This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

INTERNAL REPORT

Project Acronym: **CIVIC EPISTEMOLOGIES**

Grant Agreement number: **632694**

Project Title: **Civic Epistemologies: Development of a Roadmap for Citizen Researchers in the age of Digital Culture**

First Draft Version of D3.2 Roadmap

Revision: Internal v0.3

Authors:

Borje Justrell, (Riksarkivet)

Contributors:

Milena Dobreva (University of Malta)
Daniela Azzopardi (University of Malta)
Anders Nordström (Riksarkivet)

Project co-funded by the European Commission within the from the European Union's Seventh Framework Programme for research, technological development and demonstration		
Dissemination Level		
P	**Public**	
C	**Confidential, only for members of the consortium and the Commission Services**	

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Revision History

Version No.	Date	Author	Organisation	Description
0.1	2015-01-07	Borje Justrell	Riksarkivet	Structure of the document
0.2	2015- 01-15	Borje Justrell	Riksarkivet	Add text in most of the sections
0.3	2015-02-06	Borje Justrell	Ritksarkivet	Final text added and final structuring of the document

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

TABLE OF CONTENTS

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

GLOSSARY

Specific terms and the definitions used or to be used in the final version of the Roadmap.

Born Digital - Digital materials which are not intended to have an analogue equivalent.

Cloud computing - a phrase used to describe a variety of computing concepts involving a large number of computers connected through a real-time communication network such as the Internet.

Digital asset – the material produced as a result of digitisation or digital photography; the term includes also more complex accumulations such as online learning resources, web pages, virtual reality tours and digital/visual files.

Digital curation - has wider coverage than digital preservation and involves maintaining, preserving and adding value to digital data throughout its life-cycle.

Digital preservation - a set of activities required to make sure digital objects can be located, rendered, used and understood in the future.

Digital record – any information that is recorded in a form that only a computer can process and that satisfies the definition of a record as stated in the formal regulation and/or the policy for the cultural institution in mind.

Digital resources – encompasses both digital records and digital assets.

Digitisation – the process of converting analogue data carriers (parchment and paper records, microforms, photos, film and audio and video tapes) into digital form using scanning, digital photography, or other conversion methods.

E-Infrastructure - the term used for the technology and organisations that support research undertaken through distributed regional, national and global collaborations enabled by the Internet. It embraces networks, grids, data centres, and collaborative environments; it can also include supporting operations centres, service registries, single sign-on, certificate authorities, training, and help-desk services.

Grid computing - the collection of computer resources from multiple locations to reach a common goal.

Hub - a common connection point for devices in a network (could be of different kind).

Memory institutions - a metaphor used about a repository of public knowledge; a generic term used about institutions such as libraries, archives, museums, clearinghouses, electronic databases, and data archives, which serve as memories for given societies or mankind as a whole.

Metadata – information about data required to manage, search, understand, use, and preserve it.

Mashup - in web development, a web page, or web application, that uses content from more than one source to create a single new service displayed in a single graphical interface.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Ontology – a structural framework for organising information; used in artificial intelligence, the Semantic Web, systems engineering, library science, information architecture etc as a form of knowledge representation about the world or some part of it.

Persistent identifier - a long-lasting unique reference to a digital object, which could be a single file or set of files.

Virtualisation - refers in computing to the act of creating a virtual (rather than actual) version of something, including a virtual computer hardware platform, operating system (OS), storage device, or computer network resources.

Visualisation - any technique for creating images, diagrams, or animations to communicate a message. Visualisation today has ever-expanding applications in science, education, engineering (e.g., product visualisation), interactive multimedia, medicine, etc.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

ABBREVIATIONS

Acronyms used or to be used in the final version of the Roadmap.

AAI	Authentication and Authorization Infrastructure
API	Application Programming Interface
AQuA	Automated Quality Assurance Project
CHI	Cultural Heritage Institution
COPTR	Community Owned Preservation Tool Registry
CLARIN	Common Language Resources and Technology Infrastructure
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DIP	Dissemination Information Package
DCHH	Digital Cultural Heritage and Humanities
DCH-RP	Digital Cultural Heritage – Roadmap for Preservation
DP	Digital preservation
EC	European Commission
e-IRG	e-Infrastructure Reflection Group
EU	European Union
EUDAT	European Data Infrastructure project
4C Project	4C (Collaboration to Clarify the Costs of Curation)
GRID	See Grid computing
ICT	Information and Communication Technologies
HPC	High Performance Computing
HW	Hardware
IaaS	Infrastructure as a Service
ICT	Information and Communication Technology
INDICATE	International Network for a Digital Cultural Heritage e-Infrastructure
MW	Middleware
NGI	National Grid Initiative
NARA	National Archives and Records Administration (US)
NREN	National Research and Education Network
OAIS	Open Archival Information System
PaaS	Platform as a service
PB	PetaBytes
PEST	Political, Economic, Scientific, Technological
PraaS	Preservation as a Servic
PSNC	Poznań Supercomputing and Networking Center
RAID	Redundant array of independent disks (earlier: Redundant array of inexpensive disks)
SaaS	Software as a Service
SCAPE	SCAlable Preservation Environments
SOA	Service Oriented Architecture
SW	Software
TDR	Trusted Digital Repository
TB	TeraBytes
VPN	Virtual Private Network
VRC	Virtual research Community
VRE	Virtual Research Environment
VRO	Virtual Research Organization
WP	Work Package

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

EXECUTIVE SUMMARY

This Executive Summary will in the final Roadmap provide a brief description of the document, written in a way that is understandable and meaningful also when extracted outside the context of the deliverable, as a standalone abstract.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

1 INTRODUCTION

1.1 STRUCTURE OF THE DOCUMENT

This is the first draft version of the Roadmap that the Civic Epistemology project has been committed to design. The aim of the Roadmap is to broader e-Infrastructure deployment in order to support citizen researchers in digital culture. This first draft version, submitted as an internal document, is organised as follows:

Section 1 (Introduction) - sets out the structure and the objectives of the document, including a short review of the relationship with other deliverables and tasks in the project;

Section 2 (Setting the Scene) - offers an overview of the general context for the draft version;

Section 3 (A Roadmap for citizen science) – introduce a broader discussion about the Roadmap and it's different parts;

Section 4 (An action plan) – propose major areas of the Roadmap where actions should take place when planning for the use of e-infrastructure to support citizen research;

Section 5 (A Web-space for the Roadmap) – presents an idea to dedicate a web page on Digitalmeetsculture to preserve, maintain, update, discuss and keep alive the Roadmap;

Section 7 (Conclusions) - summarises on a general level the results in previous sections;

Annex 1 - contains an initial set of critical system requirements;

Annex 2 – describe issues on license agreements and terms of usage.

1.2 OBJECTIVES OF THE FIRST DRAFT

The Civic Epistemologies project is about the participation of citizens in research in Digital Cultural Heritage and Humanities (DCHH). The term "Civic Epistemologies" is taken from the study of Sheila Jasanoff in 'Designs on Nature' (2007) in which she defines civic epistemologies as "the institutionalized practices by which members of a given society test knowledge claims used as a basis for making collective choices."

The Civic Epistemologies project aims to investigate these practices in the research domain of DCHH. The projects main objective is to produce a validated Roadmap for the use of e-infrastructures to support the participation of citizens in the research processes and the participation of creative industries in the exploitation of digital cultural content.

