Initial Presentation

1. Consortium & People Involved
2. Scope of the Project

DPF Manager
TI/A initiative
Community

Dr. MIQUEL MONTANER
CTO at Easy Innova
miquel@easyinnova.com

EASY INNOVA
www.easyinnova.com
Consortium & People Involved

**Easy Innova, S.L. (Spain)**  
Spin-off of the University of Girona

- **Certified by:**
  - TECNIO
  - AENOR

- **Member of:**
  - Generalitat de Catalunya

- **Know-how:**
  - R&D Projects
  - Platform Architectures
  - Digital Preservation
  - IPR
  - Standards
  - Artificial Intelligence
  - Open Source Projects

- **Dr. Miquel Montaner**  
  CTO

- **Robert Sallo**  
  R&D Manager

- **Xavi Tarrés**  
  Project Manager

- **Antonio López**  
  Senior Developer

- **Dr. Víctor Muñoz**  
  R&D Manager

- **Prof. Dr. Josep Lluís de la Rosa**  
  Full Professor and Researcher

- **Dr. Albert Trias**  
  R&D Manager

- **Prof. Dr. Lukas Rosenthaler**  
  Full Professor and Researcher

- **Dr. Peter Fornaro**  
  Managing Director

---

**University of Girona (Spain)**  
Agents Research Lab

- **Know-how:**
  - R&D
  - Open Source
  - Digital Preservation

**University of Basel (Switzerland)**  
Digital Humanities Lab

- **Know-how:**
  - Cultural Heritage
  - Medical Image
  - TIFF
  - Image Formats
  - Digital Preservation
Scope of the Project

DPF Manager
Open source software
Main features

Standardization process
TI/A Standard Initiative
Research

Early adopters
Developers
TIFF Experts
DPF Manager

1. What is DPF Manager?
2. DPF Manager Features
3. How it Works?
4. Use Scenarios
5. DPF Manager current version
6. What’s Next?
7. Open Source Project
In considering the suitability of particular image for the purposes of digital preservation, relies on the use of a stable, open and well documented file format as well as some data object properties acceptance criteria.

- TIFF Baseline 6.0
- TIFF/EP
- TIFF/IT
  - Tiff/IT P1
  - TIFF/IT P2
- TIFF for Archival (TI/A) to be accepted as ISO Recommendation
- Custom policy rules
DPF Manager Features

Multi-platform

Flexible

Modular

Technologies

- Maven
- Java 8
- Spring
- Travis CI
- JavaFX
- JACPFX
How it Works?

Select the files to process

1. [Image with file types: image, folder, URL, ZIP]

Define your acceptance criteria

2. TIFF baseline 6.0
ISO TIFF/EP and ISO TIFF/IT
User defined policy rules
How it Works?

3 Manage files metadata and apply fixes

4 Get the Report

Multiple report format
Acceptance criteria result
Metadata information
Tiff file structure

palette-1c-8b.tiff

This file does NOT conform to conformance checker

<table>
<thead>
<tr>
<th></th>
<th>Errors</th>
<th>Warnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tiff/EP</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>Tiff/IT</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Policy Checker</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>
New configuration
- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports

Home Screen
New configuration > Implementation Checker

New configuration
  - Implementation checker
  - Policy checker
  - Reporter
  - Metadata fixer
  - Periodical checks
  - Summary

Check files
  - Multiple file report
  - Single file report

View historical data
  - List of previous reports

NEW CONFIGURATION

IMPLEMENTATION CHECKER

Standards
- TIFF/IP (ISO 12369)
- TIFF/IT-P1 (ISO 12639:1998)
- TIFF/IT-P2 (ISO 12639:2004)
- TIFF/A (To be standarized)

Continue
New Configuration > Policy Checker

- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports
New Configuration > Reporter

- New configuration
  - Implementation checker
  - Policy checker
  - Reporter
  - Metadata fixer
  - Periodical checks
  - Summary

