**Review of OAIS 2012 with addition of Preservation Objectives (DG 20170728)**

DMS 7 August 2017

I’ve organized this response by first stating my conclusions, followed by an analysis of OAIS 2012 without Preservation Objectives for context, then by my analysis of what I see as the result of this addition of Preservation Objectives, followed by a view of preservation actions by the Archive that is more detailed than we have given in OAIS to date. Whatever words we end up with, I’d like them to be at least not inconsistent with this more detailed view. I look forward to your responses.

**Summary Conclusions**

Since Preservation Objectives are not mandatory and ‘use’ can take the place of ‘understanding’, the result, as I see it, is that:

1. Responsibility #4 (Independently Understandable to a Designated Community) has been virtually eliminated as a significant criteria to be an OAIS Archive,

2. Independently Understandable need no longer have anything to do with the common understanding of ‘understand’,

3. It has allowed other categories of information to be preserved which previously were excluded by what would be a normal reading of OAIS. This is a positive result.

My analysis leading to the above is as follows:

**Related analysis of OAIS 2012 (without Preservation Objectives)**

*Relevant definitions are as follows:*

Interpret (dictionary): explain the meaning of (information, words, actions); or understand (something) as having a particular meaning or significance.

Understand (dictionary): perceive the intended meaning of (words, a language, a speaker); or infer something from information received.

Use (dictionary): take, hold, or deploy (something) as a means of accomplishing a purpose or achieving a result; or describe an action or state of affairs that was done repeatedly or that existed for a period of time.

Independently Understandable (OAIS 2012): a characteristic of information that is sufficiently complete to allow it to be interpreted, understood, and used by the Designated Community without having to resort to special resources not widely available, including named individuals.

(Note: ‘Interpreted’ and ‘understand” are very much the same thing as given by common dictionary usage, while ‘used’ implies there is access to the information so that it can be ‘used’. I would now argue that ‘used’ in the above definition should not be there as it is not related to the common understanding of ‘Understandable’.)

Designated Community (OAIS 2012): An identified group of potential Consumers who should be able to understand a particular set of information. (…may be composed of multiple user communities)

Consumers (OAIS 2012): the role played by persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail.

*From section 3 (OAIS Responsibilities):*

OAIS Responsibility #4: Ensure that the information to be preserved is Independently Understandable to the Designated Community.

*OAIS Responsibility #4 examples (section 3.2.4) analysis:*

The focus of the examples in 3.2.4 is on Content Information together with PDI, and it implies that persons should be able to understand this set of information. There is no suggestion that client systems should be able to understand the set of information, although the OAIS definitions allow this interpretation. It also suggests that over time the Knowledge Base of the Designated Community may evolve and that important aspects of the preserved information may no longer be readily understandable, and therefore the Archive may need to take steps to enhance the Representation Information to recover the understandability. Responsibility #4 is not limited to understandability only at the time of receipt by the Archive. This aspect of #4 has always been controversial to some.

*Implications from all the above:*

1. For example, the Archive is to ensure that research results submitted for preservation are Independently Understandable to a Designated Community and should be maintained as such.

2. Some categories of information, such as a typical recording of a musical performance, appear to be incompatible with Responsibility #4 because it does not involve understanding by a Designated Community. Another category of information that is incompatible with Responsibility #4 is information that is intended to be understood as submitted to the Archive but no updates to maintain that understanding are allowed, such as copyrighted material for which the author has not given permission for any changes or information that is to be preserved ‘as is’ for its historical form and value.

**Analysis of OAIS 2012 with Preservation Objectives (DG 20170728)**

The new Independently Understandable definition is “A characteristic of information that is sufficiently complete to allow it to be understood and/or used by the Designated Community as exemplified by the associated Preservation Objectives without having to resort to special resources not widely available, including named individuals.” Since there is now the option to replace ‘understood’ by ‘used’ I see the following:

- Used can be satisfied by virtually any criteria and is thus a very low bar, almost nonexistent, for the Archive to meet in Responsibility #4. Being such a low bar it no longer excludes those categories of information discussed above. This inclusion is a positive result but at a high cost, in my opinion.

