Open
Planets
Foundation

OPEN PLANETS FOUNDATION
ANNUAL ACTIVITY REPORT

2010-2011

Contents

Chairman's Statement

In
its first year,
the OPF has already
become recognised as
an important feature
of the digital
preservation
landscape.

The Foundation has made substantial progress towards establishing a sustainable framework. The membership has grown well beyond the initial complement. The OPF has energised the exchange between practioners and developers through hackathons and brought this approach to new communities in the JISC-funded AQUA project. Augmented by resources from both charter and affiliate members, the small OPF team has been able to set up a best-of-breed environment for collaboration, communication, software development, testing, and deployment. It has been able to thoroughly review the extensive outputs of the Planets project, and take key steps to refactor and simplify them. The results of this effort will be high quality supportable software components that directly support the needs of OPF members.

It is clear from the first year that digital preservation is an emerging practice in member organisations and beyond. This means that practioners benefit even more by engaging with a wider community.

I know that the work that we have done with the OPF has raised our awareness of alternate approaches, helped us to develop new skills, and has enabled us to address our digital preservation challenges in new ways with new tools. Our return on working in collaboration through the OPF multiplies our investment in membership.

The OPF looks forward to a strong and successful second year with new projects and new initiatives responding to the needs and interests of its members. Working together, we can make a lasting impact on the practice of digital preservation.

The Open Planets Foundation (OPF) was established with a clear goal: to help its members meet the challenge of ensuring access to growing digital collections both today and in the future. Its members recognise that they can neither afford to meet these challenges alone, nor acquire all of the needed skills and knowledge without working together. Furthermore, they recognise that substantial R&D investment has advanced our understanding of how to address the challenges, but does not provide a sustainable framework to build and mature solutions.

In its first year, the OPF has already become recognised as an important feature of the digital preservation landscape. In part this is due to the OPF Board Members, who continue to contribute substantial amounts of their time and energy to provide essential guidance, oversight and support to the fledgling Foundation.

Adam Farquhar, Chairman

About

Background

The real challenge of digital preservation is long-term access to digital content. Since the beginning of the digital era, formats have become obsolete and content has become inaccessible. The rapid innovation that drives the digital world limits the lifespan of software. Software becomes obsolete from a technical perspective or unsupportable from a commercial perspective. Therefore, preserving just the bits that make up digital content is not enough. It is essential to take active steps to ensure that future generations will have meaningful, contextual access to the content that is locked up in today's digital material.

The Open Planets Foundation (OPF) was founded on 1st June 2010. It is a not-for-profit organisation inspired by and building on the pioneering work of the EU co-funded Planets project. Planets successfully developed digital preservation best practices, prototype tools and demonstrators that were validated by the cultural heritage sector as a solid basis for a solution that addresses the challenge and accessing digital content for the long-term.

The OPF is further developing this suite of tools and practices and maturing them into a comprehensive open source range of software offerings. The OPF provides digital preservation best practice guidance and tools that cover preservation planning, costing models, format identification and characterisation.

The founders of the OPF foresee that making tools available under an open source licence where and when possible will stimulate the adoption of digital preservation practices; wider adoption will increase sustainability due to the size of the community of immediate stakeholders.

> The OPF's mission is to ensure that its members around the world are able to meet their digital preservation challenges with a best practice strategy and solutions that are widely adopted and actively developed by the cultural heritage sector and beyond.

The OPF plays an active role in further developing the open source components for its suite by means of continuous involvement with the community through requirements management, architectural guidance, configuration management support and development infrastructure support.

The OPF products and services:
- Meet digital preservation needs of archives and libraries at national scale
- Are focused on practical solutions
- Provide functional products and services, which are open source, extensible and non-proprietary
- Are positioned to deliver value for money
- Have committed international support from existing partners and growing interest from libraries, archives and industry

Objectives

" The founders of the OPF believe that in order to ensure sustainability and economy of scale of digital preservation there must be support by value-added services and digital preservation expert tools. The OPF founders also believe that this will require at least 80% of national heritage organisations to adopt a digital preservation practice by 2015.

