International Council on Archives

Experts Group on Archival Description

2017 Annual Meeting – Rome, Italy
23 to 25 October 2017

Host: Archivo Centrale Dello Stato

Summary Report
Prepared by Gavan McCarthy

Note: All photographs were taken by Gavan McCarthy

Contents:

1. Agenda for the meeting (with resources and links)

2. Day 1 (23 October 2017) overview of discussions and decisions

3. Day 2 (24 October 2017) overview of discussions and decisions

4. Day 3 (25 October 2017) overview of discussions and decisions

Section 1.   Agenda for the Meeting

The agenda for the meeting with the resources and links was compiled by Kat Timms and Daniel Pitti (Chair).

Draft Agenda for EGAD Rome meeting, Oct. 23-25, 2017

1. Introductions & updates (Oct. 23)
    a. EGAD Work Plan, timelines / deadlines
    b. RiC-O progress / status
    c. Status of digest strategy (publicly responding to community feedback)

2. RiC-CM entities (Oct 23/24)
    a. Address community comments per each RiC entity
    b. Consider suggestions for new entities
    c. Determine useful presentation style – e.g., potential method for organizing/grouping entities (primary and secondary); illustrative diagrams
    d. Treatment of digital records in RiC (entity definitions and examples should accommodate digital)

3. RiC-CM properties (Oct 24)
    a. Address community comments about RiC properties
    b. Determine useful presentation style / illustrative diagrams
    c. Treatment of digital records in RiC (consider adding additional properties; modifying existing properties' definitions and examples to better accommodate digital)

4. RiC-CM Relations (Oct 25)
    a. Address community comments about RiC relations
    b. Determine useful presentation style / illustrative diagrams

5. Other topics to be addressed (Oct 25)

6. Next steps / assignments (Oct 25)

Resources / Links:

Work plan
- EGAD ICA RiC II MOU+Schedule v6.DPrev.docx

Digest strategy (recent teleconference discussions highlighting some main issues raised in the feedback)
- Draft RiC-CM Digest_Discussion questions 20171018
- https://www.dropbox.com/preview/EGAD-All/WP%204%20-%20ConceptualModel/Consultation%20feedback/Digest%20communication%20strategy/Draft%20RiC-CM%20Digest_Discussion%20questions%2020171018.doc?role=personal

Organized community feedback
- **General feedback** (may be useful to review at the same time as reviewing any specific RiC component) EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / RiC-CM Feedback- Table 1 Broad Comments.doc

- **Feedback excerpt re: digital records issues** EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / 20171004 Feedback excerpt - digital records issues.doc

- **Entities** EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / RiC-CM Feedback- Table 3 RiC Entities.doc

- **Suggestions for new entities** EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / 20170918 Feedback excerpt - suggestions for new entities.doc

- **Properties** EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / RiC-CM Feedback- Table 4 RiC Properties.doc

- **Relations** EGAD-All / WP 4 - ConceptualModel / Consultation feedback / Organized feedback / RiC-CM Feedback- Table 5 RiC Relations.doc
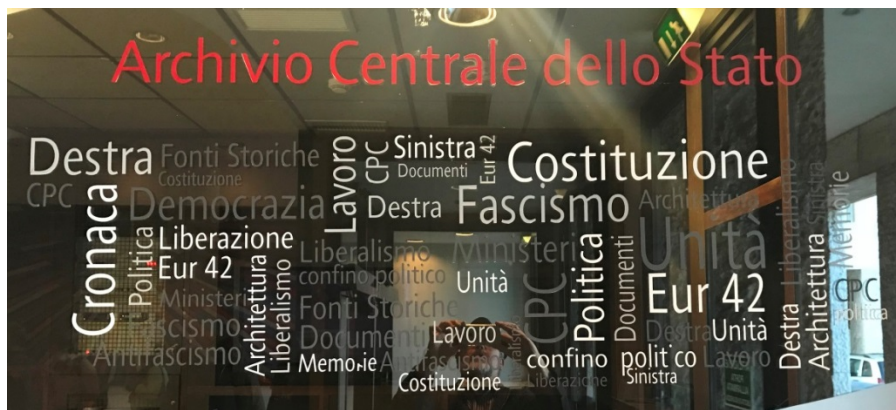
Section 2.   Day 1 (23 October 2017) overview of discussions and decisions

All photographs by Gavan McCarthy. The artworks were in the foyer of the archives building.

