Defining the Designated Community

Christian Keitel and Jenny Mitcham

DPC Technology Watch Guidance Note

July 2023

dpc DigitalPreservationCoalition

1   Introduction

There are numerous decisions to be made by those working in digital preservation: What should be preserved? Which formats will enable future reuse? What documentation and metadata needs to be kept alongside the content to enable its understanding? Making appropriate decisions on how to preserve digital content can be a challenge and can only be successfully carried out with an understanding of who the content is being preserved for. Archives serve both their current and future users. By identifying the users of archives more specifically, we can begin to understand their needs and ensure our preservation activities meet them.

This Guidance Note considers the concept of the 'designated community' as defined in the Reference Model for an Open Archival Information System (OAIS) (CCSDS, 2012). It examines how this concept has been implemented in practice, its effectiveness as a tool for digital preservation practitioners and looks at alternatives to this approach.

2   Why is the designated community important?

Preserving and providing access to digital content over long time periods is a challenging task. Preservation is not a means to an end. It is performed to enable access, whether that is immediately, or at some point in the future. It is therefore logical that an understanding of who we are preserving the content for, and what their requirements are, will inform preservation activities (and sometimes even decisions at the point of content creation). Questions that might be informed by this knowledge include:

- Is the content submitted to the archive worthy of preservation? Is it in an appropriate format? Does the quality meet the needs of those who might use it?
- Can the content be successfully re-used? Does it make sense on its own? Is additional documentation needed to allow users to make sense of it?
- Is the content in a suitable format? Would users be able to work with the content in the file format it is supplied in? Would they have access to suitable software and the necessary skills to interact with it?
- Is there enough descriptive metadata to allow the content to be located and retrieved by the people who need to use it?
- Are the methods and interfaces used to access the content appropriate for the end users?

Knowledge of the designated community is important to help inform these decisions. Rather like a compass, this knowledge provides essential navigational support as to the direction of travel for a digital archive (Keitel, 2018).

3   Designated community and related concepts in OAIS

The term 'designated community' was first coined at the first US Workshop of the standardization process, which eventually led to OAIS. In October 1995, it was agreed that 'an archive is a repository that preserves information for use by a designated community' (Sawyer, 1995). The standard that later emerged from this work defines a designated community as follows:

> An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time. (CCSDS, 2012, p.1-11).

The designated community remains key to the definition of an archive which is described as:

An organization that intends to preserve information for access and use by a Designated Community (CCSDS, 2012, p.1-9).

This concept of the designated community is so central to the OAIS Reference Model that defining it is listed as one of the six mandatory responsibilities that an organization must discharge in order to operate an OAIS archive (CCSDS, 2012, p.3-1).

Another key concept in OAIS defines how archived digital content should be 'independently understandable' to that designated community. This concept is described as:

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. (CCSDS, 2012, p.1-12)

If we know who digital content is being preserved for, the level of knowledge and understanding of that community could be identified. OAIS refers to this as the 'knowledge base' of the community (CCSDS, 2012, p.2-3). The knowledge base of the designated community effectively sets the scope of additional information and metadata that archive users will need (CCSDS, 2012, p.2-3).

Where information needed to understand the content falls outside of the knowledge base of the designated community, this information should be included alongside the digital content as 'representation information'. Representation information is the documentation, metadata and/or software necessary to allow the designated community to make sense of, render or otherwise use the digital content (CCSDS, 2012, p.1-14). Structural representation information might, for example, aid the user in transforming the bitstream of a JPEG file into the sequence of coloured pixels that the user can interpret as an image. Semantic representation information might, for example, reveal how a list of numbers in a CSV file correspond to readings of the height of a river, measured in metres. Crucially '...the OAIS must understand the Knowledge Base of its Designated Community to understand the minimum Representation Information that must be maintained.' (CCSDS, 2012, p.2-4). For example, if the designated community for an archive is composed of geography researchers, the knowledge base of the designated community would include familiarity with and access to Geographic Information System (GIS) software, which would not need to feature in associated representation information.