The participation of Europe's citizens in scientific research development has just started to be exploited while it represents a big potential for improving European competitiveness. The case of DCHH is particularly relevant because

- of the major cross-cutting role that the Humanities play in European research and innovation - acknowledged in Horizon2020;

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

- Cultural Heritage and Humanities is a subject area in which citizens are particularly active (recording, cataloguing, and discussing things on an individual, group/voluntary/amateur basis);
- the case of broadening e-Infrastructure deployment to support the participation of citizens to DCHH research, even if holding a strong impact potential for social cohesion and job development, is not yet fully explored.

The objective of this first draft version of the Roadmap is to provide a description of what a Roadmap for the use of e-infrastructure to support citizen research could look like. It is submitted early in the project (month 6) as an internal document for discussion. The preliminarily of it derives from the fact that most of the input to the roadmap will occur later in the project as output from different activities carried out in the other WPs. The final version of the Roadmap will be presented in deliverable D3.2 in the end of 2015.

The aim is that the Roadmap will lead to an implementation of an e-Infrastructure that will

- enable creation, access, use and re-use of DCHH content;
- provide learning resources;
- provide communication services to multidisciplinary research teams located in different geographic places;
- enable citizens to participate in a range of research goals established at European level together with cultural institutions and universities;

The ultimate aim is to address the scientific processes in DCHH and to bring citizens, through their associations, into the process of planning research.

1.3 RELATIONSHIP WITH OTHER TASKS AND DELIVERABLES

The design and validation of the Roadmap is a task given to WP3 (Roadmap development). An important input when executing this task will be the identification of requirements of different actors involved in citizen science: e-Infrastructure providers, researchers, cultural heritage institutions, citizens and their organisations, and creative industries. This will be done by WP 2 (Identification of requirements) and documented in the following deliverables:

- Deliverable D2.1 *Results of the workshop on requirements*, submitted in January 2015; it provides an initial systematic set of requirements identified within WP2 and discussed at a workshop in Malta in November 2014;
- Deliverable D2.2 *Key characteristics and requirement of e-Infrastructure for citizen scientists in digital culture*, to be submitted in February 2015; it will present the model of citizen research cycle, the typical types of users, basic scenarios and use cases, and the summarised list of functional and non-functional requirements identified.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

The Roadmap will be validated by two case studies and one pilot to be carried out by WP4 (Pilot and Case Studies). The results of these activities will be presented in two deliverables, D4.1 *Ethnographic Pilot report* and D4.2 *Case Studies Report.*

A network of common interest will be set up to enable the discovery of best practice, solutions, technologies and success stories that exist at regional, national, European and international level. This network will be composed of institutions from the cultural heritage sector, laboratories from the Humanities research sector, e-Infrastructure providers and representatives of the creative industry. The purpose is to ensure that the final Roadmap is not simply a desk study, but has a solid base in the domain that will implement it.

WP3 will beside this internal report and deliverable D3.2 *Roadmap* also produce:

- Internal reports on *Tools, services and established standards, Innovation policies for cultural heritage institutions, Strategic Research Agenda,* and *E-Infrastructures sustainability models* - to be used as background documents for the discussion at a workshop on the Roadmap in February 2015;
- Deliverable D3.1 *Registry of Services* that will look into existing projects and initiatives as well as standards, tools, workflows, approaches, solutions, demonstrators and applications which are useful keys while approaching the citizen researchers' issues; this deliverable will include the description of a registry of services which will be implemented online.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

2 SETTING THE SCENE

2.1 THE CONCEPT OF CITIZEN SCIENCE

Citizen science has gained substantial popularity and is becoming a new outlet for people who are not professionally trained to be researchers but have the possibility to contribute to a wide range of research. The modern technological environments allow for innovative ways to involve vast groups of such voluntary researchers in different ways; however citizen science is not a modern phenomenon being particularly prominent in 19th century.

The increasing popularity of citizen science is demonstrated by the growing number of publications in this area.[1] However, there are also substantial differences in the understanding what citizen science actually is. The definition of the term "Citizen Science" differs across the various papers written on the subject. The most common aspect seems to point out the nature of citizen scientists as being amateur researchers who gather scientific information on a voluntary basis, their only incentive being their participation to the project.

The form of these projects could differ but is not in any way unique. Oomen and Aroyo[2] highlight six different typologies of Citizen Science Projects, each one linked to a different kind of study:

1. Correction and transcription - the citizen is given access to a database (this is usually a text-based database like scanned manuscripts) and then he gets the task of transcribing or making corrections to the text which was already transcribed electronically via a computer programme.
2. Contextualization - Citizens submit data such as letters, stories, films, photographs or other documentary material in order to gather a meaningful context.
3. Complementing Collection – Citizens are asked to submit data into databases with the ultimate aim of completing them or making the collection grow.
4. Classification – Citizens tag the data, or label it, in order to easily group similar data and make the information more easily retrievable in the future.

[1] Dobreva, M., and D. Azzopardi "Citizen Science in the Humanities: A Promise for Creativity". In: G.. Papadopoulos (ed.) Proceedings of the 9th International Conference on Knowledge, Information and Creativity Support Systems, Limassol, Cyprus, November 6-8, 2014, ISBN: 978-9963-700-84-4, pp. 446-451.

[2] Oomen, J., L. Aroyo. "Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges." Proceedings of the 5th International Conference on Communities and Technologies (C&T '11). ACM, New York, NY, USA, 138-149.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

5. Co-curation – This practise occurs mostly with projects involving the aesthetic arts. Citizens interact with institutions and voice their opinions when it comes to choosing articles or items for publication.
6. Crowdfunding – Citizens are asked to gather together money and/or resources in order to support efforts initiated by others. Popular platforms used specifically for this purpose are: Kickstarter (https://www.kickstarter.com/) and Indiegogo (https://www.indiegogo.com/).

There is difference between citizen science and crowd sourcing which is to some extent blurry. Both concepts refer to activities which include contributions from multiple volunteers who are not professionals in the tasks they contribute to. However, there is one aspect which allows differentiating between them:

- Citizen science involves activities and people with a specific research focus and they are organised in a project which is led by a professional researcher.
- Crowd sourcing benefits instead from the time and skills investment of volunteers but does not aim necessarily in the first place to achieve research outcome and it is often not coordinated by a professional researcher.

2.2 CITIZEN SCIENCE AND CROWD SOURCING IN THE CULTURAL HERITAGE DOMAIN

Currently citizen science is very popular in the sciences but is not equally prominent in the domains of DCHH. However, in the cultural heritage sector participation of multiple contributors to a common task is a well established common practice through crowd sourcing projects. More exotic varieties of crowd sourcing work are the transcribe-a-thons (special sessions organised to produce transcription of a text, either modern handwritten or historical); translate-a-thons (where multiple contributors help to translate major work or a corpus of works) and review-a-thons (where users make systematic review of a translation or transcription).

