

SHERPA-LEAP: a consortial model for the creation and support of academic institutional repositories.

Moyle, M., Stockley, R. and Tonkin, S.

Case study

Abstract

Purpose

To introduce SHERPA-LEAP, a model for the consortial development, population and support of eprints repositories.

Design/methodology/approach

The organisational and technical structures of the consortium are described, including a brief summary of central and local resource responsibilities. Some positive and negative aspects of a consortial approach to institutional repository development, and of the SHERPA-LEAP model in particular, are identified. Outstanding issues and future plans for the consortium are outlined.

Findings

SHERPA-LEAP is found to be succeeding in its aims of developing and supporting eprints repositories within the federal University of London. Some lessons learned from the SHERPA-LEAP approach are identified, but the SHERPA-LEAP consortial model is found to have been mostly beneficial to the participating institutions. In particular, the networking and experience-sharing opportunities which any consortial solution will facilitate are highly-valued by the SHERPA-LEAP partners.

Value

The case study is intended to help to inform the decision-making of institutions and consortia which are considering consortial solutions to the establishment and maintenance of institutional repositories.

Keywords: institutional repositories, open access, consortium models, consortial approaches, self-archiving, academic libraries, eprints

Background

SHERPA-LEAP¹ (London Eprints Access Project, a partner in SHERPA²) was established in February 2004 as a consortium of seven Higher Education institutions. All were members of the federal University of London, whose Vice-Chancellor generously funded the project. The aims of the project were to create eprints repositories, hosted centrally by UCL (University College London), for each of the partner institutions, and to populate those repositories through collaborative advocacy. The seven development partners were:

- Birkbeck
- Imperial College London
- King's College London
- London School of Economics and Political Science (LSE)
- Royal Holloway
- School of Oriental and African Studies (SOAS)
- UCL (Lead partner)

Funding was subsequently awarded for the expansion of the project to all the institutions of the University of London (there are twenty-one in total). SHERPA-LEAP currently has thirteen partners, the following institutions having so far joined SHERPA-LEAP during its second phase:

- Goldsmiths
- Queen Mary
- The School of Pharmacy
- School of Advanced Study
- The Institute of Cancer Research
- The Institute of Education

A third and final phase of SHERPA-LEAP will begin early in 2007; this will see the implementation of a cross-repository searching service for the consortium.

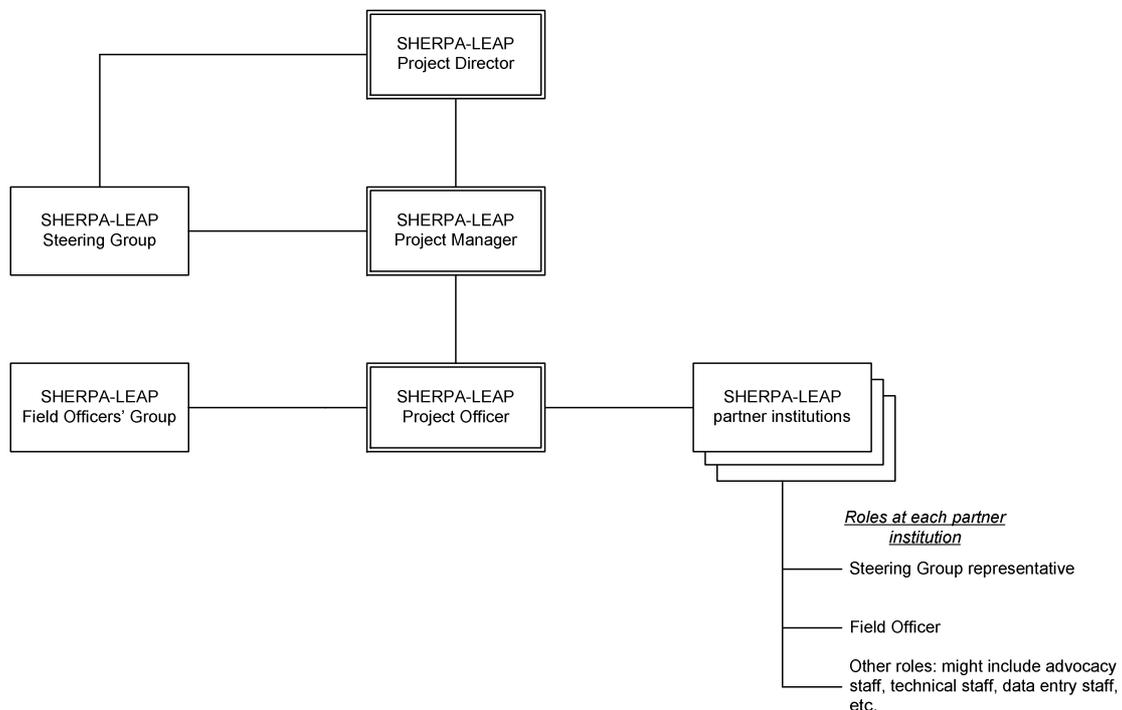
Although all partners are members of the federal University of London, SHERPA-LEAP is an *ad hoc* consortium, which was created specifically to move forward the repositories agenda in London. A UCL-hosted repository is not a condition of SHERPA-LEAP membership: it is open to any institution from within the University with an eprints repository, or with plans to develop one, regardless of repository platform and physical location. Within the partnership there is substantial diversity: the partner institutions represent a mixture of size and mission, ranging from the large, multi-disciplinary and research-led to the smaller and highly-specialised.

