

## Shaping our future memory standards (1/2)



Source: [Open Preservation Foundation blog post by Becky Mc Guinness](#)

The final conference of the PREFORMA project '[Shaping our future memory standards](#)' was held at the National Library of Estonia in Tallinn last week. It was well-attended and had a nice balance of presentations, panels and demonstrations. [Presentations](#) from the conference are already online, and the event was recorded.

### Day 1



Participants were welcomed by Kristel Veimann, National Library of Estonia and Tarvi Sits, Estonian Ministry of Culture. Borje Justrell, PREFORMA coordinator gave an overview of the project which aims to give memory institutions full control over the digital files they receive by ensuring they conform to the format specification. PREFORMA is now in the final phase of testing the three conformance checkers for PDF/A ([veraPDF](#)), TIFF ([DPF Manager](#)) and Matroska, Linear Pulse Code Modulation and FF Video Codec 1 ([MediaConch](#)). As well as developing software, each supplier has contributed to the creation or improvement of the respective standard specification. The conformance checkers check that files conform to the standard and allow memory institutions to apply their own policy restrictions. Borje remarked that there are still a large number of other formats for which there is no conformance checker available.



Julia Kim from the American Folklife Center at the Library of Congress gave a talk called 'Why can't they (files, formats, software, users, standards?) all just get along?' She feels the value in the PREFORMA project is not just in the tools and papers published, but also development of a community.

Institutions increasingly rely on tools where documentation is not available while standards are not necessarily widely followed or adopted and can be interpreted differently. This can be exacerbated by a lack of clarity as to where responsibility for these tools and standards lie. A [study by the National Digital Stewardship Residencies](#) titled 'What makes a digital steward?', ranked standards and best practices skills last in importance.

Julia described the language of PREFORMA as revolutionary; 'taking full control' is at odds with the neutrality statements often used by memory institutions. Archivists are usually at the end of the lifecycle and it is difficult for some organisations to enforce submission guidelines after content has been created.



PREFORMA provides a unique opportunity to determine the rules of the game. Many want a yes or no answer to their questions. The Federal Agencies Digital Guidelines Initiative (FADGI) has similarities to PREFORMA. If they want new tools, standards or technical guidelines they collectively commission a third party contractor to create them. FADGI also helps to create long lasting relationships amongst the agencies, and promotes widespread adoption of the standards and outputs.

The talk ended with a question about what happens after PREFORMA ends. It is important that the tools and relationships are sustained and there an opportunity to expand the model for other formats.

The morning session concluded with pitches for the validator [demos](#) and [posters](#).



After lunch the conference split into two streams. There were hands-on demonstrations for those who wanted to find out more about the conformance checkers. Delegates were shown how to install, configure and use the PREFORMA software. Meanwhile, the conference programme continued with a talk about the digital preservation landscape and the challenges and opportunities by Raivo Ruusalepp from the National Library of Estonia.

Estonia currently hosts its [first presidency of the Council of the European Union](#). This involves a large number of meetings, discussing topics such as culture, education and e-government. Raivo noted that the words efficiency, access and trust are often mentioned when addressing these topics, but preservation is not. A contributing factor to this is that the standards in our community are not as widely adopted as standards in other communities.



Raivo compared digital preservation to cooking, and described how we have our own traditions (processes), but adopt ingredients (standards, tools) and make them our own.

At the beginning of 2017 a new digital legal deposit law was introduced in Estonia. The library now receives a copy of anything serialised or printed deposited as a digital version. Digitisation is a finite project. It is a new era and they do not need to digitise contemporary material. There has been an anticipated step up in the volume and complexity of files they receive and, in response, the library updated its digital preservation policy, revised its list of accepted file formats and has defined new service levels based on the NDSA matrix of digital preservation services.

The library has implemented more automated processes in their workflows to check the quality of files they receive. They have reconceptualised their digital preservation system to encompass all content types. Information needs to outlive the system(s) that produced it and standards act as tools for systems' interoperability.

They need to make appraisal decisions for the future. This is a question of judgment when building collections, helping inform the decision regarding the type of repository required. The library is focusing on ingest to reduce the preservation workload, and is pushing responsibility upstream to the creators. They plan to use veraPDF to not only to defend position of library, but to raise awareness amongst publishers about the quality of the content they are producing.