Although the use of citizen science in the DCHH is less popular than in the Sciences, this does not mean that these do not exist. Examples follow:

- "Letters of 1916" project, a website which gathers letters to or from Irishmen submitted from all around the world. These letters can also be translated or transcribed by anyone on the website. This project helped shed light on that ear's lifestyle, thus bringing academics and enthusiasts of those times closer to that era.
- "Georeferencing: help us place our digitized maps" is another project which makes use of an online interface and of citizen scientists to decode their data. This project's aim is to help the British Library identify their historic maps and correctly place them in their modern day location.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

2.3 MAIN CHALLENGES

The participation of Europe's citizens in scientific research development has just started to be exploited while it represents a big potential for improving European competitiveness. The use of e-Infrastructures to the participation of citizens could provide relevant support, but e-Infrastructures need to broad their deployment with new services targeted to this scope. The research on DCHH can play an important role in the development of the european research area, and can take lead in the discovery of new directions of cross-disciplinary research; but this opportunity has not yet been fully developed.

The solution proposed by the Civic Epistemologies project is to empower the existing e-Infrastructures with new services, targeted to the needs of specific research domains, in order to broad the communities of users, including citizen scientists as an integrated part of the communities. It should be possible to tailor the new services to the requirements of each research community; at the same time, it should be necessary to identify common layers and standards that can be shared among different domains. This scalable and modular approach to the e-Infrastructures deployment will allow to serve better the research and to reduce costs of development.

The design of the new services should be planned with a concrete approach, based on practical case studies and pilots with real users who should provide experimental proofs of the concepts defined in the theoretical sphere. To this purpose, the DCHH domain is proposed. The new deployment should be carefully planned by indicating the steps that each stakeholder must take:

- decision-makers, policy-maker and programme owners to make available the necessary financial resources;
- research communities to identify the protocols of interaction with citizen scientists, citizens to associate and organise themselves into representative bodies;
- e-Infrastructure providers to plan for the future deployments.

Actions that are needed can on a general level be grouped in three stages:

- A **preparatory stage**, vital for the overall development of the project, and including activities such as:

  - Defining factors like essential PEST conditions, vision, scope and boundaries (physical, national, IP, other), short-, mid-, and long-term measures for awareness/leadership

  - Mapping to EC and national research agendas

  - Establishing key partnerships with relevant e-Infrastructures

  - Analysing innovation drivers (economic, technical, other drivers)
- A **development stage** focused of the content of the Roadmap; the activities are described in section 4.2.2 below.
- A **deployment and monitoring stage** with activities like

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

- Boosting work on filling the gaps
- Promoting available toolkits
- Implementing the new services in a range of sites (providing guidance/support), including the development of the repository of tools, services and digital cultural content
- Evaluating the performance – planning for improvement and enlargement.

The Civic Epistemologies project will focus primarily on the first two stages.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

3 A ROADMAP FOR CITIZEN SCIENCE

3.1 FORM AND CONTENT OF THE ROADMAP

When writing a Roadmap, there are some basic questions that have to be addressed:

WHY - overall objectives for making a Roadmap;

WHO - who to address: target groups and/or user groups, stakeholders in general, members of society;

WHERE- where to go (specific objectives and goals for the roadmap to be a basis for requirements like improved access, enhanced quality of holdings/collections, social inclosure etc);

WHEN - when shall these specific objectives and goals be reached (time line for implementing the Roadmap);

WHAT - what to produce: a Roadmap, but what are the basic drivers and the added values of citizen research and crowd sourcing as a method and what are benefits of using distributed e-infrastructure**;**

HOW - how shall the Roadmap be structured (address each targets groups and/or user group or be structured on general level).

These questions are for the moment only covered in patches in following sections.

3.2 THE ROADMAP AS AN INSTRUMENT

The "map" in the roadmap draws the landscape of citizen science for the DCCH domain based on the current situation, but needs also to take into account how the situation will change in the future.

The landscape is also changing at different levels, technical, political and legal. Distributed solutions like government clouds are becoming increasingly prevalent and some cultural heritage institutions may be obliged to make use of them. New data infrastructures with a

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

portfolio of services are constantly being built. Societal changes have also to be taken into consideration.

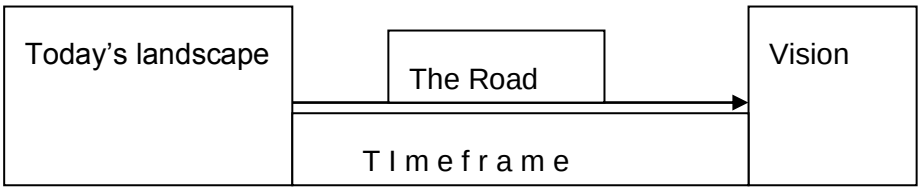
The "road" in the roadmap points to an action plan, and actions are needed in a number of areas: tools, services, authentication, trust, governance models, user requirements, funding models and business models, skills / training / awareness etc.

The Civic Epistemologies roadmap will integrate three domains of necessary intervention (business change, policy framework and better tools) with the major PEST factors (political, economic, scientific, and technological). The compilation of the roadmap will also need integration of a multitude of viewpoints and aspects, both those foreseen in the planning of the project and new ones discovered during the project's lifetime.

The roadmap is built on two implicit assumptions: firstly, that existing e-Infrastructures for research and academia are efficient channels also for the delivery of services to be used by the DCHH domain for supporting citizen science and, secondly, that it will be possible to establish common policies, processes and protocols which will allow the DCHH domain to access e-Infrastructures, despite the fact that e-infrastructures often are national entities, sometimes with different policies and procedures for access and usage.

Many cultural heritage institutions have in-house solutions for processing their digital collections and holdings but also their research applications. When comparing in-house solutions with e-Infrastructure services, it is inevitable that some discrepancies will appear, such as incompatibility of purposes or scope, lack of technical or semantic interoperability, reliance on different standards, and jurisdictional and legal barriers, etc. Therefore, the Civic Epistemologies roadmap has a strong focus on what to do and on the usability of services and technologies.

In order to achieve this, the project has adapted the following - very simple - structure of the roadmap:



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

3.3 DESCRIBING THE LANDSCAPE – TO MET STAKEHOLDERS NEEDS

3.3.1 Focus groups

The Civic Epistemologies project has carried out three focus groups targeted at different stakeholder communities. The results are in summation:

Policymakers

The general positive attitude towards citizen science was very strongly felt in this focus group. The participants discussed at length that the involvement of volunteers in the projects undertaken by institutions helps to establish a long lasting relationship and as such is a powerful way of engagement with the general public. Satisfied citizen scientists can help in future projects and might also serve as an effective "word-of-mouth" advertising, which would in turn bring more people to the institution. This could also help create dialogue with the community in terms of shared memories.

People do not attract more people with their enthusiasm only – one important point made was that if the people care, so will the governments. Political awareness might get the instruction more help from the government, making them dedicate more time and resources to the institutions.

To create a better communication with the communities, one must also get to the source of it – children. These young members of our societies are often not aware of what is going on in their own communities, let alone on a national level. By creating a better bridge between the community and the children, this would help nurture individuals who would grow up showing more interest in the cultural heritage domain and thus be more willing to volunteer their help and services in the future.

The general feeling that seemed to stem from the discussions was that citizen science was a highly valued method which could be an immense source of data, but at that point was not necessarily accessible for the institutions to make use of. While it seemed easier to use citizens in a scientific research, the participants were finding it hard to clearly see a path one could take to make use of such an encompassing resource in the cultural heritage setting.

