Open
Preservation
Foundation

END OF YEAR HIGHLIGHTS 2019

CONTENTS

WELCOME

The past twelve months have been a busy and exciting time for the Open Preservation Foundation. Guided by our strategy, we have been joined by new members, expanded our collaborative partnerships, and successfully delivered a diverse programme of activities to support digital preservation. This report showcases just some of our milestones and achievements for 2019, from establishing best practices to ensuring support for the tools relied on by the community.

A fresh approach to release management has enabled us to prioritise the user experience and improve consistency and transparency. Meanwhile, the popularity of our two online hack weeks in April and October indicates the growth of our contributor community from strength to strength.

We've welcomed the increasing popularity of our webinars, interest groups, workshops and blogs as we continue to cultivate knowledge sharing platforms for a community at the forefront of digital preservation.

The results of our member survey, together with feedback from our Annual General Meeting has fed directly into our planning and ensures our efforts align with our members' needs. Through our open community survey, we aim to produce a detailed picture of the digital preservation landscape today and use this knowledge to support its future.

Looking back on the years' events brings into focus the important contributions made by our supporters around the world, and we would like to thank each and every one of our members for demonstrating their commitment to the future of digital preservation.

As we enter our tenth year, we look forward to welcoming new members who share our vision for open sustainable digital preservation. Find out more at openpreservation.org/join.

SHARED SOLUTIONS FOR DIGITAL PRESERVATION

Our open source reference toolset addresses common issues encountered in the ingest and pre-invest stages of digital preservation workflows. The products and documentation under OPF stewardship are made freely available under open licences and can be adapted to suit the needs of almost any organisation.

We provide technical leadership and coordinate product development through software quality testing and delivery, leveraging open infrastructure, and following open source best practices.

Our product board govern the roadmap for the OPF reference toolset and play a key role in influencing strategic changes and adoption of new tools. During their bi-monthly meetings, they oversee the development prioritisation for each of our open source products and ensure members' requirements for bug fixes and new feature requests are prioritised.

INTRODUCING OUR PRODUCT RELEASE CYCLE

Over the last 12 months, we have restructured our product release cycle to make our development plans and progress more transparent.

Every product in the OPF reference toolset undergoes a regular triage session where the OPF team and tool maintainers meet to review the list of issues raised on GitHub. Each issue is identified as a bug or feature, assigned a priority level and given an estimated number of development days. Together, we agree a set of work and associated milestones for the next releases and present them to our product board for review and approval at the next meeting.

This improved process is designed to help users plan for adopting new software updates into their workflow as well as identify priorities for contributors. The success of the new process has been demonstrated in 2019 by our delivery of five new releases, which have been developed with contributions from members and the community.

User raises an issue	OPF triage sets development priorities	Product Board approves development plan	Development by OPF & community contributors	Testing & packaging	Release

WHAT'S NEXT FOR THE OPF REFERENCE TOOLSET?

JHOVE2020 is a programme of work intended to enhance and consolidate the OPF reference toolset and ensure its ongoing sustainability as we move into the 2020s.

The programme has several areas of focus including the creation of a single front end GUI for OPF tools, extended reporting capabilities, module review and instrumentation. This work will be delivered in stages over the coming year, each one providing vital support for OPF as we manage a growing portfolio of functional products.

OUR PRODUCTS

JHOVE

Latest release: 1.24-RC (December 2019)
⬇ 4,253   1.22 (April 2019)

JHOVE is an extensible software framework for performing format identification, validation, and characterisation of digital objects.

Find out more:
jhove.openpreservation.org
github.com/openpreserve/jhove

veraPDF

Latest release: 1.14 (July 2019)
⬇ 7,193

veraPDF is an open source, industry-supported conformance checker that validates all current parts and levels of ISO 19005 (PDF/A).

Find out more:
verapdf.org
github.com/verapdf

jpylyzer

Latest release: 2.0 (November 2019)

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

Find out more:
jpylyzer.openpreservation.org
github.com/openpreserve/jpylyzer

fido

Latest release: 1.4.0 (December 2019)

Format Identification for Digital Objects (Fido) is a command-line tool to identify the file formats of digital objects. It is designed for simple integration into automated workflows.