There are lots of tools, widgets and services available, too many. They recently assessed over 200 digital preservation tools to find out if they are still alive or supported and if documentation is available. They found that only a fraction actually are maintained. The community still needs similar software to the conformance checkers that PREFORMA has commissioned. Memory institutions do not have specialist staff for every format they need to preserve. Returning to his cooking metaphor, Raivo said there is a maturity in the field; organisations can choose to be in the 'kitchen' and create their own products, or they can commission a service or tool they need 'go to a restaurant. By investing in open source tools, organisations can understanding what it happening inside them.

Raivo called on organisations to become more resilient. We should not just be looking at longevity, sustainability, and automation but we should 'have the capacity to prepare for disruptions, recover from shocks and stresses, and adapt and grow from a disruptive experience'.

He concluded that as digital preservation continues to mature, we as a community need to:

- Learn to be resilient
- Make good use of standards
- Build new competencies and skills



Natasa Milic-Frayling from the UNESCO PERSIST Programme was up next talking about 'Safeguarding Digital Heritage through Sustained Use of Legacy Software'. She explained that innovation happens because there is demand in the market. The biggest danger to software sustainability and access to content that is rendered on that machine, happens when the company who created the software no longer exists. Open source is a good way of eliminating this dependency, however, there is still little understanding about the standards.

The digital ecosystem is complex with a large number of dependencies and different layers. As well as the technologies used such as the operating system and servers, and the digital assets that are created such as data streams and documents, the user experience has become more important. Digital assets cannot be used without a programme to interpret the bits and consistency is important.

Both digitised and born digital copy are subject to the same issues of obsolescence, and they still need a reader to view e.g. PDFs. The question is, how do we enable prolonged use of software? We are facing different threats: the hardware may still be available, but the expertise about that system may disappear if the creating organisations folds. Standards evolve, we need to be prepared for the next versions and think about how to manage the lifecycle of a document. IT companies can bring technology to the community and take it away again? their goal is to attract customers, they do not necessarily share a vision for long term preservation.

There are a number of digital preservation strategies used today. Migration means you say goodbye to the original file, and perhaps lose some interaction within it. Another option is to keep the original file and port the application to the new environment. You could virtualize legacy software environment through a virtual machine using the old computing stack and run the original files and software. Natasa recommended a hybrid strategy combining virtualization and format transformation.

She introduced the UNESCO memory of the world programme and their objective to:

- Ensure that documentation is available to all, without barriers and obstructions
- Embarked on digitization of physical artefacts to preserve and disseminate information and preserve cultural heritage

They plan to establish a foundation to host legacy software and negotiate licences with industry e.g. Microsoft.



The first day finished with a panel exploring experiences from memory institutions using the PREFORMA tools. Five organisations from the PREFORMA consortium were represented on the panel, chaired by Bert Lemmens from PACKED vzw. As an introduction, panelists were asked to raise their hands to some quick questions about the PREFORMA tools:

- All had used the tools, and had convinced colleagues to try them
- 3 of 5 are deploying the tools in their production environment
- 3 of 5 would be willing to pay for development of the tool

One of the issues with adopting the tools is that investigating the standards they are using internally takes time. Adoption of tools comes after the adoption of standards.

Although the PREFORMA project has been running since 2014, the tools have only reached a production-ready level within the last year. It is important to test the tools, and integrating a tool into a production workflow takes time. They need to consider if it is efficient and how it works with what they already have in place.

The panelists commented that they had found the three suppliers very responsive to their questions and willing to help and makes changes to the software. Feedback from both sides has been really important and has created very productive relationships.

There are still some issues with trust with regards to open source from management. There is a perception by some that 'you get what you pay for' so if the software is free, or the organisation does not pay a lot, then it is seen as low quality. In the past, software was adopted by institutions through a procurement process. Without this framework, institutions are having to rethink the process of adopting the software they need. With the PCP model, much more engagement is needed. However, the positive aspect of this is the ability to influence and get exactly what you want.

The next step is to bring the tools to a wider market. One of the key offerings the panel would like to see is expertise. They would like help to analyse files and understand the issues and what they mean for the archival process.

In terms of lessons learned, the panelists all felt they had had an impact on evaluation of the results. Overall they found it a positive experience as vendors, developers and archivists pooled their expertise to create the final products.

The panel remarked that it is very important to think about file formats and standards internally. PREFORMA has laid the groundwork for increasing knowledge about file formats within their institutions and helped them to prepare for the future.

Notes from day 2 coming next week...