Citizen scientists

All participants in this group agreed they had citizen science experience. Each member of the group had been involved in various projects and investigations – ranging from EU-Funded projects, international, national and local projects as well as Lifelong Learning Council of Europe initiatives. All of the participants were using museums, archives and library collections 'for personal use' but this was directly related to their activist organisation activities. Half of the participants highlighted that the local museums and cultural heritage institutions in the area, all played a key role in shaping their personal activities which directly linked with their activist role and work.

There was a unanimous agreement among the participants that the experts, technicians and specialists have to incorporate the people and average citizens into the work carried out in

cultural heritage institutions. The use of citizen engagement enhances the work and the quality of data collection, leading to a more enhanced project that is responding to the shift taking place in technology obsessed society.

The best way to include the voices of the unprofessional researchers is to ask them to get involved. Oftentimes the unprofessional researchers, those at a grassroots level, are not included and do not feel they can participate or be included in such projects. The participants highlighted that many of the projects that they have been a part of, asked them and others to be a part of the research. There was an active recruitment process from that took place and an effort made in trying to include them.

The types of digital technologies that citizen scientists can use, was discussed. There was an agreement that TV and media contribute a lot to the cultural heritage of a community and country. Through their programming and commissioned projects, media offers a way for citizens to engage with cultural heritage content. There is a risk when using these tools. A couple of participants stated that there must be a balanced approach to digital technologies. There is an agreement by all that tools can help and allow many people from various backgrounds, even those that are often excluded and marginalized, get involved and offer an opinion or contribute to the investigation(s). The digital tools cannot be the central point. The participants all stressed that there has to be a balanced approach to gathering data as it can either isolate or create community. One participant said that writing, letters and telegrams were a huge way that people shared their knowledge and were included in previous projects.

Best practice digital technology tools that can be used or that they personally use in their own work are:

a) Computers, phones, music CD's, DVD's, informal talks that incorporate digital technologies, Internet, specifically YouTube and Skype. Skype allows people to share knowledge, engage and contribute.

b) Social media can be used to recruit and engage with various citizens from various socioeconomic backgrounds and ages, and offers an immediate way to contribute.

Activists

The general pattern, as the participants saw it, is that citizens normally participate in research activities through their local or regional societies. The cultural heritage institutions are seldom first on stage in these topics.

The most useful outcomes of organising citizen science projects are, from the perspective of a cultural heritage institution:

- increased interest in the institution and its collections/holdings;
- more work will be done;
- an opportunity to engage competences that are normally not available internally.

The discussion showed that it is obvious that the activist organisations (genealogical societies) in Sweden see themselves as an important part of the knowledge society with an ability to participate in citizen research projects, mainly crowd sourcing initiatives. If there are no cultural heritage institutions in place (or not willing) to support them, they have the strength to organise

and run some of these projects themselves using cultural heritage institutions as "a source for crowding".

It did not become clear if this is the case in other Member States as well or in other countries around the world. In any case, genealogical societies and other activist organisations represent a strong movement that is using different strategies for reaching their goals: in Sweden by organising themselves in a nation wide federation strong enough at a political level to be recognised as an important partner to cooperate with or to listen to; in some other countries by using media or connecting themselves to research projects or programmes at universities with high level of awareness

The drivers behind private persons taking part in citizen science projects are normally:

- reward of some kind (could be small, symbolic and of less monetary value);
- personal interest;
- idealism (helping the local society in some way, religious duty, etc.);
- that the results could be used in the person's private research.

The conditions for organising citizens' research activities (becoming obstacles if they are not fulfilled) are mainly:

- the results of the activities have to be open for all to use ("open source");
- the technical facilities have to be in place from the beginning and also easy to use;
- the planning of the activities has to be made in cooperation with citizens research representatives, in order to incorporate their knowledge right from the beginning.

In earlier days most of the knowledge and expertise connected to the cultural heritage institutions holdings and collections were held by the institutions' own staff members. Today, with more and more of these the institutions' data and metadata available on the Internet, important parts of this knowledge and expertise are located outside the institutions, in the hands of users who also advance it by using different kinds of IT tools. An important issue for the cultural heritage institutions therefore is how to harvest this increasing external knowledge and expertise and make use of it in their internal work.

3.3.2 Other methods for catching today's landscape

Cultural heritage institutions

An online questionnaire has been designed and developed to evaluate the involvement of cultural heritage institutions with projects that include citizen scientists and, to a lesser extent, crowd sourcing activities. The online questionnaire confirmed the confusion between citizen science and crowd sourcing and showed a generally positive attitude towards the use of citizen science in the digital cultural heritage sector. The results will be presented in more detail in the coming deliverable D2.2.

Digital heritage stakeholders in general

The Civic Epistemologies project organised in November 2014 a workshop in Malta were local Maltese cultural heritage professionals and policy makers were present. They were also eager to voice their opinions and give out their recommendations to improve our research. Some of the key points that were raised in the discussion were as follows:

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

- Real accessibility needs to be available, not a theoretical one. The findings and results need to be shared with the community, with which a connection needs to be built and maintained;
- When presenting the data, or connecting to the general public, one must not be too technical, specific, or academically snobbish since this might repel the people one is trying to connect with;
- A key factor to remember is that cultural heritage belongs to the people – the job of cultural heritage institutions is to protect the embodiment of our culture and present it back to the people;
- An issue that arose with great enthusiasm during the workshop was the element of FUN. Fun has the capability to make an activity a good experience which would help increase the popularity and would encourage people to take part in it and to disseminate it;
- Getting the commitment from the government in the aid of these institutions would also be a plus;
- For any endeavor making use of citizen science to succeed, three key factors need to be connected:
    1. Research
    2. Institution
    3. Citizens
- Artifacts or data which embody a community's cultural heritage need to be equally accessible to everyone. No curators or directors should deem themselves the exclusive owners of such a collection;
- Citizens should never be considered as a subject in the research, or as a source. Their role should be that of an active participant in the research;

Amongst these and other comments that were voiced during the discussion, multiple people shared one common thought; that they were all eager to see the results of the Civic Epistemologies project.

A discussion on the technological infrastructure took place during the workshop, throughout which these issues were brought up:

- A basic framework needs to be developed which can then be adapted and reworked depending on the nature of the citizen science project that is being undertaken. This needs to be produced as a software or application;
- Constant support for the software needs to be available to whoever is using it;
- A serious issue that arose was the way users should be authenticated. A simple login via Facebook might not be enough, but users generally dislike creating and using additional accounts.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

E-infrastructures

One of the basic assumptions for the Civic Epistemologies project is that grid and clouds approaches (e-infrastructures) can offer a stable and reliable storage and computing platform to the DCHH domain. In general it seems that this domains first priority, when it comes to citizen science activities, is a flexible and stable technical environment. Other identified priorities are computer capacity for integrity checks and access to advanced virtualisation services. One conclusion is, therefore, that at least two main approaches to services supporting citizen science must be in place for distributed solutions. We can refer to them as the "kiosk" model where and the "turn-key" model respectively.

This "kiosk-model" could contain supplementary services like federated authentication, audit and certification, persistent identifiers distribution, which are typical network services that would make work easier for institutions or networks of institutions that manage digital preservation "on their own".

The "turn-key" model could contain cloud or grid based services that offer the entire process covering all the phases and functions needed in citizen science activities model, eventually with a particular focus on storage, curation services and other organisational aspects like trust.