Find out more:
https://github.com/openpreserve/fido

OUR PRODUCT BOARD

Ashley Blewer
Artefactual Systems

Karen Hanson
Portico

Johan van der Knijff
Koninklijke Bibliotheek

Thomas Ledoux
Bibliothèque nationale de France

Michelle Lindlar
Technische Informationsbibliothek

Peter May
British Library

Sheila Morrissey
Portico

Jack O'Sullivan
Preservica

Tomasz Parkola
Poznańskie Centrum Superkomputerowo-Sieciowe

Sarah Romkey
Artefactual Systems

BUILDING OPEN SOURCE COMMUNITIES

Contributors from around the world came together for two online OPF hack weeks during 2019. We introduced this initiative to celebrate and encourage meaningful contributions to open source software, enhancing OPF's reference toolset for the whole community.

In April, our development hack week focussed on JHOVE. The issues addressed by participants improved the quality of the code base and the reliability of the software, paving the way for smoother future development, including the addition of new features.

Following the success of the event, which drew contributors from around the world, we held a second hack week in October as part of Hacktoberfest. With a focus on documentation for OPF tools, the event was accessible to those without development experience, and we received some really valuable contributions including translations of JHOVE error messages into Dutch, German, French, Danish and Portuguese, and documentation for EPUB, WARC and WAVE modules.

Between these two weeks, we estimate that our participants donated around three months worth of effort to the maintenance of OPF tools and resources. What's more, these contributions went on to form the JHOVE community release. We'd like to extend a huge thank you to everyone involved!

Hacktoberfest
presented by ⊘ DigitalOcean and. DEV

CONTRIBUTING TO THE GLOBAL DIGITAL PRESERVATION AGENDA

We have presented at a select number of international conferences this year to raise awareness of our work and showcase collaborative project initiatives.

4-7 FEB — IDCC 2019

12-14 FEB — PASIG 2019

14-17 MAY — Archiving 2019

8 JULY — Preservation Action Registries

16-20 SEP — iPRES 2019

14-15 NOV — DLM Forum

3-4 DEC — Meet e-Archiving

TECHNICAL LEADERSHIP & PRACTICAL SUPPORT

Whether an experienced user or setting up digital preservation environments and systems for the first time, we are here to help.
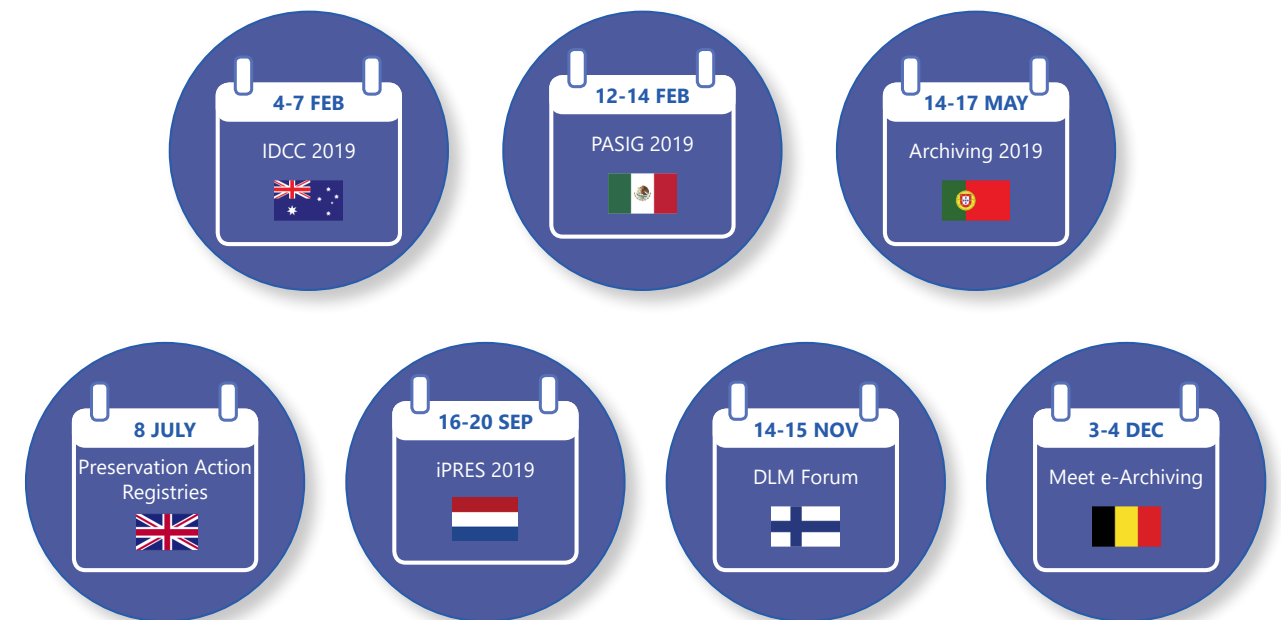
Our monthly tech clinic provides members with personalised, practical support. We offer help using OPF products, support with implementing new systems, and guidance for understanding open source requirements and best practice.

