



Workshop: Digital Preservation of Cultural Data – Introduction

Antonella Fresa

Technical Coordinator of DCH-RP

Promoter Srl

Agenda



- ❑ 14:30 – 14:45 Welcome and introduction (Antonella Fresa, Promoter Srl)

First Part: DCH and the e-infrastructures

- ❑ 14:45 – 15:10: Using EUDAT services to replicate, store, share, and find cultural heritage data in Poznań Supercomputing and Networking Center (Maciej Brzeźniak, Poznan Supercomputing and Networking Center – Damien Lecarpentier, CSC – IT Center for Science)
- ❑ 15:10 – 15:35 Authentication and Authorisation in the Cultural Heritage community (Roberto Barbera, Istituto Nazionale di Fisica Nucleare)
- ❑ 15:35 – 16:00 Scalability in preservation of cultural heritage data (Simon Lambert, Scientific Computing Department – STFC)
- ❑ 16:00 – 16:30 Break

Second Part: OAIS model, standards, provenance and authenticity

- ❑ 16:30 – 16:50 Standard models and formats for digital preservation (Börje Justrell, Swedish National Archives)
- ❑ 16:50 – 17:15 Implementation of authenticity evidence record model for supporting preservation scenarios (Luigi Briguglio, Engineering R&D Lab)
- ❑ 17:15 – 17:30 Coordination of digitisation, digital access and digital preservation in Sweden (Sanja Halling, Digisam)
- ❑ 17:30 Conclusions (Antonella Fresa, Promoter Srl)



Scope of the workshop



- ❑ To bring together projects and initiatives in the domain of the digital preservation, with a specific focus on digital cultural heritage, digital arts, digital performances and digital humanities
- ❑ To identify common goals and strategic approaches, to find synergies, to discuss opportunities for cooperation, starting from concrete use cases
- ❑ To progress towards establishing a Virtual Research Community for DCH and to prepare for Horizon2020

CH needs and requirements



- ❑ The investment in the production of DCH data is extremely high because the description of each object requires the **human intervention** of experts in the sector
- ❑ DCH content is made of several different kind of information and context data are **complex** and **precious**: 2D images and 3D models, metadata, publications, digital exhibitions, virtual reconstructions, etc.
- ❑ Even if growing very rapidly, the actual **size** of this content is still very much **smaller** than the amount of data produced by experiments and observations of the “hard sciences”.
- ❑ For cultural data, the digitisation process **cannot be** just **replicated** (as a physical experiment, for example)... therefore, preservation is even more important

Expenditures



- ❑ The estimated total cost of digitising the collections of Europe's museums, archives and libraries, including the audiovisual material they hold is approximately €100bn, or €10bn per annum for the next 10 years
- ❑ The cost of preserving and providing access to this material over a 10-year period after digitisation would be in the order of €10bn to €25 bn, provided that centralised repository infrastructure is made available for the purpose
- ❑ Sources
 - NUMERIC Study Report: http://cordis.europa.eu/fp7/ict/telearn-digicult/numeric-study_en.pdf
 - ENUMERATE Survey Report on Digitisation in European CH Institutions 2012: <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2012.pdf>
 - EC Comité des Sages Report on Cost of Digitising Europe's CH: http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/annexes/digital_report.pdf

What CH can ask to the e-infrastructures (1)



- ❑ To support the **permanent identification** of digital cultural objects and providers
- ❑ To facilitate **storage** and **preservation**, ranging from short-medium- and long-term
- ❑ To improve **search facilities** to manage **semantic search** and **linked open data**
- ❑ To enhance **processing** and **visualisation** of **complex** cultural data (e.g. 3D modelling and VR representations) through the computing resources offered by research e-infrastructures (both grid and cloud)

What CH can ask to the e- infrastructures (2)



- ❑ To enable dynamic distributed virtual organisations, facilitating **collaboration** with **information** and **resource sharing** (e.g. virtual conferences, document sharing, blog and cooperation platforms, etc.)
- ❑ To allow for **cost reduction** in digitisation, cataloguing and metadata generation by substituting expensive human workforce with cheaper machine processes
- ❑ To contribute to **standardisation** in the data world, e.g. by developing a common reference model for the DCH sector