Close to the "kiosk-model" is an approach called "micro services" presented just a few years ago. The key idea with "micro services" is that they allow flexible combinations of specialised solutions depending on the requirements of a DCH institution. "Micro services" for digital preservation are currently used in the open archival information system Archivematica.[3] However, if various micro services are to be used, they must be orchestrated in a way that assures that requirements for authenticity and integrity of digital objects are not compromised.

3.4 A VISION

The consortium of the Civic Epistemologies project shares a commitment to the values of openness, collaboration and wide participation in research within the field of DCHH to expanding the range and scope of a common set of civic values and understandings related to this field.

The project's over-riding strategic objective is to support the development of a policy on the role e-Infrastructures can play in encouraging and facilitating the mediation process of citizen science in the area of DCHH, in order to bring about a closer alignment between the private and public spheres. It seeks to identify and deploy new services and protocols enabled by e-Infrastructures, which will in turn support Europe's citizens, its creative enterprises and its wider cultural industries to enter into productive technology-enabled dialogue with cultural heritage institutions and Humanities research.

[3] http://archivematica.org/wiki/index.php?title=Development_roadmap

The Civic Epistemologies project is informed by the consortium's awareness that new technologies are very powerful tools in the processes of creativity, co-creation and innovation; but that the creative and cultural sectors are both highly segmented and small-scale (many SMEs, micro-enterprises) and are lacking in technical know-how. Further, much Humanities-based scholarship is both mistrustful of new technologies (for example much scholarship is still highly traditional in terms of the means of publication and dissemination of written outputs) and faces difficulties in engaging with wider audiences. The notion of the 'prosumer' – the reader of published research who also contributes interactively with that research via new technologies – has not yet fully penetrated either the academy or the cultural heritage sector.

Further, new skills are needed in our changing society. Underinvestment in skills renewal and knowledge/technology transfer and the loss of traditional skills leads to the risk of innovation deficit and of a general lack of diversity and choice across design, production and markets, resulting in missed employment and commercial opportunities. A Roadmap which offers new understandings and ways of grasping opportunity can also lead to economic as well as social benefits.

Finally, the consortium considers it vital to ask the following question: how can Humanities-based research in which the citizen is invited to play an active role, support re-conceptualization of the ways in which cultural heritage can reflect, construct and enrich individual and collective identities, and represent these increasingly fluid identities more fully, within a context of continuing social change?

3.5 A TIMEFRAME

The Civic Epistemologies roadmap should make it possible for each institution in the DCCH domain to define its own practical action plan with a realistic timeframe for the implementation of its stages.

- Short-term (2015 - 2016)

  The purpose of proposing a short-term action plan (2015) is to initiate the development of e-infrastructure services on a level that will be self-sustainable and continue to progress on its own. This further progress is defined in terms of two further proposed time spans:

- Medium-term (2017-2018), i.e. two years after the end of the Civic Epistemologies project), and
- Long-term (2019 and beyond) for the logical continuation of the work.

3.6 E-INFRASTRUCTURE SERVICES FOR CITIZEN SCIENCE

3.6.1 Important actions to cover in a Roadmap

Different parts of the DCHH domain have different needs, depending on if they are small or large, the kind of projects they have etc. The conditions (e.g. resources) for managing these

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

projects differ also quite much. Services for supporting citizen research and crowd sourcing, therefore, have not only to be flexible, but also easy to adapt and utilise, and address several areas. That is a clear message from most stakeholder groups.

E-infrastructure services for citizen science and crowd sourcing are normally structured around development of tools, but need also to involve policy instruments necessary to achieve efficient intervention in the DCHH sector. Important actions to be covered by a Roadmap will, therefore, be:

- Defining the product that the Roadmap focuses on
- Defining stakeholders, and their needs, scenarios, business models etc;
- Defining critical system requirements and map them to technology drivers as well as key performance indicators;
- Analysing current technological offers and gaps;
- Defining drivers for making a shift in institutional practices in the cultural heritage sector and the Humanities research.

Supporting actions to the Roadmap will be:

- Planning the structure of a registry of services;
- Developing and agreeing on a common Strategic Research Agenda
- Assessing new approaches in pilot experiments and case studies

3.6.2 Defining the product that the Roadmap focuses on

This will be e-Infrastructure enabled services for citizens and creative industries

- to make cooperation possible with cultural institutions and research teams, and
- to support citizen science and crowd sourcing activities.

3.6.3 Defining stakeholders, and their needs, scenarios, business models etc

These tasks will involve several WPs in the Civic Epistemologies project and be performed during 2015.

3.6.4 Defining critical system requirements and map them to technology drivers as well as key performance indicators

Some basic conditions will be

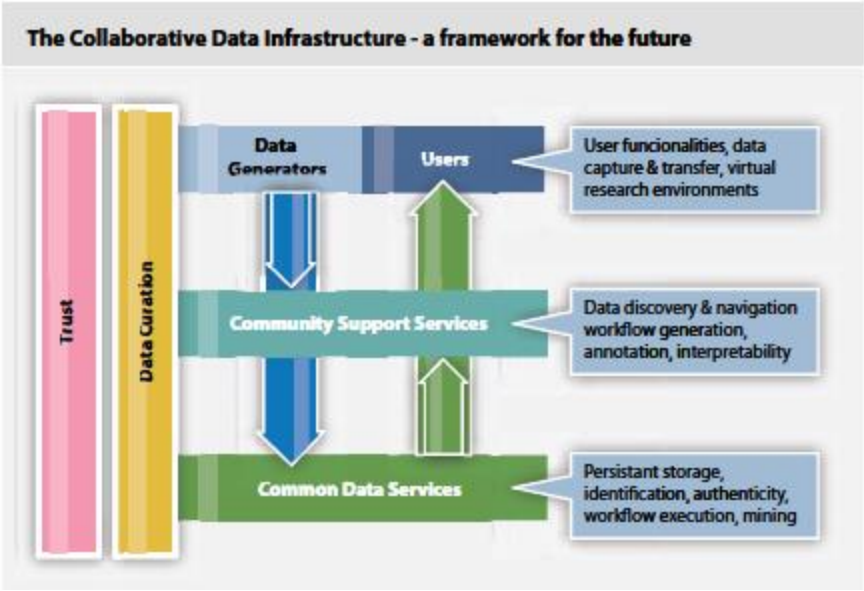
- Type of service architecture
- Conditions for infrastructure federation
- Current technological offers and gaps

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Type of service architecture

The EUDAT project has presented the architecture of a conceptual model that integrates various infrastructures with vast amounts of research data, and adds services for curation and trust in addition to the interface to users.

As it stands, this model represents basic stakeholder needs in the research area: ensure the trustworthiness of data, provide for its curation, and permit an easy interchange among the generators and users of data. These needs could also be said to be basic ones in the DCHH domain, and the EUDAT projects conceptual model can, therefore, serve as a base for further development.