Carl Wilson, our Technical Lead, sits on the RATOM (Review, Appraisal, and Triage of Mail) project Advisory Board and Becky McGuinness, our Community Manager, serves on the PASIG Steering Committee.

RATOM

PASIG 2019
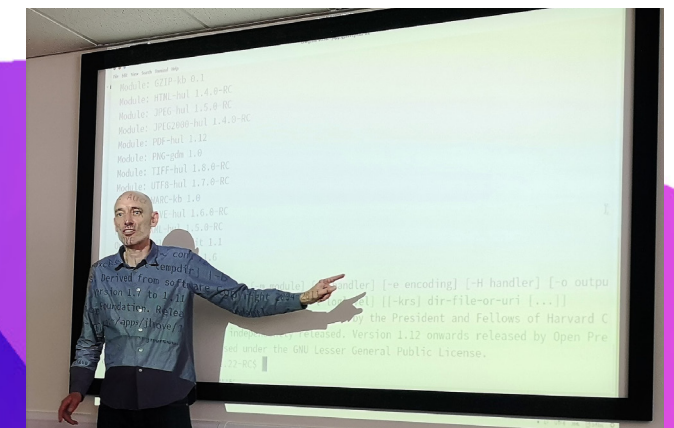February 12–14, 2019 | El Colegio de México | Mexico City

GETTING STARTED WITH OPF TOOLS

We were pleased to welcome members and non-members alike to the training workshops we delivered during the year. Our 'Getting Started with JHOVE' workshops included an interactive demonstration showing how to install and configure JHOVE and how to understand the results. The workshop also covered the importance of format identification as a basis for preservation planning.

In a similar workshop designed for new users of veraPDF, participants were introduced to the PDF/A specifications and learned how veraPDF can improve digital preservation workflows. Participants were also shown ways that they could contribute to the improvement of OPF's open source tools by using GitHub.

Participants at 'Getting Started with JHOVE, The Hague

Carl Wilson presenting JHOVE at Boston Spa, UK

"I have a better understanding of how to analyse the results and what to do next"

"I found the background and history of the tools very interesting."

"I particularly liked the enthusiasm of the speaker and his depth of knowledge"

WORKING AND LEARNING TOGETHER

Knowledge sharing is a key area of strategic focus, and we facilitate a range of services and activities which enable the dissemination of ideas, insights and good practices throughout the community.

ARCHIVES INTEREST GROUP: STUDYING SIGNIFICANT PROPERTIES OF SPREADSHEETS

Three years ago, the OPF Archives Interest Group (AIG) was formed by three of the OPF's archive members to collaborate on shared everyday challenges. The group has grown to include new archive, library and vendor members, but remains driven by archive priorities.

During 2019, the AIG focussed on preserving spreadsheets in born-digital collections. Several archives restrict the file formats they ingest, and this can mean that the spreadsheets they receive are migrated to PDF or TIFF and therefore many of the dynamic elements of the file are lost.

The group started by researching previous work that has been done on spreadsheet preservation and significant properties. They built a knowledge base of their research, spreadsheet specifications and example files and decided to use the InSPECT methodology as a framework.

Beginning with the object analysis, the AIG identified functional, behavioural and structural properties by studying the specifications and using property extraction tools. Few tools exist for extracting spreadsheet specific properties, so Remco van Veenendaal, Nationaal Archief, developed the spreadsheet complexity analyser, a prototype open source tool command line tool to measure properties such as the number of worksheets, cells used, formulas, comments, hyperlinks, revision history, macros etc. in Excel files.

