

A maturity and resourcing survey for building digital preservation business cases

Somaya Langley
Cambridge University Library
University of Cambridge
Cambridge CB3 9DR
szl20@cam.ac.uk

Lee Pretlove
Cambridge University Library
University of Cambridge
Cambridge CB3 9DR
ljp62@cam.ac.uk

William Kilbride
Digital Preservation Coalition
Glasgow G12 8QQ
william.kilbride@dpconline.org

Sarah Mason
Bodleian Libraries
University of Oxford
Oxford OX2 0EW
sarah.mason@bodleian.ox.ac.uk

James Mooney
Bodleian Libraries
University of Oxford
Oxford OX2 0EW
james.mooney@bodleian.ox.ac.uk

Edith Halvarsson
Bodleian Libraries
University of Oxford
Oxford OX2 0EW
edith.halvarsson@bodleian.ox.ac.uk

ABSTRACT

Benchmarking against peer institutions can be a powerful tool for building digital preservation business cases. A particularly challenging area to seek funding for is the staff resource required to run digital preservation programs. Data about current resourcing levels in non-US institutions is still not readily available to organizations looking to make the case for digital preservation. Instead, information about resourcing levels is often gathered informally, through conversations between practitioners or via straw polls. This approach forms a potential barrier for institutions not already involved in the digital preservation community.

The following abstract outlines the design of a new survey created by the Digital Preservation at Oxford and Cambridge (DPOC) project in collaboration with the Digital Preservation Coalition (DPC). The purpose of the survey was to gather information about the maturity levels of organizations doing digital preservation as well as collecting current statistics about full-time equivalent (FTE) staffing resources dedicated to digital preservation activities, and how these FTEs are arranged. The intention is to provide up-to-date information for organizations looking to build business cases and advocate for digital preservation programs. It also fills an immediate need for the DPOC project: to be able to benchmark against similar institutions. Anecdotal evidence from the DPC indicates there is an appetite for this information within their membership.

KEYWORDS

Maturity modelling, resourcing, staffing, policy, strategy, business case, capacity building, advocacy, digital preservation, survey

1 BACKGROUND

The DPOC project is a two-year collaboration between Bodleian Libraries, Oxford and Cambridge University Library (CUL). One of the main aims of the project is to deliver business cases that will facilitate both organizations' need to shift from their current digital preservation 'project model' approach to a more mature and sustained 'programmatic model' [3].

Segueing from project to program is known to be a particularly challenging step for organizations looking to enhance their digital preservation maturity [8]. Learning from past project outcomes,

the DPOC project utilized the Three-Legged Stool model for digital preservation. This model should also underpin future programs at Bodleian Libraries and CUL, in order to achieve sustainability when the DPOC project ends in December 2018. The Three-Legged Stool model reasons that the organizational infrastructure (including staffing), resources framework and technological infrastructure are all equally necessary for running a digital preservation program [7]. The DPOC project has found that two 'legs' of the stool (the organizational infrastructure and resources framework) are the hardest pieces of the digital preservation puzzle to 'solve' in an economic sense. The authors argue that this is due to the funding models that United Kingdom (UK) academic and research libraries operate within. Capital funding such as 'one-off monies' for hardware and software is easier to obtain. However, technology only corresponds to one 'leg' of the stool. Alone, technological infrastructure cannot provide a successful digital preservation program.

Short cyclical and technologically-focused funding models makes a holistic digital preservation program business case a challenge to sell to senior decision makers. This is not a hurdle only limited to the digital preservation community [9]. Another barrier for developing the DPOC business cases has been a lack of data on maturity levels and staffing resources for digital preservation programs in the UK and Europe. The DPOC project found they required concrete data from like-institutions that operate within similar economic environments, to strengthen the argument for specialist staff resources for managing and preserving digital content.

2 CONTEXT AND PURPOSE OF SURVEY

After reviewing previous studies, the authors found that basic information about maturity and resourcing—particularly for organizations based outside of the US—was hard to come by. As a remedy to the lack of data available for business cases, the DPOC project began collaborating with the DPC in August 2017 to develop a maturity and resourcing survey [11]. The survey compliments data collected in the latest NDSA Digital Preservation Staffing Survey (2017) [1], by attempting to gather figures relating to more non-US based institutions. One of the goals of the initiative was to create a survey template that could be sent out to DPC members (as well as

non-members) on an annual basis. In order to encourage participation and to sustain future surveying, the survey is relatively brief and intentionally excludes questions about organizations' digital collections.

Other than the NDSA surveys (2012, 2017), there have been other studies that cover maturity, resourcing, strategy and policy. These include Mind the Gap (2006) [13], Surveying Digital Preservation Readiness (2004) [12], Indiana University's Digital Collections and Preservation Survey (2016) [6], the EArK Project's Maturity Survey (2015) [5], and Axiell's Archives - Digital Preservation Survey (ongoing) [2]. However, these surveys have not specifically gathered detailed data about FTE resources for non-US based institutions. The NDSA dataset is the exception; it included data for 26 non-US respondents. Yet the NDSA survey does not specify if the FTE figures provided are estimates or exact. This is an important gap to fill, as this level of detail is required for DPOC business cases. For dates or numerical values requested, the DPOC maturity and resourcing survey also requested clarification as to whether the data provided is an estimate or exact figure.