Another relevant concept from OAIS that we need to understand is that of a 'consumer'. This term is closely related to (and sometimes confused with) the designated community.

Consumer is the role played by those persons, or client systems, that interact with OAIS services to find and acquire preserved information of interest. (CCSDS, 2012, p.2-3)

OAIS notes that the designated community is a specific class of consumer, the subset of consumers who should be able to understand the preserved information. So, whilst an archive may have a wide user base, their designated community is typically a narrower set of users with specific knowledge and skills that they are committed to serving.

4   The designated community in practice

OAIS is a high-level reference model that doesn't explain *how* an archive should go about defining its designated community or *what* such a definition would look like. It does give some examples of the sorts of information that might go into the specification, but these examples are very brief and tend to focus purely on the level of education or knowledge base of the community. For example:

For some scientific information, the Designated community of Consumers might be described as those with a first year graduate level education in a related scientific discipline (CCSDS, 2012, p.3-4).

So how are organizations actually defining their designated communities in practice? This is a difficult question to answer as few organizations publicly advertise their designated community. A quick poll of attendees at a DPC event to discuss the concept of designated communities in 2022 demonstrated that the vast majority did not have a designated community definition (Mitcham, 2022). There are however some good examples within the community including designated community statements published by NARA's Electronic Records Archive (which defines agency and NARA staff as the designated community) and NASA's Space Science Data Coordinated Archive (which defines the designated community to be those who have completed a college-level degree program or have equivalent experience in one of the relevant scientific disciplines) (NARA, 2022; NASA, 2016). A more detailed statement is published from the Netherlands Institute for Sound and Vision, identifying both internal and external communities and documenting their needs (Verbruggen, 2016).

A number of digital preservation certification standards require evidence of a documented definition of the designated community as part of their assessment criteria (CCSDS, 2011; CoreTrustSeal, 2022; nestor Certification Working Group, 2013). Applications for CoreTrustSeal are made publicly available in the interests of openness and transparency and include a definition of designated community (CoreTrustSeal, 2023). The CoreTrustSeal application form, asks for the description of the designated community to be 'brief' and the resulting definitions are frequently only one or two sentences in length.

For some organizations (in particular those that have a discipline specific remit), their designated community definition is, at first glance, quite tightly defined. The Norwegian Marine Data Centre for example states that:

> The repository's Designated Community is researchers within Norwegian marine research. The research is within marine sciences such as management of fish stock, salmon lice development, causes of increase in ocean temperature (Norwegian Marine Data Centre, 2022).

The Language Archive states that it:

> …mainly targets scholars of language-related research disciplines. This includes linguists, anthropologists and psychologists, but also neurobiologists and geneticists who study language-related phenomena (The Language Archive, 2022).

The National Geoscience Data Centre states that its designated community:

> …consists of a wide range of users of subsurface geoscience data and models. Users are typically in academia, local authority organisations, and industry e.g. hydrocarbons industry, environmental consultants, as well as the geotechnical and site investigation sector (The National Geoscience Data Centre, 2022).

Most definitions appear to focus on the professions or research areas of their designated community, rather than drawing out specifics of the knowledge base or skill set. However, the National Library of Scotland references the knowledge base of their community of web users with a brief discussion of the technical skills they would expect this community to have:

Web Users must have internet access and sufficient skills to interact with browser-based web services written in English and using text, images, and maps as the main navigational aids (National Library of Scotland, 2022).

Historic Environment Scotland takes a different approach with the inclusion of a list of those groups that make up its designated community (from its staff to school children) and an acknowledgement that a wide range of skill levels and knowledge bases are represented (Historic Environment Scotland, 2021).