To raise awareness of this work and to get feedback from the community, the AIG submitted a poster to iPRES 2019 - and won the Best Poster Audience Award.

Following the conference, the AIG are now working on the stakeholder analysis and writing up their findings in a report. The practical nature of this work has helped the AIG members gain a much deeper understanding of the spreadsheets they preserve and in some cases has led to a review of their preservation policies.

Kati Sein, National Archives of Estonia, Frederik Holmelund Kjærskov, National Archives of Denmark, Remco van Veenendaal, National Archives of the Netherlands, Jack O'Sullivan, Preservica, and Martin Wrigley , OPF at iPRES 2019, Amsterdam

"Without international cooperation, digital preservation might sometimes feel like lonely work. There was one more outcome of the AIG that cannot be measured. We share all kinds of news and information during the monthly calls and turn to each other for advice. Often we get the advice and even if not – it is oh so valuable to feel shoulder to shoulder with virtual colleagues." - Kati Sein, Estonian National Archives

SHARING EXPERIENCE THROUGH WEBINARS & VIRTUAL PANELS

Our popular webinar series, which has been running since 2012, features expert speakers on hot topics in digital preservation and draws participants from around the world. Our webinars are free to attend and open to everyone, and this year, we increased the number of places available in order to accommodate our growing audience.

OPF members receive priority booking for our webinars, and we encourage all members to attend and share their work as part of the series. Members also receive exclusive access to the archive of recordings and slides which they can revisit in their own time.

Our first digital preservation virtual panel took place in September, focussing on digital preservation policy. OPF membership is growing and the virtual panels provide an inclusive space for members to connect with one another and find out what is happening in other member organisations across the world.

This members' only event sparked some lively debates among our audience about how memory institutions can begin to develop digital preservation policies and some of the challenges that they may face in doing so. We look forward to running further virtual panels in the coming year.

15	589	47	324
webinars	registrations	attendee countries	attendees

OPF BLOGS: JOIN THE CONVERSATION

Our community blog showcases the fantastic digital preservation work being carried out around the world. You can find information on a wide range of topics such as tools, project news, case studies and best practice. Blogging on our site is a great way to share your ideas and receive feedback from the community.

22,495
reads

management · best practice · community · OPF · hackathon · archive · emulation · heritage · veraPDF · born digital · preservation · virtual · meet · development · collection · PRONOM · library · service · community · JHOVE · file format · sustainability · significant properties · database · conference · metadata · JPEG · workflow · tool · advice · documentation · hack week · characterisation · open source

OUR MEMBERS MAKE US

Charting the course for OPF's activities and outputs, our members give a voice to the global digital preservation community. In 2019, we welcomed three new members. We now have 28 members in 16 countries.

Albert-Ludwigs-Universität Freiburg
Arcsys Software
Artefactual Systems
Austrian Institute of Technology
Biblioteca Nacional de Portugal
Bibliothèque nationale de France
Bibliotheque nationale de Luxembourg
British Library
CSC - IT Centre for Science
Det Kgl. Bibliotek
Ex Libris
Goportis
Harvard Library
International Atomic Energy Agency Archives and Records Management Section
Jisc
Koninklijke Bibliotheek
Nasjonalbiblioteket
Nationaal Archief
Österreichische Nationalbibliothek
Portico
Poznańskie Centrum Superkomputerowo-Sieciowe
Preservica
Rahvusarhiiv
Rigsarkivet
The University of North Carolina at Chapel Hill
The National Archives, UK
Yale University Library

ANNUAL GENERAL MEETING 2019

All OPF members are invited to attend our Annual General Meeting (AGM) to discuss our work from the past year, connect with other members and influence our activities going forward. Our 2019 AGM took place in the Netherlands and was kindly hosted by the Nationaal Archief and the Koninklijke Bibliotheek.

The AGM is a valuable opportunity for members to share their feedback on OPFs outputs and vote in elections for the board. Following the official business, we gave a round-up of OPFs activities over the year, including our busy programme of knowledge sharing activities, the projects we've been working on, and what our next steps are.

