

Lindlar, Michelle; Pohlkamp, Svenia; Zarnitz, Monika; Bähr, Thomas; Strathmann, Stefan

Conference Paper — Published Version

Mapping the landscape of Digital Preservation Networks. The nestor Digital Preservation Community survey

Suggested Citation: Lindlar, Michelle; Pohlkamp, Svenia; Zarnitz, Monika; Bähr, Thomas; Strathmann, Stefan (2023) : Mapping the landscape of Digital Preservation Networks. The nestor Digital Preservation Community survey, In: Proceedings of the 18th International Conference on Digital Preservation 2022, iPRES, Glasgow, pp. 359-365,
<https://www.dpconline.org/docs/miscellaneous/events/2022-events/2791-ipres-2022-proceedings/file>

This Version is available at:
<http://hdl.handle.net/11108/609>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.



<https://creativecommons.org/licenses/by/4.0/>

MAPPING THE LANDSCAPE OF DIGITAL PRESERVATION NETWORKS

The nestor Digital Preservation Community survey

Michelle Lindlar

TIB – Leibniz Information
Centre for Science and
Technology
Germany
michelle.lindlar@tib.eu
[0000-0003-3709-5608](tel:0000-0003-3709-5608)

Svenia Pohlkamp

DNB
Germany
s.pohlkamp@dnb.de

Monika Zarnitz

ZBW Leibniz Information
Centre for Economics
Germany
m.zarnitz@zbw.eu
[0000-0001-9229-1877](tel:0000-0001-9229-1877)

Thomas Bähr

TIB – Leibniz Information
Centre for Science and
Technology
Germany
thomas.baehr@tib.eu
[0000-0002-9337-7127](tel:0000-0002-9337-7127)

Stefan Strathmann

Göttingen State and
University Library
Germany
strathmann@sub.uni-goettingen.de
[0000-0001-5328-1174](tel:0000-0001-5328-1174)

Abstract – "Digital preservation is people" and "Digital preservation cannot be done alone" are often heard statements within our domain. Nevertheless, no exhaustive survey of digital preservation communities had been done. The nestor Digital Preservation Community survey closed this gap and the nestor working group "Community Survey" is currently working on a publication of the survey results. This short paper presents the survey design and process, gives an insight into some of the findings by using the survey results to answer questions about the landscape of digital preservation communities and gives a brief outlook on further work.

Keywords – digital preservation networks, digital preservation communities, survey

Conference Topics – Community; Exchange.

I. INTRODUCTION

"Digital Preservation is People" has become one of the most fervent slogans of our domain. It has been used to highlight the contextualization of our work within an institutional framework [1], the skills needed by people to do the job [2] and the relationship between people and technology [3].

Another universally accepted statement about our domain is that digital preservation is an enormous task - one that is too big to be tackled alone [4]. Therefore, it comes as no surprise that digital preservation networks have been around for almost as long as national programs addressing the risks of digital information loss. Within the first decade of the new millennium, networks like the Open Preservation Foundation¹ and the Digital Preservation Coalition² were built upon the momentum of EU funded projects and continue to grow and flourish until today. Other networks, like the US-based Digital Preservation Network (DPN) have ceased to exist [5].

It would be naive to assume that the "usual suspects" of networks contain all institutions worldwide who deal with digital preservation. Furthermore, we have to acknowledge that our knowledge of networks and communities is naturally limited by the geographical and domain-based framework that we ourselves interact in. Looking at iPRES, figures of authors [6] or

¹ OPF - <https://openpreservation.org/>

² dpc - <https://www.dpconline.org/>

attendees by country [7] show that digital preservation is tackled by institutions across the globe. It is therefore safe to assume that the landscape of digital preservation networks is larger than we know.

While surveys about digital preservation topics are not uncommon, they mainly target individuals and institutions, asking for input on resources or requirements. Examples for this are the OPF Community Survey [8] or the NDSA Storage Survey [9]. There had not been an extensive survey aimed at mapping the landscape of digital preservation networks³. Within nestor⁴, the German Competence Network of Digital Preservation, the working group "Community Survey" was founded and tasked with closing this gap. The survey presented here was drafted and conducted in 2019 - 2020, the results were analyzed in 2020 - 2021 and results are currently being finalized to be published in the second quarter of 2022⁵. Section II of this paper presents further background information about the design of the survey and the way in which it was conducted. Section III briefly showcases some of the survey results and how they can be used to answer questions about the digital preservation community landscape. We conclude this paper with an outlook to further work in Section IV.

