OPENPRESERVATIONFOUNDATION

ANNUAL REPORT
2014-2016

VISION
Shared solutions for effective and efficient digital preservation

MISSION
The Open Preservation Foundation sustains technology and knowledge for the long-term management of digital cultural heritage. We provide our members with reliable solutions to the challenges of digital preservation

CONTENTS

DR. ROSS KING, CHAIR
AIT AUSTRIAN INSTITUTE OF TECHNOLOGY

My name is Ross King and I am a Senior Scientist at the Austrian Institute of Technology (AIT) in Vienna. AIT is one of the founding members of the Open Preservation Foundation, and I have served as Chairman of the Board of Directors since February 2013.

The past two years of the Open Preservation Foundation have seen many milestones and highlights. Towards the end of 2014 (and beyond), the OPF completed its contributions to the SCAPE project by providing a sustained archive for the project website, deliverables, and software artefacts. This was one of the primary reasons for establishing the Foundation – to provide a sustainability channel for the results of European projects. Since the de-prioritisation of digital preservation in the H2020 program, this aspect of the Foundation has been undervalued, but I hope to see this part of our mission become significant again in the years ahead.

Another important step was the re-branding of the organisation and launching a new web site. The name Open Planets Foundation reflected our roots, but failed to convey our core mission. Making the small but significant change to Open Preservation Foundation was an important step that still retained our branded initials "OPF."

The veraPDF PREFORMA project has been a crucial technical success by providing the definitive implementation of a PDF/A validation tool, which was released in version 0.14 earlier this year. Such a tool is eagerly awaited by organisations around the world and will certainly have a huge impact in the preservation community and beyond. The project also supports important on-going collaboration with the PDF Association.

Since February 2015, the Foundation has taken over stewardship of the JHOVE validation and characterisation codebase. JHOVE has been consistently identified as one of the most important open source tools used in production by our members and the digital preservation community at large. Ensuring on-going development and support of such tools is a key aspect of the Foundation's mission.

For this reason, the OPF has created a new type of membership, the "Software Supporter," designed for those institutions whose primary concern is keeping open source tools available for production and who can make use of yearly IT budgets to support these tools. The Software Supporter option is part of our on-going efforts to provide flexible membership models with clear benefits to a wider variety of potential member institutions.

Of course, the past two years have also been somewhat tumultuous with regard to our Executive Director, first with the resignation of our founding Director Bram van der Werf in September 2014 and then also his successor, Ed Fay, a year and a half later. We were very fortunate to be able to recruit Joachim Jung who has taken over as Executive Director in December 2015. Since then Joachim has been working tirelessly to strengthen the membership, strategic plan, and community network of the OPF.

Joachim's experience and communication skills are already serving the organisation well. For example, since taking office, he has worked on securing cooperation with related national organisations such as the NCDD in the Netherlands.

I look forward to working with Joachim, Carl, Becky, and the OPF board in the coming year. Together we will continue to provide knowledge, support, and reliable solutions for the challenges of digital preservation.

TECHNOLOGY

KNOWLEDGE

ADVOCACY & ALLIANCES

" This was one of the primary reasons for establishing the Foundation – to provide a sustainability channel for the results of European projects. "

- Dr. Ross King, Chair

STRATEGIC PLAN 2015-2018

A major milestone of the past two years was the launch of the strategic plan 2015-18. This was produced in close consultation with our members from archives, libraries, research and technology organisations. We established three core strategic themes: technology, knowledge and advocacy and alliances. The plan defines the aim of each theme, describes member benefits, and demonstrates the Foundation's support for the digital preservation community.

We surveyed our members in October 2014 to build an evidence base of their needs, and identify shared priorities. Analysis of the results directly informed our annual plan which is circulated to members.

Strategic Plan

Annual Plan

NEW NAME, BRAND, & WEBSITE

In November 2014 we changed our name from Open Planets Foundation to Open Preservation Foundation. The name change reflects the Foundation's core digital preservation aims and activities while retaining our familiar acronym, OPF.

The Foundation was established to sustain the results of PLANETS, an EU-funded digital preservation project which ended in 2010. We now sustain the results of other projects including SCAPE and SPRUCE. Our name change reflects our broader role.

We launched our new brand with a re-designed website, providing a better user experience with improved navigation and functionality. The new design reflects our strategic themes and was tested with a focus group comprising Foundation members and the wider community.