Agenda Item 1: (and the meeting flowed organically from that point on)

Daniel, as the Chair of EGAD, assembled the group, welcomed everyone to Rome and thanked our host.

Present at the meeting – going around the table from Daniel's right: Daniel Pitti (USA); Victoria (Scotland); Kat Timms (Canada); Miia (Finland); Gavan McCarthy (Australia); Beatrix (Spain); Javier (Spain); Helene (France); Florence (France); Martin (Austria); Bogdan Popovici (Romania); Vitor (Brazil); Padre (Africa); Stefano (Italy); Salvatore (Italy).

Apologies: Bill Stockting (England); Aaron (USA)

Initial introduction included acclaim for the work done by Kat in drafting the agenda but more so for her work on collating and analysing the feedback from the first draft released for comment. Daniel also covered the work on the 'Project Charter', originally conceived as a Memorandum of Understanding. Most recent feedback from the ICA indicated that the draft Charter has been agreed upon and is ready for signing. Daniel will do this. Daniel took the meeting through the timeline for the project as laid out in Schedule A of the Charter.

The key deliverables are:

RiC – Conceptual Model (CM);

RiC – Ontology (O)

RiC – Application Guidelines (AG)

The ICA is keen on training, recognising the issues raised in the feedback, so the communication of RiC and 'training' in the most general sense was seen as critical to the success of the project. The goal of the training will not be to teach them how to describe archival materials but to reveal to them the deeper concepts that underpin what they are doing. We need a range of 'training materials' and approaches depending on the audience (archival students, professionals, systems developers, National Archivists and senior policy people.

Daniel noted that there are already implementations based on the thinking behind RiC, for example, SNAC, the Spanish, the Finnish, the work of the eScholarship Research Centre, the French, others member of EGAD but also others not represented at the table but watching these developments very closely.

Florence noted that is quite advanced in having a RiC-O proof of concept ready for exposure.

Vitor noted that ICA PCOM has stressed the importance of training – in a sense all our presentations are a form of training, for example the workshop on 26 October 2017 in this building and then the panel session and workshop at the ICA conference in Mexico at the end of November 2017.

Workplan: Daniel anticipates four meetings in the run up to the end of this Project Charter which is the third quarter of 2020. Meetings organised or planned are: Rome 2017 (this one); Berlin; London; Romania (Brasov?) and perhaps even a fifth meeting at the University of Virginia if funding could be found.

There was discussion about extra days added to the planned meetings to deal with RiC-O. there are arguments for and against – against being the b the end of our general (CM) meeting we are all completely exhausted.

Florence presented an up-date on RiC-O status and progress and referred us to various documents and web addresses. She has articulated to the principles that drive the design and building of RiC-O and she articulated these in her presentation at the workshop on 26 October 2017. As a rule this is looking positive and certainly the principles make sense. She is working on a typology of Relations, an area in the CM and more generally that was underdone and elicited much feedback. She noted that she would be looking for volunteers to test the proof of concept more thoroughly once it was available. Others, including a Professor in Spain and the Swiss are developing their own ontologies demonstrating that the time is right for this type of development and that there is a community to help in the assessment, validation and continuing development.

Just prior to morning coffee there was reflection on the 'digest strategy' as a means of communicating our work with the broader community. It was noted that there were pros and cons to 'openness' but it was agreed that the members of EGAD could not engage in continuous or dynamic discussion with individuals in the community and that we had to find a means by which worked within the constraints of the EGAD members.

After morning coffee it was noted that e need to do work on the EGAD documents in Dropbox as the current file namng and folder structure is proving unworkable. (I thought is was just me!). "We have a room full of archivists but we are not very good at records management!"