II. SURVEY DESIGN AND PROCESS

The working group was kicked off in February 2019 and consists of 5 members from nestor partner institutions. Having had its first meeting just shortly before the global pandemic came into light, members' available resources frequently changed and subsequently the time plan had to be shifted and adjusted several times. Due to this, the overall process took over 3 years. The entire survey project can be broken down in 4 phases, which are as follows:

- *Phase 1: Definition and Preparation* February 2019 - May 2019
- *Phase 2: Survey* September 2019 - May 2020

- *Phase 3: Analysis* May 2020 - July 2021
- *Phase 4: Preparing Publications* August 2021 - April 2022

The following subsections briefly describe key issues that needed to be addressed organizationally as part of the project. These form necessary background information for the outcome presentation in Section III.

A. *Definition of "Community" and of the survey's goal*

The first step towards the preparation of the survey was finding a shared definition of "Community". The goal was to find a framework in which no network would define itself as "too small" or "too broad", thus feeling it does not fit into the boundaries of the survey. The intention of the definition was therefore to include rather than to limit. After much discussion, the working group reached the following shared definition for "digital preservation community" [10]:

- an open community of persons and/or institutions who engages with the subject of digital preservation as its sole or one of several subjects
- a community whose members are committed to digital preservation in a manner that goes beyond pure self-interest, in particular it goes beyond the sole or central purpose of supplying a product or providing a commercial service
- a platform for discussing digital preservation practice and research, including the development of tools
- a community can be
 - local, regional, national or international
 - large or small
 - product-related or not product-related

In parallel, we needed to formulate what we wanted to achieve by conducting this survey. Through discussion within the working group it

³ Communities" and "networks" are used interchangeably through- out this paper

⁴ <https://www.langzeitarchivierung.de/Webs/nestor/EN/Home/>

⁵ At the time of writing, the results have not been published yet. However, this is scheduled to happen before iPRES2022. In the case of acceptance, all references in this paper will be changed to the published versions.

became clear that we wanted to create a "map" of digital preservation communities - a map in a geographic as well as a subject-based sense of the word. The survey results should include a registry, which interested practitioners and researchers as well as other networks could use to identify networks that cover issues they are interested in. Such a registry could also allow for identification of targets for cross-community collaboration, hence creating synergies and making best use of our limited resources in digital preservation. Based on this, the working group formulated two types of output for the survey results: a registry of community profiles on the nestor website as well as a report on the survey data set, which summarizes anonymized results.

B. Questionnaire

The questionnaire⁶ was designed as an online questionnaire using the "Mailingwork" survey tool⁷. It consisted of 40 questions which were divided into the following categories:

- *Formal aspects* (10 questions)

Rationale: Understanding of where community is located, what it defines as success factors, how long it has been operating and how it can be reached.

- *Governance structure & financing* (5 questions)

Rationale: Understanding of community's legal status, financing sources and internal governance bodies (e.g., Board).

- *Organizational structure* (12 questions)

Rationale: Understanding of membership types, membership numbers and distributions across organization types; Understanding of geographic and subject scope as well as key services; Understanding of personnel resources (FTEs) and collaborations with other communities.

- *Communication* (10 questions)

Rationale: Understanding of outreach activities in width and depth; Understanding of collaborative work spaces used.

- *Events* (3 questions)

Rationale: Understanding of events organized for members / other target groups

C. Distribution of the survey

In a first step, the working group collected a list of known digital preservation communities as well as of mailing lists via which the survey announcements were circulated. Contacts from known communities were contacted directly and asked to take part in the survey, but also asked to suggest networks that they thought should be included in this survey. These named candidates were then also approached directly. Two follow-up emails were written if no response had been received. In addition to the direct contacts and mailing list distribution, the working group members used their social media channels and international practitioner networks asking to amplify the project. The survey ran for 8 months. While this may seem like a long time, it seemed necessary to receive the best amount of responses during global lock-downs.

III. RESULTS

Overall we received 73 responses. After deduplication and data cleansing of entries that did not match the given community definition, 54 valid responses formed the basis for all result analysis.

The data set presents a unique information resource about digital preservation communities. In this section we briefly describe the structure of the community profiles and showcase how the survey result presented in the forthcoming nestor publication can be used to answer questions about the current digital preservation community landscape.