The community needs decision practices and tools that support the process of determining when, what and how to safeguard this long term access to digital content. They need to be alerted when digital content is endangered, decide what needs to be preserved and test how this can be done. Content holders traditionally are very well organised to decide about "what" to preserve, but it is the missing "when" and "how" where OPF practises and tooling will prove their value. "
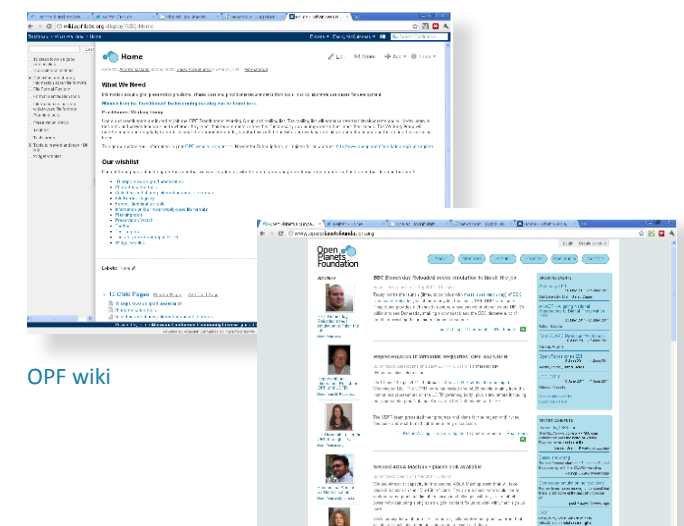
Highlights

Practitoner Community

The community established by the OPF brings together practitioners and developers to build practical solutions to address the challenges of long-term access to digital content.

We work closely with our members to identify and understand their digital preservation problems and ascertain their requirements for tools and services. We have set up templates to help practitioners to further define and focus these requirements to create concrete use cases for development work. By documenting this information in the OPF wiki (http://wiki.opf-labs.org/display/REQ/Home), members with similar problems can share ideas and experiences, avoid duplication of effort and reduce costs.

The OPF hosts hackathons throughout the year which give the opportunity for practitioners and developers to meet face-to-face and work directly with one another to develop practical solutions for specific digital preservation issues. The resulting tools are available to the OPF members to test and provide feedback for further development work.

Developer Collaboratory

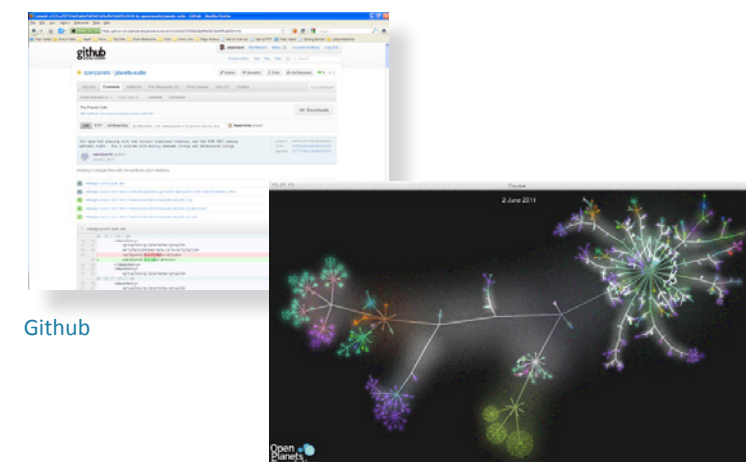Guided by the OPF Technical and Architecture Advisory Board, a best-of-breed development infrastructure has been set up to support the digital preservation technical community. Distributed source control repositories from GitHub are combined with the powerful Atlassian tool set, including the Confluence wiki, the JIRA issue tracker, and the Bamboo integrated build server. This infrastructure has been adopted by the AQuA and SCAPE projects (see page 5), as well as being used by the OPF to manage our own code and resources.

The original Planets software has been revised and refactored in order to make it more sustainable, as part of an on-going process of reviewing research outputs to generate production tools. We have made the original Planets code smaller, more modular, and easier to deploy. We continually simplify the components of our software suite, focussing it more keenly on our members' needs.