The **Technology** section sets out our open source principles and practices, and showcases our digital preservation products and services. We have added a section describing corpora which are potentially useful for testing and evaluating software.

The **Knowledge** section offers access to our blogs which attract international contributions, including case studies and tool developments, as well as our latest news and community events calendar.

We publish member-only content on the website, including our archive of webinar recordings, the member survey results, materials from our events and hackathons and reports from external conferences.
www.openpreservation.org

" We believe that openness is a fundamental pillar of digital preservation. "

- Joachim Jung, Executive Director

TECHNOLOGY

We provide stewardship for leading open source digital preservation tools in our software portfolio, ensuring long-term sustainability.

In February 2015 we took over stewardship of JHOVE (JSTOR/Harvard Object Validation Environment). This is a practical example of our strategic priority to build confidence in existing tools. JHOVE is an extensible software framework that performs format identification, validation, and characterisation of digital objects. It is a widely used tool in the digital preservation community.

JHOVE is the first product for which we are offering software support. It is open source, but supporters and members can steer maintenance and development activities through the JHOVE Product Board, as well as receive technical support and training. We have concentrated on ease of installation, branding and stability for the 1.12 release. JHOVE 1.14 released in May 2016, contains three new modules: WARC, gzip, PNG and new regression testing tools.

jpylyzer is a validator and feature extractor for JP2 images. JP2 is the still image format defined by Part 1 of the JPEG 2000 image compression standard (ISO/IEC 15444-1).

jpylyzer tells you if a JP2 image really conforms to the format's specifications (validation). It also reports the image's technical characteristics (feature extraction). It was part-funded by the SCAPE Project and is maintained by Johan van der Knijff at the Koninklijke Bibliotheek.

fido (Format Identification for Digital Objects) is a command-line file format identification tool. It returns a PUID (PRONOM Persistent Unique Identifier) which specifically identifies a file format and version. This PUID can be used in conjunction with file format registry software to determine actions needed to successfully preserve a file or how to access it.

Artefactual maintains fido as their affiliate member contribution. fido is included in Archivematica, but is also provided as a command line tool or it can be incorporated into other digital preservation systems.

TECHNOLOGY

SOFTWARE MATURITY MODEL

We are an experienced leader in open source digital preservation software development. We advocate open source principles and follow open source practices to further collaborative, agile development.

Our software maturity model comprises five complementary approaches: quality, visibility, training, open licensing, and community integration.

We use open, public infrastructure to host the tools in our software portfolio and work with our members and the community to establish shared requirements. All of our software is released under open licences to encourage adoption, reuse, and to ensure the contribution of further enhancements back to the community.

Our approach has been adopted by some of the leading digital preservation software tools produced by the research and development community in recent years.

EXPERTISE

We have been commissioned to carry out code reviews for BitCurator, and the EU TIMBUS project. Carl, our Technical lead, also sits on the BitCurator Advisory Board.
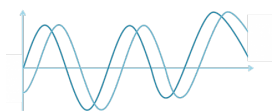
QUALITY

VISIBILITY

TRAINING

OPEN LICENSING

COMMUNITY INTEGRATION

xcorrsound

xcorrsound compares sound waves. It packages three tools to address different digital audio preservation scenarios:
- finding the overlap between two audio files.
- detecting all occurrences of a shorter wav within a larger wav.
- ascertaining whether two audio files are similar by analysing them for correlation.

xcorrsound is maintained by the Statsbiblioteket.

SUSTAINING TECHNOLOGY

The Foundation offers a central, sustainable home for open source digital preservation software projects. We assess the maturity of the software and ensure the code is hosted and discoverable. The sustainability of project results can be funded directly from projects or through new memberships.

STEWARDING TECHNOLOGY

We steward open source software which we consider to be mature and ready for use in production workflows. Software under our stewardship is actively maintained, developed and tested; it is given a safe sustainable home.

We adopt tools for stewardship in response to the needs of our members. Stewardship is funded by membership and software supporter fees.

KNOWLEDGE

We facilitate open best practice knowledge exchange through our webinars, events, blogs and interest groups.

WEBINARS

We host monthly webinars on a range of digital preservation topics with international speakers in the digital preservation community. Our webinars focus on four themes:
- Preservation in practice - digital preservation case studies and experiences
- From the toolbox - exploring digital preservation tools and technology
- Open source with OPF - best practice software development
- Community projects - showcasing current digital preservation projects

Our members receive priority registration and can access the webinar recordings and slides through our website. The live webinars are open to the community, with as many as 75 registrations for a single webinar.

