iPRES PREFORMA Workshop

Between the mountain passes and river bends north of the Alps mountain range lies the de facto capital of Switzerland: **Bern**. It's a UNESCO world heritage site and just outside of it lies BernExpo, a square, meta-coloured building that's home to toy fairs as well as this year's iPRES conference.

iPRES is an annual gathering of preservationists, including archivist, librarians, data storage gurus, and digital art curators. What makes the crowd special is that it brings together a range of people with various backgrounds, storing various kinds of materials for the long term, be it scanned books, video games from the 1980s, net art works, or digital born film and video. The conference brings a mix of talks and workshops. The [PREFORMA](#) team organised one of those workshops. We gathered round on Wednesday afternoon to (1) dive into the larger narrative of the project (giving preservationists the tools and control to check their files' conformity), (2) show the three different tools in development (VeraPDF, DPFManager, MediaConch), (3) detail the three standardisation strands (PDF/A, TI/A and CELLAR) and (4) get a conversation started between (potential) users of the tools with the people who are hard at work making them happen.

After an introduction of the project and development work, we split up in four working tables. In these small groups we discussed organisations' individual needs and questions. At table 1, led by **Börje Justrell** and **Erwin Verbruggen**, we discussed i**ntegration opportunities and future challenges** with people who work on projects such as the UK National Archives's technical registry [PRONOM](#), the [checkit_tiff](#) conformance checker for baseline TIFFs and the [Arcsys](#) record management software. The conversation touched upon how to integrate the PREFORMA suite of tools with any and either one of these tools, some of which have overlapping functionality. One of the participants indicated to be interested in writing a wrapper for the DPF Manager to fit into the Rosetta preservation system they use. All agreed that it's important to give extensive information about what to do with errors: if an institution does not have the technical expertise to judge a conformance error message, the error messages should help them decide what to do with that information - ignore the issue, repair the metadata, or reject the file. This type of error explanation could potentially be tied to preservation service levels - as files and collections treated under a certain level might incur a different approach than files and collections in a higher, more stringent category. The big challenge to normalising these error warnings is one

that is present within the PREFORMA, namely in how a standard API and policy structure can bring together the three toolsets.

At the table that discussed **audio-visual formats**, **Ashley Blewer** and **Jérôme Martinez** discussed with their table mates the same topic of knowing which policy to apply. The archives present discussed the idea of having a platform for sharing policies, where users could interact and exchange knowledge about what policy to use in what case, a feature forthcoming in the [MediaConch](#) online platform. The MediaConch team is seeking people interested in supporting further development for different formats such as JPEG 2000, or MP4. The team furthermore discussed the possibility of using MediaConch as a simple qc-tool for all media formats, with which doing a simple headcheck is a possibility. All around the table agreed on one thing: preserving video is hard work!

The table that discussed the [DPF Manager](#) and the **TI/A standardization work**, led by **Erwin Zbinden**, talked about appropriate standards and how welcome the TI/A work is after the many confusing TIFF that exist. The group concluded that the DPF Manager is a useful tool, but that some of the TIFF tags can currently not be evaluated - it can for instance not yet tell what images are tiled or stripped. Other topics included the openness or propriety of compression algorithms in TIFF and the wish for a formulated policy check that complies with future standards & recommendations.

At the fourth table, **Boris Doubrov** and **Joachim Jung** received a lot of input on how to proceed with the [veraPDF project](#). Users indicated that the tool is nice but that files in use often aren't PDF/A. veraPDF is therefore looking into strategies on how to extend the tool ? yet the PDF standard is of massive volume, and Joachim indicated that perusing version 1.7 would cost 10x as much time as what's put in already. The team therefore indicated it might be useful to divide up the standard to progressively do work around it in waves. Besides technical discussions on what the hardware needed for the tool looks like, the group further discussed the potential of the tool to give feedback to the developers through providing extensive statistics. Also at this table, the topic of errors significance came up: it's nice to have warnings, but what do they mean for your actions or for what software allows you to open the document? The table proposed using a rating system to indicate the gravity of these warnings.

All in all, the workshop gave us not just food for thought, but also concrete challenges to further provide solutions for in the framework of the PREFORMA project!

Related links:

- Workshop Notes for ?Sharing Using and Re-Using Format Assessments?]

Other iPres blogs:

http://wp.lancs.ac.uk/highly-relevant/2016/10/10/ipres-2016-international-digital-preservation-conference-bern-switzerland/

http://digital-archiving.blogspot.nl/2016/10/some-highlights-from-ipres-2016.html

http://www.dpconline.org/publications/whats-new/1797-ipres16-oais#.V_d8tH1aZFc.twitter

http://blog.dshr.org/2016/10/software-art-and-emulation.html