The collaborative data infrastructure - a framework for the future; from *Riding the Wave*, p. 31

Improvements and adjustments of the model have already been made in, for example, the area of research data. The Data Archiving and Networking Services (DANS) in the Netherlands has developed based on the EUDAT conceptual model a federated data infrastructure with three layers of roles and responsibilities for the various stakeholders (The Front office – Back office model) [4]

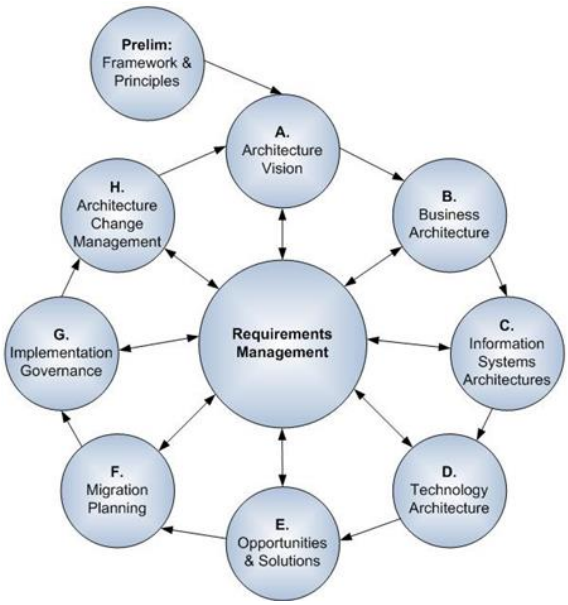
Cultural heritage institutions as well as research centers of different kind have sometimes built up their own infrastructure or handling digital resources. But it is undoubtedly true, that continuing investment in in-house solutions for citizen science will contribute to the lack of interoperability and fragmentation of resources into "digital silos". Stand-alone solutions that are not transferrable and interchangeable lead to fragmentation and do not offer economies of

[4] See www.dans.knaw.nl

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

scale. Instead, shared solutions for creation, storage and use of digital resources, including the e-Infrastructures, will become the major component of the future knowledge economy.

In order to move ahead from the current state into shared, decentralised solutions, it is important to define key institutional requirements in a standardised way. The use of enterprise architecture models is one possible approach because enterprise architectures seek to address system complexity while aligning technological developments with the institutional needs. There are a number of approaches for defining enterprise architectures; one of the popular ones is the Open Group Architectural Framework (TOGAF)[5] and its eight-stage Architecture Development Method that help to manage requirements within complex systems.



Architecture Development Method, TOGAF.

Service architecture as a technical area is very close to service-oriented architecture (SOA), which is a software design and software architecture design pattern based on pieces of software that provides functionality as a service easy to combine into different kind of applications. Services mean in this case not services for the users but services in terms of written functions ready to be used by programmers, and by other applications.

SOA can be seen in a continuum: from older concepts of distributed computing and modular programming on to current practices of mashups, SaaS, and cloud computing, which some see as the offspring of SOA. In the context of the Civic Epistemologies roadmap, aiming at the use of e-Infrastructure, SOA can clearly be regarded as a concept to get inspiration from.

[5] http://www.opengroup.org/togaf/

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Conditions for infrastructure federation

The needs to access networked applications and remote/distributed data is evolving dramatically. Authentication and authorisation are often separated from the application and the data themselves: authentication of the users is done by the users Identity Providers while the authorisation is done by the services based on the information received by the Identity Providers.

Access that follows this model is known as federated access and has advantages for both users and application developers. However, the usage of federated access requires that some technical and trust issues have to be solved.

For the Civic Epistemologies project federated access is a key element, both in terms of using federated storage to handle preservation data distributed all over Europe and in terms of user management. Federated access is in fact particularly desirable in a situation where services are offered across institutions and to users that do not belong to the institution that offers the service or technical facilities.

Federated access provides the technical and policy framework to allow for services to be shared in a trustworthy fashion across borders. How authentication is carried out by the institutions and how rights management is carried out by the service provider is left up to the respective parties.

When deciding whether to offer federated access, e-Infrastructures offering services should assess their potential user-base: whether they expect many local users or many users coming from different institutions. Federated access caters for the latter use-case and brings the following benefits:

- *Users* will be able to log in once (single sign-in) using their institutional credentials and access multiple services (sign on), Single Sign-On, whilst having the assurance that their personal data will not be disclosed to third parties.
- *Digital cultural curators and cultural institutions participating* will be free of the burden of user name and password administration, and will have access to more tools for managing data. On a large scale of users this means reduced administration and service provisioning costs; and it avoids duplications of identity stores.
- *Collaboration* among different parties becomes easier.

The first step to join a federation is to talk to the federation operator in a specific country. The list of existing federations is available online at: https://refeds.org/resources/resources_list.html

Current technological offers and gaps

Citizen science is composed of various elements such as applications, performers ("workers"), and institutions, which need to work harmoniously together in order to reach the project's goals. Various infrastructures are employed to make the process run as smoothly and seamlessly as possible. There are previous studies that have explored for example the technological devices used in citizen science initiatives. These devices will be presented and discussed in more detail in later in the project but here we are providing some examples:

2. Smartphones/mobile apps – Applications used for a variety of purposes, such as logging or providing data, tracking the citizen's movements, etc.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

3. Websites – These can provide information on the projects as well as act as points to input the data.
4. Video for training – Video to showcase the method for gathering data and submitting it.
5. Online data entry – This can be done via the application or the website.
6. Data analysis tools – Tools used to glean more information from the given data.
7. Social media – Can be used to disseminate information about the project and keep the users updated on the project, thus increasing the public's awareness of the project.
8. Mapping capabilities – Mapping the data gathered to triangulate common patterns, etc.
9. Database improvements – More storage, options, etc.

3.6.5 Defining drivers for making a shift in institutional practices in the cultural heritage sector and the Humanities research

In addition to the technological challenges, innovations around the internal workflows of the organisations operating in the DCHH domain is of great importance for the achievement of the vision of the Civic Epistemologies project. Internal workflows currently encountered among DCHH players imply that a number of actions need to be taken by many institutions that are engaged in citizen science, in order to make their digital resources more usable. Firstly, roles inside the organisation have to be re-organised to guarantee that citizen science is accepted as a method of work. Secondly, in order to create new skills and competences, practitioners have to be trained in both understanding and the handling of the new conditions associated with citizen science in a digital context i.e. the changing forms of artefacts and metadata, the changing methods of work, and the rapid changes in technology itself. Furthermore, decisions have to be taken about the procurement of services related to citizen science and computing resources. All these actions require time to be performed and financially resourced. Advocacy of the need for citizen science is, therefore, another important action in order to create the conditions required for understanding, acceptance, and endorsement by decision makers.