EVENTS

We invite expert speakers to support our events focussing on specific digital preservation topics such as digital forensics and emulation, or concentrate on preservation of a particular format e.g. PDF. Our events have a practical and hands-on approach, where participants can get experience of using digital preservation tools and services.

We support all relevant digital preservation events on our website's events calendar, and welcome contributions from the community. **openpreservation.org/events**

Preserving PDF: identify, validate, repair, Hamburg, 2014

"   With OPF we get direct networking with similar institutions solving similar problems within the area of digital preservation. "

- Bjarne Andersen, Head of IT development, State and University Library

INTEREST GROUPS

We have established interest groups in response to feedback from our members: documents, emulation and archives. Their purpose is to facilitate collaboration on specific digital preservation topics, selected by our members. Each group is led by one of our member organisations. We connect people with common interests so that they can share experiences and expertise. The groups produce practical outputs including requirements for collaborative development, a knowledge base, use cases and blog posts.

The Document Interest Group (DIG) is our longest established group. Participants' efforts have focussed on:

- Capturing PDF related JHOVE errors on our knowledge base.
- Using this information to develop standardised JHOVE errors through the JHOVE GitHub wiki.

The DIG plans to continue with JHOVE enhancements and explore possibilities for collaborative software testing against local collections and publication of results.

The emulation and archive interest groups are in the early stages of formation. We are working on establishing leadership, and identifying goals with Foundation members.

DOCUMENTS

EMULATION

ARCHIVES

SUSTAINING KNOWLEDGE

In addition to software, we also preserve supporting material from digital preservation projects including websites, reports, documentation and training materials. We publicise these results through webinars, blog posts, mailing lists, and social media, and provide links from our website.

ADVOCACY & ALLIANCES

We work with international partner organisations to promote synergies in the development of digital practices and tools and reduce duplication of effort.

NCDD Nationale Coalitie Digitale Duurzaamheid (since 2016)
The NCDD and OPF signed a memorandum of understanding in April 2016 to facilitate co-operation on educational activities and to work together to further digital preservation software sustainability. The MOU also commits the organisations to share plans and support each other's outreach activities.

PDF Association (since 2015)
In January 2015 the PDF Association and the OPF signed an agreement formally recognising their on-going partnership. Olaf Drümmer, Chair of the PDF Association was the keynote speaker at our event 'Preserving PDF: identify, validate, repair', and we collaborate with the PDF Association on the veraPDF project.

IMPACT Centre of Competence (since 2014)
Announced at the 'SUCCEED in Digitisation, Spreading Excellence' conference at the Bibliothèque nationale de France in December 2014, OPF and IMPACT Centre of Competence in Digitisation signed an Agreement of Alliance to strengthen their collaboration and enhance their respective membership benefits. The Agreement brings together expertise in digital preservation and text digitisation.

Digital Preservation Coalition (since 2014)
We signed a memorandum of understanding with the DPC to share knowledge and expertise, deliver joint events, and to support the development of tools and best practices. The Foundation sponsored one of the 2014 Digital Preservation Awards categories.

SPONSORSHIP OF THE DIGITAL PRESERVATION AWARDS 2014

Created in 2004 to raise awareness about digital preservation, the bi-annual Digital Preservation Awards celebrate the achievement of those people and organisations that have made significant and innovative contributions to digital preservation.

The OPF Award for Research and Innovation recognised excellence in practical research and innovation activities. The shortlisted candidates were:

- Jpylyzer by the National Library of the Netherlands and partners
- The SPRUCE Project by The University of Leeds and partners
- bwFLA Functional Long Term Archiving and Access by the University of Freiburg and partners (winner)

We will serve on the judging panel for the upcoming 2016 awards.

RESOURCES

We support a number of community digital preservation initiatives. We host digipres. org which includes COPTR (Community Owned Tools Registry) and the digital preservation Q&A site.

Our knowledge base captures digital preservation datasets, issues and solutions, contributed by event participants and the community. Other project resources preserved on our wiki, include the SCAPE Catalogue of Preservation Policy Elements and the SPRUCE business case examples. We are working to make information easier to discover and access by moving to MediaWiki. We hope you'll make use of and enhance our knowledge base. **wiki.opf-labs.org**

RAISING AWARENESS

The Foundation presents at international conferences raising awareness of the importance of digital preservation. These include: iPRES, PASIG, and Open Repositories. We were part of the organising committee for PASIG 2016 in Prague.