The SHERPA-LEAP consortial model

Organisational structure

Figure one shows the organisational structure of the SHERPA-LEAP Consortium.

Figure 1: SHERPA-LEAP: consortium structure



The project is overseen by the SHERPA-LEAP Project Team, consisting of a Project Director, a Project Manager, and a Project Officer. The Project Director (the Director of UCL Library Services) chairs a Steering Group for the project, which meets two or three times per year. This Group is responsible for monitoring the progress of the project against its objectives, and for discussing and developing project policy. Each partner institution is represented on the

Steering Group by a senior member of Library staff. The Steering Group ensures that the project partners cohere at a strategic level.

The Project Manager devotes roughly 0.1 FTE to SHERPA-LEAP, although this figure was higher during the early months of the project. He participates in the Steering Group, to whom he is accountable for the progress of the project and the management of its funds. The Project Manager also oversees the work of a full-time Project Officer, whose post was introduced as part of the second phase of the project. The Project Officer has a general remit for support activity within the SHERPA-LEAP community, providing advice and guidance to institutions on all matters relating to eprints, such as software configuration, metadata and copyright. An important part of the Project Officer's role is in facilitating the sharing of experience between the partner institutions.

Each partner has a designated Field Officer, responsible for coordinating day-to-day aspects of repository administration, such as advocacy to the local community, the management of repository workflows within their institution, the design and implementation of collecting policies, and IPR matters. The SHERPA-LEAP Field Officers all work closely with the Project Officer. The Field Officers also meet regularly as a Group, to share experience and discuss any issues to have arisen locally in the management of their repositories.

Working with the Field Officers at each institution may be any number of intermittently-contributing technical staff, data entry officers, staff with responsibilities for advocating the service to potential depositors, and so on, depending on institutional needs and preference.

The Project Team also organises outward-looking events, which supplement the locally-focused work of the Field Officers, for the benefit of the wider community. A one-day conference³, aimed at academics from the University of London's institutions, was held in mid-2005, and a second conference will follow in mid-2007. A repositories workshop, aimed at library and related staff, is currently being planned for early 2007; this workshop will be open to all institutions in London and South-East England, and will enable the SHERPA-LEAP partners to share some of their learning and experience with others in the region.

Technical Structure

The seven repositories in the first phase of the project were hosted by UCL. GNU EPrints was selected ahead of other open source platforms, primarily because of the availability of support from the SHERPA Technical Officer (who indeed proved to be very helpful during the setup phase). The feeling among the early partners was that GNU EPrints offered a quick and direct route to repository creation, and that migration in future to other platforms, should that be desired by any partner, would be unproblematic. UCL continues to host repositories for most of the partners: at the time of writing, ten out of thirteen repositories are UCL-based EPrints implementations, two partners run locally-hosted DSpace repositories, and one partner has a local EPrints installation.

