Telecon Points of Discussion:

(Original notes from DMS;)

1. Representation Information can end with a Data Object that needs no further Representation Information (e.g., a hardcopy document) . Such a Data Object is therefore not strictly an Information Object as defined at the conceptual level. Representation Information and PDI are all Information but not always Information Objects.

(I believe Don and John agree; maybe Steve and Mark?)

JGG: I agree that a data object (data, Representation Information, PDI object) that does not have Representation Information associated with it is not an Information Object.

2. In response to a question from Mark, Don alleged that it was sufficient for an Archive to describe its storage in terms of the AIP concept for it to be considered OAIS RM conformant.

JGG: I agree.

Also OAIS-RM conformant is not ISO 16363 conformant.

At the same time it did not need to implement all aspects of the AIP concept for every AIP.

JGG: I think I disagree with this. What is criteria for what you need to have and what you don’t need to have? Is a raw data object with just a Reference and Fixity Information an AIP? Or a Physical Object (a hard-copy document) with just a Reference an AIP? Why or why not?

DMS: Taking the OAIS RM as primarily a communication framework, there is no specification of ‘what you do or do not need to have’. If you say you have an AIP with just a data object and Reference, then I may ask why you don’t have Fixity or don’t need Rep. Info. You may of may not have a satisfying answer but at least we’re likely to be communicating about preservation in the same general domain. Further discussion may be enlightening.

This was not how Mark and John saw it. Don argued that this view is needed because some AIP implementations may not need the full AIP concept.

For example, if the Content Information was taken to be a hardcopy document written in English, then it does not need additional information to “map the Data Object to more meaningful concepts” (i.e., Representation Information, 2012 OAIS RM).

JGG: The example has a bit of a problem. Is there any Representation Information with the hardcopy document? Example says it is the Content Information (which, in my understanding, it can only be if it has Representation Information). I think the point is that if the Content Information is understood then additional (Representation) information is not needed to complete the Representation Information Network, i.e. that hardcopy document is a lea of the Representation Information Network.

DMS: The example is a possible real case so if there is a problem it is either with the modeling or the interpretation of the modeling. If you think the current modeling specifies that an implementation of Content Information must also have an implementation of Representation Information, then the real example falls outside the current modeling and thus the modeling is deficient. If you take the view that the modeling is only for communication purposes, then it allows someone to say that their hardcopy document without Representation information is their Content information as provided by some Producer and no additional Representation Information need be present because it is already understandable or presentable to some Designated Community.

Section 2 says the Content Information is the original target of preservation. It consists of the Content Data Object (physical or digital) and its associated Representation Information needed to make the Content Data Object understandable to the Designated Community.

JGG: The new proposed text is “The Content Data Object Object (Physical Object or Digital Object, i.e., bits) is that information which is the original target of preservation. It has associated Representation Information which is needed to make the Content Data Object understandable to the Designated Community.”

DMS: However as you are interpreting it, it doesn’t allow a CDO to not have Representation Information even if that is what the external Producer has specified. It also downgrades Representation Information, when needed, as not being part of the ‘original target of preservation’ despite the fact that the CDO in this case is useless without the Representation Information. I find this to be a step backward in promoting the critical importance of Representation information for encoded data objects.

If the Content Data Object is already directly understandable to the Designated Community (e.g., physical document in English), Representation Information is not needed. One can argue that technically it is not Content Information because it does not have Representation Information, but it is in the sense that it is from the Producer and it is “Information that is the original target of preservation”.

JGG: I think it is simpler and clearer to think of the CDO as the “original target of preservation”.

DMS: Again, I prefer the emphasis of the OAIS RM as a communication vehicle, not a mandatory for any aspect of implementation. Then these issues do not arise.

The broadly based OAIS RM concepts can not be expected to model the details of all appropriate implementation cases.

JGG: I think I disagree. I thought the OAIS conceptual model was intended to provide a model that all implementations could be mapped to.

DMS: The intent was to provide a communication framework for people to discuss their implementations. Clearly it can not model implementations in detail. Unless it is perfect, even the general level of modeling it provides may not fit all reasonable cases. The hardcopy document above is a case in point.

A balance has been struck between providing widely applicable, high level, concepts and an attempt to conceive of all possible cases with much finer granularity. That balance can be altered in the 5-year review. One proposed alteration is the suggestion to explicitly show that PDI can be applied to the preservation of other PDI and this would add more detail to the conceptual modeling capability.

JGG: I also disagree with this proposal. I think it complicates the model without adding much additional value and understanding.

DMS: It has been our colleague David who is pressing to ensure PDI is understood to be recursive. I consider it a reasonable extension of the current concepts.

This is also related to my interpretation that all the current AIP components CDO, Representation Information, and PDI are required for every AIP. If everyone else agrees that this is not the correct understanding, then I would be more likely to change my mind. But if everything is not required, then I think we have massively diluted what an AIP is and what an OAIS archive is.

DMS: That is how the OAIS RM has defined it. Actual implementations may not implement all of a given concept, but they still want to talk about their implementations using the common framework, such as saying ‘our version of the AIP is as follows…’. Again, a communication vehicle not a rigid implementation criteria. ISO 16363 is a different animal.

It can be inferred from the current models, but this proposal would make it explicit. At the same time, one shouldn’t expect all AIPs to need all the PDI components, nor to implement PDI on PDI in all cases. The conceptual model says all the components are expected, but real implementations may find that this is too broad a set of concepts to be implemented fully in all or maybe even most cases. Auditors may or may not agree depending on local circumstances.

3. Mark asked the question: “what is the harm in leaving Figure 2-3 as it is, with PDI applicable to the Content Information” as shown and discussed in the text? Don saw none. Don did not hear a response on this from John or Steve.

JGG: My interpretation of Figure 2-3 is that Content Information and PDI are in the same package. I do not interpret Figure 2-3 as saying anything about the relationship of Content Information and PDI (other than them being part of the same package.

Text that accompanies the Figure may make statements about the relationship of Content Information and PDI, but I don’t think the Figure shows it.

DMS: Yes for the figure, but the text clearly says “The PDI applies to the Content Information and is needed to preserve the Content Information…” This is the only model (apart from Information Object) that an Archive is required to ‘support’, from an information modeling perspective, to be an ‘OAIS’.

I believe there are other Figures that show PDI applies to Content Information before we agreed to change PDI to the CDO rather than the Content Information.

I do not believe making that change means that you can’t also have PDI information about the Representation Information.

DMS: There are figures in Section 4.2 but they are not required to be ‘supported’. Figure 4-13 shows the Information Package where Content Information is further described by PDI, and Figure 4-15 shows the AIP where Content Information is further describe by PDI. If you are proposing no change to Figure 2-3, but some change to 4-13 or 4-15, note that these latter two are not required to be supported to be an OAIS. Again, I see the proposed ‘PDI applied to CDO’ as being a step back in terms of the importance of Representation information and providing a broadly applicable and consistent communication framework.

4. Steve provided a complex PDI example consisting of a number of digital data objects, each with associated Representation Information, and specific relations among these Information Objects. All this was needed to be scientifically useful, and not just to display the image. Don saw this as making up a proper AIP, although a very complex one. We did not specifically address PDI for this case.

JGG: I agree that the multiple data objects with the Representation Information and specific relations somehow covered and proper PDI for all of data objects would be a single proper AIP if so desired.

5. Other?