

Digital Preservation Services for Institutional Repositories (IR) in India

Dr. Smita Chandra*

* Librarian

Indian Institute of
Geomagnetism, Navi Mumbai,
Maharashtra, India

QR Code



Abstract: - Two approaches to the successful implementation of third-party digital preservation services within institutional repositories has been suggested in the paper. The ways the two approaches could be implemented by institutions has been proposed. Finally, without stressing on choosing any one of the approaches to digital preservation, the option to choose any one has been left on the individual repository based on needs assessment. The paper concludes that Digital preservation community based solutions are not likely to be a free-and-easy cure to all the problems, but it would certainly be a viable option worth considering. Vendor-supported solutions do not have a monopoly on services or good support as there are many service providers in the field, but they may be a faster and easier choice especially for simple needs.

Keywords: Digital Preservation; Preservation Services, Distributed Preservation Models, IRs

Introduction

While there has been some confusion about the role of institutional repositories (IRs) for digital preservation. Lynch (2003), in his essay clearly defines the role of institutional repositories as,

"A set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials,

including long-term preservation where appropriate, as well as organization and access or distribution."

Digital preservation within institutional repositories is beginning to form a comprehensive area of research. Institutional repositories have begun to address issues related to a digital preservation policy, a content policy, and issues related to sustainability, etc. (Li, 2011). However, the same cannot be said for IRs in India. Since IR is still in a stage of development at many institutions, lack of sustainable funding and

adequate staffing could present an obstacle in implementing successful digital preservation programs. Despite the challenges it is very encouraging to see a number of digital preservation policies being developed, as listed at the Directory of Open Access Repositories (DOAR).

Of the increasing number of digital preservation strategies being implemented for IRs, the need and role of a third-party digital preservation service provider has been stressed upon. Two of the most common approaches to implementing a third-party digital preservation service are listed in the paper. Without going into the specifics of a particular product, the way the two approaches to institutional repositories can be addressed has been stressed upon, followed by the option to choose the best approach based on needs assessment.

Need for Third-party Digital Preservation Service (DDPS) provider

Repository staff is often occupied with the task of advocating its use in an institution, handling submitted data and performing regular maintenance. In many circumstances, they do not have sufficient time to actively manage the preservation process. The argument is supported by Lavoie and Dempsey (2004), who suggest "*long-term stewardship may be beyond the means of an individual institution*". The MetaArchive Organization (MetaArchive), a recently

established and growing inter-institutional alliance, asserts that cultural memory organizations can and should take responsibility for managing their digital collections, and that such institutions can realize many advantages from taking the services of service providers for long term preservation and access strategies. This assertion is based both on the shared convictions of its members and on the successful results that MetaArchive has achieved in recent years through coordinated activities with third-party preservation services. Third-party digital preservation holds out the promise of indefinite preservation of materials without degradation, but this needs to establish effective processes, both technical and organizational to enable it.

In the context of the critical need for a new consensus on how to preserve digital collections, a growing number of cultural memory organizations have now come to believe that the most effective digital preservation efforts in practice succeed through some strategy for distributing copies of content in secure, distributed locations over time. This conceptual strategy is a straightforward carry-over of the practices that the old world of handwritten records, enabled scholars to preserve content through millennia. But in the digital age this strategy requires not only the collaboration of like-minded individuals, but also an investment in a distributed array of servers capable of storing digital collections in a pre-coordinated methodology. A single cultural memory

organization is unlikely to have the capability to operate several geographically dispersed and securely maintained servers. Taking help from third-party preservation services is essential, and this requires both organizational and technical investments. Not only a pre-coordinated technological solution, but also strong, long-term inter-institutional agreements must be put in place, or there will be insufficient commitment to act in concert over time. The following quote from a joint National Science Foundation (NSF)/Joint Information Systems Committee (JISC) study captures the opportunity presented by this situation succinctly:

The increased number and diversity of those concerned with digital preservation—coupled with the current general scarcity of resources for preservation infrastructure—suggests that new relationships that cross institutional and sector boundaries could provide important and promising ways to deal with the data preservation challenge. These enabling methods could potentially help spread the burden of preservation, create economies of scale needed to support it, and mitigate the risks of data loss.

Approaches to third-party Digital Preservation Service

When considering the option of taking help of a third-party preservation service provider, the following two proposals are available with an institutional repository,

- i. Commercial vendors for digital preservation
- ii. Software and services of the digital preservation community

i. Considering commercial vendors for digital preservation services

The concept of services provided by a vendor is not a new one. It is hard to generalize about different types of services provided by them because they fit into a huge range of categories. The services could be packaged into a one-size-fits-all package that institutional repositories in India, could easily install on a desktop. Such systems are configurable, that takes time and care to roll out across an organization. Such services by commercial vendors comes at a cost which might need continuous audits as the annual maintenance contracts might need to be reviewed or as newer versions of a software become available or as consultation fees for problem solving or training purposes. Such third-party digital preservation services might be solely supported by the vendor, or the vendor might be only one piece in a broad ecosystem of providers, consultants and implementers that could help organizations set up and running its digital preservation services.