OPF members at the AGM

We heard all about the brilliant work taking place in our members' organisations during the 'member showcase' session, and it was particularly interesting to hear from two of our newest members about the exciting developments that are taking place at their institutions, and how they plan to progress their digital preservation programmes over the coming years.

On day two, we introduced the JHOVE2020 programme of work, discussed the future of the OPF reference toolset and gathered feedback from our members during breakout sessions on the Preservation Action Registries (PAR) project. To close the AGM, we brought together a selection of digital preservation practitioners from our member organisations for a fascinating panel discussion on the future of digital preservation.

JOIN US AT OPFCON TO CELEBRATE A DECADE OF OPF

2020 marks 10 years of the Open Preservation Foundation. To honour this occasion, we invite you to join us at a global celebration on 10 June at the Austrian National Library in Vienna. OPFCON celebrates a decade of OPF, considers our impact on the community, and reflects on our role in the future of digital preservation.

We would like to thank the members of our new working group, who have been helping us to plan an exciting event that truly celebrates this special occasion.

The call for contributions is now open! The theme of OPFCON is **open sustainable digital preservation** and we invite contributions that consider the following areas:

• Knowledge sharing, policy and best practice
• Innovation and technology
• Sustainability through community collaboration

More information about OPFCON is available at https://openpreservation.org/event/opfcon/.

OPFCØN

CELEBRATING 10 YEARS OF THE
OPEN PRESERVATION FOUNDATION

PEOPLE WE WORK WITH

The digital preservation community has a number of collaborative bodies, some centered on fixed term project work, others with longer term constitutions. OPF works collaboratively with partners to align efforts, benefitting our members and the wider community. We currently have memorandums of understanding in place with the following organisations:

ENGAGING WITH NEW COMMUNITIES

In November, we became an affiliate member of The Open Source Initiative® (OSI) the global non-profit organisation formed to educate about and advocate for the benefits of open source software, development and communities. In this capacity, OPF can engage with a community of the foremost open source software leaders, projects and businesses, and will gain opportunities to participate in incubator projects and working groups that support software freedom.

DEVELOPING COLLABORATIVE PARTNERSHIPS

Jan Dalsten Sørensen, and Remco van Veenendaal sign the MOU

November 2019 saw the signing of a new memorandum of understanding between OPF and the DLM Forum. The memorandum was signed at the DLM Forum Members' Meeting in Helsinki by Remco van Veenendaal, Director, OPF Board, and Jan Dalsten Sørensen, Chair of the DLM Forum. The agreement facilitates both parties to build a collaborative partnership to address existing and emerging challenges in long-term archiving and digital preservation, and to ensure a sustainable future for tools and knowledge.

Marcel Ras, DDHN, and Martin Wrigley, OPF, sign the MOU

Earlier in the year, we also updated our partnership with the Dutch Digital Heritage Network (DDHN). This renewal of our shared commitment followed the merging of the Netherlands Coalition for Digital Preservation (NCDD) into the DDHN in 2018. The new agreement was signed in January 2019 at the National Library of the Netherlands during an OPF Board meeting.

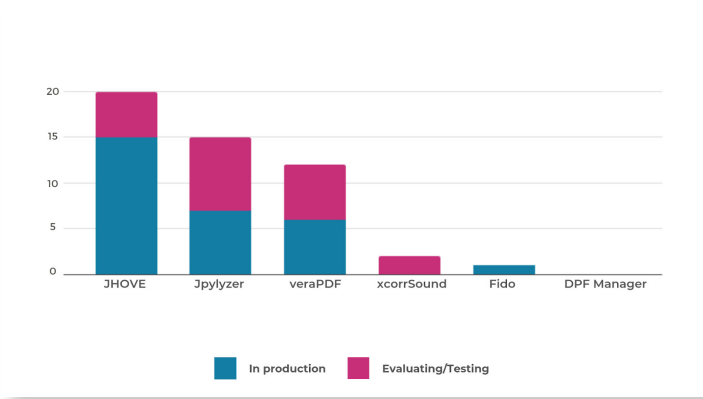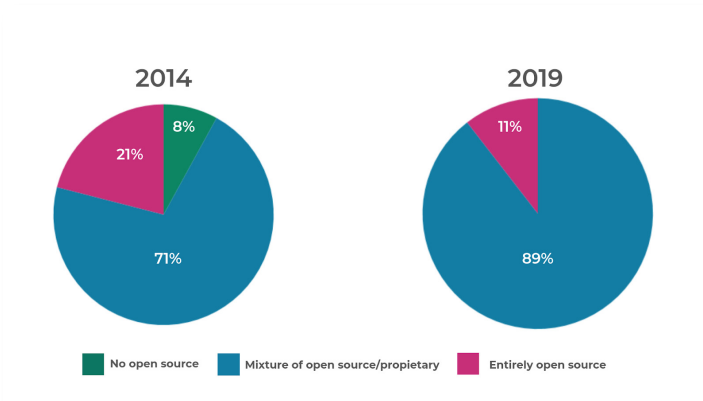# SURVEYING THE DIGITAL PRESERVATION COMMUNITY