An infrastructure dedicated to the cultural heritage



- ❑ The longer term plan is to implement a dedicated CH infrastructure able to provide **distributed safe storage, seamless powerful access**, effective **preservation services** to a “continuum” of data coming from the largest number of small and big cultural institutions
- ❑ This infrastructure should be able to **aggregate and interoperate** existing features and services that are under development, often (unfortunately) following separate ways
- ❑ Implementation of the **preservation** services of such infrastructure is among the first priorities for the DCH sector

Digital Preservation



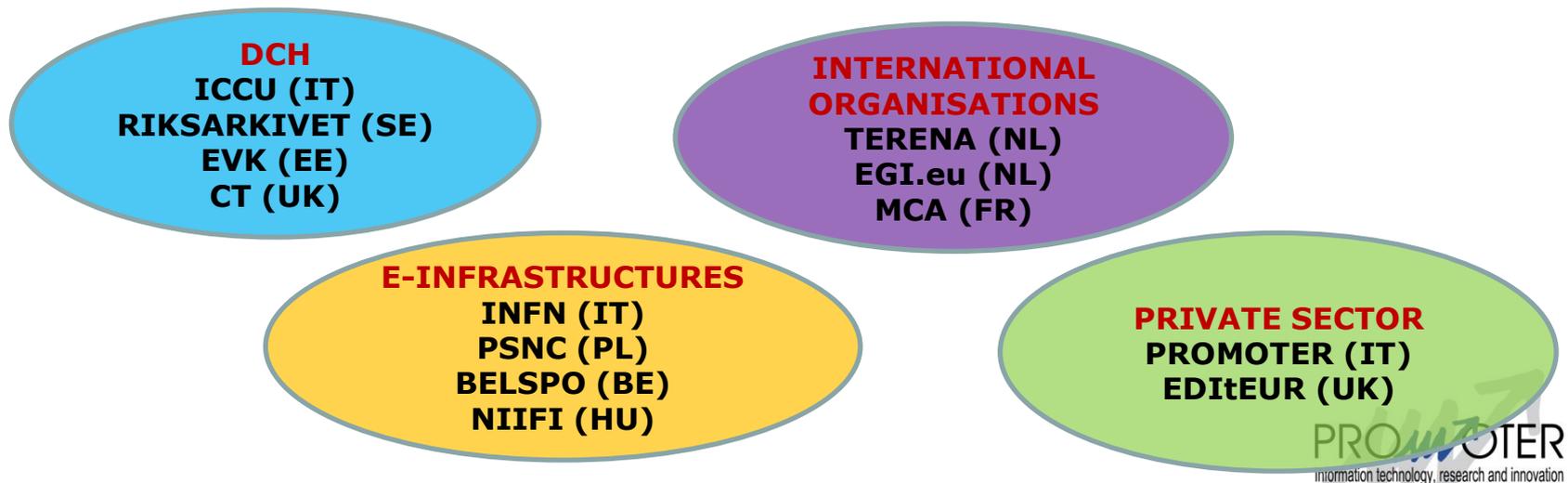
- ❑ The digital cultural content is defined in this context as the combination of:
 - **digitised content** (the result of digitisation processes)
 - **born-digital content** (databases, catalogues, files, etc.)
 - **metadata**

- ❑ Each digitisation programme is currently addressing the issue of preservation in a separate manner, while a **shared implementation** of common e-Infrastructure layers could be beneficial and cost effective to all

E-infrastructures for Digital Preservation: DCH-RP case



- ❑ Project background: built upon the knowledge generated by **DC-NET** and **INDICATE**
- ❑ Consortium: **multidisciplinary**, composed of 13 partners from 7 European Countries
- ❑ Aim: to develop a validated **Roadmap** for the implementation of a preservation infrastructure for DCH



