

# PREservation FORMAts for culture information/e-archives

*PREFORMA*  
*Future Memory Standards*

# Project Identity Card



- ❑ PREFORMA is a **Pre-Commercial Procurement** project co-funded by the European Commission under FP7-ICT Programme.
- ❑ **Start date:** 1 January 2014
- ❑ **Duration:** 48 month (end date: 31 December 2017)
- ❑ **Total budget for the procurement:** 2.805.000 EUR
- ❑ **Website:** [www.preforma-project.eu](http://www.preforma-project.eu)
- ❑ **Contacts**
  - Project Coordinator: Borje Justrell, Riksarkivet, [borje.justrell@riksarkivet.se](mailto:borje.justrell@riksarkivet.se)
  - Technical Coordinator: Antonella Fresa, Promoter Srl, [fresa@promoter.it](mailto:fresa@promoter.it)
  - Communication Coordinator: Claudio Prandoni, Promoter Srl, [prandoni@promoter.it](mailto:prandoni@promoter.it)



# Project Partners



- ❑ RIKSARKIVET, Sweden **Project Coordinator and memory institution**
- ❑ PROMOTER SRL, Italy **Technical and Communication Coordinator**
  
- ❑ **Technical partners**
  - PACKED EXPERTISECENTRUM DIGITAAL ERFGOED VZW, Belgium
  - FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V, Germany
  - HOGSKOLAN I SKOVDE (University of Skovde), Sweden
  - UNIVERSITA DEGLI STUDI DI PADOVA, Italy
  
- ❑ **Memory institutions**
  - STICHTING NEDERLANDS INSTITUUT VOOR BEELD EN GELUID, Netherlands
  - Koninklijk Instituut voor het Kunstpatrimonium, Belgium
  - GREEK FILM CENTRE AE, Greece
  - LOCAL GOVERNMENT MANAGEMENT AGENCY-AN GHNIOMHAIREACTH BAINISTIOCHTA RIALTAIS AITIUIL, Ireland
  - STIFTUNG PREUSSISCHER KULTURBESITZ, Germany
  - AYUNTAMIENTO DE GIRONA, Spain
  - Eesti Vabariigi Kultuuriministeerium, Estonia
  - KUNGLIGA BIBLIOTEKET, Sweden



# Pre-Commercial Procurement (PCP)



- ❑ Pre-Commercial Procurement (PCP) is a competition-like procurement method, which enables public sector bodies to engage with innovative businesses and other interested parties in development projects to arrive at innovative solutions that address specific public sector challenges and needs.
- ❑ These innovative solutions are created through a phased procurement of development contracts to reduce risk.
- ❑ Pre-Commercial Procurement is becoming more and more common within the public sectors of the European Union.



# Project Concept



- ❑ Memory institutions are facing **increasing transfers** of electronic documents and other media content for long term preservation.
- ❑ Data content are normally stored in specific **file formats** for documents, images, sound, video etc., and these files are usually produced by software from different vendors.
- ❑ Even if the transferred files are in standard formats, the correct **implementation of standards** cannot be guaranteed:
  - The software used for the production of the electronic files is not in control neither by the institutions that produces them nor by the memory institutions.
  - Conformance tests of transfers are done by memory institutions, but are not totally reliable; different software for testing could end up in different results.
- ❑ This poses problems in **long-term preservation**. Data objects meant for preservation, passing through an uncontrolled generative process, can jeopardise the whole preservation exercise.



# Project Aim and Objectives



- ❑ **The aim:** to implement good quality files in various standard formats for preserving content in a long term.
- ❑ **The main objective:** to give memory institutions full control of the process of conformity tests of files to be ingested into archives.
- ❑ **The main objective of the PCP launched by PREFORMA:** to develop an open source software for the management of the whole conformance test process, supporting a range of standards, addressing the needs of any memory institution or other organization with a preservation task.



# Open Source approach



- ❑ PREFORMA is following an **open source approach**, with the aim of establishing a sustainable research and development community comprising a wide range of contributors and users from different stakeholder groups.
- ❑ The open source nature ensures long-term availability of the software, beyond the memory institutions and suppliers involved in PREFORMA.
  
- ❑ Licenses
  - All **software** developed during the PREFORMA project will be provided under two specific open source licenses: “GPLv3 or later” and “MPLv2 or later”.
  - All **digital assets** developed during the PREFORMA project will be provided under Creative Commons CC-BY v4.0, and in open file formats (SOU 2009:86).



