

Open
Preservation
Foundation



END OF YEAR HIGHLIGHTS

2018

CONTENTS

3	Best Practice and Knowledge Sharing
8	Open Source Digital Preservation Tools
11	National and International Projects
12	Meet our Members
13	Governance and Board
14	Staff
15	Outreach and Communications
17	Join us

VALUES

Open
Member driven
Collaborative & Inclusive
Innovative

VISION

Open sustainable digital preservation.

MISSION

Enabling shared solutions for effective and efficient digital preservation; the Open Preservation Foundation leads a collaborative effort to create, maintain and develop the reference set of sustainable, open source digital preservation tools and supporting resources.

This set of tools (including software and standards) enables organisations to evaluate, validate, document, mitigate risk, and process digital content to be preserved in line with desired policies and community best practice.

WELCOME

Welcome to the Open Preservation Foundation's 2018 end of year highlights. With the launch of our new strategy, vision and mission alongside the increase of our team from three to five, it has been an eventful year for the OPF. As our strategic plan approaches the end of its first six months in operation, this report allows us to reflect on the Foundation's diverse projects and activities, all of which represent our continuous efforts to reach our vision of open sustainable digital preservation.

Our work to address common digital preservation challenges has been underpinned by the ongoing development of the OPF reference toolset. This is driven by the needs of our members, and we are grateful for their valuable insights, guidance, and collaboration.

2018 has also provided opportunities to celebrate the global nature of the Foundation. We have taken part in a number of international conferences and events, made plans to open a subsidiary office in the Netherlands, and welcomed new members - increasing the number of countries across the OPF membership to fifteen.

The Foundation's belief in openness steers our efforts to develop shared standards and solutions to common digital preservation problems. Our programme of webinars, blogs, and workshops enables us to ensure that the expertise available to us is of maximum benefit to the community, while our outreach activities allow us to share these valuable resources with the widest possible audience. The topics covered across these platforms in 2018 have been particularly diverse and innovative, and we would like to thank all contributors for sharing their work.

OPF LAUNCHES NEW STRATEGY

One of our key achievements of 2018 was the development of our new strategy. We defined new core values which encompass the Foundation's fundamental beliefs and underpin everything we do and we refreshed our vision and mission.

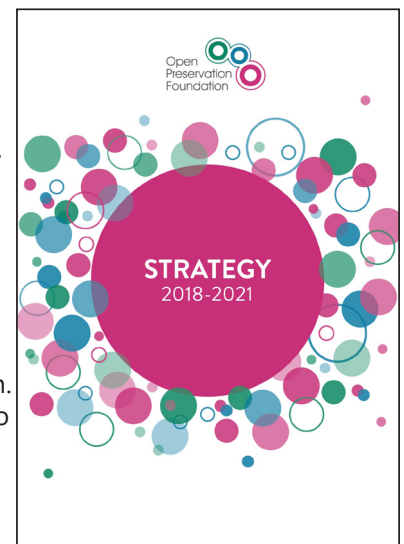
The Foundation has grown and evolved over the last eight years, and our strategy brings together the different strands of work we do.

The strategy sets out our long-term goals and plans for the future and helps guide the tasks, tools, and projects that the Foundation will take on. It outlines the practical actions to achieve these goals and the measures to evaluate its effectiveness.

Central to the strategy is the stewardship of our open source digital preservation tools and our best practice and knowledge sharing programme of activities.

We welcome all organisations who share our goals to [join us](#) to help realise the strategy's success.

[Download a copy of the Strategy.](#)



BEST PRACTICE & KNOWLEDGE SHARING

OPF runs a series of initiatives to promote and facilitate knowledge sharing and best practice amongst members and the digital preservation community. This year's webinars and blog posts demonstrate the wide range of innovative work being done within our field, while our interest groups provide an inclusive and collaborative environment in which to share practical insights on a wide range of digital preservation topics. 2018 has seen the growth of our blog readership and webinar attendance, not just in terms of numbers, but also regarding the international scope of the individuals who contribute to and engage with these series.

OPF Interest Groups

Our member interest groups foster collaboration on specific digital preservation topics and provide practical benefits members can apply to their day-to-day work.

