JISC

JISC Final Report

Project Information			
Project Identifier	*To be completed by JISC*		
Project Title	Sustainable PReservation Using Community Engagement (SPRUCE)		
Project Hashtag	#sprucedp		
Start Date	1/10/2011	**End Date**	29/11/2013
Lead Institution	University of Leeds		
Project Director	Bo Middleton		
Project Manager	Paul Wheatley		
Contact email	m.m.middleton@leeds.ac.uk		
Partner Institutions	University of Leeds, British Library, Digital Preservation Coalition, London School of Economics, and Open Planets Foundation		
Project Web URL	http://wiki.opf-labs.org/display/SPR		
Programme Name	Infrastructure Portfolio grant funding 12/11		
Programme Manager	Neil Grindley		

Document Information			
Author(s)	Paul Wheatley		
Project Role(s)	Project Manager		
Date	29/11/2013	**Filename**	SPRUCE Project Final Report v1.0
URL			
Access	This report is for general dissemination		

Document History		
Version	**Date**	**Comments**
1	29/11/2013	Draft for approval by JISC

Table of Contents

1   Acknowledgements

2   Project Summary

The stated aims of the SPRUCE Project were – to create a digital preservation business case, to deliver comprehensive support, and to widen the pool of appropriate expertise by applying a community driven approach to digital preservation.

The SPRUCE Project has engaged with the grass-roots digital preservation community in order to deliver support to HE organisations and beyond in preserving their digital assets. By putting those at the forefront of digital preservation activity in the driving seat, SPRUCE has been able to target its support where it is most needed. It has assisted over 70 different organisations as part of its face to face Mash-up and Hackathon events and delivered support via a host of crowd sourced and collaborative online initiatives. A key output from the project is a comprehensive Digital Preservation Business Case Toolkit which provides practitioners and middle managers with the tools they need to secure funding for their preservation activities. Innovative collaborative approaches have been developed and documented, with the most successful, such as COPTR, acting as demonstrators for more efficient and effective ways of delivering support to the digital preservation community.

3   Project Details

3.1   Project Outputs and Outcomes

Output / Outcome Type (e.g. report, publication, software, knowledge built)	Brief Description and URLs (where applicable)
SPRUCE Mashups	SPRUCE delivered 3 mash-up events and 1 hackathon, providing hands on support and knowledge sharing for practitioners and developers working to preserve their digital assets. http://wiki.opf-labs.org/pages/viewpage.action?pageId=13041673
Digital Preservation and Data Curation Requirements and Solutions	A comprehensive collection of current digital preservation practice detailing over 150 preservation challenges. This acts as a growing reference resource for those tackling digital preservation challenges as well as providing an indication of the needs of practitioners in guiding the focus of future preservation developments. http://bit.ly/spruce-results

Digital Preservation Business Case Toolkit	A comprehensive toolkit to help practitioners and middle managers build business cases to fund digital preservation activities. http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit
Getting Started With Hands on Digital Preservation	Guides to the first steps practitioners can take in preserving their digital assets. http://wiki.opf-labs.org/display/SPR/SPRUCE+Project+Digital+Preservation+Results+and+Outputs
The Community Owned digital Preservation Tool Registry (COPTR)	COPTR describes tools useful for preserving digital information for the long term. COPTR is also an initiative to collate the knowledge of the digital preservation community on preservation tools in one place, and address the multitude of small and outdated existing tool registries. http://coptr.digipres.org/
FITS and C3PO preservation tool enhancements	SPRUCE identified a key practitioner need for better characterisation tools and then targeted small development projects at enhancing the FITS and C3PO tools and developing the processes and infrastructure to make these useful but outdated tools more easily maintainable by the community. http://wiki.opf-labs.org/display/SPR/SPRUCE+Project+Digital+Preservation+Results+and+Outputs
How to run your own digital preservation mash-up	A SPRUCE style mash-up is a 3 day workshop for around 30 practitioners and developers that provides a way of sharing digital preservation expertise, identifying and understanding preservation challenges and developing preservation solution to those challenges. This successful event format has been documented to enable others to run events of this kind. http://wiki.opf-labs.org/pages/viewpage.action?pageId=13041673
Collaborate with the digital preservation community	A broad selection of community or crowd sourced initiatives that SPRUCE collaborated in, that provide both support to digital preservation practitioners and mechanisms to encourage community working. http://wiki.opf-labs.org/display/SPR/Collaborate+with+the+digital+preservation+community
SPRUCE Publications	Formal SPRUCE publications that were presented at conferences, including 3 posters and 2 papers. http://wiki.opf-labs.org/display/SPR/SPRUCE+Publications

3.2   Project methodology

SPRUCE has applied a community oriented approach to delivering support for digital preservation. This approach is based upon a number of key principles:

- Capturing and then targeting the needs of the digital preservation practitioner community

- Sharing data, requirements, and outcomes in a completely open way to facilitate knowledge sharing and re-use

- Building connections and communication across the community in order to maximise awareness of existing helpful solutions and resources and minimising wasteful duplication and blinkered development

- Developing in increments with user/practitioner input

- See also the SPRUCE Mash-up Manifesto[1]

A core part of meeting SPRUCE's aims has been delivered through the use of agile events, including Mash-ups and Hackathons. The events brought together practitioners (who contribute digital data and preservation challenges) and developers (who apply tools to solve the practitioners' challenges). They supported expert attendees in expanding their understanding and tackling complex challenges, and helped staff from organizations taking their first steps in digital preservation activity. These events provided main focus of the SPRUCE Project and enabled the targeting of additional supporting activities onto clearly identified practitioner and user needs. Requirements, approaches, software tools and other information gathered during the events proved invaluable in developing subsequent SPRUCE outputs.

In addition to the face to face events, SPRUCE provided a variety of contributions to an array of online collaborative initiatives that aimed to develop better supporting resources for preservation practitioners while encouraging community and collaborative working. A number of these initiatives were quite experimental by their nature, including techniques such as crowd sourcing and the delivery of collaborative events via online tools. Results were (as expected) mixed. Some were unsuccessful and others were both ground breaking and delivered strong results (such as COPTR). An evaluation of these initiatives can be found in SPRUCE's recent iPRES paper[2].

SPRUCE delivered the Digital Preservation Business Case Toolkit (DPBCT) which built on a wealth of results from a series of exercises conducted with practitioners during the mash-up events, as well as lengthy research and literature search. Rather than writing up this work in typical fashion and then circulating the report for comment, SPRUCE organised a 3 day sprint event in which 13 people co-authored the toolkit on a wiki. This innovative approach enabled the maximum input and refinement of the toolkit by the informed experts at the sprint.

The key aims for the project were to support digital preservation practitioners and organisations in preserving their digital assets, develop the digital preservation community with the aim of helping it to help itself (during and beyond the project) and to support sustainability by helping practitioners and middle managers make the case for their work to be funded. These objectives did not change throughout the project, but the range of activities undertaken during the project expanded as opportunities became available. For example, the high impact COPTR initiative was implemented in a relatively short time period at the end of the project despite not being in the original project plan.

---

[1] http://wiki.opf-labs.org/display/SPR/The+SPRUCE+Mashup+Manifesto
[2] "Supporting practical preservation work and making it sustainable with SPRUCE" http://purl.pt/24107/1/iPres2013_PDF/Supporting%20practical%20preservation%20work%20and%20making%20it%20sustainable%20with%20SPRUCE.pdf

3.3   Lessons Learned

- Evaluation activities were built into many of the activities that were undertaken with user led activities that were typically evaluated or trialled by practitioners as they were developed. Specific evaluation activities included:

    o Anonymous event evaluation forms revealed consistently high levels of attendee satisfaction, with extremely positive feedback in almost all cases (see quotes here, http://wiki.opf-labs.org/pages/viewpage.action?pageId=13041673).

    o Independent review by Charles Beagrie LTD of the DPBCT, commissioned by the DPC, stated "Overall an extremely useful resource and well written – I think it will be valuable for the community and support anyone starting out thinking about a business case for digital preservation."

    o Evaluation of the technical developments and online collaboration initiatives is discussed in detail in the SPRUCE iPRES paper, http://purl.pt/24107/1/iPres2013_PDF/Supporting%20practical%20preservation%20work%20and%20making%20it%20sustainable%20with%20SPRUCE.pdf.

    o SPRUCE achievements were described and then reflected upon in the final SPRUCE event "Fund it, Solve it, Keep it (with SPRUCE)" at which Neil Grindley stated during the final panel session, in the context of JISC and digital preservation: "SPRUCE is one of the best things we've done for many years"

- User testing and experiences in applying SPRUCE techniques and approaches are openly documented here (http://wiki.opf-labs.org/display/REQ/Digital+Preservation+and+Data+Curation+Requirements+and+Solutions). This includes content from over 100 practitioners from over 70 different organisations.

- An independent internal report commissioned by JISC to assess the state of play in digital preservation highlighted two projects as the highest profile activities in the field: SCAPE and SPRUCE.

- 12 SPRUCE Awards to practitioners and organisations who had participated in SPRUCE events enabled event work to be taken further and for SPRUCE outputs, such as the DPBCT, to be trialled.

3.4   Immediate Impact

- Digital preservation support has been provided to over 100 practitioners in over 70 organisations (including events run by the previous AQuA project and related hackathons).