The eleven centrally-hosted repositories sit on a single server. They are currently configured as eleven discrete archives running under a single copy of EPrints. SHERPA-LEAP did not implement a shared repository in the strictest sense - compare, for instance, the White Rose Consortium Eprints Repository, another successful UK consortial initiative [Proudfoot 2005]. The SHERPA-LEAP technical model was intended to allow most technical and policy decisions to be taken locally: the aim was to support diversity, and to focus on identifying and disseminating good practice, rather than to impose consistency. Each partner institution is responsible for configuring its own EPrints archive, to implement any locally-required functionality, and to achieve an institutional 'look and feel', without any need for SHERPA-LEAP branding. Partners also implement their own local metadata sets, decide on allowable document types and formats in support of their collecting policies, and develop their own deposit processes. Some guidance in EPrints configuration and metadata policy, mostly

collated from existing sources, was available to the partners in the first phase, but there was a substantial onus on each participating institution to resource the customisation of its repository. The appointment of the Project Officer at the start of the second phase of the project has meant that improved central support for repository configuration is now available, although the emphasis is still on supporting partners, rather than carrying out technical work on their behalf.

Costs

The fixed costs of server management, software maintenance, and the consortial infrastructure are absorbed by the project. The project also makes some funds available for local advocacy materials and events, and funds central events such as the Conference. Beyond this, each partner is responsible for resourcing its own repository. As with any IR, the costs vary from implementation to implementation [Crow 2002]. Examples of policy decisions which need to be taken by each SHERPA-LEAP partner and which influence operating costs include:

- Will authors upload papers, or will library staff mediate for them?
- Will library staff contact publishers whose policy on self-archiving is unknown or ambiguous on behalf of prospective depositors?
- Will the repository accept material in any format? Will some or all formats be migrated to a standard?
- How will supplementary data, charts, and so on be handled at ingest?
- Unless deposit is mandatory, the service will need marketing to potential depositors: who will do that?

Review

In this section some of the advantages and disadvantages of the consortial approach to repository development, especially relating to the SHERPA-LEAP consortial model, are considered.

Advantages

It is clear that SHERPA-LEAP is achieving its mission of developing and populating repositories, and, in this over-riding respect, the consortial approach clearly has something to offer. The centrally-hosted repository model offers partners a quick way to initiate an institutional repository [Rumsey 2005], with significantly reduced resource overheads compared to those that would be required to host a repository locally. It makes feasible the possibility of launching an institutional repository in pilot mode, which is often an essential step towards permanent funding. It is evident that at certain smaller institutions within the partnership eprints repositories had not been identified as priorities for funding, and those repositories would not be in place without SHERPA-LEAP. The central service has also helped some of the larger partners, who have used the first phases of the project to learn through experience and to develop their own detailed set of requirements for a locally-managed repository. As was anticipated, the migration of existing content away from the EPrints-based central service to other platforms is easily supported, and one of the original partners has already implemented migration to a local repository.

The technical configuration of the central service, with the hosted repositories running under a shared copy of EPrints, brought some advantages in the early days of the project. While SHERPA-LEAP did not at the outset produce technical documentation to add to what was already in the public domain, the configuration files from each archive are visible to others with server access privileges: at setup, UCL undertook the first repository customisation, and its configuration was available for copying by other partners. The shared copy of EPrints keeps the burden of central maintenance down, and mid-project enhancements, such as the addition of Analog reporting, have easily been propagated across the repositories. The administrative separation of each archive allows each partner to tailor its IR to suit the needs and expectations of the institution: for instance, if heavy institutional branding is important,

that can be achieved; if a "plain vanilla" pilot is more appropriate, that is easily achievable; and document types and file formats can vary between institutions, with metadata to match.