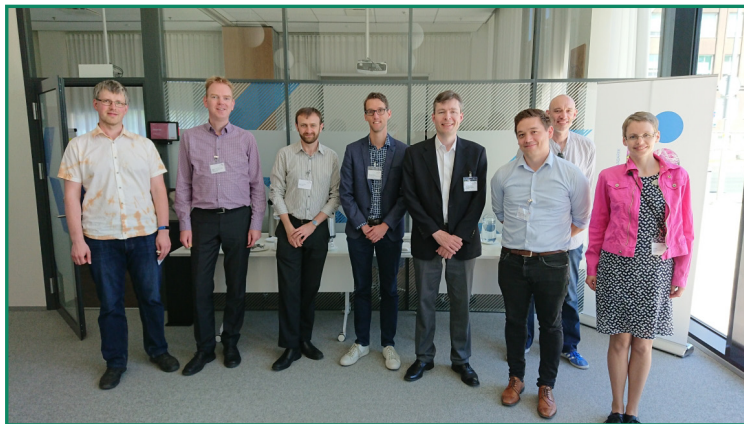
Format Interest Group

- Sharing member activities or challenges
- Building knowledge about file formats and associated processes (e.g. validation)
- Evaluating software in relation to file formats
- Driving software tool improvements (e.g. JHOVE, veraPDF)
- Using format knowledge to push for improvements to software and documentation

Archive Interest Group (not just for archives!)

- Adopting the SCAPE Catalogue of Policy Elements to agree priorities
- Exchanging knowledge and best practice about required and preferred formats for digital archiving
- Investigating significant properties in spreadsheets for long-term preservation recommendations
- Developing a prototype spreadsheet analyzer to support this work

OPF members can join the interest groups at any time or can put forward ideas to establish new groups.



OPF Archive Group meeting in Tallinn, May 2018

News highlight: New faces at OPF

In October, we were delighted to welcome two new starters at the Open Preservation Foundation, growing our team from three to five members of staff.

Our Project Manager, Martin Speller, has been leading the internal project to improve the management of OPF's open source digital preservation tools, as well as providing project management for the Jisc PAR project.

As Project Officer, Charlotte Armstrong has been supporting our outreach and community activities, helping to manage the OPF office, and working on internal projects such as the upcoming member survey.





OPF Webinars

Our webinar series has been running since 2012, providing an open and collaborative platform on which members and the wider community can share their digital preservation work. OPF members have the ability to suggest the topics to be covered in our webinar series, and can point to relevant specialists in their fields for us to approach and ask to share their insights and expertise.

Our webinars cover four categories:

- Preservation in practice** - case studies and user experiences
- From the toolbox** - exploring digital preservation tools
- Open source with OPF** - best practice software development
- Community projects** - showcasing current research and development projects.

This year, several of our members have shared their work with a global community of 313 unique attendees from 22 countries.

The OPF has also shared significant internal developments in this year's webinar series.

Our members-only webinar in January provided a detailed walkthrough of our new vision, mission, and strategy, including an overview of the planning and implementation process and insights into Martin Wrigley's first months at the foundation.

In September, we explored the OPF's open source reference toolset and shared new roadmap plans for the tools under our stewardship. While an archive of all webinars is available for members to view online, this webinar was opened up to the wider digital preservation community, as it provides a unique insight into who OPF are and what we do.

OPF members are offered priority registration for the webinars, and any remaining places are released to the community. Members receive exclusive access to the archive of recordings and slides. We welcome new and returning speakers to share their digital preservation work - please get in touch!

The OPF blogs have had 13,958 reads this year! We've had exciting contributions from both regular bloggers and newcomers sharing a diverse range of hot topics in digital preservation:



- Yvonne Tunnat, 'Sherlock Carriage – PRONOM's blind spot on (some) PDFs from 2010 to 2014' (1586 reads)
- Remco van Veenendaal, 'Let's play Preservia: the digital preservation board game' (1678 reads)
- Johan van der Knijff, 'Resurrecting the first Dutch web index: NL-menu revisited' (1295 reads)

These posts drew a large readership and generated some fruitful conversations. They also represent the broad scope of the OPF blog, which provides an inclusive space for the digital preservation community to share their news and engage with a vast network of experts from the field. New authors are always welcome, and the OPF can provide help with proofreading and editing, so get in touch if you'd like to join the conversation!



News highlight: World Digital Preservation Day 2018

OPF joined the community in celebrating world digital preservation day on 29 November. This was an opportunity to celebrate the work of our members and showcase some of our internal activities to the wider community.