- ‘Used’ is very different from ‘understood’ and not something one would expect from adding the qualifier ‘Independently’ to “Understandable”. As a result of the ‘and/or’, the definition can now be read as “A characteristic of information that is sufficiently complete to allow it to be used by the Designated Community as exemplified by the associated Preservation Objective without having to resort to special resources not widely available, including named individuals”. Since Preservation Objectives are not mandatory, ‘Independently Understandable’ can be satisfied simply by making the information available so it can be ‘used’ and thus there is virtually no requirement that Responsibility #4 puts on the Archive. The result, as I see it, is that:

1. Responsibility #4 has been virtually eliminated as a significant criteria to be an OAIS Archive,

2. Independently Understandable no longer need have anything to do with the common understanding of ‘understand’,

3. It has allowed other categories of information to be preserved which previously were excluded by what would be a normal reading of OAIS.

The definition of Designated Community, which changes ‘understand’ to ‘understand/use’ implies to me that both are acceptable, and the phrase ‘as exemplified by the Preservation Objectives’ seems to imply that such objectives will be present when there is no mandatory requirement for this. As a result, I see the Designated Community as no longer putting any real constraint on what an Archive needs to do to be considered an OAIS Archive.

I’m not comfortable with this approach given the negative implications of 1 and 2 above. However I’m strongly in support of not excluding certain categories of information from preservation, under the OAIS definitions, as noted in 3 above.

One approach might be to make Preservation Objectives mandatory. However this still leaves the incongruity between what ‘Independently Understandable’ would seem to mean and the proposal that it can simply be a matter of ‘used’. My feeling is that many will interpret this particular incongruity as somewhat unnatural and a bit (or considerably) contrived. I’m not comfortable with this.

I was thinking that possibly adding a separate definition somewhat like ‘Independently Useable’ along with making Preservation Objectives mandatory could be the basis to accomplish the desired results more naturally. However I think ‘use’ is anything the Consumer does with the information, which in OAIS context is about using the DIP. This does not put the focus on preservation by the Archive, where it belongs.

I think that whatever words we end up using, I would like to have them compatible (not inconsistent) with what I see as a more detailed view of the preservation process by the Archive. I see this as follows:

For information to be preserved by the Archive it must first ensure that the information it has received is meaningful. It must do the following:

1. For a given digital object, it must understand the associated Rep. Info. If it does not, then it is not in a position to claim that the Rep. Info. provides meaningful information nor that the digital object plus the Rep. Info. provides meaningful information. Alternatively, it may enlist the knowledge of experts to validate that the Rep. Info. is meaningful or it may recognize that the Rep. Info. is part or all of a recognized standard for which experts exist.
2. For a given digital object, it must apply the Rep. Info. to the digital object to ensure that the result provides meaningful information. For very large digital objects, spot checking and validation by the Producer should be sufficient to gain a reasonable level of confidence. For the case that it does not understand the Rep. Info. but has employed an expert to determine that it is understandable, it must have the expert also verify that the application of the Rep. Info. to the digital object provides meaningful information that through spot checking, or other approaches, can be validated by the Producer. Alternatively, if the Archive has relied on the fact that the Rep. Info. is a recognized standard, then there must be associated recognized software, that incorporates knowledge of Rep. Info, that the Archive can apply to the digital object to generate the resulting information that is in a form that can be sufficiently validated by the Producer. This must be possible regardless of the nature of the information, including research results, music files, copyrighted documents, historic artifacts, etc.

Once the Archive has assured itself that the information received is meaningful and has been validated by the Producer, then it must ensure that this information is preserved over time. Depending on the preservation objectives, this may range from ensuring that the digital object and Rep. Info. are unchanged over time to needing to be proactive in ensuring that experts from a Designated Community can continue to understand the information and, when questionable, make updates to the Rep. Info. and possibly the digital object as needed to meet the preservation objectives. There are many variations in Archive scenarios between these two extremes. Perhaps generating a range of acceptable Preservation Objectives would be a useful next step.

Cheers-

Don