# Target users and stakeholders

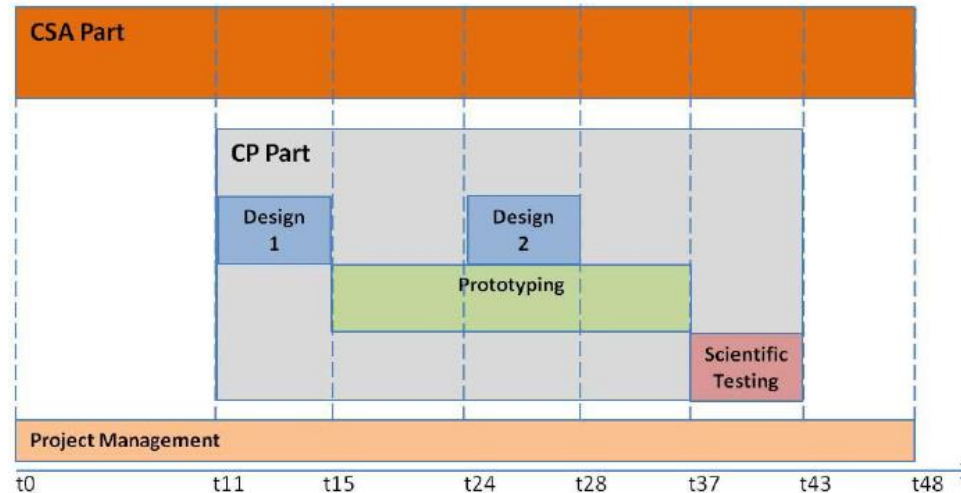


- ❑ **Memory institutions** and cultural heritage organisations, which are involved in (or planning) digital culture initiatives.
- ❑ **Developers** contributing code for the PREFORMA open source tools.
- ❑ **Research organisations** providing technical advice to cultural stakeholders.
- ❑ **Standardisation bodies** maintaining the technical specifications of the preservation formats covered in PREFORMA.
- ❑ **Funding agencies**, such as Ministries of Culture and national/regional administrations, that own and manage digitisation programmes and may endorse the use of the PREFORMA tools in the digitisation process.
- ❑ Other **projects** in the digital cultural heritage domain.





# Project implementation schedule



- ❑ **Design phase** (4 months): November 2014 – February 2015
- ❑ **Prototyping phase** (22 months): March 2015 – December 2016
  - First prototypes: March 2015 – October 2015
  - Re-design: November 2015 – February 2016
  - Second prototype: March 2016 – December 2016
- ❑ **Testing phase** (6 months): January 2017 – June 2017