3 SURVEY DESIGN

The survey consisted of five sections covering maturity, resourcing, strategy, policy and general information. A range of logic was embedded into the survey to ensure participants were only presented with questions relevant to their previous responses. Few questions were mandatory with the intention to discourage 'drop-outs'. A decision was also made to exclude asking any questions about digital content or collections (type, size, extent etc.) in organizations' custody. The survey ran from April to May 2018 and was promoted by the DPC to the international digital preservation community. A selected anonymized dataset [10] and short analysis of initial findings has been made available to the digital preservation community via the DPOC website [3].

4 LESSONS LEARNT

Self-reporting surveys are easy to design and obtain data for, yet the quality of the data is often low due to biases of the respondents [4]. This has led to error-prone data from the survey. Although the survey asked respondents to clarify if figures were 'exact' or 'estimates' there were still inconsistencies in the data. For future surveys, the authors would instead suggest filling in the questions as part of an interview, where immediate clarification can be sought. This approach would result in a smaller dataset, but with stronger data to underpin arguments in business cases.

Issues with the survey data are partly due to survey respondents being required to input numerical data, rather than select from pre-generated number or year ranges, or drop-down lists. While this may have saved time in generating the interactive survey, the data proved difficult to use or draw correlations from. Furthermore, the intention of creating a brief survey with few mandatory questions to encourage participation still led to a 54% 'drop-out' rate. Those that did complete the survey skipped a considerable proportion of the questions. This suggests that optional questions still yields lower participation. The language used also discouraged participation; comments from some respondents stated the length and wording of questions were confusing.

The authors have learned that in future surveys, questions should be clear, concise and consistent; questions should also be mandatory if the data is important for analysis. Despite being predominantly restricted to numerical data, free-input fields should be used sparingly to improve consistency and analysis.

5 FINDINGS AND CONCLUSIONS

Despite several issues with the survey and resulting data, it was encouraging to still observe some strong results. Sixty-nine percent of organizations with digital preservation policies also maintained a strategy to guide future planning. There was also some correlation between having a digital preservation policy in place and a higher number of FTEs than average. The survey showed that at present, there is a clear preference for a single digital preservation team (50%) over distributing individuals throughout an organization (34%). The authors believe that the survey data will in itself be valuable when making the business case for digital preservation programs, as well as benefiting the wider digital preservation community. However, there is still an ongoing data gap regarding digital preservation staffing levels, where an interview format would yield more reliable results.

6 ACKNOWLEDGEMENTS

The authors wish to thank the DPC who promoted the survey, the contributions of their DPOC colleague Dave Gerrard, the DPOC project managers, and The Polonsky Foundation.

REFERENCES

- [1] Winston Atkins, Carol Kussmann, Katherine Kim, and Aliya Reich. 2017. Staffing for Effective Digital Preservation 2017: An NDSA Report. (2017). DOI: <http://dx.doi.org/10.17605/OSF.IO/3RCQK>
- [2] Axiell. 2018. Archives - Digital Preservation Survey. (2018). <https://www.surveymonkey.co.uk/r/digitalpreservation>
- [3] Digital Preservation at Oxford and Cambridge. 2016. Digital Preservation at Oxford and Cambridge. (2016). <http://www.dpoc.ac.uk/>
- [4] Donaldson, Stewart I. and Grant-Vallone, Elisa J. 2002. Understanding Self-report Bias in Organizational Behavior Research. *Journal of Business and Psychology* 17, 2 (2002), 245-260.
- [5] EArK Project. 2015. EArK Maturity Survey. (2015). <http://earkmaturitysurvey.dlmforum.eu/0/intro>
- [6] Indiana University. 2016. 2016 Digital Collections and Preservation Survey Results. (2016). <https://wiki.dlib.indiana.edu/display/DIGIPRES/2016+Digital+Collections+and+Preservation+Survey+Results>
- [7] Anne R. Kenney and Nancy R. McGovern. 2003. Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems. (2003). <http://dpworkshop.org/dpm-eng/conclusion.html>
- [8] Anne R. Kenney and Nancy R. McGovern. 2003. *The Five Organizational Stages of Digital Preservation*. Michigan Publishing, University of Michigan Library. DOI: <http://dx.doi.org/10.3998/spobooks.bbv9812.0001.001>
- [9] Rob Kitchin, Sandra Collins, and Dermot Frost. 2015. Funding Models for Open Access Repositories. (2015). DOI: <http://dx.doi.org/10.3318/DRI.2015.4>
- [10] Langley, Somaya and Pretlove, Lee. 2018. Digital Preservation Maturity, Resourcing, Policy and Strategy Survey (2018) - Selected Anonymised Raw Data. (2018). <https://doi.org/10.17863/CAM.26222>
- [11] Langley, Somaya and Pretlove, Lee and Kilbride, William. 2018. Digital Preservation Maturity, Resourcing, Policy and Strategy Survey (2018) - Survey Questions. (2018). http://www.dpoc.ac.uk/wp-content/uploads/sites/181/2018/04/20180423_DPOC_DPC_MeasuringDigitalPreservation_PolicyStrategyStaffResources_Survey.pdf
- [12] Northeast Document Conservation Center. 2014. Surveying Digital Preservation Readiness. (2014). <https://www.nedcc.org/free-resources/digital-preservation>
- [13] Martin Waller and Robert Sharpe. 2006. Mind the gap: Assessing digital preservation needs in the UK. (2006). <https://www.dpconline.org/docs/miscellaneous/advocacy/340-mind-the-gap-assessing-digital-preservation-needs-in-the-uk/file>