There was further discussion on the Digest Strategy. The community already has access to all the feedback but it is so dense and voluminous that very few have engaged with it, therefore, the role of the digest. Beatrix suggested an FAQ approach. There was discussion about the nature of the feedback, its role in reflecting the variability and diversity in the archival community, and the types of personalities that influence thinking and attitudes.

If the goal is that we have a second draft of RiC-CM out to the community in early 2018 and have a draft that is stable enough to support RiC-O development then there is a lot of work to be done. There is both some

conceptual and reframing work following the feedback as well as re-drafting and extending the introductory and explanatory narrative that leads the broad range of readers to the CM itself.

The workshop and panel session in Mexico (Daniel, Bogdan, Vitor and Gavan) will be our first opportunity to return to the community with evidence of the work that has been over the last 12 months and an indication of the things to come. The workshop in Rome 26 October 2017 was not a bad dress rehearsal for this.

Daniel noted that our earlier presentation strategy was not entirely successful and there was then a discussion on returning to the five fundamental questions that underpin the quest for meaning and interpretability, namely:

Who? What? When? Where? Why?

After some discussion it was agreed that we have three high level groupings of Entities that would allow a more gentle and measured introduction into the world of entity-relation models. Namely:

Archival Resource(s); Agent(s); Function(s) – Space-Time (Date-Place) – and perhaps another group that encapsulated the notion of extensibility (Rest of the World).

It was agreed that 'Records Component' would be relabelled 'Record Part'.

There was also discussion about the term 'Property' in the CM which was at odds with the technical use of the term in OWL ontologies. 'Characteristic' was suggested but later the decision was to run with 'Attribute' but on the understanding that some 'properties' us currently listed need to be promoted to the level of 'Entity'. This was returned to on subsequent days.

After lunch Daniel reported that the Charter was finalised and we returned to the discussion of a communications plan. Florence noted that we need a section on 'design principles' – this was agreed. We again reflected on the three products (CM,O and AG) and that faults and oversights had been revealed in the original 'Introduction', for example missing substantive parts of archival history both in time across the globe.

Daniel agreed to write up something that defines:

> What is an entity?
> What is a 'property' (or attribute)?
> What is a relation? etc and to do this as soon as possible after this meeting.

There was a substantial discussion about all those important archival constructs (entities) that were neither Archival Resources nor Agents (in the form of Persons, Groups etc). There was eventual consensus that all those activities that included activity, process, transaction, function, mandate, would all be grouped together under the broad heading of Function(s) [GJM: my inclination is to pluralise these higher level terms so it is clear that we are talking about these collective groups.]

There was also discussion about the entity 'Function (Abstract)' – a label not well received – but there was agreement that we need an entity that served the function of describing a concept that applied universally rather than being specific to a particular archival jurisdiction. Indeed it is these types of entities along with other more general classification-type entities that will enable the bridging and interconnectivity between realms of descriptions (the things that we share as opposed to the things that keep us apart). Without these types of entities there is probably little point investing in RiC-O and the Linked Open Data approach.

Martin, as he often did during the meeting, distilled our discussions into what would they mean in the archival classroom and training sessions. He noted: 'The description of a Function (concept) is very thin which means it is a place of inter-jurisdictional connectivity'. There was general agreement but the meeting continued to unpack this issue in some detail as this was an area of significant confusion in the feedback so we needed to worked it through in more depth.

There was reference to the Hurley: **Doers, Deeds and Documents** which we have given the more formal names of **Agents, Functions and Archival Resources**. Other concepts tossed about included: Ends-Means; Objective-Task; Function-Activities.

It was felt that much of the confusion was due to the inadequate and poorly developed examples in the first public-release draft of RiC-CM. In terms of Deeds (Functions) there was some discussion of the concept of 'Transaction' but this was dealt with in any detail. My guess is that in terms of time available in the meeting we were not going tob e able to tackle this and we all knew. However, it will need to be tackled at some point. I noted that AGRIF (Australian Commonwealth Government recently released recordkeeping ontology was silent – or not apparent – on this issue although it seemed to be at the essence of what needed to be tackled). It is the 'point' of transaction defines the requirement for evidence (and thus records) for much of what we do either personally or in our jobs.