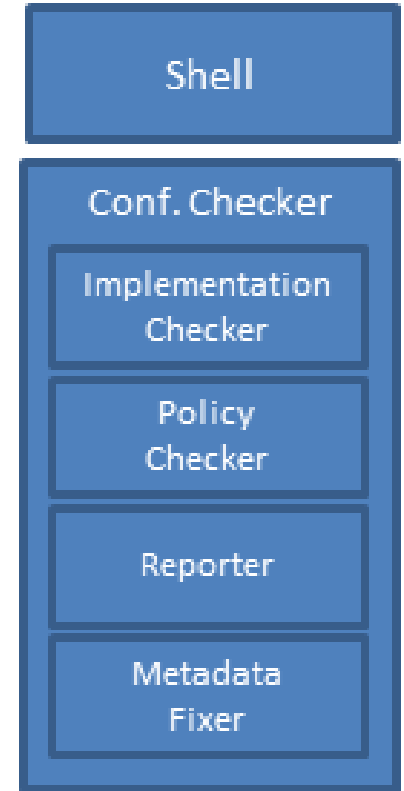
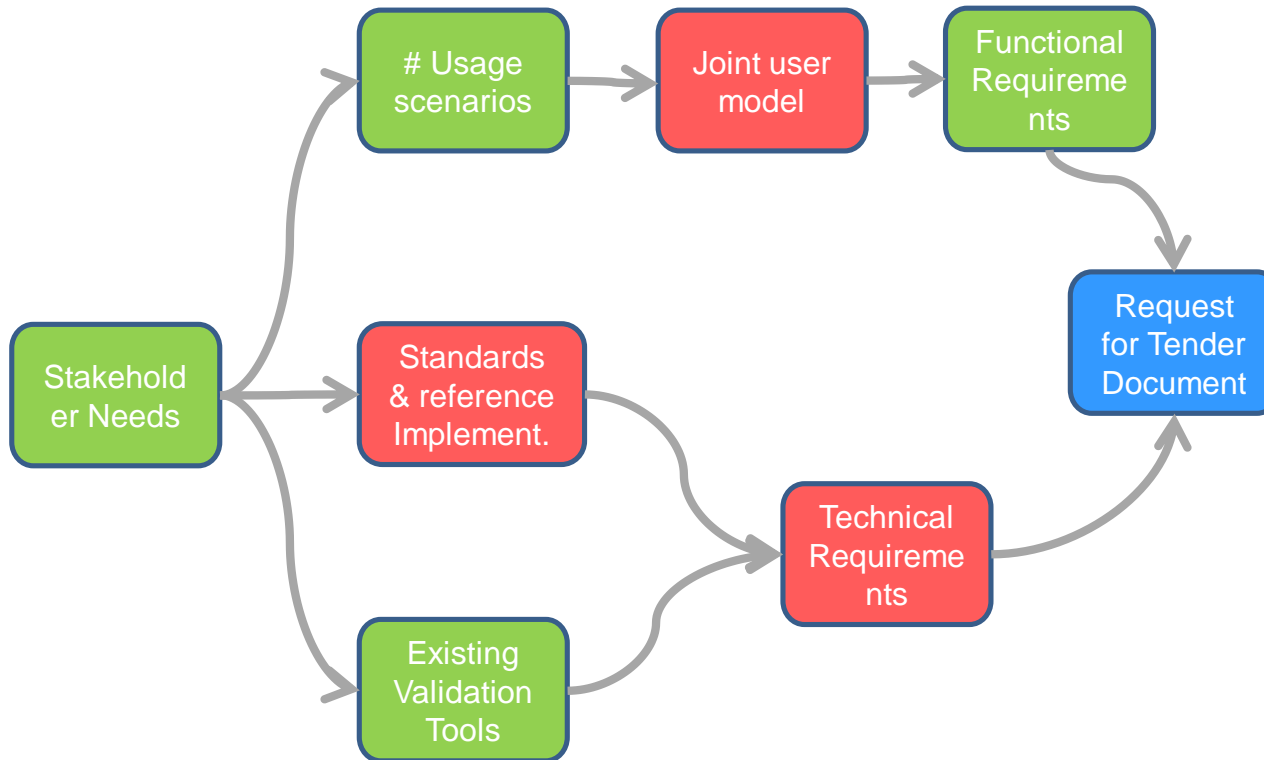
# Overall R&D Objective (The PREFORMA Challenge)



- ❑ Develop an **open source conformance checker** that:
  - checks if a file complies with standard specifications
  - checks if a file complies with the acceptance criteria of the memory institution
  - reports back to human and software agents
  - perform simple fixes
  
- ❑ Establish an ecosystem around an **open source reference implementation** that:
  - generates useful feedback for those who control software
  - advances improvement of the standard specification
  - advances development of new business cases for managing preservation files



# Conformance Checker



# Reference Implementation (standard specifications)



Content type	Standard specifications
TEXT	<b>PDF 1.7</b> (ISO 32000-1:2008) <b>PDF/A-1</b> (ISO 19005-1:2005) <b>PDF/A-2</b> (ISO 19005-2:2011) <b>PDF/A-3</b> (ISO 19005-3:2012)
IMAGE	<b>TIFF/EP</b> (ISO 12234-2:2001) <b>TIFF/IT</b> (ISO 12369:2004)
AUDIOVISUAL	<b>MKV</b> ( <a href="http://www.matroska.org/technical/index.html">http://www.matroska.org/technical/index.html</a> ) <b>Ogg</b> ( <a href="https://xiph.org/ogg/doc/">https://xiph.org/ogg/doc/</a> ) <b>Lossless JPEG2000</b> (ISO/IEC 15444-1:2004) <b>Lossless FFV1</b> <a href="http://www.ffmpeg.org/~michael/ffv1.html">http://www.ffmpeg.org/~michael/ffv1.html</a> <b>Lossless Dirac v2.2.3</b> ( <a href="http://diracvideo.org/download/specification/dirac-spec-latest.pdf">http://diracvideo.org/download/specification/dirac-spec-latest.pdf</a> ) <b>Linear PCM</b> (IEC 60958-1 ed3.1 Consol. with am1: 2014)



# PREFORMA Suppliers in the design phase



1. veraPDF Consortium (led by Open Preservation Foundation and PDF Association) – The PDF/A conformance checker accepted industry-wide (PDF/A)
  2. Preservica - Digital Preservation Validation Framework (PDF/A)
  3. EasyInnova – Digital Preservation Formats Manager (TIFF)
  4. LIBIS - LIBIS/Aware checker for TIFF (TIFF)
  5. MediaArea – PREFORMA MediaConch - CONformance CHECKing for audiovisual files (MKV|FFV1|LPCM)
  6. UCL - OpenMediaCheck (MKV|JPEG2000|LPCM)
- In red those who have been selected for the prototyping phase