There is some variation in how organizations apply the designated community concept. Some university libraries distinguish between primary user communities such as faculty members, staff, and students, and secondary user communities encompassing other researchers, the media, and the general public (York University, 2022; McMaster University, n.d.). An organization may also choose to define separate designated communities for specific collections or types of content that they preserve. For example, one of the authors of this Guidance Note reports that although the archival holdings at the State Archives of Baden-Württemberg are primarily aimed at the general public, they also include seismographs from an earthquake monitoring station, which are targeted only at historians and geologists with an assumed understanding of this type of digital content.

5   The general public as designated community

There has been some debate centring on whether 'the general public' is an appropriate designated community (Bettivia, 2016; L'Hours, Kleemola, and de Leeuw, 2019; Underdown and Talboom, 2019). OAIS itself doesn't actively discourage this approach, though it should be noted that such a broad definition will likely require additional work to ensure adequate metadata and documentation is available to enable this community of users to understand the digital content. As a result of this additional work, OAIS suggests that it may be better to apply a broad definition of designated community from an early stage rather than start with a narrow definition and later find it should be expanded (CCSDS, 2012, p. 3-4).

Many of the organizations mentioned previously (see section 4) that appear to have quite a specific designated community also broaden the definition out to encompass anyone who has an interest in consulting their collections. For example the Norwegian Marine Data Centre ends its designated community definition with 'In addition to the research community, NMD also makes data available to the public' and The Language Archive states that 'The rich collections of recordings of languages and cultures in their natural context however are of interest to a much broader range of scholars, e.g. ethnomusicologists and ethnobiologists, as well as to the language communities themselves and the general public' (Norwegian Marine Data Centre, 2022; The Language Archive, 2022). The general public is frequently used as a catch-all within a much more targeted definition.

Large memory institutions such as national archives or libraries often state that their records and objects are preserved for the 'general public'. A community so broad brings particular challenges around understanding and documenting specific skill levels and requirements (see for example Underdown and Talboom, 2019).

An alternative approach is that taken by NASA's Space Science Data Coordinated Archive, who states that:

In general, the representation information associated with the individual data component or AIP is designed to make the data usable by a scientist with the appropriate college-level education or equivalent. Although NSSDCA distributes the data to the general public, the

general public is not considered to be part of the designated community for these data. NSSDCA staff is not generally available to support use of the data by individuals without the appropriate background. (NASA, 2016)

This statement acknowledges that whilst the general public may use their digital content, they are not considered to be a part of the designated community and thus additional information and support is not available for them.

6   Defining a designated community

As mentioned previously, there is little guidance on *how* to define a designated community. Where organizations plan to define their designated community, the following questions should first be considered:

1. What is the organizational driver for defining a designated community?
2. Who will approve or sign off on the definition?
3. Which documents within the organization should link to or reference the definition?
4. How will the definition be actively used to inform decision making?
5. How will the designated community be monitored?
6. What will the schedule be for update and review?

The following questions may help in defining the designated community:

1. Does the organization's overarching mission or remit mention specific communities that should be served?
2. Does the organization's funder have stipulations about the communities that should be served?
3. Are there communities that should be prioritised due to legislative or regulatory considerations?
4. Do the donors and depositors of the digital content have expectations about who the content is being preserved for?
5. Are there other key stakeholders for the organization who should be considered or consulted?

Remember that defining a designated community is *not* the same as documenting your users. A designated community does not need to include all categories of user currently served by an archive.

Also note that whether your designated community definition is broad or narrow, there will be implications and actions that follow on from this definition. For example, does appropriate 'representation information' exist to ensure that digital content is 'independently understandable' to the community that has been defined? The broader the definition of designated community, the greater the likelihood that additional documentation and metadata and support will need to be provided to ensure this community has the necessary knowledge to interpret and understand the digital content.

Once the designated community has been defined and documented, it can then be used to aid decision making in digital preservation activities (helping to answer those questions described in section 2) as well as bringing a level of consistency and transparency to those decisions.