We joined the twitter conversation with a series of tweets inviting our followers to 'get to know OPF'. Our project officer, Charlotte, also shared insights into her first weeks at the OPF in a [blog post](#) which was published on the day.

The evening's DPC Digital Preservation Awards ceremony was attended by our executive director, Martin, along with Boris Dubrov of Dual Lab. The pair represented the consortium behind veraPDF, which was a finalist for the Software Sustainability Institute's Award for Research and Innovation.



*Martin Wrigley, OPF
and Boris Dubrov,
Dual Lab*

OPF Annual General Meeting 2018

This year's AGM was held at the Joint Ministries Building in Tallinn, organised in collaboration with the OPF member the Estonian National Archives. Remco van Veenendaal from the National Archives of the Netherlands was elected as a new board member. Stephane Reece of the Bibliothèque nationale de France retired from the board after three years of service.

After the official business, we reviewed the new OPF strategy, with members participating in valuable discussions to provide feedback and help finalise it. Day 1 ended with member showcase presentations.

We introduced the OPF Reference Toolset and explored the process with members to discuss how OPF tools map to a typical digital preservation workflow. Following this, we presented the new JHOVE roadmap. This year we had a new open debate session where members suggested topics on the first day and then voted for the most popular for discussion. They included Epubcheck, emulation, communications and branding.

[Presentations from the AGM](#) are available to OPF members on our website.



OPF Annual General Meeting, Tallinn, May 2018

OPF Workshops

Access to Digital Collections that Take Bitstreams Seriously: Hands-on with BitCurator, BitCurator Access and BitCurator NLP Software

In June 2018 we held our BitCurator workshop at the British Library in London. It was fully booked with 25 participants from across Europe and one from Canada. Participants were introduced to the BitCurator, BitCurator Access and BitCurator NLP (Natural Language Processing) projects by Cal Lee and Kam Woods from UNC Chapel Hill. Over the two days, participants learned about and worked through practical exercises applying the tools on test data.



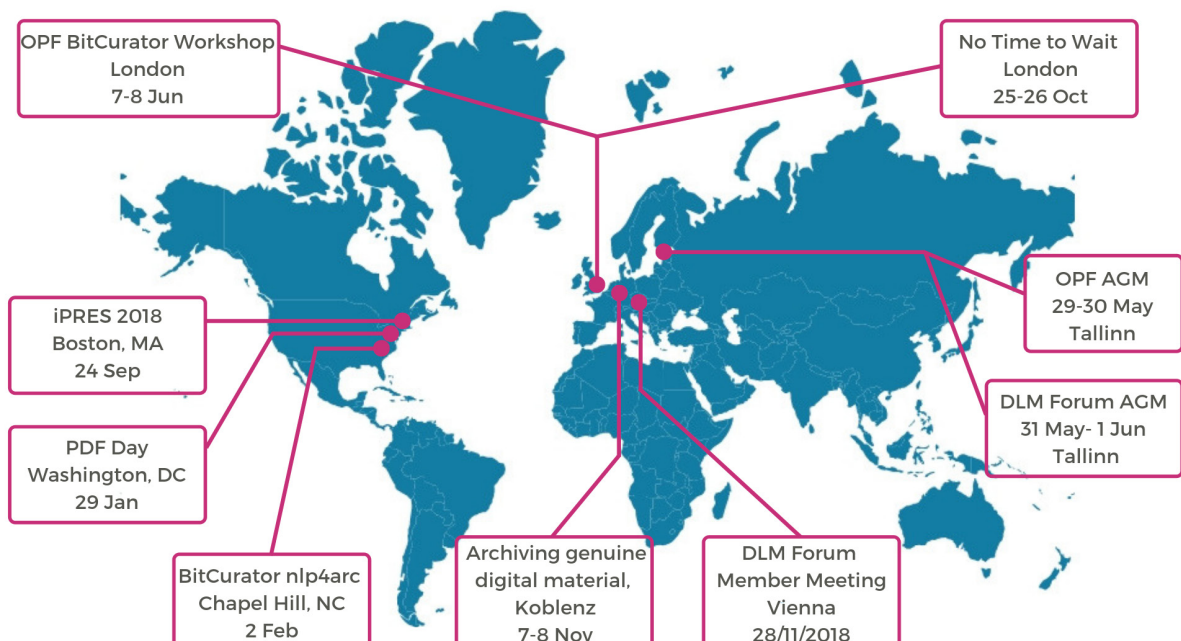
OPF BitCurator workshop, June 2018, London

JHOVE 101 tutorial at iPRES 2018

Carl Wilson, the OPF's technical lead, gave a JHOVE tutorial during the ad hoc sessions at iPRES 2018. The tutorial started with an overview of what JHOVE does and a brief history of its development (the OPF adopted JHOVE in 2015). Carl demonstrated how to install and run JHOVE and explained how to interpret the results.