- Check files
  - Multiple file report
  - Single file report

- View historical data
  - List of previous reports

---

**NEW CONFIGURATION**

```
Format

- HTML
- PDF
- XML
- JSON

Select HTML template
Select PDF template

Expert mode
```

---

[Continue]
New configuration

- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files

- Multiple file report
- Single file report

View historical data

- List of previous reports

NEW CONFIGURATION

- Automatically fix metadata errors 🔄
- Remove private data 🔄

- Automatically add metadata to files
  - Copyright
  - PREFORMA project
  - Add new field

- Automatically remove metadata from files
  - Author 📈 Date 📈
  - Select field 📈 Add new field to remove

Continue
New configuration > Periodical Checks

- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports

- Run this check once
- Run this check every:
  - Week
  - Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday

- At:
  - 12:00 h

- Send the report by email after each check finishes:
  - admin@archive.org

Continue
New Configuration > Summary

**Check files**
- Multiple file report
- Single file report

**View historical data**
- List of previous reports

**New configuration**
- Implementation checker
- Policy checker
- Reporter
- Metadata fixer

**Summary**
- Periodical checks

<table>
<thead>
<tr>
<th>Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy checker</td>
<td>Policy: 7200px; Color profile = sRGB; Image width = 7200px</td>
</tr>
<tr>
<td>Report</td>
<td>Format: HTML; Default template, expert mode</td>
</tr>
<tr>
<td>Metadata fixer</td>
<td>Fix errors automatically; Remove fields; Author, Date; Add fields; Copyright: PREFORMA project</td>
</tr>
<tr>
<td>Periodical Checks</td>
<td>Every week (Tuesday, Thursday at 12:00 h); Send report by email: <a href="mailto:admin@archive.org">admin@archive.org</a></td>
</tr>
</tbody>
</table>
Check Files > Multiple File Report

- New configuration
- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports
New configuration
- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports

Check Files > Single File Report

### TFF / EP conformance

<table>
<thead>
<tr>
<th>Error</th>
<th>Description</th>
<th>Possible solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tag 279</td>
<td>Color profile is missing</td>
<td>Re-calculate value</td>
</tr>
<tr>
<td>Tag 294</td>
<td>Compressed images are not allowed for archival purposes</td>
<td>Uncompress data</td>
</tr>
</tbody>
</table>

### Warning

<table>
<thead>
<tr>
<th>Tag 298</th>
<th>Description</th>
<th>Possible solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution Unit is missing</td>
<td>This tag is important if no scale is visible in the scene</td>
<td>Check if scale is visible in the scene, and add tag with correct value if not</td>
</tr>
</tbody>
</table>

### Policy checker

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color profile</td>
<td>sRGB</td>
<td>✔️ The color profile is sRGB</td>
</tr>
<tr>
<td>Image width</td>
<td>&gt; 3000 px</td>
<td>✔️ Image width = 3000 px</td>
</tr>
</tbody>
</table>

### Automatically updated tags

<table>
<thead>
<tr>
<th>Type</th>
<th>Issue</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>Tag 33432 Copyright missing</td>
<td>Added tag 33432 with value = PREFORMA project</td>
</tr>
<tr>
<td>Info</td>
<td>Tag 306 Date/Time present</td>
<td>Removed tag 306 Date/Time</td>
</tr>
</tbody>
</table>
New configuration
- Implementation checker
- Policy checker
- Reporter
- Metadata fixer
- Periodical checks
- Summary

Check files
- Multiple file report
- Single file report

View historical data
- List of previous reports
Use Scenarios

1. Standalone
   - Graphic Interface
   - Command line

   C:\>dpfmanager.exe
   - files=/archive/*.tiff
   - config=config.xml
   - reports_folder=/tmp/reports
   - fixed_files_folder=/tmp/files

2. Client - Server
   - Client
     - client:\>dpfmanager.exe
     - server=80.45.32.45
     - port=80
   - Server
     - server:\>dpfmanager.exe
     - mode=server
Web application

Browser Client

Server

High availability

Client

Server
Use Scenarios

Integration with OAIS

Document lifecycle
Conformance checking at:

- Creation
- Transfer
- Migration
- Digitalization

Reference framework
QA, AIP generation

[Diagram of OAIS integration with various components like SIP, Ingest, Data Management, AIP, Access, DIP, Policy, Budget, and Administration]
DPF Manager current version

VERSION 2.0 – Released on April 4th 2016

• Conformance checker
  • Baseline version 6.0
  • TIFF/EP
  • TIFF/IT
    • TIFF/IT P1
    • TIFF/IT P2

• Policy checker
  • Create Policy rules from GUI
  • Extend developing your own policies

• Metadata fixer
  • Add, edit or remove metadata
  • Apply auto-fixes in TIFF files
  • Extend developing your own auto-fixes

• Interfaces
  • Command-line interface
  • Graphical user interface

• Modular architecture
  • Develop your own modules for DPF Manager
What’s Next?