I tried to capture the connection of the ideas that were being discussed in the following sentence: 'In undertaking an **activity** in pursuit of a **function**, an **agent** does things resulting in **transactions** for which there must be **evidence** in the form of **records**.' The discussion continued with some intensity despite the lateness of the day.

Following some housekeeping announcements 9denner etc), Daniel noted that we had two days left but that we were still dealing with issues at a fairly high level, reflecting that we often made our best progress in the private conversations, in the breaks at meals, walking to and from the venue. He noted that we needed to progress to the Digest at which point he would hand the lead to Kat.

Key outcomes for today seemed to be:

High Level Groupings of Entities – mostly as an explanatory and human comprehension device (I noted that I had intuitively created these cluster points in RiC-IM (Informatic Model);

Primary and Secondary classification of RiC entities that brought to the fore the critical archival entities (in the groups Agents, Archival Resources, Functions) but allowed for a broader range of supporting entities that were necessary but not peculiar to the archival mission;

Redrafting, Re-structuring and Extension of the Introduction.

Still trying to put flesh on the RiC Beast

Section 3.   Day 2 (24 October 2017) overview of discussions and decisions

A beautiful clear day in Rome – perfect for clear thinking

In recapping the decisions from yesterday we agreed to have at least three high level groupings of 'archivally' critical entities:

Archival Resource(s): for Record, Record Set, Record Part to which others could be added for those that may need to extend the model for local purposes;

Agent(s): for Person, Group, (Corporate Body, Family, etc) and again to which others could be added for those that may need to extend the model for local purposes; [*It remained unclear to me whether we had promoted Agent as an Entity, as in the first draft to a clustering role, with Person and Group as the primary entities in this group – what we decide here makes a difference to me (GJM) in re-building the RiC Informatic Model*]

Function(s): (end / means) for purpose, goal, objective, responsibility: Function (local), Function (concept), Activities (process), transaction.

Space-Time (Date-Place): was parked for the moment.

A discussion then got underway about the relationship between the concepts and understanding of **Data, Data Sets and Databases** and their relationship to records, evidence and the notion of recordness. [*It was fairly confused but as the discussions continued over lunch, dinner and walking to the meeting for Day 3 it became clear that there seemed to be more in common with how we saw these issues than would have been apparent from the conversation. I am not sure why it often takes us such a long time and with seeming disagreements before we finally realise that we actually have a common understanding.*]

We then reviewed the remainder of the top level entities as described in RiC-CM discussion draft. The world at this level seemed to divide into those entities that were inclined towards a controlled, defined or emergent classification scheme (e.g. Documentary Form) and those that would be driven by more unpredictable diachronic and synchronic locally unpredictable variables and would therefore have to reflective rather classificatory (e.g. Mandate, some instances of Concept). It was recognised that this distinction, though important in some ways, is not clear cut.

The connection with SKOS controlled vocabularies was identified.

Discussion was wide ranging and philosophical and hinted at things such prescriptive and proscriptive rules but there seemed to be a conclusion that:

RiC, as a network in and of itself, is (needs to be) extensible within certain prescriptive and proscriptive rules.

Discussion returned to Archival Resources and the record set of entities. It was proposed and agreed that **'Record Component'** be re-labelled **'Record Part'** due to the linguistic and semantic issues raised in the feedback about this label. In essence we seemed to agree that what is and what is not a 'record' is contextually determined – that is it is determined by its contingent and environmental relations. So for example the same document may stand as a Record in one context but may be a Record Part in another context.

Morning Coffee was held at about 11am.

In resuming the discussion we seemed to be dancing around the idea: "The whole is greater than the sum of its parts" which is the catch phrase for mereologists. For archivists we might transform this to be: "The record is greater than the sum of its parts."