Our member survey, which ran this Spring, aimed to gather information about open source approaches, organisational and collection capacities and digital preservation policy and strategies among our members. From universities, libraries, and archives to digital preservation vendors and government institutions, we received 21 responses from organisations of all types. The high response rate enabled us to build a detailed picture of approaches to digital preservation among our members. The results were made available to our members in August, and we published selected highlights openly on the OPF blog for non-members.

Our member survey, which ran this Spring, aimed to gather information about open source approaches, organisational and collection capacities and digital preservation policy and strategies among our members. From universities, libraries, and archives to digital preservation vendors and government institutions, we received 21 responses from organisations of all types. The high response rate enabled us to build a detailed picture of approaches to digital preservation among our members. The results were made available to our members in August, and we published selected, anonymised highlights openly on the OPF blog for non-members.

The survey included questions on organisational capacity, strategies, collections and workflows, and storage and technology. Where possible, we were able to compare the results with our 2014 survey, enabling us to draw some interesting comparisons and trace developments over time. Some of the most interesting insights gathered were in relation to the use of our reference toolset. When we conducted a similar survey in 2014, 64% of our members responded using JHOVE, which led to its adoption by OPF. Five years later, JHOVE is the most popular tool in the toolset, used by over 95% of respondents. Jpylyzer came in second, and is used by 71% of our members, whilst veraPDF is used by 57%.

We were interested to discover the variety of activities being carried out across our members' organisations. Software development/maintenance remains a popular activity across the organisations. Other widespread activities include format validation and identification, metadata creation/extraction and storage/bit preservation. While the 2014 survey indicated that emulation remained a less widely adopted preservation strategy than migration, the 2019 results demonstrate that several institutions are now researching or developing the capacity to carry it out. Digital forensics,

Use of OPF tools in production and for evaluation/testing

Use of open source

PARTICIPATING IN PRACTICAL PROJECTS AT THE FOREFRONT OF DIGITAL PRESERVATION

We work on selected research & development or consultancy projects that align with our strategic goals.

Wikidata for Digital Preservation
The digital preservation community relies on metadata to describe file formats, software, operating systems and hardware. This metadata is used to inform and drive digital preservation processes. The vision of Wikidata for Digital Preservation is to leverage Wikidata as a technical registry of metadata related to file formats, software and computing environments. Facilitating collaborative curation of this metadata, and making it available as linked open data will reduce the amount of redundant work digital preservation professionals perform to describe resources. A reliable source of machine-readable, linked open data that describes the digital preservation domain can be used by software applications and information systems, reducing the overhead when building new tools.

Connecting Europe Facility eArchiving Building Block
The E-ARK consortium is working with the European Commission's Connecting Europe Facility to develop the eArchiving Building Block. This comprises specifications and software with accompanying training and outreach to help tackle the challenge of implementing and maintaining electronic archiving solutions. OPF lends its technical expertise to ensure the accuracy and consistency of the specifications and develop accompanying validation software. We also contribute to the project's extensive outreach programme, working to promote engagement with, understanding of, and access to eArchiving services among enterprises in both the public and private sectors, national, regional and local authorities and administrations, aggregators and content providers, as well as individual citizens.