We have been invited to contribute news items to publications including DLib Magazine, and the Information and Records Management Society Bulletin. We are also on the Editorial Board for the Preservation, Digital Technology & Culture journal.

20 May 2016

✉ 694 subscribers

🐦 1,271 followers

 76 collaborators

in 348 members

🏠 16,067 visitors
(since 1 Jan 16)

COMMUNITY SURVEY

We ran an online survey In December 2014 - January 2015 to assess the current state-of-the-art in digital preservation software and practices. The survey explored the adoption of digital preservation approaches and technology. Building upon the OPF member survey this was open to all institutions allowing analysis of trends across the sector.
openpreservation.org/ knowledge/surveys

MEMORY INSTITUTIONS

INDUSTRY
ISO TC 171 SC 2

SERVICE PROVIDERS
REPOSITORY SYSTEMS

veraPDF consortium

Led by the OPF and the PDF Association, the veraPDF consortium is building a purpose-built, open source, file-format validator covering all PDF/A parts and conformance levels. Designed to meet the needs of digital preservationists, veraPDF is supported by the PDF software developer community.

veraPDF is funded by PREFORMA –PREservation FORMAts for culture information/e-archives, a Pre-Commercial Procurement (PCP) project co-funded by the European Commission's FP7-ICT Programme. The project addresses the challenge of implementing standardised file formats to support the long term preservation of digital objects, giving memory institutions control over the acceptance and management of digital content.

The consortium is delivering a definitive conformance checker for PDF/A. To achieve this it is also establishing authoritative ground truth by creating corpora covering all requirements of all versions and levels of the PDF/A specifications. The conformance checker includes a policy checker allowing users to define institutional acceptance criteria beyond the format specifications. A policy profile registry will be launched enabling the sharing of policy profiles and best practice between institutions.

Phase 1 (design) of the project ran from November 2014 to February 2015. We worked closely with memory institutions and the PDF industry gathering requirements for the PDF/A conformance checker and producing functional and technical designs and specifications.

veraPDF is now in Phase 2 (prototyping), which began in April, 2015. The first proof of concept was released in September 2015, our seventh release was made in May 2016 (v 0.14). veraPDF is now capable of full PDF/A 1b validation and supports some validation features for all flavours of PDF/A. The test corpus is growing rapidly and will cover all test cases by the end of 2016. PDF parsing currently uses a PDFBox engine. The consortium is developing a purpose built parser supported by the capability to swap parser implementations. The software has an established Java API with command line and graphical interfaces. A prototype REST API and browser interface has been released but will not be completed before the last quarter of 2016.
verapdf.org

PROJECTS

VERAPDF

2014 - 2016
PCP (EU FP7)
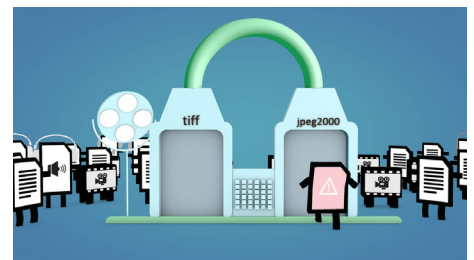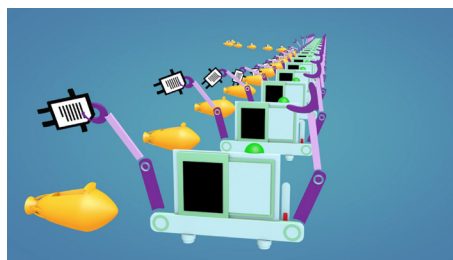5 partners

SCAPE

2011 - 2014
EU FP7
20 partners

30 September 2014 marked the end of the four-year EU FP7 SCAPE Project. The aim of the project was to produce scalable solutions for the preservation of large scale and heterogeneous datasets. Solutions were developed to address real requirements from the content-holding partners in the consortium.

The project results were well-received by the European Commission, gaining an 'excellent' rating in all its annual reviews.

OPF acted as co-technical lead for the project, and led the training and sustainability work packages. We were also involved in the development of characterisation tools, and dissemination.

SCAPE both extended existing preservation tools and services to improve their scalability performance, and developed new tools. The range of open source tools and services are inter-connected and encompass the whole digital preservation lifecycle.