Organisationally, the consortium works well. The Field Officers network is especially appreciated: it offers the opportunity to share problems and solutions, ideas, and good practice, through the regular Field Officer meetings, the project Web site and mailing list, and informal contacts across the consortium. Naturally, the scale of SHERPA-LEAP and the diversity of its member institutions means that across the partnership there is engagement with a wide range of academic disciplines, from performance art to hepatology, and substantial collective wisdom is held by the partners in terms of advocacy to different audiences. The appointment of the full-time Project Officer has undoubtedly helped to build on the experience-sharing potential of the consortial structure, to the advantage of all concerned. The facilitation of opportunities for mutual support is probably the day-to-day aspect of SHERPA-LEAP most valued by its members, and a key strength of the consortial approach.

Since it began, SHERPA-LEAP has been linked with a number of other development projects. For instance, six of the partners are involved in SHERPA DP⁴, which is demonstrating a model for the distributed digital preservation of eprints. The original seven partners participated in the ShibboLEAP project [Moyle 2006], which saw the implementation of Shibboleth Identity Provider software at each of the participating institutions. SHERPA-LEAP, because of its size, the diversity of its membership, and because the partners collectively control a significant body of repository content, is a potentially interesting testbed for new projects. Several partners have benefitted from new development opportunities which have become available to them through membership of the consortium.

Disadvantages

Most of the perceived disadvantages of the SHERPA-LEAP consortial model relate to the technical structure. During the early setup phase, in which the first seven institutions made the initial configuration of their archives, only limited technical support was available from the centre. A dedicated technical post would have used up most of the project resources, and it was felt that, taking advantage of the published EPrints documentation and its technical support network, and with some mutual help within the project, it was realistic to expect most institutions to be able to configure their repositories reasonably speedily. In practice, the size and skills sets of relevant technical departments varied considerably between partner institutions; the EPrints documentation, which, not unexpectedly, was not up to the standard set by commercial systems suppliers, was found to be patchy in places (it is since much improved!); and the result was that some partners found the initial setup more problematic than had been anticipated. A related lesson learned, with hindsight, is that the project team provided insufficient clarity at the outset over local and central technical responsibilities. The technical help available to later entrants to the partnership from the centre is significantly better.

Sharing a copy of EPrints across institutions helped to keep the costs down, but is not recommended in the long-term for two reasons: firstly, under EPrints version 2, there is a risk that all the archives go down together if a configuration change made to one archive fails to compile; secondly, as partners' requirements became more sophisticated, it occasionally became desirable to edit the shared code which sits "above" the archives, and a change made at this level will generally affect each participating institutions whether it cares for it or not. In practice, the problems caused by the shared architecture have been minimal; but, if the project were starting afresh, it would seek to build better isolation for the constituent repositories.

Future plans and outstanding issues

The consortium has funding to mid-2008. The third phase of SHERPA-LEAP, the creation of a cross-searching service, will begin early in 2007. This will serve as a research showcase for

the members of the University of London, perhaps stimulating further inter-institutional research collaboration and new partnerships with the industrial and commercial sectors. The aggregation of search results may help to smooth over some of the discrepancies in local coverage of different disciplines, since research in the Arts, Humanities and Social Sciences is under-represented in comparison with Science, Technology and Medicine subjects. In the long-term, it is expected that this work will become aligned with the national repository search infrastructure⁵ which the JISC (Joint Information Systems Committee) is funding for the UK.

Some partners have used their engagement with SHERPA-LEAP to help them to crystallise the ideal functional requirements for their institutional repository, to consider what functionality is needed to cement the repository in the wider institutional research infrastructure, and in some cases to consider their strategy regarding digital asset management in its broadest sense across the institution. One of the original partners has migrated to a local DSpace implementation; other early partners are now considering moving their repositories away from the centrally-hosted, EPrints-based service. Two migrations, to local Fedora and DigiTool platforms respectively, will take place in the near future, and others are expected to follow. The consortium will incorporate a number of repository platforms in future, and this will serve to enrich further the collective experience and knowledge of the partners.