3.5 THE MAIN COMPONENTS OF THE ROADMAP

Empowering existing e-Infrastructures with new services: targeted to the needs of specific research domains;

Tailoring new services to the requirements of each research community;

***Improved interoperability*:** includes better integration of internal and external digital resources within the overall workflows for handling research data; in a way this is a set of measures to avoid building 'digital silos' within the organisation;

Establishment of conditions for cross-sector integration: a key condition for maximising the efficiency of successful solutions, transferring knowledge and know-how; a scalable and

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

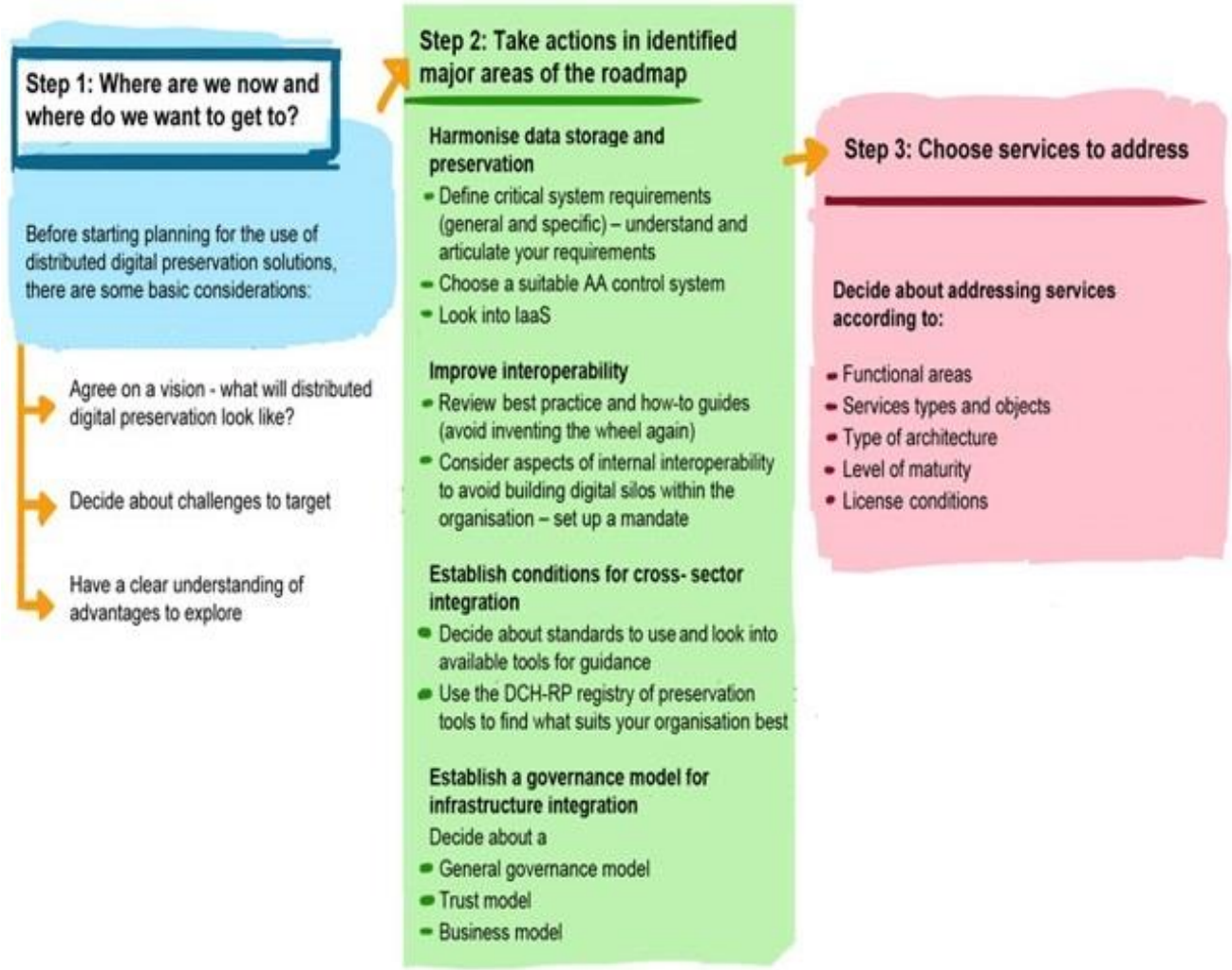
modular approach to the e-Infrastructures deployment is needed that will allow serving research better and reduce costs of development.

Governance models for infrastructure integration: a necessary condition for successful institutional participation in larger e-Infrastructure initiatives, and aggregation and re-use of digital resources.

For each area a set of prioritised actions are suggested.

3.6 STRUCTURE OF THE ROADMAP

3.6.1 General version



Example of a Roadmap from the DCH-RP project

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

3.6.2 Addressing target groups

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Lobby management into proper resourcing of selection and appraisal practice and focus on cost-effective digital curation activity	●	●			
Curation Researchers	Conduct research into automatic appraisal and selection techniques based on codified value criteria	●	●	●		
Data Users	Content experts to work with technologists to establish value criteria and represent 'designated communities'	●	●			
Managers	Incorporate the concept of 'value' into strategic and tactical decision-making		●	●	●	
Member Organisations	Help establish relationships between organisations to facilitate the transfer or 'handoff' of digital assets	●	●	●	●	●
Policy Makers	Establish requirements for digital asset value assessment as part of data management and curation planning			●	●	●
Solution Providers	Build on existing tools (e.g. file format registries) to provide automated selection & appraisal tools		●	●	●	

Example of a part of a Roadmap from the 4C project

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

4 AN ACTION PLAN

4.1 ESTABLISH A VALUE CHAIN

The Civic Epistemologies project will look into other domains, to see if there are experiences concerning value in distributed services that are transferrable to the DCHH domain. Apparently, very little has been done so far, but looking into a close related service, distributed digital preservation, the e-journal preservation community has achieved much in terms of evolving mechanisms and organisations to look after services in their field of interest. The technical, organisational and financial challenges have been proved to be solvable, given strong commitment from the communities involved. The key issue appears to be the ways in which these communities have organised themselves to bring about long-term agreements and infrastructures to make services happened.

Cost will clearly be a key variable when deciding whether or not to contract out services to an external service provider (e-infrastructure). But there are also other factors to consider, and the advantages and disadvantages of each of them need to be balanced against the overall mission of the institution. For example, legal provisions due to privacy or confidentiality may influence whether outsourcing is appropriate or not. The extent to which the potential advantages of using e-infrastructure services can be maximised and the potential disadvantages minimised is also dependent on the possibilities for dedicating staff resources to citizen science activities. The costs for these staff resources need to be added to the overall contract costs when calculating the cost benefit of using distributed services offered by e-infrastructure. However, one have to be aware of that most of these costs will be or should be received even when citizen science services are not outsourced.

Digital Preservation Coalition has listed a number of issues and potential advantages and disadvantages of using distributed services in digital preservation activities.[6] They can to some extent be applied on services supporting citizen science.

[6] See Preservation Management of Digital Materials: The Handbook, p. http://www.dpconline.org/advice/preservationhandbook /

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Issue	Potential advantage of using 3rd party services	Potential disadvantage of using 3rd party services
Limited practical experience	Avoids the need to develop costly infrastructure (particularly important for small institutions). Allows the institution to focus on other aspects of service provision. Provides specialist skills and experience which may not be available within the institution If there are economies of scale, outsourcing may well be cost effective. Allows action to be taken in the short to medium term, pending development of infrastructure.	Without some practical experience and expertise, it will be difficult to develop and monitor effective contracts. Without practical experience it will also be difficult effectively to communicate the requirements of the organisation (or to assess whether they are technically feasible or not). Danger of either not developing or losing skills base. There is no established bench marking. It is still too new an area. Risk of business failure Until the market increases there may be an overdependence on one contractor. Unless there are adequate exit strategies, may be locked into an outsourcing contract longer than intended.
Access considerations	Monitoring usage may be more efficient (assuming the contractor has a demonstrated ability to deliver meaningful usage statistics). There may be synergies and cost savings in outsourcing access and preservation together.	Difficult to control response times which may be unacceptably low and/or more costly, especially for high-use items.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Rights Management	Avoids what is often a resourceintensive activity for the institution.	May significantly increase the cost of the contract and/or complicate negotiations with rights holders.
Security	Contract can guarantee security arrangements required by the institution.	Lack of control, especially for sensitive material.
Quality control	A watertight contract will build in stringent quality control requirements.	Risk of loss or distortion may still be unacceptably high for highly significant and/or sensitive material.