# How to contribute – developers & researchers



- ❑ Access to the PREFORMA Open Source Portal at <http://www.preforma-project.eu/open-source-portal.html>
- ❑ Get in contact with the PREFORMA suppliers to contribute to the open source projects

A screenshot of the PREFORMA Open Source Portal website. The page has a dark grey background with white and yellow text. At the top left is the PREFORMA logo. To its right are the European Union flag and the Seventh Framework Programme logo. Further right is a text box stating: "This project has received funding from the European Union's Seventh Framework Programme under grant agreement no 618566". Below this is a navigation menu with links: HOME, PROJECT, PARTNERS, TENDER, ACTIVITIES, OPEN SOURCE PORTAL, COMMUNITY, DOWNLOAD, CONTACTS. The main content area is titled "OPEN SOURCE PORTAL" and contains an introductory paragraph: "This section provides an overview and references to each open source project that is currently working in the prototyping phase. It acts as an entry point for all interested suppliers and memory institutions allowing easy navigation to all externally hosted resources." Below this are three project entries:

- PROJECT N. 1. VeraPDF: THE PDF/A CONFORMANCE CHECKER ACCEPTED INDUSTRY-WIDE**  
by *Open Preservation Foundation, PDF Association, Digital Preservation Coalition, Dual Lab, KEEP SOLUTIONS*  
A unique collaboration, the VeraPDF Consortium brings together an end user community and a software industry rooted in the principle of interoperability based on ISO standardized technology... [access project page >>](#)
- PROJECT N. 2. DPF MANAGER: DIGITAL PRESERVATION FORMATS MANAGER**  
by *Easy Innova*  
DPF Manager is an open source modular TIFF conformance checker that is extremely easy to use, to integrate with existing and new projects, and to deploy in a multitude of different scenarios... [access project page >>](#)
- PROJECT N. 3. MEDIACONCH - CONFORMANCE CHECKING FOR AUDIOVISUAL FILES**  
by *MediaArea.net*  
MediaConch is an extensible, open source software project consisting of an implementation checker, policy checker, reporter and fixer that targets preservation-level audiovisual files for use in memory institutions... [access project page >>](#)

On the right side of the page, there are sections for "PREFORMA OPEN SOURCE PROJECTS" and "OTHER EXISTING TOOLS". Under "PREFORMA OPEN SOURCE PROJECTS" are links for "PDF/A CONFORMANCE CHECKER", "DPF MANAGER", and "MEDIACONCH", followed by a link: ">> View all the successful proposals that participated to the design phase". Under "OTHER EXISTING TOOLS" are links for "ARCHIVEMATICA", "JPPLYZER", and "MEDIA FILE CHECKER".

# How to contribute – Memory institutions



- Provide datasets to be used to test the prototypes that are being developed by the PREFORMA suppliers (including valid/conform files as well as corrupted files):
  - **Training data**  
To be used by the suppliers for training the software components during the prototyping phase
  - **Test data**  
To be used by the suppliers and by the PREFORMA partners for testing the software components during the prototyping and testing phases
  - **Demonstration data**  
To be used by suppliers, by the PREFORMA partners and by third parties for demonstrating and endorsing the PREFORMA solution



# Lesson Learned 1



## The added value of working together

- ❑ Working together for the implementation of a **joint procurement** proved to be a good practice for memory institutions, as it leads to a clearer definition of the requirements and to a better control on the process as a whole and particularly on the results
- ❑ The combination of **different expertise** represents an added value for the project
  - Memory institutions are in charge of defining the requirements, starting from what they actually need
  - Academic partners support the memory institutions with scientific and technological advice
  - Technical partners provide their legal and management expertise to implement and monitor the tender process
- ❑ Collect all the **contact details** of the people which show interest to contribute to the project as a starting point for the establishment of a **sustainable community**





# Lesson Learned 2



## Avoid self-referential approaches

It is important to involve **experts outside the project**, to ensure a wider impact beyond the project itself:

- The **Advisory Board** will contribute to the definition of the challenge
- The **External Evaluators** will complement the expertise of the partners
- The **Associate Partners** will participate to enlarge the Network of Common Interest, through Cooperation Agreements
- The **Associate Projects** will cooperate through dedicated Memorandum of Understanding, e.g. to disseminate together information to common audiences
- **Public events**, and **public presentations** will spread the knowledge and engage in an open dialogue, all around the EU