- Impact in the wider community on the web and in the literature has been significant and is detailed here (http://wiki.opf-labs.org/display/SPR/SPRUCE+on+the+web).

- COPTR received strong coverage at the Aligning National Approaches to Digital Preservation Conference II, and was touted as a model for collaborative working in the digital preservation field.

4   Conclusions and Recommendations

A number of SPRUCE blog posts (http://openplanetsfoundation.org/blogs/paul) and presentations have highlighted the challenges of communication and coordination, and what goes wrong when there is inadequate support for these mechanisms that are essential to a healthy community[3]. Duplication and the waste of precious resources are particularly concerning outcomes.

Through its focus on community and collaborative solutions, SPRUCE has made some valuable contributions to the communication required to break away from these negative outcomes. At the lowest level this may simply involve connecting community members with relevant contacts based on an awareness of activity right across the community. For example connecting a user experiencing a particular preservation challenge to an appropriate tool they weren't aware of; making a software developer aware of sources of feedback published elsewhere on some of their code; joining up developers or projects with common aims; heading off new developments, where existing solutions already exist. Connections of these kinds can be important but low key, although they can establish the foundations for far greater partnerships.

The COPTR initiative demonstrated that with limited effort and resourcing it is possible to tackle some of the biggest collaborative failures in this community, and bring organisations together so they are pooling their effort and expertise instead of actively competing with each other. It is the approach, connections and focus on a collaborative solution that made COPTR a success.

SPRUCE argues that there is a case for a dedicated "digital preservation community manager". SPRUCE has experimented with playing this role and has shown how valuable it is in coordinating activities across the community and in different projects/initiatives. But, as is typical in digital preservation, the role has been funded by a project with a finite lifespan. Ideally this role therefore needs to be adopted by a more sustainable, long term organization such as the OPF, the DPC, or perhaps the ANADP initiative.

Key recommendations:

- Practitioners and developers working at the forefront of digital preservation must be adequately supported if digital assets are to be secured.

- A Community Manager responsible for aligning international digital preservation activities would reduce waste and duplication and lead to higher quality results.

- Better support is needed for small organisations working on digital preservation, who are struggling with inadequate guidance and a lack of products/solutions suitable for them.

- New digital preservation developments must build on existing resources rather than reinventing the wheel.

- Outputs from new projects must be made more sustainable. Key lessons learned are outlined in the SPRUCE poster here (http://wiki.opf-labs.org/download/attachments/35127416/ANADPII+new+approach+poster+07.pdf?version=1&modificationDate=1385033917000).

- This community depends on project funding, but many projects result in outputs that do not embody long term preservation principles and quickly die. Funders need to encourage better development practices, and place greater emphasis on sustainability.

---

[3] Bacon, J, The Art of Community, O'Reilly, http://www.artofcommunityonline.org

5 Implications for the future

Future development recommendations building on SPRUCE's work:

- Many community created digital preservation software tools are out of date, sub standard, difficult to use, difficult to install or generally of a poor quality[4]. SPRUCE has piloted an approach via SPRUCE Awards focused on FITS and C3PO for the development of test infrastructure, a development process, and associated governance for preservation software tools with the aim of making them community maintainable. This should be rolled out further to other key tools on which significant digital preservation products depend

- Further development of supporting resources is required to enable practitioners to better fund and implement their digital preservation

    - Further collaborative resource building following the COPTR model

    - Better guidance documentation for those starting off in digital preservation and those working with minimal resources

    - Continued development of the DPBCT, specifically developing a benefits framework that realises potential benefits into evidence based hard benefits

- Ongoing support is required to continue to build the practitioner community that SPRUCE has worked with. This community requires coordination and help to operate effectively in preserving digital assets.

Sustainability:

- SPRUCE has incorporated considerations for sustainability from the beginning of the project.

- By engaging closely with users and practitioners SPRUCE results have been received well by these users and so have a greater chance of survival with this engaged community.

    - The best SPRUCE results have been relocated to the SPRUCE partners DPC and OPF for maintenance and further development (while retaining SPRUCE and JISC attribution). Web based resources are on the respective organisation's wikis.

- The SPRUCE website has been redesigned to emphasise the main SPRUCE outputs.

- All outputs have been made available under CC-BY-SA or Apache2 licenses making re-use easier.

- Governance groups have been established for key outputs including the FITS toolset and the COPTR initiative. These groups will meet every 2 months, beginning in January 2014.

- SPRUCE approaches (Mash-up format, Mash-up manifesto, community approaches, etc) have been carefully documented and widely disseminated so that others can reproduce them.

6 References

See: http://wiki.opf-labs.org/display/SPR/SPRUCE+on+the+web

---

[4] Based on feedback at ANADPII Action Session 1 - Building Community Owned Resources for Digital Preservation: http://educopia.org/events/ANADPII/program