OPF in the Community

We have attended and presented our work at industry conferences around the world and sit on the PASIG (Preservation and Archiving Special Interest Group) Steering Committee:



OPF Tech Clinic

The OPF Tech clinic provides our members with the opportunity to book a one-to-one session with Carl Wilson, our technical lead. We appreciate that the digital preservation experience of our members varies substantially, and as such, the scope for discussion topics and level of detail in these sessions is left relatively open. We are particularly eager to provide assistance with specific issues relating to OPF strategic activities such as use of tools in the reference toolset, or participation in the open-source development process. This might include:

Getting started with OPF tools, e.g. installation and basic usage
Assistance with open source development best practices

Further topics of discussion have included:

Help with integrating open source tools into local automated workflows and systems
Investigating problems and issues with open source tools
General software development advice including tutorials for beginners

Code Reviews & Consultancy

OPF has extensive experience carrying out code reviews and software maturity assessments for open source projects. Most recently we were commissioned to carry out a code review for the BitCurator NLP project in December 2018. Our Technical Lead, Carl Wilson, sits on the BitCurator Technical Advisory Board. We also work with members and selected partners to carry out digital preservation consultancy work.

OPEN SOURCE DIGITAL PRESERVATION TOOLS

We are an experienced leader in open source digital preservation software development. Our software maturity model is designed to provide assurances about the quality of software under our stewardship.

Funds from membership fees go directly towards hosting, maintenance, and improvement of software, resulting in robust and reliable tools which are used by members, and across the digital preservation community.

We believe in an open source approach to our software development. It helps our users to understand how the tools work and enables them to download and modify them to meet their requirements. Users can also report bugs, submit fixes and make feature requests. The software can be widely and publicly tested to help ensure it is robust and reliable.

Our maintenance and development work is guided by our members through the OPF Product Board. We have developed roadmaps and a new release management process to prioritise fixes and feature requests and to schedule in maintenance work and supporting documentation. The JHOVE roadmap was developed first and defines several components: fixes, features, maintenance, testing & QA, and documentation. By publishing the roadmaps openly, we aim to encourage contributions from members and the community.



Latest release: 1.20 (March 2018)

Number of downloads: 4,158

[Download JHOVE 1.20](#)

JHOVE (JSTOR/Harvard Object Validation Environment) is an extensible software framework for performing format identification, validation, and characterisation of digital objects.

JHOVE 1.20 contains improvements to the WAVE module including support for RF64 files, better reporting and minor bug fixes. There are some new validation checks in the PDF module for the header, document catalog and page dictionary. Image format reporting has been added to NISO MIX 1.0 output and PDF image MIME types detection has been fixed.

Find out more:

jhove.openpreservation.org

github.com/openpreserve/jhove



Latest release: 1.12 (May 2018)

Number of downloads: 10,494

[Download veraPDF](#)

veraPDF is an open source, industry-supported conformance checker that validates all current parts and levels of ISO 19005 (PDF/A). The software is accompanied by a comprehensive, atomic test corpus, covering each clause in the various parts of the PDF/A specifications.

Version 1.12 features multi-threaded support and parallelization in the command line interface, significantly improving performance.

Find out more:

verapdf.org

github.com/verapdf

News Highlight: veraPDF a finalist in the Digital Preservation Awards

The prestigious Digital Preservation Awards is a world-wide celebration of significant and state-of-the-art contributions to securing our digital legacy. September brought the exciting news that veraPDF, one of our core open source digital preservation tools, had been announced a finalist in the Software Sustainability Institute's Award for Research and Innovation.