7   Other approaches

Uncertainty on how to define and document a designated community effectively, and the lack of examples of it being used in practice, have been a cause for concern within the digital preservation community. Designated community is described by Kim as an 'elusive concept' for the digital curation community (Kim, 2015). In Bettivia's words: 'This term is a source of frustration to some in digital preservation, particularly librarians who find the need to specify a particular community to be at odds with their professional, and sometimes legal, mandate to serve broad populations' (Bettivia, 2016). Furthermore, Underdown and Talboom suggest that the concept of a Designated Community is one of the least developed aspects of the OAIS model, with many archives being unsure how to articulate it clearly (Underdown and Talboom, 2019). Building on this, Moles discusses the implications of an absence of a clear designated community for open government data (Moles, 2021). A recent study exploring the tension between the designated community's centrality to Trusted Digital Repository Certification and the uncertainty about how to use it in practice, highlighted this as an area of risk for repositories (Frank and Rothfritz, 2022). Interviews with the standards developers quoted in this study flagged up concerns that repositories found it hard to differentiate their designated community from current service users (Frank and Rothfritz, 2022). A recent online discussion on the theme of designated communities hosted by the Digital Preservation Coalition also highlighted a disconnect between the theory (as stated in OAIS), and what actually happens in practice, with some concern that engaging with this concept might be primarily a 'tick box' exercise for organizations who are seeking certification rather than an essential practice that is more widely embedded in our work (Mitcham, 2022).

Whether we define and document our designated communities or not, it is clear that those of us preserving information for the long term cannot effectively do so without some understanding of the potential users of that information. The Digital Preservation Coalition's Rapid Assessment Model (DPC RAM), a maturity model which encapsulates digital preservation good practice, highlights the importance of embedding user requirements into digital preservation policy and decision making (Digital Preservation Coalition, 2021). Other DPC Guidance Notes in this series provides advice on understanding user needs and highlights a number of methodologies for doing so (McMeekin, 2021), and a case study from the National Library of Scotland provides a helpful example of how consultation with users of their online maps service is helping to inform future decisions around preservation and access (Fleet, 2022).

In practice, it seems that many organizations publish a definition of their 'users' as part of a wider policy on access to collections without reference to the OAIS concept of designated community. See for example the Archives and Special Collections Access Policy from the University of Leicester which includes a bulleted list of the groups that make up the community that they serve (University of Leicester, 2017). Some organizations have found it helpful to document user stories or personas to help guide their work. Those created by the Software Preservation Network, and the Rockefeller Archive Centre are interesting examples and appear to provide a more detailed picture of potential users, their needs and their digital skills and knowledge than many of the examples referenced in this Guidance Note (Rockefeller Archive Centre, 2019; Software Preservation Network, 2017).

It may also be appropriate to focus on types of use of digital content rather defining specific groups of people. The nestor working group on Preservation Planning came to the conclusion that most user requests to access digital content fall into four types of 'designated use':

- Perceiving the complete object, e.g.,viewing a picture or film, listening to an audio stream.

- Evaluating/obtaining information, e.g. searching for an individual piece of information or a complex of information within an information object.
- Further processing after extracting a use package, e.g. describing a text or using a film clip for a documentary.
- Running the object, e.g. playing a computer game.
  (nestor Working Group Preservation Planning, 2014, p. 19)

Considering which if these designated uses should be supported may be a more concrete way of aiding digital preservation decision making than an understanding of designed communities.

Similarly, to overcome some of the challenges of their statutory duty to serve 'everyone' the National Archives UK have chosen to focus on the methods that those users might use to access and interrogate the digital content and thus their digital skill sets and knowledge base. Those who simply 'read' the information are considered in a different category to those 'data users' who use more advanced computational methods to work with the digital content. A further group of those who are 'digitally curious' is composed of individuals who aspire to be 'data users' but don't yet have the necessary digital skills (Underdown and Talboom, 2019).

8   Conclusion

This Guidance Note has described why knowledge of the users of digital content is valuable for informing digital preservation activities. More specifically, it has summarised the OAIS designated community concept and provided examples and guidance on defining this in practice. It has also discussed some of the challenges of the term and its use within the digital preservation community and given an overview of other approaches that organizations have taken to understand and document their users.