A. Community Profiles

As described above, one of the targeted outcomes of the community survey is a registry of digital preservation communities. For this, a community profiles template was created, which includes 32 criteria that can be extracted from the survey questions. These criteria are grouped into the sections "General characteristics", "Mission and scope", "Governance structure and financ-

⁶ The full questionnaire will be made available in March 2022 as part of the nestor materials publication [10]

⁷ <http://mailingwork.de/software/features>

ing", "Organizational structure", "Cooperation", "Modes of communication - scope", "Events organized by the network". For those survey respondents who had indicated that they would be willing to include their data in a publicly available registry, profile sheets were generated and sent to the named contact asking for corrections and approval of the profile as well as for a logo to be included in the registry. At the time of writing this paper, 33 networks have agreed to be included and have approved their profile.⁸

B. *Aggregated results for the nestor material publication*

While the published community profiles give an in-depth insight into single communities, there is not a profile for every respondent to the survey. In contrast, the nestor material publication [10] includes the anonymized results of all 54 valid responses, making it an excellent resource for quantitative analysis. Since a discussion of the entire data set is not possible within the limits of a paper, we will showcase the results using 5 sample questions that can be answered using the data presented in the report.

1. *Where are digital preservation networks located?*

Despite the working group's efforts to spread the survey as wide as possible, the majority of the responding communities (80%) are located in Europe or North America. Table I shows the distribution of all respondents by geographic region. However, it needs to be noted that 7.4% (4 cases) listed "International" or "World" as their location. Other respondents stressed that their membership is indeed international, their offices, however, are all located in Europe or North America. Table I can therefore only be used to make a statement about the main location of the community, not of its geographic reach.

Table I
Surveyed communities clustered by geographic regions

Region of the world	% of answers
Asia	3.7%
Australia	5.6%
Europe	51.9%
North America	29.6%
World	7.4%

⁸ At the time of writing this paper, the profiles are not yet published via the nestor website. They are scheduled to go

2. *Is there a correlation between a community's founding year and the size of its membership?*

One might think: "The longer the network has been around, the more members it has". But is that really true? While "founding year" allows for a comparable answer, membership number is not quite as straightforward to interpret. This is largely due to different types of memberships that exist, such as personal or institution based categories. One respondent may have 20 organizations as members, whereas another community counts 1,000 individuals as members. Nevertheless, a quantitative analysis of founding year against membership numbers can offer some insights. Figure 1 shows the distribution of founding year and number of members for all respondents who supplied data for both questions (n=47). Especially the older networks, which were founded between the years 1945 - 1995 may come as a surprise as digital preservation was not really a topic back then. These responses can be contextualized by cross-checking the results in Figure 1 against answers to the question whether digital preservation is one of several topics of the community (10 cases). Since digital preservation research only dates back to the 1990s, it is safe to assume that communities found prior to that cover digital preservation as one of several topics. This, in return, needs to be taken into consideration when looking at the membership numbers of these communities - broader services and fields of interest, such as several library-relevant topics in addition to digital preservation, may have an impact on the overall number of members.

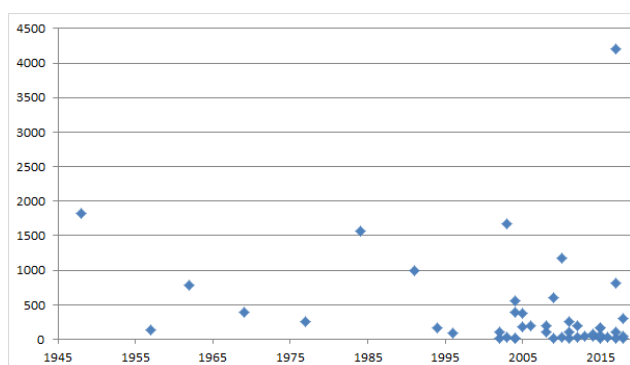


Figure 1 Number of members by networks' founding year

online later in March. In case of a successful, review the link will be added to the publication.

The clustering of communities found between 2000 and today underlines the previous statement that digital preservation research only dates back to the 1990s. It is thus no surprise, that the majority of responding communities were found after 1995. Three interesting observations can be made by looking at the numbers of communities found between 2000 and today: there is no decrease in intensity, suggesting that new communities have been created at a somewhat consistent rate over the course of the past 20 years; while there appears to be a peak in member-rich communities around 2005, the overall numbers suggest that comparatively young communities can still reach high numbers of membership; new communities are still found at a rather consistent rate, in other words: there appears to be no "over-saturation of digital preservation communities".