Executive Director of OPF Martin Wrigley said:

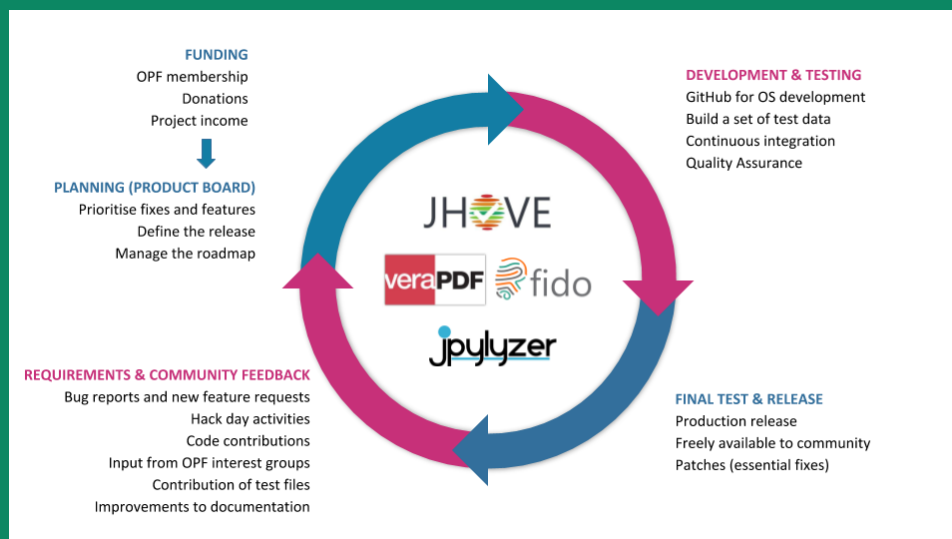
"I'm thrilled to see the recognition of the great work that has gone into veraPDF to make it such an essential digital preservation tool and would like to thank the Digital Preservation Awards judges and the DPC for this honour."

Sponsor of the Award for Research and Innovation, Neil Chue Hong of the Software Sustainability Institute said: "I am delighted to be able to share the work of our finalists. Each one demonstrates world-class innovation and research which has the potential to benefit the entire digital preservation community."

The nomination was a testament to the hard work and dedication of the veraPDF consortium and a wonderful opportunity to celebrate the success of the tool.

HOW DO OPF PROJECTS WORK?

The diagram below shows the OPF development cycle. All of our tools are open source but we rely on funding from membership fees, donations and project income to give them a central and sustainable home and to carry out maintenance and development work.



Latest release: 1.18 (August 2017)

[Download jpylyzer](#)

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).

The most current release included updated build scripts, improved handling of codestream comments with non-text data, and restructured repo to make packaging more straightforward, along with various code clean-ups based on PEB8 and PyLint analysis.

Find out more

jpylyzer.openpreservation.org

github.com/openpreserve/jpylyzer



Latest release: 1.3.12 (September 2018)

[Download fido](#)

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. Fido uses PRONOM signatures converted to regular expressions.

Find out more

github.com/openpreserve/fido

NATIONAL & INTERNATIONAL PROJECTS

The Foundation has worked on selected research and development or consultancy projects that align with our strategic goals.

Connecting Europe Facility eArchiving Building Block

E-ARK4ALL is working with the European Commission's Connecting Europe Facility to develop the eArchiving Building Block. The Building Block will provide the core specifications, software, training and knowledge to tackle the challenge of data management and reuse. The OPF lends its expertise to developing the validator to ensure that information packages comply with the requirements set in E-ARK specifications and producing the sample software portfolio which provides an open source reference toolset. We are also supporting the outreach and training activities.



The project's most recent output is a series of enhanced and stabilized eArchiving specifications. The Common Specification for Information Packages (CS IP) and the Common Specification for Electronic Records Management Systems (CS ERMS) are now available for review. The second group of specifications will be released for review towards the end of 2018. This group includes the E-ARK SIP, E-ARK AIP, E-ARK DIP, the review components of the Common Specification SIARD and Geodata specifications.

Preservation Action Registries

The Preservation Action Registries (PAR) project delivered an initial proof-of-concept platform for the exchange of recommended preservation actions, as part of the Jisc Research Data Shared Service (RDSS) pilot program.

In a new initiative supported by Jisc, Preservica, Arkivum, the Open Preservation Foundation and Artefactual are working together to deliver a PAR which aims to exchange preservation action best practice within the preservation community and between different preservation systems. This exchange of preservation knowledge and its practical implementation has the potential to vastly accelerate the pace of digital preservation research and the deployment of best practice to live systems.