The designated community concept is not the only way to answer the question "who are we preserving this information for?" and it is important to consider whether a literal implementation of this theoretical concept will be the most effective way of meeting the needs of an archive in practice, particularly for those organizations serving very broad communities. What is clear is that an organization should have some understanding of who it is preserving its digital content for, what their requirements might be, and an active and continuing dialogue with these users as a means to inform and steer decisions around preservation and access both now and in the future.

9 References

Bettivia, R. S. (2016) *The Power of Imaginary Users: Designated Communities in the OAIS Reference Model*, Proceedings of the Association for Information Science and Technology, 53(1) pp.1-9. Available at:
https://web.archive.org/web/20230725151144/https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pra2.2016.14505301038

CCSDS: Consultative Committee for Space Data Systems (2011) *Audit and Certification of Trustworthy Digital Repositories*. Magenta Book (CCSDS 652.0-M-1). Available at:
https://web.archive.org/web/20230606025647/https://public.ccsds.org/Pubs/652x0m1.pdf

CCSDS: Consultative Committee for Space Data Systems (2012) *Reference Model for an Open Archival Information System (OAIS)*. Magenta Book (CCSDS 650.0-M-2). Available at:
https://web.archive.org/web/20230327234400/https://public.ccsds.org/Pubs/650x0m2.pdf

CoreTrustSeal (2022) *CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025 Extended Guidance*, Available at:
https://web.archive.org/web/20221221204355/https://zenodo.org/record/7051096#.Y6NwEnbP1D8

CoreTrustSeal (2023) *CoreTrustSeal certified data repositories*, Available at:
https://web.archive.org/web/20230725151644/https://amt.coretrustseal.org/certificates/

Digital Preservation Coalition (2021) *Digital Preservation Coalition Rapid Assessment Model (DPC RAM)*, Version 2. Available at:
https://web.archive.org/web/20230601123416/https://www.dpconline.org/docs/miscellaneous/our-work/dpc-ram/2433-digital-preservation-coalition-rapid-assessment-model-v2/file

Fleet, C. (2022) *Understanding user needs: a case study from the National Library of Scotland*, Digital Preservation Coalition Technology Watch Guidance Note. Available at:
http://doi.org/10.7207/twgn22-01 [accessed 25th July 2023]

Frank, R. and Rothfritz, L. (2022) *Designated Community: uncertainty and risk*, Journal of Documentation, 79(4), pp. 880-897. Available at: https://doi.org/10.1108/JD-07-2022-0161 [accessed 25th July 2023]

Historic Environment Scotland (2021) *CoreTrustSeal application form*, Available at:
https://web.archive.org/web/20230725152544/https://www.coretrustseal.org/wp-content/uploads/2021/10/historic-environment-scotland_2021-10-12.pdf [accessed 25th July 2023]

Keitel, C. (2018) *Der einzige Kompass den wir haben. Zur Kritik der Designated Community*, Informationswissenschaft: Theorie, Methode und Praxis, 5(1), pp. 25-37, Available at:
https://web.archive.org/web/20230725152637/https://bop.unibe.ch/iw/article/view/4245 [German only]

Kim, Y. (2015) *"Designated communities": through the lens of the web,* International Journal of Digital Curation, 10(1), pp.184-95. Available at:
https://web.archive.org/web/20230725152908/http://www.ijdc.net/article/view/10.1.184

L'Hours, H., Kleemola, M. and de Leeuw, L. (2019) *CoreTrustSeal: From academic collaboration to sustainable services,* IASSIST quarterly, 43(1), pp. 1-17. Available at:

https://web.archive.org/web/20230725153120/https://iassistquarterly.com/index.php/iassist/article/view/936

McMaster University (n.d.) *Designated Community*, Available at: https://digitalarchive.mcmaster.ca/node/54 [accessed 25th July 2023]