Discussion also covered the idea that things are 'social constructs' even those things that seem awfully tangible like materials records and people. So we ended up with this notion of:

$$\text{Social Construct} \quad / \quad \text{Physical Artefact}$$
$$|_____|$$
$$\text{Bridging Properties}$$
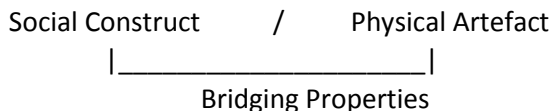
So we spent some time trying to unpack a whole lot of issues with regard to the idea of 'Bridging Properties'. We agreed to re-visit Properties and the work done by the Spanish. With the decision that we ditch Properties as a label for these things, some of which need to treated as entities (and some that do not in the archival description context). The proposal was that we adopt the term 'Attributes' (*but I think I would prefer something more explicit such 'Attribute Entities' which I may have to use in the RiC-IM*).

We seemed to hit on a principle: Humans do archival description – they make assertions based on evidence (that is the documents they see in front of them and have access to). The feedback would indicate that there is still a concern in some parts of the broader community that the archivist is expected to make 'truth' statements rather than being refractive of the world as they find it (see Karen Barad).

It was decided that a Sub-Group (of no more than 4) would be formed to unpack this decision in more detail – I am not sure who ended up volunteering for this task.

And thankfully it was lunchtime (c.1.30pm) – which went until 2.15 – with the goal of the afternoon session to go until 5 so that we had time to get back to the hotel and rest before heading out to the hosted EGAD dinner.

Following lunch we defined three working groups with specific tasks:

Group 1: Physical – Intellectual (Control) [social construct]

They will cover things like carrier and encoding format on the one hand and the processes that determine the unit of record and therefore description on the other.

Group 2: The major RiC entity groupings

So they will refine which are the primary and secondary entities in each of Archival Resources, Agents and Functions – and how each of these could be extended in local implementations.

Aside: Daniel is keen to get rid of the term 'type' – it is a cheap way of extending the semantics. Type should be reserved for a variety of entity that share all attributes [I *can see how this makes sense for the ESRC with the OHRM and EAC implementations as they all share the same underlying informatics. However, I am struggling to see how I 'type' and group things in the RiC-IM without resorting to type, or is that 'reverting to type'.*]

Group 3: Relations

The main challenge here is coming up with one or more sensible and explainable groupings of relations so it makes it easier to explain (in a training sense) but also have them in sets that could be used in context sensitive drop-down selection tools within implementations. I already have one grouping schema established in RiC-IM and Florence has had a first go at a grouping in RiC-O.

Membership of the group is: Florence, Gavan, Javier and Victoria (with Daniel as ex-officio)

At this point the membership of the other groups was not finalised although Daniel was taking an ex-officio membership role in each.

The rest of the afternoon was handed over to Kat to take us through those sections of the digest that had not been covered in the teleconferences. We found that much of the content had been covered in the meeting so far – which was a relief. I did not take any extra notes in this session as it as being well captured by others. Also, I was loading RiC-IM data to the UoM ESRC web server so we could use the graph visualisation tool to help us understand the existing structures and get ideas for what new structures might look like. As the internet bandwidth was week that data had to be loaded in small batches which meant it was quite time consuming.

Section 4.   Day 3 (25 October 2017) overview of discussions and decisions

A feast for the eye, a feast for the mind and a feast for the stomach

The intention for the day was to start with Section 3 of Kat's Digest and work through to the end of the document. It had been planned that we would have a 45 minute tour of the archive after lunch but due to the amount have work we had to get through this was cancelled.

But first there was a return to the discussion about classification which still had issues to be played out. The general challenge is that network models are inherently complex and only suit a small percentage of the audience. So we need a staged or phased introduction – a narrative or story that will enable us to take most archivists with us and hopefully reducing the number of calls for 'more entities'. By looking at the higher level clusters – that may help the bulk of the audience feel comfortable with RiC-CM. RiC-O is a separate challenge and will probably be best presented by demonstration rather than as a technical, ontological system which tends to reinforce the notion that it is magic black box that provides answers.