Web-based document scanner

The digital government service of the Netherlands Ministry of the Interior and Kingdom Relations commissioned the OPF to build an open source web crawler with partner Dual Lab. It is based on Heretrix and veraPDF and crawls Dutch government websites to find and count attachments including different types of PDFs, Microsoft Word and ODF files.



Logius
Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

Additional work has followed to develop a web application for converting Office documents supported by LibreOffice / OpenOffice to PDF/A-1 Level A documents with optional follow-up validation by veraPDF. The collaboration will continue in 2019 to extend the crawler functionality.

File format identification study (RDSS)

OPF was appointed as a supplier on the Jisc Research Data Shared Service (RDSS) framework. Our role was to carry out a scoping study of file format identification issues, requirements and solutions based on a test corpus of research data supplied from RDSS pilot institutions and elsewhere.



The scoping study comprised two parts. We investigated different sources of signatures used by format identification software. We then ran selected file format identification tools over the test corpus to examine the results and compare the differences. This helped to identify the coverage gap between the tools. We then carried out work to determine if that gap can be bridged by an automated process and also show how many files and formats cannot be identified by any of the tools.

MEET OUR MEMBERS

Our global network of members currently represent national libraries and archives, universities and research organisations, and software vendors. Membership is open to all types and sizes of organisations who share our goal to safeguard digital materials for the long term.

Austrian Institute of Technology
British Library
Bibliothèque Nationale de France
Bibliothèque Nationale de Luxembourg
Det Kgl. Bibliotek
Ex Libris
Goportis
International Atomic Energy Archives
Jisc
Koninklijke Bibliotheek
Nasjonalbiblioteket
Nationaal Archief
Österreichische Nationalbibliothek
Preservica
Rahvusarhiiv
Rigsarkivet
The National Archives UK
Yale University Library
Arcsys Software
Albert-Ludwigs Universität
Artefactual
Biblioteca Nacional de Portugal
Portico
PSNC (Poznan Supercomputing & Networking Centre)
University of North Carolina

OUR PARTNERS

We believe that collaboration is essential to advance digital preservation. We work with partner organisations to share expertise, coordinate planning and deliver joint activities.



GOVERNANCE

Board of directors

Our board are elected by our members at our Annual General Meeting each year. They represent member needs and meet once a month to guide the Foundation's activities. Our current board members as of December 2018 are:



Barbara Sierman, (Chair)
Koninklijke Bibliotheek



Peter May, (Deputy Chair)
British Library



Max Kaiser, Österreichische
Nationalbibliothek



Ross King, AIT Austrian
Institute of Technology



John Kaye,
Jisc



Michelle Lindlar, Technische
Informationsbibliothek



Remco van Veenendaal,
Nationaal Archief

News highlight: Barbara Sierman awarded DPC digital preservation fellowship

OPF chair, Barbara Sierman, was awarded the DPC fellowship in recognition of her sustained personal contribution to the field of digital preservation and her generously shared insights and ongoing collaboration for the widest possible benefit. Barbara is known to many in the community for her work with the Koninklijke Bibliotheek (National Library of the Netherlands) and the Open Preservation Foundation, as well as her development of the ISO standards 16363 Audit and Certification of Trustworthy Digital Repositories and ISO 16199 Requirements for bodies providing audit and certification of candidate trustworthy digital repositories, among many other digital preservation projects.

STAFF

Martin Wrigley, Executive Director

Martin was appointed as our new Executive Director in August 2017. He is responsible for working with our members and the board to develop the future strategy and to enhance the Foundation's portfolio of open source digital preservation tools and supporting resources.



Becky McGuinness, Community Manager

Becky manages our programme of webinars, events and interest groups. She also is responsible for the OPF's marketing and communications and looks after the day-to-day running of the office.



Carl Wilson, Technical Lead

Carl oversees all of the OPF's technical activities. He leads the sustainability, maintenance and development work for the OPF reference toolset and provides open source technology expertise to our external projects. He is an experienced software engineer with a focus on software quality through testing.



Martin Speller, Project Manager

Martin is the OPF Project Manager overseeing the project delivery tasks associated with the OPF national and international grants, projects and re-engineering internal processes.