DCH-RP: first outcomes

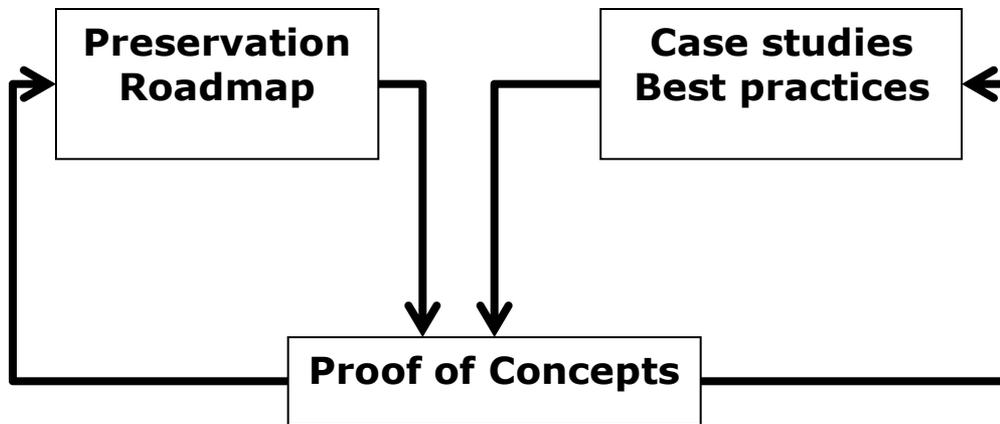


- ❑ A **registry of Services** was implemented and published online
- ❑ The first **Proofs of Concept** is just been completed: cultural institutions experimented the actual use of distributed computing and storage infrastructures (such as the **e-Culture Science Gateway**: <http://ecsg.dch-rp.eu/>) to store and manage cultural digital resources, starting from real scenarios and use cases
- ❑ The first version of the **Roadmap** will be released next month

139

Name	OAIS stage	Functionality	Web
AccessToSiard	data mgmt	A collection of scripts to automatically convert MS Access files to the SIARD format.	http://sourceforge.net/projects/accessosiard/
ACE (Audit Control Environment)	archival storage	Set of tools to help archives monitor the integrity of collections. It provides a mechanism to allow a 3rd party to independently verify a collections integrity.	https://scm.umiacs.umd.edu
Acrobat XI Pro	ingest, data mgmt, conversion	Create, modify, convert from/to PDF documents. Convert a Web page or an entire Web site to a single PDF, offline viewing.	http://www.adobe.com/products/acrobatpro.html
AFF	ingest, acquisition	Tools for the creation of disk images, used in conjunction with the AFF open and extensible file format	http://afplib.org/

DCH-RP: the results of the first Proof of Concept



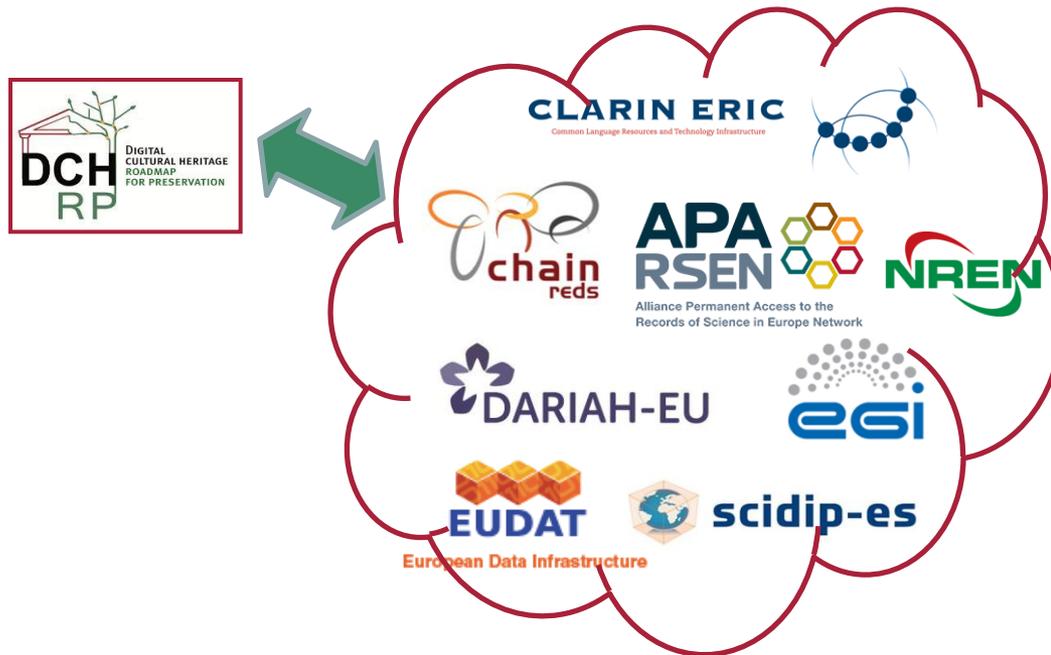
- ❑ Cultural data are curated by many different persons: data **management** and **administration** + **user access control** are very important.