ii. Considering software and services of the digital preservation community

In contrast, the open source solutions and free services for digital preservation are typically developed, marketed and distributed by a loosely organized community of individuals that comprise the digital preservation community. Considering such software and services of the digital preservation community means that the source code itself, the instructions that cause the application to do what it does, can be easily viewed, modified or downloaded by anyone with technical expertise. Because of the inherent nature of Free software or FOSS (Free and Open Source Software), one is free to run these software packages for any purpose—the institutional repository would generally not pay anything to acquire them. Secondly, the source code is free—so with the help of expertise the code could be studied and understood as to how it works. Thirdly, one is free to copy and redistribute the package to anyone. And finally, one is free to modify the software however one likes, and to release those modifications. Community is the key concept to understand how digital preservation is enabled within organizations--this is an organized, albeit distributed group of people with three main things in common: shared goals and interest, community rules and practices and the ability to execute as a team and create software and services. The term ‘community’ here is an informal term used for users and developers as well as supports of digital preservation. A community could operate as a cooperative that could comprise of say an enterprise in which the infrastructure of the endeavor would be owned

and operated by its users. The Private LOCKSS Network (PLN), forming a cooperative for digital preservation purposes is one example that offers a simple, low-cost mechanism for digital preservation. Based on the PLN, a cooperative formed to implement OAIS reference model based digital preservation within institutions would comprise of domain specific institutions that have an institutional repository. In such a cooperative, the responsibility to enable each of the Open Archival Information System (OAIS, 2012) functional model and/or information model components for the other institutions would be distributed. Therefore there would be one or many geographically dispersed member(s) of the cooperative acting as Distributed Digital Preservation Service (Reich, 2009) provider(s) for other members of the same cooperative. Thus, for example, for one of the institutional repositories ingest could be one functional component being handled by another institution with the technical expertise to do so and that also acts as the archival storage service provider. This would mean that a group of repositories could come together to form a cooperative and one of the repositories could act as a third-party DDPS provider to the other.

It can be said that both these options (discussed above) for taking help as third-party preservation services have the same advantages, disadvantages and similarities as those between commercial software and open source software. It is therefore necessary for an institutional repository to assess its needs. It is critical to look at the features,

attributes and costs for each--and then weigh the unique benefits each model can bring. By choosing to use a community based open source system, organization could tap into a community of peers and take more control of its digital preservation software and service needs, but by choosing this, it would also be choosing a course of action that will require more work to support--down the road.

Conclusion

In the end, there is no easy answer. Digital preservation community based solutions are not likely to be a free-and-easy cure to all the problems, but it would certainly be a viable option worth considering. Vendor-supported solutions do not have a monopoly on services or good support as there are many service providers in the field, but they may be a faster and easier choice especially for simple needs.

Finally reaching a decision on which of the above two options to select should be based on the following criteria. More could be incorporated if the need arises.

1. Needs assessment of the institutional repository on how highly a digital preservation service will be customized or tailored to suit a repository. It might for example be more expensive to customize a vendor provided service than a free community service.

2. Assessment based on how much training and support is needed to enable third-party preservation service for the institutional repository. Since there is currently no trained manpower in the area, it is expected that there would be need for lots of training and hands-on experience. It has been seen from experiences that traditional vendor-provided solutions are likely to be a better fit in this case.

3. There should be an assessment of the features of the services and software provided by vendors as compared to free solutions provided by the digital preservation community. For example, community based service might not be able to handle the security risks prone to digital preservation within the institute, as compared to commercial vendors.

4. Cost is a major factor that needs due consideration. A commercial vendor, for example, can develop a software package and then release it under an open source license--typically then making money from support or implementation work. Conversely, many communities provide services, support and even packaged implementation options for digital preservation. Which of the two suits an institutional repository in the long run should be studied in depth.

Bibliography

Lynch, Clifford A. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. Washington, D.C.: Association of Research Libraries, Retrieved from <http://www.arl.org/bm~doc/br226ir.pdf>.

Li, Yuan & Banach,Meghan. (2011). Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries. *D-Lib Magazine*, 17(5/6). Retrieved from <http://www.dlib.org/dlib/may11/yuanli/05yuanli.html#5>

Directory of Open Access Repositories (DOAR). Retrieved from <http://www.opendoar.org/>

Lavoie, B., and L.Dempsey. (2004, July/August). Thirteen ways of looking at... digital preservation. *D-Lib Magazine*, 10,(7/8). Retrieved from, <http://www.dlib.org/dlib/july04/lavoie/07lavoie.html>

MetaArchive Cooperative. Retrieved from <http://www.metaarchive.org/>

Private LOCKKS Network. Retrieved from, [http://www.lockss.org/lockss/Private LOCKSS Networks](http://www.lockss.org/lockss/Private_LOCKSS_Networks)

Consultative Committee for Space Data Systems Secretariat. (2012). Reference model for an open archival information system (OAIS): Recommended practice (CCSDS 650.0-M-2: Magenta Book), CCSDS, Washington, DC.

Retrieved from
<http://public.ccsds.org/publications/archive/650x0m2.pdf>

Reich, Victoria & Rosenthal, David. (2009). Distributed Digital Preservation: Private LOCKSS Networks as Business, Social, and Technical Frameworks. *Library Trends*, 57(3). Retrieved from 10.1353/lib.0.0047