McMeekin, S. (2021) *Understanding User Needs*, Digital Preservation Coalition Technology Watch Guidance Note. Available at: https://web.archive.org/web/20230620011508/https://www.dpconline.org/docs/technology-watch-reports/2474-twgn-userneedsanalysis/file

Mitcham, J. (2022) *Designated communities - a disconnect between theory and practice?* Available at: https://web.archive.org/web/20230725153839/https://www.dpconline.org/blog/designated-communities-applying-the-theory-in-practice

Moles, N. (2021) *Open government data (OGD): challenging the concept of a "Designated Community"*, Records Management Journal, 31(1), pp. 18-33. Available at: https://doi.org/10.1108/RMJ-10-2019-0065 [accessed 25th July 2023]

NARA (2022) *Designated Community Statement.* Available at: https://web.archive.org/web/20230725154702/https://www.archives.gov/preservation/digital-preservation/designated

NASA (2016) *NSSDCA's Designated Community*. Available at: https://web.archive.org/web/20230608210623/https://nssdc.gsfc.nasa.gov/nssdc/designated_community.html

National Library of Scotland (2022) *CoreTrustSeal application form*. Available at: https://web.archive.org/web/20230725154953/https://www.coretrustseal.org/wp-content/uploads/2022/03/20220325-National-Library-of-Scotland-OAIS-Archive_final.pdf

nestor Certification Working Group (2013) *Explanatory notes on the nestor Seal for Trustworthy Digital Archives*, nestor-materials 17. Available from: https://web.archive.org/web/20230725155204/https://d-nb.info/1047613859/34

nestor Working Group Preservation Planning (2014) *Guideline for Preservation Planning: Procedural Model and Implementation*, nestor materials 15, Available at: https://web.archive.org/web/20230725155620/https://d-nb.info/1060129876/34

Norwegian Marine Data Centre (2022) *CoreTrustSeal application form*. Available at: https://web.archive.org/web/20230725155845/https://www.coretrustseal.org/wp-content/uploads/2022/09/20220929-norwegian-marine-data-centre_final.pdf

The Language Archive (2022) *CoreTrustSeal application form*, Available at: https://web.archive.org/web/20230725155949/https://www.coretrustseal.org/wp-content/uploads/2022/03/20220322-the-language-archive_final.pdf

The National Geoscience Data Centre (2022) *CoreTrustSeal application form*, Available at: https://web.archive.org/web/20230725160032/https://www.coretrustseal.org/wp-content/uploads/2022/03/20220303-national-geoscience-data-centre_final.pdf

Rockefeller Archive Center (2019) *Rockefeller Archive Center Personas*. Available at: https://web.archive.org/web/20230725160301/https://github.com/RockefellerArchiveCenter/personas

Sawyer, D. (1995) *ISO Archiving Standards - First US Workshop – Minutes*. Available at: https://web.archive.org/web/20090620204619/http://nssdc.gsfc.nasa.gov/nost/isoas/us01/minutes.html

Software Preservation Network (2017) *Collections as Data Personas*. Available at: https://web.archive.org/web/20230725161037/https://collectionsasdata.github.io/personas/

Underdown, D. and Talboom, L. (2019) *'Access is What we are Preserving': But for Whom?* Available at: https://web.archive.org/web/20230725161200/https://www.dpconline.org/blog/access-what-we-are-preserving

University of Leicester (2017) *Archives and Special Collections Access Policy*. Available at: https://web.archive.org/web/20230725161309/https://le.ac.uk/library/about/policies/archives-access

Verbruggen, E. (2016) *Monitoring Designated Communities: Wat we weten en willen weten over de gebruikersgroepen van Beeld en Geluid*. Available at: https://web.archive.org/web/20230725161535/https://publications.beeldengeluid.nl/pub/406 [Dutch only]

York University (2022) *Designated Community.* Available at: https://web.archive.org/web/20230609012148/https://www.library.yorku.ca/web/about-us/library-policies/digital-preservation-policy/designated-community/