OPF wiki

OPF website

Website

The OPF launched its 2.0 website (http://www.openplanetsfoundation.org/) in November 2010. It offers information about the OPF and its members, the projects it is involved in, and publicises relevant conferences and exhibitions. In addition, it provides a platform for digital preservation practitioners and developers to post blogs on the challenges they face, tool developments, and to invite feedback from the community.

Digital preservation experts from around the world have contributed blogs and have commented on a wide-range of topics such as bit rot, migration and normalisation, emulation, format identification, format registry eco-systems, JPEG2000, and digital preservation market research.

Github

Visulisation of first year's development actvity

Highlights

Publications and Papers

A New Registry for Digital Preservation: Outline Proposal
Bill Roberts, September 2010
http://www.openplanetsfoundation.org/new-registry-digital-preservation-outline-proposal

Digital Preservation and Long Term Access Functionality
Cambridge Judge Business School, December 2010
http://www.openplanetsfoundation.org/blogs/2010-12-16-market-research-cambridge-judge-business-school-part-2

A New Registry for Digital Preservation: Conceptual Overview
Bill Roberts, January 2011
http://openplanetsfoundation.org/sites/default/files/OPF_A_New_Registry_Conceptual_Overview_v1%201.pdf

Planets Project: Preserving Delicate Digital Materials
Multimedia Information & Technology, Vol 37, Number 1
Bram van der Werf and Clive Billenness, February 2011
http://www.cilip.org.uk/get-involved/special-interest-groups/multimedia/journal/pages/journalcurrentissue.aspx

Introducing the Open Planets Foundation
ARC Magazine
Bram van der Werf, June 2011
http://www.archives.org.uk/images/documents/arc/arc%20june%202011.pdf

Conferences and Events

Oracle Preservation and Archiving Special Interest Group (PASIG) Meeting Madrid, July 2010
Presenting OPF: a pragmatic community supporting practical tools and practices

iPres Vienna, Spetember 2010
Introducing OPF to the global digital preservation community

JISC Dev8D Developer Days London, February 2011
Outreach to the UK Higher Education developers community to address long term access challenges

Oracle Preservation and Archiving Special Interest Group (PASIG) Meeting London, April 2011
An overview of digital preservation over the last years and development of architectures

AQuA Hackathon Leeds, April 2011
Practitioners and developers working together to solve practical digital preservation problems

UDFR Stakeholder meeting Washington, April 2011
Requirements and technical involvement with a global format registry development

EU meeting 'The Future of the Past' Luxembourg, May 2011
Meeting to establish priorities for research in digital preservation for the next Framework Programme (FP8)

Oracle Preservation and Archiving Special Interest Group (PASIG) Meeting California, May 2011
Developing digital preservation architectures and future role of PASIG

Aligning National Approaches to Digital Preservation
Estonia, May 2011
International collaboration, discussions on standards, certification and accreditation

JHOVE2 Workshop Rome, May 2011
Direct involvement with JHOVE2 characterisation tool and its future development and maintenance

OPF Hackathon Amsterdam, November 2010
The OPF hosted a 3-day event for both practitioners and developers to work together. The first day provided an overview of the OPF's requirement and development process and the delegates worked together in groups to define requirements for a selection of OPF tools. Days 2 and 3 were spent defining the requirements into solid use cases and collaborating and hacking the tools.

Delegates commented:
"Mixing developers with practitioners gave me a very useful insight in to how tools should be developed."

OPF Hackathon at the International Institute of Social History, Amsterdam

Projects

Participating in projects helps the OPF to expand its knowledge in digital preservation R&D and build a wider network to benefits its members. The OPF will also play a role in influencing future digital preservation projects. The OPF is currently a consortium partner in SCAPE and AQuA.

The **SCA**lable **P**reservation **E**nvironments (SCAPE) project is co-funded by the European Commission under FP7 and led by the Austrian Institute of Technology. The OPF joins experts from memory instititions, data centres, research labs, universities, and industrial firms in researching and developing scalable preservation systems.

SCAPE will develop infrastructure and tools for scalable preservation actions; by providing a framework for automated, quality-assured preservation workflows and by integrating these components with a policy-based preservation planning and watch system. These concrete project results will be validated within three large-scale Testbeds from diverse application areas: Digital Repositories from the library community, Web Content from the web archiving community, and Research Data Sets from the scientific community. Each Testbed has been selected because it highlights unique challenges.