Charlotte Armstrong, Project Officer

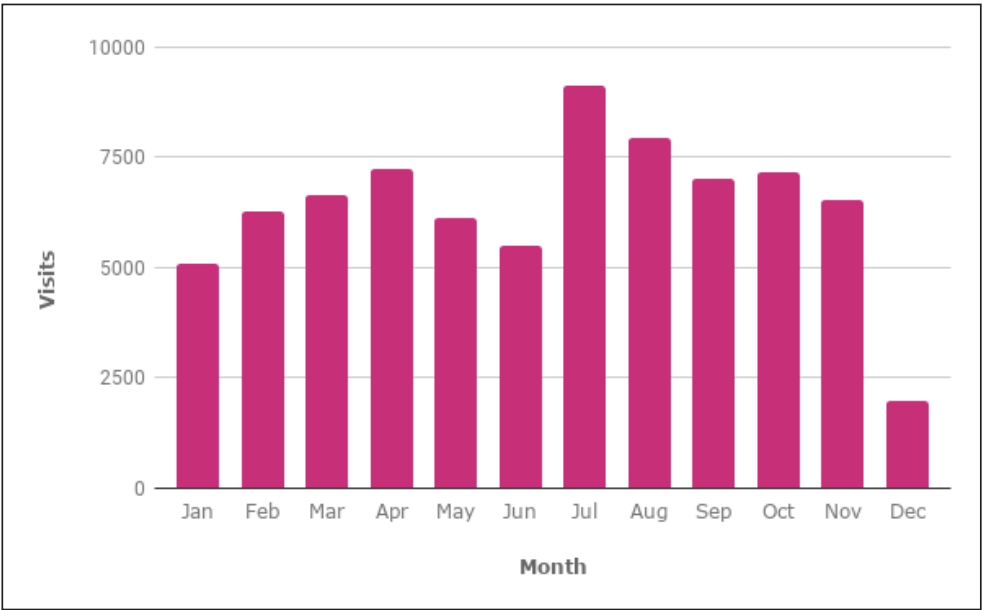
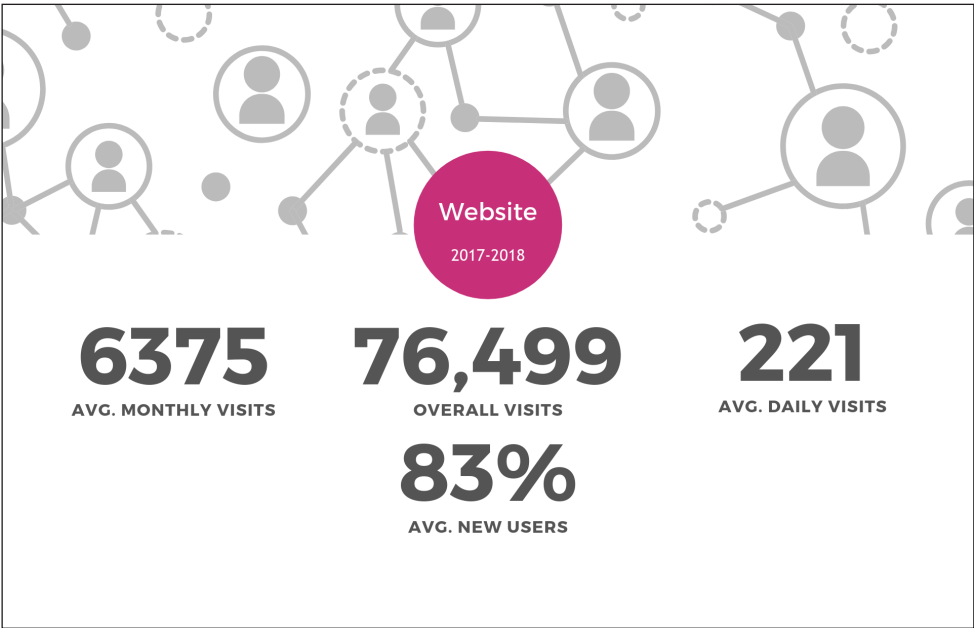
Charlotte is the OPF's Project Officer, helping to manage the OPF office as well as supporting our outreach and community activities including workshops and social media.