Preservation Action Registries (PAR)
Digital preservation systems generally lack a common and consistent way to describe and execute preservation policies and actions. Initiated by Jisc, the PAR project aims to address this by capturing and sharing good practices for the preservation of digital objects. In collaboration with Artefactual, Arkivum, and Preservica, our work on the initial phase produced a conceptual model accompanied by an API and schema based upon the model. These were delivered as a working proof of concept that allows the description of a preservation action for a given file format to be exchanged between two previously incompatible action registries.

BOARD OF DIRECTORS

Barbara Sierman, (Chair)
Koninklijke Bibliotheek

Peter May, (Deputy Chair)
British Library

David Clipsham,
UK National Archives

Max Kaiser,
Österreichische
Nationalbibliothek

Michelle Lindlar,
Technische
Informationsbibliothek

Paul Stokes,
Jisc

Remco van Veenendaal,
Nationaal Archief

Our directors represent the voice of our membership. The Board plays an important role in making strategic decisions and monitoring ongoing activities. This year, we said goodbye to John Kaye from Jisc after four years of service, as well as Ross King from AIT, who had been a dedicated leader for OPF since it was first established almost a decade ago. In turn, our members unanimously voted to welcome two new faces to our board, Paul Stokes from Jisc and David Clipsham from the National Archives, UK.

OPF ON THE MOVE

In June, we made the move from our long-time home at the British Library at Boston Spa to accommodate our growing team. The OPF team are now happily situated in Leeds city centre. We would like to take this opportunity to thank the British Library for hosting the OPF since it was founded in 2010.

2019 also saw the opening of our new subsidiary in the Netherlands. In uncertain political times, Stichting Open Preservation ensures our continued participation in international digital preservation initiatives.

OPF HQ, Leeds

TAKING TEAM BUILDING TO NEW HEIGHTS

The OPF team raised over £2000 for Macmillan Cancer Support in August by embarking on the Wharfedale 3 Peaks Challenge. A much less travelled alternative to the Yorkshire 3 Peaks, the challenge covers the three summits of Great Whernside, Buckden Pike, and Birks Fell. We walked for 21 miles with an ascent of around 5,000 feet across some very rough and boggy terrain, completing our challenge in 12 hours. Many thanks again to everyone who sponsored us on our adventure!

The OPF team at the Great Whernside summit

2,205
followers

1,880
subscribers

709
subscribers

93
members

MEET THE TEAM

Martin Wrigley
Executive Director

Becky McGuinness
Community Manager

Carl Wilson
Technical Lead

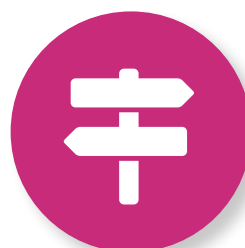Charlotte Armstrong
Project Officer

Martin Speller
Project Manager

JOIN US

The Open Preservation Foundation exists for and is sustained by its members. We put our members at the heart of what we do, and every decision we make is taken with your interests in mind. While we work for you, membership is also an opportunity to make a vital difference to the field of digital preservation and contribute to a thriving digital preservation community.

Networking   **Governance**   **Knowledge**   **Support**   **Impact**

OPF maintains open source digital preservation software that is used by thousands of organisations worldwide. By joining OPF, your institution can help to ensure the long-term sustainability of these tools. Part of the funding from membership fees goes directly towards the hosting, maintenance, and improvement of dependable software that is relied on across the digital preservation community. As a member, your organisation's priorities will become our priorities. You can participate in the ongoing development of the tools under our stewardship, have your say about further tool adoption, and help to steer the overall development of the OPF reference toolset.

WHY OPF? ASK OUR MEMBERS

What motivated you to join OPF?

'The range of members and shared experience.'

'Networking within the international digital preservation community.'

'Development of tools to support the work we are doing.'

'OPF is doing important work in maintaining and developing open source solutions for digital preservation. We see this as a key element in our efforts to produce national digital preservation.'

What do you see as the main benefits of OPF membership?

'The ability to get technical feedback on our own projects.'

'Collaboration is an important motivation to join and work with the OPF. Digital preservation is a task not to tackle on your own, but with a community to learn from each other'

'To engage with a network of experts and share and bounce ideas off each other.'

'As a board member, we can influence the direction OPF is going.'

'Tool recommendations, and best practice awareness. Support in future activities.'

'A source of information and news, support for the use of open source tools. '

FOLLOW US