The OPF is responsible for sustaining the project's outputs.
scape-project.eu

Stills from the SCAPE project video

MEMBERS

We have 21 members from 12 countries comprising archives, libraries, research and technology organisations working together to preserve our digital heritage.

We introduced a new membership and software supporter model in June 2015 making participation more affordable to organisations of all sizes. The Foundation is a small not-for-profit organisation and relies on the commitment of, and contributions from our members and supporters to continue our work. To become a member or read about other ways you can support the Foundation visit:
openpreservation.org/join

CHARTER MEMBERS

AIT **AUSTRIAN INSTITUTE OF TECHNOLOGY**

DET KONGELIGE BIBLIOTEK

Preservica Digital Preservation

{BnF Bibliothèque nationale de France

KB Koninklijke Bibliotheek **Nationale bibliotheek van Nederland**

RIGSARKIVET

BRITISH LIBRARY

L N B LATVIJAS NACIONĀLĀ BIBLIOTĒKA

RAHVUSARHIIV THE NATIONAL ARCHIVES OF ESTONIA

GOPORTIS

National Archives of the Netherlands *Ministry of Education, Culture and Science*

STATS**BIBLIOTEKET**

Jisc

Österreichische Nationalbibliothek

THE UNIVERSITY OF ALABAMA | University Libraries

AFFILIATE MEMBERS

UNI FREIBURG

BNP BIBLIOTECA NACIONAL DE PORTUGAL

PSNC

artefactual

PORTICO

UNC SCHOOL OF INFORMATION AND LIBRARY SCIENCE

SOFTWARE SUPPORTERS

Software supporters help us ensure the sustainability of open source digital preservation software. Revenue from software supporters is spent on hosting, testing, maintaining and improving the software resulting in robust, reliable tools used by the digital preservation community.

Our supporters join a product board which oversees financial resources and roadmaps future development. Supporters receive technical support as well as access to tool evaluations and training materials.

At present we only offer software support for JHOVE. Other products will be considered for support in response to the community's needs.

PEOPLE

BOARD

We are governed by our Board of Directors, elected by our members. The Board holds monthly teleconferences and meets in person at least once a year to guide the organisation's strategic direction.

We would like to thank all Foundation members who have served on the Board.

STAFF

We welcomed Joachim Jung as our new Executive Director in December 2015. He takes over the role from Ed Fay who has moved on to the post of Associate Director (Collections, Content, Discovery) at the University of Southampton Library.

The Foundation's job roles have been aligned with our 2015-18 strategy:

Joachim Jung - Executive Director
Joachim is the Executive Director of the Open Preservation Foundation. He has a long term interest in open communities, and brings nearly 10 years experience of working in cultural heritage organisations. He was Project Manager for the EU FP7 SCAPE (Scalable Preservation Environments) project from 2011-2014, overseeing the development and delivery of scalable, open source services for digital preservation. He has been involved in a number of national (German and Austrian) as well as EU funded projects, Planets and IMPACT among them.

Becky McGuinness - Community Manager
Becky joined the organisation in 2010, after working for The British Library on the EU Planets project. She is responsible for the Foundation's webinars, training and events, internal and external communications, and the day-to-day running of the office. She has a Professional Postgraduate Diploma in Marketing, and a degree in English Literature. Her background includes marketing and branding roles in the publishing and engineering consultancy sectors.

Becky contributes to the dissemination and take up activities for veraPDF. She was responsible for the SCAPE project training programme.

Carl Wilson - Technical Lead
Carl is the Technical Lead for the Open Preservation Foundation, overseeing all of OPF'S technical activities. He is an experienced software engineer with a focus on software quality through testing. He is an open source enthusiast, both as a user and developer. His professional interest is using virtualisation, automation and continuous delivery techniques to improve the software development process.
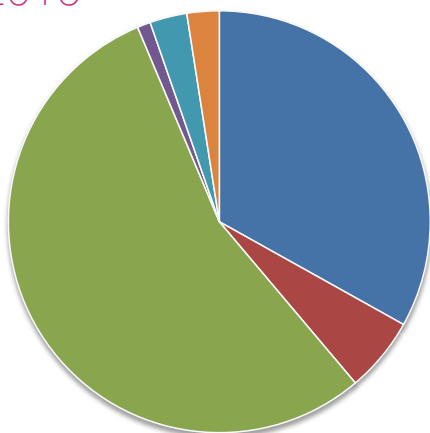
Carl is leading the development team for veraPDF and is responsible for the software quality, website development and continuous integration / delivery. Before this he was responsible for OPF's technical contribution to the SCAPE project.
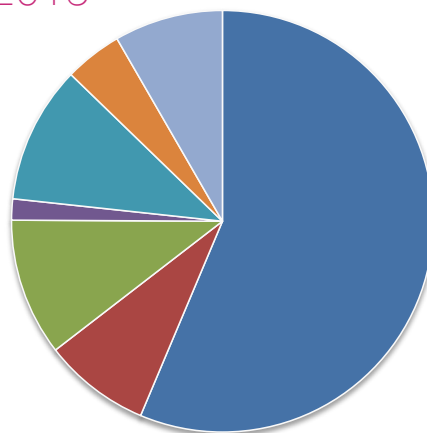
Prior to joining OPF Carl worked for The British Library's Digital Preservation Team on internal and external projects.