OUTREACH & COMMUNICATIONS

The Foundation maintains a number of outreach and communications platforms to ensure effective and open communication that begins with OPF members and extends throughout the global digital preservation community. This year, we have seen an increase in our social media following and mailing list subscriptions, and we welcome the opportunity to continue developing our communications activities in line with this growth. Through these channels, we aim to share our vision and mission; plan, coordinate and publicise our programme of events and webinars; and gain valuable feedback and insights on our projects and activities.

Our website (www.openpreservation.org)



Comms
2017-2018

1313
MAILING LIST
SUBSCRIBERS

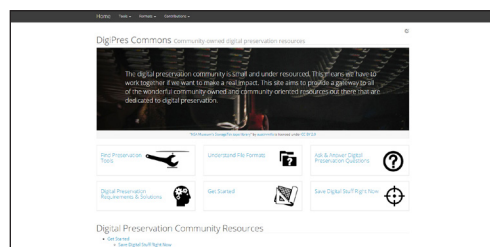
368
LINKEDIN GROUP
MEMBERS

1874
TWITTER FOLLOWERS

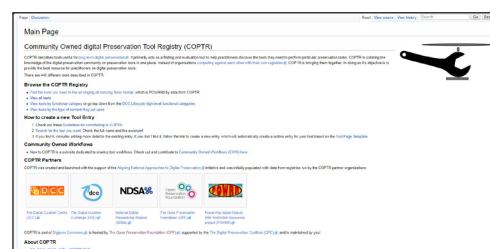
Join the conversation!

Did you know?

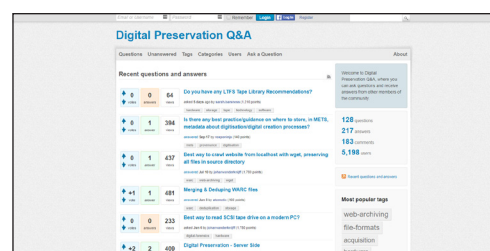
The OPF provides hosting services for valuable community resources including:
Digipres.org



COPTR



Digital Preservation Q&A



News highlight: Bibliothèque nationale de Luxembourg joins the Open Preservation Foundation

The Open Preservation Foundation (OPF) is pleased to welcome the Bibliothèque nationale de Luxembourg (BnL) as our most recent charter member.

The BnL has been digitising documents from the Luxemburgensia collection since 2002. Its mass digitisation program falls within the context of the European Union's objectives to make Europe's cultural and intellectual heritage accessible online. Through the development of their Digital Preservation project, the BnL seeks to ensure the management, availability and long-term preservation of digital publications, resulting from a digitisation process, collections of born digital material or coming through the digital legal deposit.

Roxana Maurer, Digital Preservation Coordinator at the BnL said:

"The Bibliothèque nationale de Luxembourg is pleased to join other worldwide member institutions of the Open Preservation Foundation to address the challenges of digital preservation together. As a member of the OPF, the National Library will not just follow international standards and best practices, but also participate in their development through specialist projects, initiatives and task forces."

Martin Wrigley, Executive Director of the OPF said:

"We are very happy to welcome the Bibliothèque nationale de Luxembourg to the OPF and bring the perspective from another new country to our membership. We look forward to working with them and facilitating our members to share experience and best practice."

JOIN US


Joining the Open Preservation Foundation helps us continue our work to sustain and develop open source digital preservation tools that are relied upon by our members and the community. Membership fees go towards both developing software and to enhancing our programme of best practice and knowledge sharing activities such as our webinars, interest groups and workshops.

Members join an international, collaborative network and benefit from a practical and technical approach to digital preservation. Members influence our strategy, activity planning and adoption of new tools. Charter members are eligible to serve on the board of directors and the product board. Members benefit from technical support, interest group participation, priority access to our webinars and events and exclusive access to member content on our website.


There are four ways to support the Foundation's work:

- Become a **Charter member**
- Become an **Affiliate member**
- Become a **Technical Contributor**
- Make a **Donation**


Find out more at: www.openpreservation.org/join




Join a global community




Access member content



Influence OPF's strategy, tools and resources



Get OPF guidance and support



Open
Preservation
Foundation

<http://openpreservation.org>



[@openpreserve](#)



<http://openpreservation.org/subscribe>



<https://github.com/openpreserve>



info@openpreservation.org



www.openpreservation.org

Open Preservation Foundation

c/o The British Library, Boston Spa, Wetherby, LS23 7BQ
info@openpreservation.org

Published December 2018