Meanwhile, it is expected that there will be demand for a centrally-hosted repository service for some time to come. However, funding for the project is finite, and the sustainability of the consortial model post-project is an important issue to be addressed by the Steering Group. At this stage, there are some open questions: if a hosted repository service is to continue, how will it be resourced? How many participants will, by then, still desire such a service, and how many will have moved, or be ready to move to local repositories? If the hosted service should disband, will smaller partner institutions have the in-house skills to set up their own repositories? (How far has the supporting contribution of SHERPA-LEAP prevented those skills from being developed in-house?) What, if any, central solutions need to be in place for long-term preservation? Such sustainability concerns may easily be managed when consortia are established and have the benefit of stable income streams. However, for project-specific consortia with fixed-term funding, such as SHERPA-LEAP, the question of sustainability demands attention, and an exit strategy will be high on the agenda for 2007.

Finally, an important part of any exit strategy will be to ensure that the network of mutual support and shared experience which SHERPA-LEAP has facilitated is maintained beyond the funded term of the project. The collective approach to discussion, ideas-sharing and problem-solving which evolved within the partnership is certain to continue across the London institutions.

Conclusion

The SHERPA-LEAP consortium set out to develop, support and populate eprints repositories within the federal University of London, and it is succeeding in these aims. The SHERPA-LEAP consortial approach is characterised by pragmatism. There is no shared repository - although a cross-searching service is under development - and the project supports the development of institutional repositories tailored to local needs. A central, hosted-repository service is available, but use of that service is not a condition of partnership, and it is not expected that institutions with hosted repositories will require them for ever. The organisational structure of the consortium has facilitated a culture of mutual support within the partnership. Increased technical resource at the start of the project would clearly have made for smoother initial progress, but, with a few caveats relating to aspects of SHERPA-LEAP's technical structure, the advantages of a consortial approach to the development of eprints repositories have outweighed the disadvantages for the partners and the project team. During the coming year, the Steering Group will assess the future needs of the partner institutions and consider both the form in which the SHERPA-LEAP consortial model might

be sustained beyond the lifetime of the project, and the means by which this might be achieved. The networking and experience-sharing opportunities which a consortial approach facilitates have proved to be key factors in the success of the partners' repositories, and these highly-valued features of SHERPA-LEAP are likely to continue, regardless of what becomes of the central services in the longer term.

Acknowledgement

Grateful thanks are due to those colleagues from the SHERPA-LEAP partnership who contributed opinion and insight to this retrospective appraisal of the consortium and its achievement.

References

Crow, R. (2002) "The case for institutional repositories: a SPARC position paper"
Available http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf

Moyle, M. (2006) "ShibboLEAP: seven libraries and a LEAP of faith", Ariadne issue 48
Available <http://www.ariadne.ac.uk/issue48/moyle>

Proudfoot, R. (2005) "The White Rose Consortium Eprints Repository: creating a shared institutional repository for the Universities of Leeds, Sheffield and York", ALISS Quarterly Vol 1 No 1, pp. 19-23

Rumsey, S. (2005) "The institutional repository at LSE: plotting our course", ALISS Quarterly Vol 1 No 1, pp. 15-18

Notes

¹ SHERPA-LEAP Web site: <http://www.sherpa-leap.ac.uk>

² SHERPA (Securing a Hybrid Environment for Research Preservation and Access) Web site: <http://www.sherpa.ac.uk>

³ SHERPA-LEAP Conference on Open Access to Research, 13 June 2005. Web site: <http://www.ucl.ac.uk/Library/scholarly-communication/londonad05.shtml>

⁴ SHERPA DP: Creating a Persistent Preservation Environment for Institutional Repositories
Web Site: <http://ahds.ac.uk/about/projects/sherpa-dp/>

⁵ UK Institutional Repository Search Service, Draft Project Proposal:
<http://www.intute.ac.uk/projects/uk%20ir%20search%20project%20proposal%20v5.4web.pdf>

Authors

Martin Moyle is Digital Asset Manager, UCL Library Services, and is SHERPA-LEAP Project Manager. Email: m.moyle@ucl.ac.uk.

Rebecca Stockley is Project Officer, SHERPA-LEAP, UCL Library Services.

Suzanne Tonkin is Site Librarian, Faculty of the Built Environment, UCL.