Both Stefano and Bogdan outlined the importance of this topic in their countries and Bogdan provided an excellent real-life example of inherited control codes and classifications from earlier periods in the life of stuff. These are often partial and therefore emergent as you work the description of the materials. It is rarely the case that you will find a document or documents that define these schemas [*you will have to ask him for more details!*]. Daniel and Florence brought this discussion to a point of consensus.

On the flip Chart [See 4-2, page10 Flip Chart Report] made a nice outline of the informatics of 'classification' as an entity . As I noted, this is elegant and will work in just about every case we can think of.

Javier made the observation that this type of description can also be done as a relation [*but of course this is only possible if the relation has the 'properties / attributes' of an entity – which interestingly creates a situation where everything is just an entity – or just a relation – it depends on perspective and purpose.*]. It was not clear that this issue was well understood by all in the group – 'we let it go through to the keeper'.

However, there was general consensus that all 'Attributes' will be treated as entities which includes 'name' and thus all realm of alternate names.

It was recognised that this was a significant rework and there was concern that we might not be able to get this done by January 2018. However, Daniel felt we were in better shape than others might think and was confident this could be done.

It occurred to me that this new approach better suits the work already done in RiC-IM which means that it might be a useful tool, with its ConneX graph visualisation to help communicate these ideas and concepts.

Morning Coffee at last.

From here to lunch we worked through Digest items 4 through to 13. See Kat's notes for a summary of the discussion and decisions.

Item 13 concerned 'Concept/Thing' which generated more discussion and had certainly been found challenging by a number of the feedback respondents.

There was also discussion on Date types.

Lunch – we resumed at 2.45pm and went through to about 5pm.

We planned the next phase of work and covered the rest of Item 13 and then Item 14 in the Digest. [I have more comments and thoughts in my notebook but they are more concerned with issues that I will need to consider with regards RiC-IM. One task for me is to capture a record of the graph visualisations as the stand now – with the new data that I edited over the last few days but before we start on the bigger changes based on the working groups, and in particular the work we will do on the relations. At the relation groups are paired by their reverse group (i.e Agents-Functions Relations is paired with Functions-Agents Relations but it might be more revealing to create a 'clique' of all relation groups that share the same source (or domain).

The meeting finished around 5pm as some members had travel obligations and we were all exhausted.

The following day Daniel Pitti gave an overview of our work to the well-attended workshop for the Italian archival community. Below are a few slides from his presentations that provided an overview of the Rome EGAD meeting and the plans for the immediate future.

A Few Highlights from Rome EGAD Meeting #1

• Good portion of discussion focused on

• Refining our understanding of the role and nature of the three RiC products:

Conceptual Model, Ontology, and Application Guidelines

• Reformulating the plan of presentation of the CM

• Discussion of methodology for deciding what is and is not an entity

(”significant object”), characteristics of the of the entity (attributes), inherent

traits and states of entities, relations and events …

• Reevaluation of entities: primary, secondary …


A Few Highlights from Rome EGAD Meeting #2

• Provisional evaluation of entities

  • Archival resource

    • Record

      • Record Part

    • Record Set

  • Agent

    • Person

    • Group

      • Family

      • Corporate Body

    • Position

    • Delegate Agent

  • Function

    • Goal

    • Activity

    • Transaction

A Few Highlights from Rome EGAD Meeting #3

• Supporting entities optimally maintained as controlled terms or values

> • Occupation

> • Function (Concept) includes Goal, Activity, Transaction

> • Record Classification

> • Documentary Form

• Place and Date

• Name and Identifiers

• Relations

> • Types of relations treated as a graph (broader/narrower; primarily hierarchical but …)