In the next VERSIONS…

https://github.com/EasyinnovaSL/DPFManager/milestones

• Conformance checker
  • TI/A (TIFF for Archival)

• Policy checker
  • New policies to apply

• Metadata fixer
  • New metadata
  • New auto-fixes in TIFF files

• Interfaces
  • Server interface
  • Web interface
  • Schedule checks

• Conformance checkers interoperability
Open Source Project

DPF Manager

Open source licenses
MPL v2+
GPL v3+

Open source project
http://dpfmanager.org/
https://github.com/EasyinnovaSL/DPFManager
http://preforma-project.eu/dpf-manager.html
Initial Presentation

DPF Manager

TI/A Initiative

1. File Format Preservation
2. TIFF Format
3. TI/A Standard initiative
4. Timeline

Community

Dr. PETER FORNARO
Managing Director at University Basel
peter.fornaro@unibas.ch

UNIVERSITY OF BASEL
www.unibas.ch
Even if bit-stream preservation was successful (hardware, data-carrier), the format of the files must be decoded to render the content properly.

A format is good for archiving if it is widely used, well documented, open and allows storage of appropriate quality of the content.

Migration not necessary
TIFF Format

It is there!

- Tiff was the major file (digital master) for archival purposes
  - TIFF is a final rendered image
  - TIFF is 16bit
  - TIFF is lossless
  - TIFF is professional
  - TIFF is multichannel and multilayer

- Many archives and museums store TIFF files
  - because the others do it
  - because it is of large data volume
  - because it is widely used and well documented

88% Memory Institutions

Standarization

- TTIF-Files will be checked for “correctness”.

- TIFF is a complex standard with a lot of possible features. It is not clear to say what “correctness“ means. It can depend on the application.

- However: It is necessary to define and recommend a well-defined set of features that are accepted by most memory institutions.
TI/A Standard Initiative

The approach is well-known:

• A PDF is a full featured digital document, some of them are not wanted in archives.
• A PDF/A is an adopted version, fulfilling archival needs. Some features are **forbidden** some are **required**.
• We follow a similar approach: TIFF will be precisely specified based on a feedback of an expert user community and a survey on „hot“ data.

Recommendation of TIFF for Archives is in progress

Well specified set of mandatory, forbidden and optional tags

Standardization process has started in 2016
Timeline

- Expert Community has been created
- First draft of TIFF for Archives Specification
- Discussion of approach on ICAI16 in Tokyo
- First discussion with ISO/Adobe
- Start of survey with real data
- Discussion of approach on Archiving16 in Washington
- Discussion of proceeding with Adobe in San Jose
- Survey with real data
- Standardization with ISO TC171

Collecting feedback from community

We are here
Initial Presentation
DPF Manager
TI/A Standard Initiative

Community
1. Registered Early Adopters
2. TIFF Experts Involved
3. Join us!

Prof. JOSEP LLUÍS de la ROSA
Full Professor at University of Girona
pepluis@eia.udg.edu

UNIVERSITY OF GIRONA
www.udg.edu
More than 45,000 images checked

Feedback from more than 2,500 reports
TIFF Experts Involved

69 experts in TI/A intranet from 16 countries

support of 61 memory institutions to create the TI/A standard
Join us!

visit www.dpfmanager.org

Early adopters

DPF Manager

Install now
For Windows, Mac OS X and Linux
Version 2.0 - Release (04-04-2016)

visit www.ti-a.org

Tiff experts

Get involved!

visit http://www.dpfmanager.org/community.html

https://github.com/EasyinnovaSL/DPFManager

Developers

New pull request
Thanks for your attention!
FEEL FREE TO ASK US

www.easyinnova.com