove any reference to physicality to make clear that we're only talking about intellectual. And this'll emphasize that we're dealing with intellectual representations/constructions, too.

- DP: Following from further suggestions emailed by BP, will add to the definition of Record Set something like "in the course of life or work events and activities" to provide a link to preserving the evidential values.

- Classification: FC: But any property could be used as the criterion for establishing a Set. KT: So it shouldn't be a separate property? DP: Leave it in because it's recognizable and familiar, a pedagogical bridge thing, even if classification is technically covered elsewhere (e.g. local identifier, or any other property shared in common by a set of Records that defines how they collectively belong as members of that Set).

- Classification establishes membership in a Set vs. arrangement, which is concerned with ordering members within a Set.

- SV: Difference between simplistic 'classification' (putting in buckets) vs. formal classification scheme.

- Issue of when a Record is related to more than one Record Set. Then Classification is not a property of the Record, but is representative of its relation to the Record Set.

- Is there a way to present a complex relationship, or a separate Classification entity to sit between Record and Record Set?

- FC/BS: Could Classification be a type of mandate? (Rule)

- DP: Re: having a separate Classification entity. Properties – name/term. Classification schemes exist out there in systems. Need controlled vocabulary management system to make it work. Defer adding separate Classification entity to RiC CM. Classification schemes are local to particular system.

- DP: Some classification schemes combine selection and arrangement, in which case would not need separate arrangement note.

- * Changes for DP to make here: Note that for Record, cannot arrange it, but can assign it a particular code that identifies its position in a given arrangement. JR suggested to move the second sentence of the definition of Classification ("The criterion or criteria may be represented in a RiC property or relation that is shared by the descendant record members") to the Classification property in Record as well. Could also add another statement that would permit combining classification and arrangement together as one property.

* To-do: 1) Record Set: Classification & Arrangement - will combine the two. Definition is okay already; and 2) Record: leave classification & arrangement there as well.

- Question posed: Can you discuss arrangement of a Record, or is arrangement only identifiable in relation to its membership in a Set?

Function Entity

- DP: Worked with VP when revising definitions for Function, Activity and Mandate (plus Occupation to some extent) and has tried to use common language (e.g. fulfil a function; perform an activity).

- Group discussion on "explain" vs. "determine" re: why an agent does something. Opinion that function is *what* an agent does, not why. Counter that Activity is the *what*, but then, what is Mandate?

- BS: Note that we have moved away from having a hierarchy of functions.

- JR: Also connect to Records ('as evidenced in Records') in addition to the connection already made here to Agent.

- Discussion of wordsmithing "elected" vs. "chosen" vs. tacitly authorized.

- Ended up removing authorization sentence – covered by Mandate.

Activity Entity

- The discussion focused on revisions to the definition.

Mandate Entity

- Mandate has subsumed the former Rule entity. It had become clear that one would have difficulty differentiating between a Rule or Mandate in a given case.

- DP: Discussion with Adrian C. re: not all mandates being explicitly stated.

- DP: emphasized not to confuse Mandate-the-document with Mandate-not-the-document (an abstraction).

- Discussion of Mandate not emanating from documentary sources such as customs, established practices, social mores….

- Mandate in this sociocultural context vs. in a legal context, in which highest concept may be 'Norm'.

Occupation Entity

- Note: only applies to one type of Agent (person).

- Person in context of Corporate Body occupies a position (role).

- Occupation – focus on competency (broad), position requires competency of the occupation (is more specific within a context).

- One person occupies a position at a time. Many people pursue a particular occupation.

- Occupation is broad enough to include non-professions such as hobbies or other things that are pursued that require competencies.

- Discussion of how Position fits in to this, as Position and Occupation are dissimilar. Some people holding positions hold the competencies of the associated occupation/profession but not always. DP: Agrees but idea of Position not yet accommodated and not here. Position is a relation of a person to a corporate body. Do we need another, separate entity, for Position?

- FC: In other ontologies, Position is distinguished as a separate entity. Add Position as a core entity, not down in the shared entity area.

- DP: already have person who can be related to corporate body. Would further segment by having a person related to a position that is situated within a corporate body?

- FC will produce proposal re: how to define Position and where to put it. Will include some use cases.

- DP: Position might have competency requirement or may not, or may not know. BF: not sure that it's necessary to include it as an entity.

<u>Presentation of draft to community</u>:

- Make clear that our focus was on describing 'the thing' and must keep this distinct from describing the information about the thing

- Could add repeatability/cardinality for properties throughout the document. (DP later clarified that in RiC's CM cardinality is consistently 1:1 or M:1 for properties and M:N for relations.)

Draft Outline:

Introduction

1. EGAD members
2. General statement of what this is, and what this isn't: Our charge, function, objective (community building) (objective not to homogenize descriptive practices of allied cultural heritage institutions, respecting the differences, but scope for collaboration – shared entities), scope (step 1: our focus was identifying entities of concern to archivists, their properties and relations to one another – that's it; what hasn't yet been addressed is the act of description itself – the control section)
3. Methods
4. Multi-level description – why unit of description disappeared; multi-entity, multi-dimensional description
5. Inheritance
6. Explain connection to existing 4 ICA standards – overview of how RiC changes in approach; higher level abstraction
7. Audience – for RM's and archivists.
8. Expectations
9. Bridge to other communities

Next steps for CM:

- Introduce Agent putting together the description
- Every property and relation is an assertion made by some Agent at X date and Y place based in Z evidence (source, certainty, etc.).

- Will need to include an explanation of the presentation of the CM – conventions and so forth.

- Examples – could be few for initial release. Afterwards gather and augment documentation – e.g. on a website and slowly integrate.

- Wait to translate until draft is stabilized.

- Will focus on introduction, cleaned up draft and a few diagrams. No time to prepare examples for first release by end of May.

- DP to draft introductory content; GMcC has created a hyperlinked database version of the CM data as part of his proof-reading efforts, and this could output a flat file version to be used as a consultation draft; KT and BS will also contribute to the editing and formatting efforts of the first consultation draft, due by end of May.

Editorial work to be done:

- Check for consistency of language issues (e.g. fulfilling of function, performance of activity).

- Ensure that definitions do not re-use the same term as in their label/name.