SCAPE will develop scalable services for planning and execution of institutional preservation strategies on an open source platform that orchestrates semi-automated workflows for large-scale, heterogeneous collections of complex digital objects.

The OPF has three key roles in the SCAPE project:

- It leads the SCAPE training and sustainability activities
- It provides technical integration of Planets and SCAPE technology
- It contributes to outreach and dissemination activities working with its user community to enhance the impact of SCAPE.

http://www.scape-project.eu/

Automating **Qu**ality **A**ssurance (AQuA) is a JISC-funded collaboration between the University of Leeds, University of York, The British Library and the Open Planets Foundation aiming to address quality issues in digitised content by applying tools to automate the validation process.

AQuA bought together content holders and developers at two mashup events. The content holders presented their problematic digitised content and articulated requirements for their validation. The developers were paired with the content holders to further discuss the issues and apply tools to automate the detection and identification of preservation and quality issues in the collection.

During the three-day mashup events, content holders and developers documented their activities in the OPF wiki:

- Collections: an overview of the digitised content and samples from the content holders
- Issues: a detailed account of actual or potential preservation issues in the collection
- Solutions: specific solutions to the issues addressed with links to the code

The OPF's role was to provide administrative support and technical resources through its members. It also helped to sustain the outputs from the events.

Since the events, one prototype has been turned into a fledgling JHOVE2 format module. Two other prototypes are being integrated into The British Library's open source content profiling and characterisation toolset.

AQuA attendees continue to contribute to the wiki even after the events and the resources on the wiki have received interest from the wider digital preservation community.

http://wiki.opf-labs.org/display/AQuA/Home

Board of Directors

The OPF is governed by its Board of Directors who are elected by its members at the Annual General Meeting. The Board comprises senior managers from its member organisations who have responsibility for or a special interest in digital preservation. The active Board meets monthly to review progress against the business plan and provide strategic direction.

Adam Farquhar Chairman
The British Library

Bjarne Andersen Director
The State and University Library, Denmark

Neil Grindley Director
Joint Information Systems Committee (JISC)

Hans Jansen Director
The National Library of The Netherlands

Max Kaiser Director
The Austrian National Library

Ross King Director
with Special Responsibility for Financial Matters
The Austrian Institute of Technology

Jacqueline Slats Director
The National Archives of The Netherlands

Technical and Architecture Advisory Board

To support the Board of Directors, the OPF has established the Technical and Architecture Advisory Board. Its key objectives are to make practical, technical decisions, providing direction on new technologies to support. The Technical Board comprises senior developers and architects from member organisations. They are involved in engaging the wider developer community and actively participate in decision-making processes to support the development of production quality software. The hold monthly calls to monitor development progress and ensure the OPF solutions are deployable and applicable for use in libraries and archives.

Asger Askov Blekinge
The State and University Library, Denmark

Andrew Jackson
The British Library

Andrew Lindley
The Austrian Institute of Technology

David Tarrant
University of Southampton

Carl Wilson
The British Library

Staff

Bram van der Werf
Executive Director

Rebecca McGuinness
Membership and Communications Manager
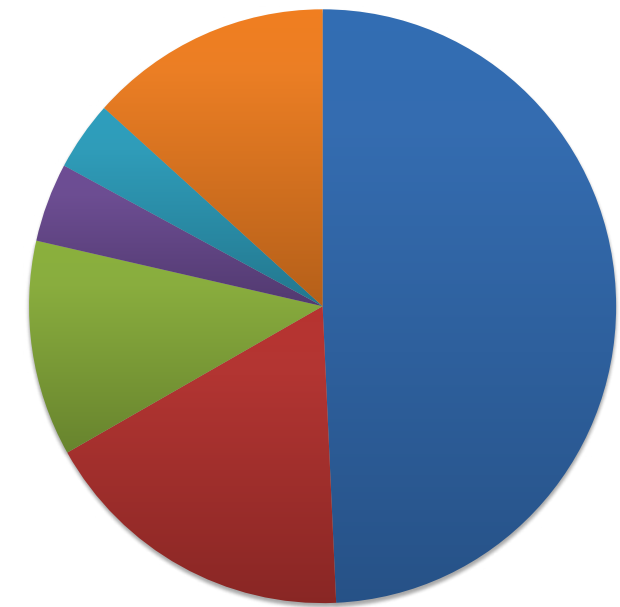
David Tarrant
Senior Developer

Technical Resources

The OPF has agreements with its affiliate members, the Vienna University of Technology and HATII at the University of Glasgow, to continue to host an R&D version of the Plato planning tool, and test copora, respectively.