2016

2015

2015-2016	Expenditure	2014-2015
£113,773	**Staff salaries**	£132,860
£19,752	**National insurance, social security and pension contributions**	£19,331
£188,440	**Project costs and recharges***	£24,846
£3,479	**Office rent, equipment & hosting (website, wiki, github and webinars)**	£3,826
£9,790	**Outreach (travel, sponsorship and events)**	£24,844
£8,487	**Professional services (accountants, legal and insurance)**	£10,420
(£4,715)	**Foreign exchange losses/gains, bank charges and hardware depreciation**	£19,662

*Income and expenditure have significantly increased in 2016 due to the veraPDF project.

JOIN THE OPEN PRESERVATION FOUNDATION

Membership

Charter members steer our strategy and activities, affiliate members have access to all the outputs. For a full list of member benefits visit our website and read our Strategic Plan 2015-18.

For more information about becoming an OPF member or software supporter visit: **openpreservation.org/join**

Software Supporter

We offer three categories of support which entitle organisations to varying levels of control over our stewardship and maintenance activities, and varying levels of technical support and training.

Annual operating budget	Charter	Affiliate
>£50m	£10,000	40 days or £5,000
£10m - £50m	£5,000	20 days or £2,500
<£10m	£2,500	10 days or £1,250

Category	Fee
Gold	£5,000
Silver	£2,500
Bronze	£1,250

JOACHIM JUNG, EXECUTIVE DIRECTOR

We believe that openness is a fundamental pillar of digital preservation. We release all of our software under open source licences, and contribute to community-driven initiatives. The Foundation makes the results of its work as openly available as possible. While these resources are provided free to the community, they are not produced and maintained without cost. To ensure ongoing funding for the Foundation, we reserve additional benefits for our members. The staff work to an annual plan approved by the members, and are always happy to respond to their requests for help.

Over the past year we've:
- Connected members with shared interests.
- Publicised digital preservation job advertisements for members.
- Supported members getting started with new technologies through code reviews and one on one help e.g. with git and GitHub.
- Offered a forum for members to present their work through our webinar series.

Guided by the priorities set out in our strategic plan we will expand and diversify our membership over the coming year. We plan to achieve this by encouraging more archives and scientific libraries to join the Foundation and by expanding the geographical spread of our membership. The Foundation's Software Supporter model is intended to provide funding for our software stewardship activities. We will work to increase cooperation among digital preservation organisations, exploring better ways to work together.

We want open source software to be more widely-used. The end of 2016 will see the complete 1.0 release of the veraPDF software.

The Foundation's other technology goals for the coming year include:
- The automated production of apt packages for Fido.
- RPM packages for veraPDF, JHOVE, Fido and Jpylyzer.
- Revised contribution guidelines for our flagship projects with accompanying regression testing tools and processes.
- Automated testing and comparison of format identification tools, e.g. DROID, Fido, Siegried, with publication of test reports showing accuracy of identification and breadth of coverage.
- Integrating new services for better QA of external software contributions, e.g. static code analysis and test coverage.

A strategic priority is facilitating the exchange of knowledge amongst peer institutions and developing best practices. We aim to increase participation in our interest groups and continue our successful monthly webinar series to support this. By curating our digital preservation knowledge base our goal is to make it easier to find and contribute information to this valuable resource. We will review and improve our website while continuing to raise awareness of existing community resources including digipres.org, the Q&A site and COPTR.

We are aware that these are ambitious aims for our small team! We welcome input from our members that will help us to successfully meet their needs. We thank you for your loyalty and hope for your continued support in 2016.

www.openpreservation.org