# Lesson Learned 3



## Sound management and reliable procedures

To ensure the smooth progress of the PCP, it is necessary to establish a **good management structure** made of:

- Different **roles/expertise** should be defined: general project coordinator, technical coordinator, scientific coordinator, legal advisor, expert on licensing and open source, communication manager
- Clear and precise **procedures** are necessary; they should be established and agreed in advance, particularly for the monitoring, assessment and evaluation of the results of each phase
- **Tools** should be defined to allow an easy and fruitful collaboration, e.g. mailing lists, documental repository, virtual room, etc.
- Exact **scheduling** of each phase should be detailed and respected to offer a sound reference to the suppliers



# Lesson Learned 4



## Methodologies and procedures

- ❑ The PCP instrument is still novel. A two-way channel of communication with EC and other projects is important to re-use existing knowledge coming from previous projects and to provide back experience to others:
  - Re-use standard **methodologies** for the evaluation of the results of the PCP. Such as those use by the **EU** (individual evaluations followed by consensus meeting, hearings and eventual panel meetings)
  - Re-use **best practices from other PCP projects**, e.g. templates for the Framework Agreement and the Tender Form
  - Share the experiences gathered in the project with other PCP, for example participating in the annual concertation meetings
- ❑ Analyse in details any possible issue related to copyrights and IPR, with particular regard, in the case of PREFORMA, to **open source licenses**



# Lesson Learned 5



## **The importance of human relationships**

Trust has a very important role in the implementation of the PCP and, in addition to sound procedures, it requires a reliable communication between the consortium and the suppliers:

- The competition should be intended as an opportunity for everybody to learn and for this reason its implementation should be fair
- The consensus model, whose implementation took 2 months in our case, is a complex exercise where the role of everybody should be constantly respected
- The communication channels between the partners and the suppliers should be reliable and effective, for the management of administrative and monitoring purposes
- Answers to possible doubts and questions must be timely and appropriate



# Forthcoming events



- 1. Training event for Open Source companies** that will take place in Stockholm in April 2016, in connection with the first Prototype Demonstration.
- 2. Experience Workshop** that will take place in Berlin in December 2016, in connection with the second Prototype Demonstration, where the PREFORMA partners will share with memory institutions their experiences of working with suppliers under R&D services agreements.
- 3. Final Conference** that will take place in Stockholm in December 2017 to present the results of the project.



# Follow us!

## PREFORMA Website

[www.preforma-project.eu](http://www.preforma-project.eu)



A screenshot of the PREFORMA website homepage as of Thursday, 23 April 2015. The page features a navigation menu with links for WEBSITE, PROJECT, PARTNERS, TENDER, ACTIVITIES, OPEN SOURCE PORTAL, COMMUNITY, DOWNLOAD, and CONTACTS. Below the navigation is a header with the European Union flag, the PREFORMA logo, and the Seventh Framework Programme logo. The main content area is divided into several sections: a login area for 'admin', a media partner section for Digital Culture, a search bar, and a 'RELEVANT NEWS' section with two articles. The first article is titled 'PREFORMA, FUTURE MEMORY STANDARDS' and discusses a new Pre-Commercial-Procurement project. The second article is 'PREFORMA Starts Prototyping Phase', mentioning the awarding of contracts to three consortia. On the right side, there is a 'COORDINATOR' section listing Riksarkivet, a 'TECHNICAL COORD.' section listing PROMOTER, and a 'PARTNERS' section listing PACKED, Fraunhofer IDMT, and the University of Skövde.

A screenshot of the PREFORMA Open Source Portal. The page features a navigation menu with links for HOME, PROJECT, PARTNERS, TENDER, ACTIVITIES, OPEN SOURCE PORTAL, COMMUNITY, DOWNLOAD, and CONTACTS. The main content area is divided into several sections: a 'VISIT THE OPEN SOURCE PORTAL' section with a call to action 'Give your contribution to the prototyping phase', a 'PREFORMA, FUTURE MEMORY STANDARDS' section with a brief description of the project, and an 'UPCOMING EVENTS' section listing a 'DAVID - FINAL USER-TEST-WORKSHOP' in Vienna on 28-29 April 2015. The page also features the European Union flag and the Seventh Framework Programme logo.

## PREFORMA Blog

[www.digitalmeetsculture.net/projects/preforma/](http://www.digitalmeetsculture.net/projects/preforma/)

PREFORMA General Presentation

