Cloud Atlas

Navigating the Cloud for Digital Preservation

Andrea Goethals
National Library of New Zealand
New Zealand
andrea.goethals@dia.govt.nz
orcid:
0000-0002-5254-9818

Maria Praetzellis
Internet Archive
USA
maria@archive.org
orcid:
0000-0001-5047-3090

Roslynn Ross
Library and Archives Canada
Canada
roslynn.ross@canada.ca
orcid:
0000-0001-6006-6034

Nicholas Taylor
Stanford Libraries
USA
ntay@stanford.edu
orcid:
0000-0002-9504-0640

Abstract – Cloud services have become the increasingly dominant paradigm for many other types of IT-based services; why not for digital preservation? Cloud services are by now a familiar, though not mature, part of the digital preservation landscape. It is worth considering further how commercial cloud services are currently or might prospectively be used for digital preservation, as well as the implications for memory institutions, individually and collectively, if digital preservation shifts primarily in that direction. This panel will offer contrasting institutional perspectives on the potential or the perils of the cloud for digital preservation, featuring case studies on how memory institutions can leverage the cloud in deliberate and mission-supporting ways, and how some are working to build alternative, community-based infrastructures.

Keywords – cloud, community, digital preservation, infrastructure, storage

Conference Topics – Exploring New Horizons; The Cutting Edge: Technical Infrastructure and Implementation

Introduction

The hallmarks of cloud-hosted services are unparalleled flexibility and scalability, features that have proven sufficiently compelling as to bring about a sea-change in the way that organizations think about their IT infrastructures. The cultural heritage community is no exception; both consideration and application of cloud services for digital preservation has accelerated, as memory institutions negotiate the evolution of their parent organizations' IT strategy, if not the shift in the macro IT environment. How to best leverage cloud services to serve digital preservation or build complementary or alternative community-based services are areas of active exploration. Application of cloud services for digital preservation should ideally consider both fitness and externalities.

Notwithstanding their dynamic impact for IT and society, the large commercial cloud service providers have lately come in for a good deal of criticism. Some in the digital preservation field have raised questions regarding the fitness of commercial cloud providers in meeting fundamental characteristics that the best practices in digital preservation necessitate. For example, David S.H. Rosenthal points out that many cloud storage services have a number of conspicuous shortcomings: continuity of storage contingent on continuity of payment, opaque data integrity assurance mechanisms, vulnerability to privileged insiders, vulnerability to operator error, attractiveness as a target of attack, steep fees for data egress, and organizational immaturity.[1]

Cultural heritage institutions would do well to mind alignment with their goals and values. The library profession has traditionally held privacy protection as a core tenet. How well do the business models of cloud services support this value, or can cloud services be leveraged in such a way that this concern is mitigated? How can memory institutions utilize cloud services to enhance their impact and relevance, and not cede either capacity or commitment to serving as information stewards? Can, or how can, memory institutions enlist cloud service providers to partner in meaningful ways to support community needs, given our rich missions and often modest resources, as well as the

iPRES 2019 – 16th International Conference on Digital Preservation
September 16– 20, 2019, Amsterdam, The Netherlands.

1

near-monopoly power of some of the key companies?

These questions highlight areas of possible mission misalignment between memory institutions and cloud service providers. The latter typically prioritize sustainability rather than profit, openness rather than capture, diversification rather than vertical integration, long-term stewardship rather than market-contingent product commitment, expansive dissemination rather than tolled access, community rather than market predominance.

At the same time, the trends in the IT environment are both unmistakable and unignorable, and cloud services - whether those specifically provided by well-known commercial companies, or the paradigm of hosted services more generally - offer advantages that definitely make them worthy of consideration for use in digital preservation. Acknowledging that cloud services play a growing and important role in digital preservation, it is crucial that those entrusted to digital preservation approach this new environment with the ability to assess trade-offs between benefits and risks, a vision for the infrastructures to help realize as a field, and an interest in mapping what alternatives the "cloud" might afford.

Session Objectives

The panel members will offer case studies of how their respective organizations, individually and in partnership with one another, are negotiating the role of cloud services as part of or in juxtaposition to their digital preservation infrastructure portfolios. Along the way, the panelists and facilitator will engage in and stimulate audience discussion on a number of important and timely questions:

How can cloud services best be utilized to enhance and extend digital preservation practice and capabilities without compromising core mission or ceding core capacities?

How can the cultural heritage community nurture complementary, principled, and practical alternatives to commercial cloud services, and why does that matter?

Where do cloud services offer the greatest potential for impact for digital preservation? Where do they pose the greatest risks?

Contributors

Andrea Goethals is the Digital Preservation Manager at the National Library of New Zealand where she manages the Preservation Research & Consultancy Team. Andrea will serve as a facilitator for this session.

Maria Praetzellis is a Program Manager with the Internet Archive where she works with the library community building and managing programs that support web archiving and digital preservation. Maria will discuss new IA initiatives to develop decentralized technologies to bolster existing institutional infrastructure and provide new tools for efficient data management and preservation.

Roslynn Ross is the Director of Digital Preservation and Migration at Library and Archives Canada. Roz will discuss the challenges of managing a large digital collection and considerations when working with diverse teams to implement a technical solution leveraging cloud capabilities.

Nicholas Taylor is the Program Manager for LOCKSS and Web Archiving at Stanford Libraries. Nicholas will discuss the LOCKSS Program's longstanding support of community-based, local, distributed digital preservation infrastructure as a juxtaposition to commercial cloud services. While the LOCKSS Program has generally maintained a skeptical orientation towards cloud services for digital preservation storage, in particular, a recent, major software re-architecture offers opportunities to better leverage cloud-like infrastructures as well as strengthen the model for LOCKSS as a hosted service.

REFERENCES

[1] Rosenthal, D. (2019). Cloud For Preservation. [online] DSHR's Blog. Available at: https://blog.dshr.org/2019/02/cloud-for-preservation.html [Accessed 20 Mar. 2019].

[2] Joseph, H. (2018). Securing community-controlled infrastructure: SPARC's plan of action. College & Research Libraries News, [online] 79(8), p.426. Available at: https://crln.acrl.org/index.php/crlnews/article/view/17246/18986 [Accessed 20 Mar. 2019].

iPRES 2019 – 16th International Conference on Digital Preservation
September 16– 20, 2019, Amsterdam, The Netherlands.

2