> • Reduce to the a basic and essential set

> • Direction of relations and inverse relations

> • Extensible

• Much, much more …

Work Timeline

• RiC-CM: Second draft for public comments early in 2018

• RiC-0: Incomplete beta draft early in 2018

• RiC-AG: First draft for public comments in early 2019

• Timeline thereafter variable depending on comments, work schedules and so on

• Objective will be for all three products to be complete and relatively stable late in 2020
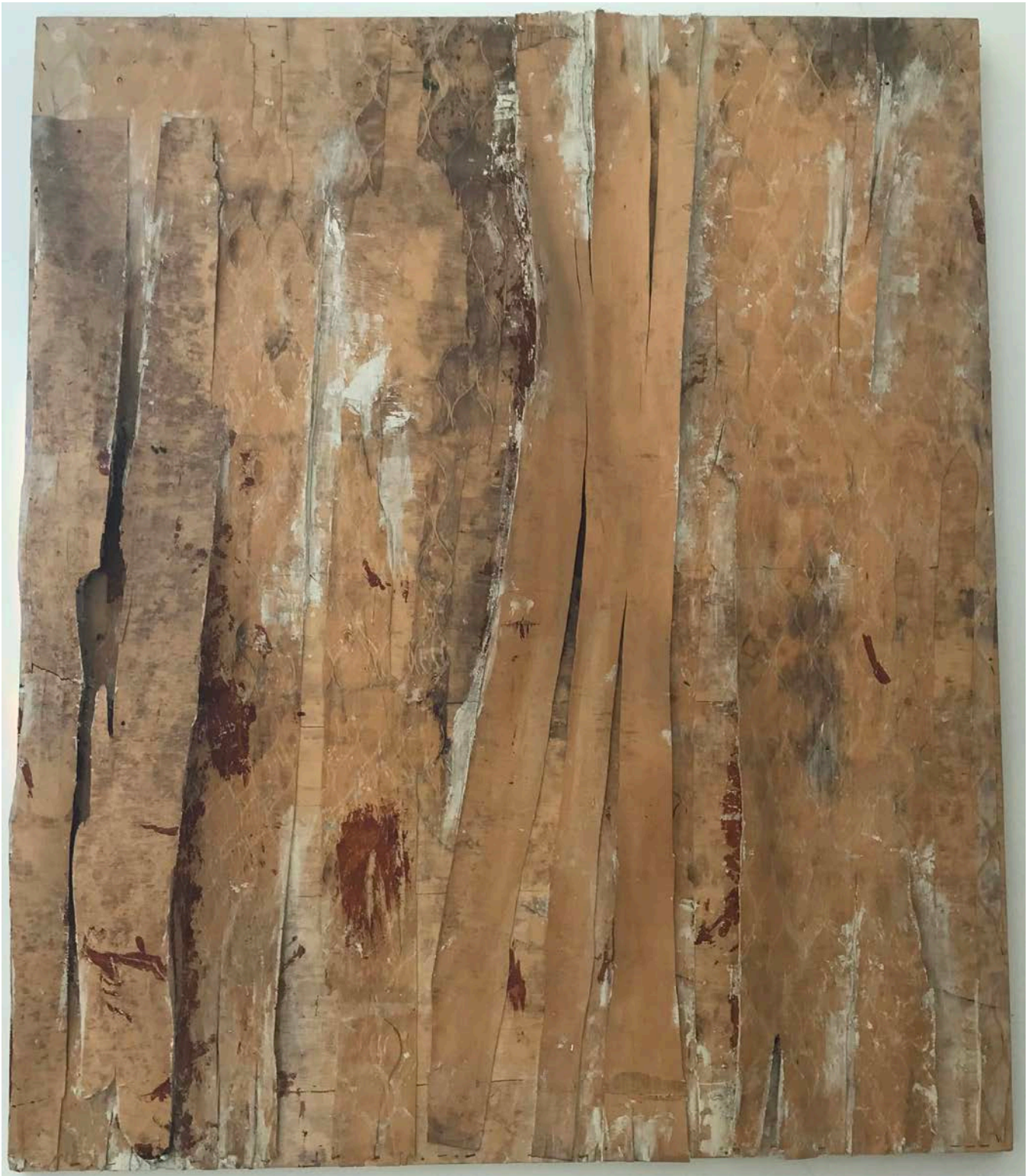
We need to start to see the wood from the tress