3. Are digital preservation communities just for archives?

When asking "who is involved in digital preservation" a first answer is often "archives", followed by "libraries". But is that all? And who is actively involved in digital preservation communities? To understand this, the survey asked about members' organization types. Out of the 54 respondents 46 supplied answers to this question. Some of those who didn't provide answers stressed that their community has no official membership model, making it hard to estimate organization types.

Table II shows the different organization types that were mentioned including how often they were mentioned and what percentage of communities have members of this organization type. It may come as a surprise that "Universities" ranks highest in the list of mentioned cases, however, we need to keep in mind that often a university library, archive, research institute or a computing center is the direct benefactor of the membership, but the university itself signed the membership agreement. Overall, the listing shows that the need for digital preservation communities exists in a broad organization base - over half of the communities have universities, libraries, archives and research institutions as their

members; over 40% additionally have museums and enterprises as members. The high number of "Others" is surprising. While one respondent classified their entire membership base as "others", 13 respondents made use of that category in addition to the named categories. The report [10] gives further information on the breakdown of different organization types across communities' membership basis. It is interesting to note that amongst the survey respondents were also "specialized networks", where one organization type makes up for 100% of the members. Such "specialized" digital preservation communities exist for libraries (n=2), archives (n=3), universities (n=1), enterprises (n=1) and government (n=1).⁹

Table II
Members' organization types across surveyed communities

Member type	Cases	% of communities with member type
Universities	39	72.2%
Archives	37	68.5%
Libraries	34	63.0%
Research Institutions	28	51.9%
Enterprises	23	42.6%
Museums	23	42.4%
Government agencies	4	7.4%
Broadcasting	2	3.7%
Individuals	2	3.7%
Others (unnamed)	15	27.8%

Table II also highlights where there is room for growth within communities. Why, for example, are broadcasting companies only mentioned twice? What can we do to get those organization types engaged in more networks?

4. What services do digital preservation communities offer?

After the previous subsections gave insight into who the communities are, another vital question is what services they provide to their respective members. Survey participants were asked to indicate which service they offer for their members and, if applicable, non-members. 10 possible services incl. the option "Other" were given. 51 participants provided answers regarding

⁹ Not included in this listing is 1 case of 100% "Others" and 1 case of 100% "Individuals", as those do not allow for organization matching.

services. The answer options as well as the number of times they were chosen can be seen in Table III.

Table III
Members' organization types across surveyed communities

Services	Cases	% of communities offering service
Knowledge transfer	44	81.5%
Community building	37	68.5%
Technology / Tool development	21	38%
Technology Watch	13	24.1%
Standardization	13	24.1%
Digital-preservation-as-a-service	12	22.2%
Lobbying	12	22.2%
Offering technical solutions / digital preservation software	10	18.5%
Certification	6	11.1%
Other	8	14.8%

The services can be grouped together in three "blocks". The two top ranking services are two classic "community" items - knowledge transfer and community building. These are offered by 68 - 81.5 % of the responding communities. Several communities who responded offered only one of these two services and no others.

A second block of services deals with technology in form of facilitating (joined) open source tool development, technology watch services or offering technical solutions such as (end-to-end) digital preservation software or even full digital-preservation-as-a-service. Between 18 - 25 % of the communities offer one or several of these technology-themed services. While this seems low in comparison to knowledge transfer and community building, it still stresses the high importance of community support around technology in digital preservation.

While the first block connects members to each other and the second block connects members to technology, the third block can be described as outward facing services for the members. These are: lobbying, standardization and certification as well as fundraising, which was the only entry made in the additional free-text field for "other". As can be seen in Table III, these types of services are less frequently offered than technology services, ranging between 11 -25%.

5. How are digital preservation communities financed?

The last question we would like to showcase in this paper is how the surveyed communities are financed. 53 communities provided information for this. We did not ask the participants to weigh their financing sources, i.e. indicated how many percent of overall funding a specific category makes up for, but to just list those that are applicable. Approximately 40% (n=22) of the responding communities are (partially) financed through membership fees, 39% (n=21 for each of the three categories) listed revenues from services, sponsoring or in-kind contributions as funding sources. The categories sponsoring (n=14), which was mentioned by 26% of the participating communities, as well as government funding, which was mentioned by 14.8% (n=8), received fewer mentions.

It is therefore safe to say that digital preservation communities are largely funded by the community members themselves – either directly in form of membership fees or fees for services or indirectly in form of in-kind contributions.

