KOST-Val released by the TI/A Standard Initiative team

KOST-Val is an open source validator for different file formats (TIFF, SIARD, PDF/A, JP2, JPEG) and Submission Information Package (SIP).

It has been developed by **KOST-CECO**, is a Swiss coordination office which is member of the **TI/A Standard Initiative** team, a group of experts focussing on the definition of a specification of a Archival TIFF Format

For futher information visit the **KOST-Val page in the Community Owned digital Preservation Tool Registry (COPTR)**.

Funtional Principle

KOST-Val complies with the following requirements.

- TIFF validation: KOST-Val reads a TIFF file and uses **JHOVE** to validate the structure, the content, and **ExifTool** to validate the key properties such as compression, colour space, and multipage. These properties can be configured.]
- SIARD validation: KOST-Val reads a SIARD (eCH-0165 v1) file and validates the structure and the content.
- PDF/A validation: KOST-Val reads a PDF or PDF/A file (ISO 19005-1 and 19005-2) and uses 3-Heights? **PDF/A Validator** by PDF-Tools or **PDF/A Manager** by PDFTron to validate the structure and the content of the PDF file. KOST-Val organises the different error messages into main categories such as fonts, graphics, and metadata. KOST-Val supplies only a limited version from 3-Heights? PDF/A Validator by PDF-Tools. Module J extracts (with **iText**) and validates the JPEG and JP2 images contained in the PDF file (depending on the configuration). It is also possible to configure whether the JBIG2 compression is accepted or not.]
- JP2 validation: KOST-Val reads a JP2 file (ISO 15444) and uses **Jpylyzer** to validate the structure and the content.]
- JPEG validation: KOST-Val reads a JPEG file (ISO 10918-1) and uses **Bad Peggy** to validate the structure and the content.]
- SIP validation: KOST-Val reads an SIP (eCH-0160 v1 as well as Swiss Federal Archives SFA v1 and v4) and validates the mandatory requirements of the SIP specification. The validated requirements are organised into groups such as folder structure, schema validation, and checksum validation. At the outset, a file format validation is performed.

The results (including information on inconsistencies and errors) are output for every step and written into a validation log. The validation steps are executed sequentially. Whenever possible the validation shall continue after an error has been detected in order to reduce the number of correction cycles.

Third-party applications

KOST-Val uses unmodified components of other manufacturers by embedding them directly into the source code. Users of KOST-Val are requested to adhere to these components' terms of licence.

- The TIFF validation module uses **JHOVE** and **ExifTool** and evaluates its output further.]
- For the PDF/A validation module **PDF-A Manager** or **3-Heights PDF/A Validator** are used.]
- The JP2 validation module uses **Jpylyzer** and translates the failed tests into appropriate error messages (DE/FR/EN).]
- The JPEG validation module uses **Bad Peggy** and evaluates the error message "Not a JPEG file" further.]
- To extract the JPEG and JP2 images from PDF/A the **iText library** is used.]

- For the file format identification DROID is used. For performance and granularity reasons an own SignatureFile is used instead of the official PRONOM registry.]

About the TI/A Standard initiative]

The TI/A Standard initiative is promoted by the Digital Humanities Lab of the University of Basel, the Agents Research Lab of the University of Girona and Easy Innova with the support of many interested memory institutions.

This standard will be created in parallel with DPF Manager, an open source TIFF format validator that, in addition to the current TIFF ISO Standards, will be the first conformance checker for the TI/A new standard.

This initiative has been boosted by PREFORMA, a PCP project that aims to address the challenge of implementing good quality standardised file formats for preserving data content in the long term.