In addition to its employees, the OPF has secured additional technical resources from partners, and its affiliate member organisations (as part of their in-kind contributions) to support its members' needs.

Year One Expenditure

- Staff salaries, national insurance and pension contributions £82, 181
- Technical resources: hire of digital preservation specialists to support members' needs £29,247
- Outreach activities: member events, visits to current and potential members, profile raising in the digital preservation community, travel sponsorship for members £19,306
- Community building and communications: Development and maintenance of the OPF 2.0 website, wiki and development collaboratory £7,186
- Professional services and financial costs: payroll, management accounts, bank charges and foreign exchange fees £6,315
- Start up costs: recruitment costs, solicitor's fees, insurance, office equipment and staff training £22,291
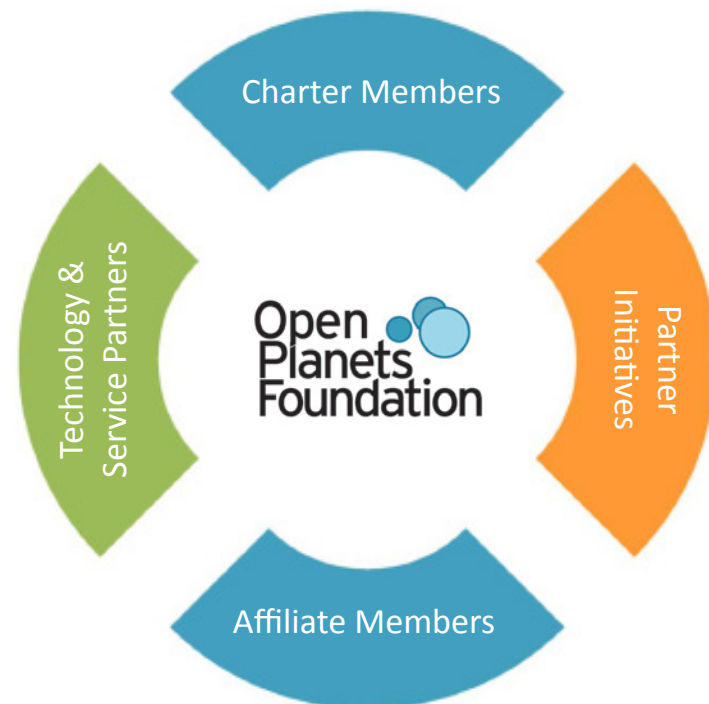
Total **£116,526**

OPF Digital Preservation Community

Charter members are organisations with a mission and a responsibility to manage substantial digital content for the long term. These include major memory institutions (archives, libraries and museums) in the public or private sector. Most of these institutes have the need, the infrastructure and resources to deploy OPF practices and tools. Charter members are eligible to nominate representatives from their organisation to sit on the OPF Board.

Affiliate members are typically academic organisations with a strong research and development focus who want to be part of the scientific network within OPF, developing digital preservation technologies and tools. They include higher education institutions that want to work closely with the OPF experts and its practitioner community to establish the digital preservation higher education curriculum of the future.

Partnerships
The OPF work in partnership with technology and services providers and also with existing digital preservation organisations and coalitions.