IV. OUTLOOK AND FURTHER WORK

As demonstrated in the Results section, the data gathered in the survey is a valuable resource. It can be used to answer questions about digital preservation communities such as how they are financed and what services they offer to their members. Since such structured information about digital preservation communities was previously not available, the nestor community survey has closed a gap in digital preservation discourse. Nestor has included the Digital Community Survey as a line-item in its product matrix and the working group is planning to re-run the survey in regular intervals, currently looking at every 3 years. Valuable lessons-learned in this first run of the survey will be reflected upon and fed into the next version of the survey.

A key issue the working group would like to improve in the next run of the survey is the time plan. Reflecting upon the time needed for the four phases as described in Section II, the time for "Definition and Preparation" (4 months) as well as for running the survey itself (8 months) seems reasonable. We may consider to keep the survey open for a short time frame – however, the longer period allowed us to individually chase known communities and ask them to participate. Already

having a contact list of networks to build upon, we may consider shortening the time the survey is open slightly while still being considerate of the high work load that many community managers face and the additional burden that a survey might pose. Without a doubt, the time needed for phases 3 “Analysis” (14 months) and 4 “Preparing Publications” (9 months) is too long and needs to be improved upon. We are hoping that the workflows we have established in this first run of the survey, in particular the templates for the community profiles and automated mechanisms to populate them as well as overall decisions regarding presentation forms, will allow for significantly shorter phases 3 and 4 in the future.

In addition to a stricter time schedule, the working group is in particular hoping to reach more communities, especially in currently non- or under-represented regions (see Table I). Within the working group itself, a higher awareness towards international communities and networks exists and communities identified throughout the year are kept track of to include in future surveys. Presentations and publications, such as this paper, help the working group in spreading the word about the value of the survey outcome and we are hoping that this will encourage more communities to participate in the future. We are currently identifying target channels to publish the results through to heighten the visibility of the community profiles and the report.

Lastly, the in-depth analysis during phase 3 has provided some feedback which will be fed into the next questionnaire to make it more concise and universally understandable. A number of questions have offered re-occurring answers in “others” free-text fields. A particularly high number (n=6) of named “other” categories could be found in the question regarding the legal status. All named “other” categories will be considered for inclusion as fixed categories in the next instance of the questionnaire. The working group will ensure that any changes made to the survey structure will still allow for comparability of the results across different survey instances over the years. Additionally, the working group will be happy to receive any feedback and comments on the survey and is hoping for the wider digital preservation community to shape this survey into a useful tool to keep on mapping our global landscape.

ACKNOWLEDGMENT

The authors would like to thank all networks that took time out of their busy schedules to participate in the survey and/or helped spread the word about the project.

REFERENCES

- [1] T. Owens, *The Theory and Craft of Digital Preservation*. Johns Hopkins University Press, 2018.
- [2] G. Hurley, Digital preservation is people: Thinking about digital skills for archivists, Personal Online Blog, 2018. [Online]. Available: <https://www.granthurley.ca/blog/digital-preservation-is-people-thinking-about-digital-skills-for-archivists/>.
- [3] J. Ranger, Digital preservation is people, AVP Online Blog, 2014. [Online]. Available: <https://blog.weareavp.com/digital-preservation-is-people>.
- [4] Libraries are facing big challenges in digital preservation: We cannot do it alone, Online IFLA Blog, 2017. [Online]. Available: <https://blogs.ifla.org/lpa/2017/11/30/libraries-are-facing-big-challenges-in-digital-preservation-we-cannot-do-it-alone/>.
- [5] DPN Nodes, The digital preservation network (dpn) to cease operations, Message published in LYRASIS online blog, 2018. [Online]. Available: <https://duraspace.org/the-digital-preservation-network-dpn-to-cease-operations/>.
- [6] M. Lindlar, Ipres authors 2004 - 2017 by country (infographic). Zenodo, 2018. [Online]. Available: <https://doi.org/10.5281/zenodo.1462500>.
- [7] iPRES2019, “Conference attendees,” in *Proceedings of the 16th International Conference on Digital Preservation*, 2019, p.20. [Online]. Available: <https://ipres2019.org/static/proceedings/iPRES2019.pdf>.
- [8] Open Preservation Foundation, “2019 - 2020 digital preservation community survey,” Tech. Rep., 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.4066912>.
- [9] National Digital Stewardship Alliance, “2019 storage infrastructure survey,” Tech. Rep., 2020. doi: DOI10.17605/OSF.IO/UWSG7.
- [10] T. Bähr, M. Lindlar, S. Pohlkamp, S. Strathmann, M. Zarnitz, “Results of the nestor community survey 2019-2020,” in *ser. nestor Materials*, nestor, 2022 (Forthcoming).