- ❑ **Security** of the data is very important for cultural institutions: **trust building** is a key factor when it is not determined where data are stored
- ❑ Functionalities and services offered by e-infrastructures should not **impact** on the outgoing traffic of the institution
- ❑ Access to the e-infrastructure services should be **simple** without requiring IT specialist knowledge

DCH-RP: what's next?



- For the **second Proof of Concept** DCH-RP is planning to experiment with services offered by other projects (EUDAT)



- We are looking for other new services to be used with data made available by the cultural institutions participating the second Proof of Concept

DCH-RP on the Web

- ❑ Official website: www.dch-rp.eu
- ❑ A constantly living showcase on Digital Meets Culture www.digitalmeetsculture.net providing:
 - easy access to the ongoing activities and to the last achievements,
 - supporting the dialogue between the partners inside the project with the discussion carried out outside the project

The screenshot displays the official website for the Digital Cultural Heritage Roadmap for Preservation (DCH-RP). The page features a colorful geometric header and a navigation menu. The main content area is divided into several sections:

- PRESENTATION OF THE PROJECT:** A central article titled "DIGITAL CULTURAL HERITAGE, ROADMAP FOR PRESERVATION" with a sub-header "A coordination action project focused on the digital preservation." It describes the project's goal to integrate efforts of DCH institutions, e-infrastructure, research, and private organizations.
- EDITORIALS, INTERVIEWS, NEWSLETTERS:** A sidebar on the right containing links to various content types.
- FILL IN THE DCH-RP SURVEY!:** A prominent call-to-action box with a red and green background, encouraging users to help define better authentication and authorization services.
- DCH-RP LATEST NEWS & EVENTS:** A list of recent activities, including a review meeting, a virtual research community, and a conference on scientific data in Europe.
- DCH-RP MAIN ONGOING SERVICES:** A list of key services such as the eCulture Science Gateway, Registry of Services & Tools, and the DCH-RP Repository.
- CONTACTS:** Information about the project coordinator, technical coordinator, and website.
- UPCOMING EVENTS:** A vertical list of future events, including the Lucca Beni Culturali 2013 conference and the eBooks on Demand Conference.
- IN FOCUS:** A section highlighting a seminar in Rome on "The digital cultural heritage towards Horizon 2020".

Logos for various partners and organizations are visible on the right side of the page, including "fondazione sistema toscana", "europeana", "HARMOSEARCH", "cultura", "europeana photography", "LINKED HERITAGE", "decipher", "Paths", "insiddo", and "d culture".

Conclusions



- ❑ DCH-RP is one step, belonging to a longer and wider process.
- ❑ In the last decades, cultural institutions started to move their cultural content on the digital world. This has implied to know new instruments, new rules of access, new standards to exhibit digital heritage, new ways of communication.
- ❑ The amount of digital cultural content is now so valuable that the issue of preserving the digital cultural heritage is becoming as much urgent as the preservation of tangible heritage.
- ❑ The national policies about cultural preservation (digital and tangible) needs to re-use best practices, to share solutions, to avoid duplication of efforts.
- ❑ The use of the e-infrastructures is a pillar in this direction and the cultural heritage sector should progress towards its full integration in the new concept of open science.

Next appointments



□ November 4th, 2013 – Vilnius

- “Digital Cultural Heritage moving toward an e-Infrastructure based approach to digital preservation” at the **e-IRG Workshop**

(info at <http://www.digitalmeetsculture.net/article/open-e-irg-workshop-to-discuss-infrastructures-and-e-infrastructures/>)

□ November 7th, 2013 – Vilnius

- Networking Session at the **ICT 2013 Conference**

(info at <http://www.digitalmeetsculture.net/article/a-virtual-research-community-for-the-preservation-of-digital-cultural-heritage/>)



Thank you!

Antonella Fresa

Promoter Srl

fresa@promoter.it

www.digitalmeetsculture.net