Member Benefits

By joining the OPF, members become a part of an interntional effort to develop digital preservation best practices, tools and technologies. Members benefit from:

- Support for tools and services
- Peer-reviews of their organisation's systems and tool developement
- Reports and surveys being carried out on their behalf into specific areas of digital preservation
- Priority booking and free attendance at OPF events
- A free API licence key for applicable services (format services)
- Exclusive access to results and documentation from OPF events

Members are also in a privileged position to:

- Influence the OPF strategy
- Influence OPF event topics
- Shape the tools and services roadmap by providing requirements and use cases
- Test and validate prototype tools and provide feedback
- Join an active community that fosters collaboration between preservation practitioners and technical experts
- Participate in meaningful discussions with digital preservation experts from around the world

Our current members are national archives, libraries, higher education and research organisations. They share a commitment to ensuring effective long-term access to digital material.

Charter Members

The Austrian Institute of Technology

The Austrian National Library

Goportis

The National Library of The Netherlands

JISC

Microsoft Research

The National Archives of The Netherlands

The Danish National Archives

The National Library of Wales

The British Library

The National Library of France

The Royal Library, Denmark

The State and University Library, Denmark

Stanford University Libraries & Academic Information Resources

Affiliate Members

HATII at the University of Glasgow

University of Southampton

Bodleian Libraries, University of Oxford

Cambridge University Library

Vienna University of Technology

King's College London (Centre for e-Research)

University of Portsmouth

Digital Curation Centre

OPF Timeline

The OPF is founded with 10 charter members

OPF website launched

Stanford University Libraries join as a charter member

June 2010

July 2010

National Library of Wales joins as a charter member

Danish National Archives join as a charter member

Oct 2010

OPF-labs set up with JIRA suite

OPF Technical & Architecture Advisory Board formed

JISC joins as a charter member

Feb 2011

AQuA project begins

OPF website 2.0 launched

OPF Hackathon, Amsterdam

OPF Github repository set up

Nov 2010

SCAPE project begins

March 2011

50th blog post on OPF website

OPF AGM held in The Hague

National Library of France joins as a charter member

June 2011

Large-scale refactoring of the Planets suite codebase begins

April 2011

JISC and OPF establish group affiliate membership for UK HE institutions

Vienna University of Technology joins as first affiliate member

2012 Preview

Over the course of our first year we have established a strong technical community. We are confident that this community, supported by our Technical and Architecture Advisory Board, is well positioned to produce and maintain tools and best practices in the OPF open source collaboratory.

Throughout the first year it also became apparent that the requirement process will be our main focus for year two. We need to increase our capacity to connect theory from R&D prototypes with the reality of production practice.

This requires active involvement from practitioners to understand the types of tools they need today. Validating prototypes and beta versions of tools and practices will be developed by working closely with practitioners. The OPF will help the, to scope their immediate needs, rather then focusing on more functional and future requirements as the approach in R&D projects.

In order to strengthen our practitioners community, we will host quarterly member events similar to the AQuA project events. Members organisations will be invited to bring examples of their collections with which they are experiencing preservation issues.

During and after these events, developers and technical experts will collaborate with hacking and deployment activities.

These events will be fully documented from initial requirements and use cases through to the hacking and coding, which we will make available in the OPF collaboratory.

In order to meet different needs across our members, each event will have a topic and specific focus, such as database archiving, emulation, bit-rot, preservation planning, and testing (corpora and testbed).

Tools and Services Roadmap 2011

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Collaboratory: the code repository, issue tracking, and code management tools	available	available	available	available
Practitioner wiki: sharing use-cases and solutions to workflows and experiences	available	available	available	available
Developer wiki: building a common understanding of our preservation software and systems	available	available	available	available
Test corpora: shared sets of realistic test files, starting with the Planets Testbed corpora	not available	available	available	available
Planning: the OPF Preservation Planning Tool, taking Planets PLATO into production	prototype	prototype	beta	supported
Fido: a lightweight command-line interface for identification, powered by PRONOM signatures	prototype	prototype	beta	beta
Pricing: making the LIFE3 cost modelling tool available through the OPF	prototype	prototype	prototype	available
Risk analysis tool (RAT): scans collections for known preservation issues and risks, reported via a traffic-light rating system	development	development	prototype	beta
Results evaluation framework (REF): sharing results from executing preservation processes on the Test Corpora	development	development	prototype	beta
Format registry editing tool: a Drupal-based web interface to the PRONOM data, allowing collaborative editing	prototype	prototype	in REF	in REF