Major advantages, specific for the DCHH domain when using distributed services offered by e-Infrastructures, could for example be the following:

- Long-term preservation (i.e., bit-level preservation) and access to digital objects of different kind, also so called "live" content (e.g., streaming audio and video collections);
- Multiple entry-points that suit a variety of user interfaces (e.g. APIs, protocols). New cloud based search engines are under development, based on multilevel nodes that can combine different data sources (documents, images, books etc) from multiple content providers;
- The DCHH domain can focus on its own areas of specialisation by deploying new services for monitoring and management tools that ensure smooth and secure running of distributed operations;
- Forming a community of practice or a Virtual Research Community that transcends discipline and national boundaries while achieving economies of scale by bringing together international communities;
- Benefitting from integration within the research and educational e-Infrastructures support framework;
- Central hosting and monitoring of middleware services;
- Simple authentication and authorisation infrastructures for large (and potentially unbounded) user groups;
- Connections to shared services in other countries and sectors. (e.g. research data centres, commercial businesses, etc.).

To summarise: it is important for institutions in the DCHH domain to have a clear understanding of what to exploit, before taking a decision about the use of distributed services to support citizen science activities.

Research and development on the use of such services built on distributed facilities instead of ones performed in-house has just started.[7] Some identified drivers that probably will underpin an enhanced the use of distributed services to support citizen science are:

- increased flexibility in digital preservation architectures based on granular or layered structures (e.g. SaaS, PaaS, IaaS) that are easy to adapt to a variety of preservation scenarios;
- clearly defined sets of metrics or benchmarks for comparing preservation tools and services and their performance;
- terminology and standards that no longer converge along professional community borderlines but instead are agreed cross-sectorial.

4.2 ACTIONS TO TAKE

4.2.1 Empowering existing e-Infrastructures with new services

This is an area not yet studied by the Civic Epistemologies project.

4.2.2 Tailoring new services to the requirements of each research community

An initial set of requirements are listed in Annex 1

4.2.3 Improve interoperability

Identify and promote best practices

An annex in the final version of teh roadmap will be dedicated to best practices, presenting an overview of the most important practical guidelines and lessons learned connected with the integration between the DCHH domain and the e-infrastructure providers.

Analyse interoperability issues

To avoid building 'digital silos' within the organisation, the following aspects need to be considered:

1. *Technical aspects*;

2. *Semantic aspects*: there are many vocabulary sources already available and it makes sense to check these out before inventing a new one.;

[7] An example is the InterPARES Trust (ITrust 2013-2018), a multi-national, interdisciplinary research project exploring issues concerning digital records and data entrusted to the Internet (http://interparestrust.ordg)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

3. Organisational and inter-community issues;

4. *Legal issues*: the transfer of personal data has to be in line with European directives on data protection and their implementation in national legislation; harmonisation of legal frameworks in general have also to be addressed, for example concerning the issue of cross boarder storage and differences in legal positions regarding preservation of master files;

5. *Political/human aspects*.

4.2.4 Establish conditions for cross-sector integration

Identify common layers and standards to be shared among different domains

One of the challenges for the DCHH domain is to choose among the vast number of standards that are already available. This may be problematic, especially for small institutions with limited knowledge in and/or resources in this field. There are also non-technical issues that have to be resolved. One is differences in the legal system between countries, especially when data is covered by copyright or classified.

The conclusion is that much work has already been done, but more efforts are still needed before these standards (including guides and tools etc.) can give substantial help to the DCHH domain. For example, many of them need to be more user-friendly in order to be understandable for non-technical personnel. Furthermore, practical tests made within the different EU financed project have shown that already developed e-Infrastructure services must be modified and/or improved in order to provide a "pan-European" solution for the DCHH domain.

In this internal report we are not bringing forward arguments for adopting or recommending specific standards, but information about standards are reported in Annex 2.

Registry of services

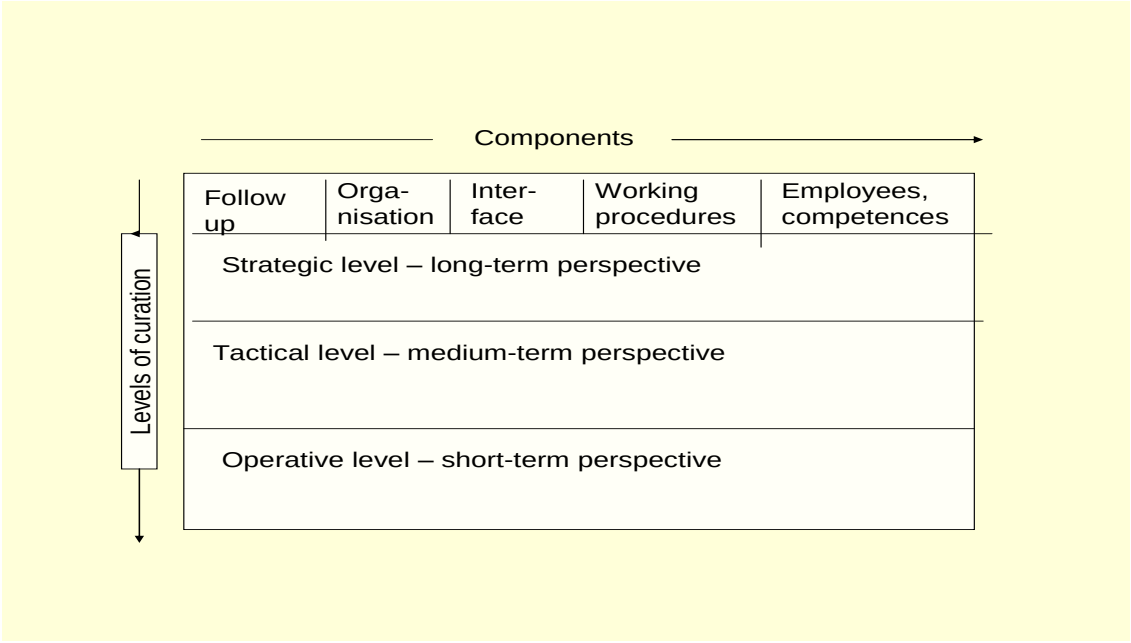
The development of the Civic Epistemologies services registry is a key step in the construction of the Roadmap. In this regard, it should be noted that the collection and summarisation of information on services is quite an onerous task. Deliverable D3.1 *Registry of Services* will look into existing projects and initiatives as well as standards, tools, workflows,

4.2.5 Establish a governance model for infrastructure integration

Analyse major information governance patterns and windows of opportunities

The model for governance to use must be tailored to the concept of distributed services supporting citizen science(and crowd sourcing) . The following framework can be seen as a role model for how to achieve good governance:

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694



A framework for the governance of distributed digital preservation services

This framework consists of five components that highlight different dimensions of governance focusing on three different levels (strategic, tactical and operative). The components are:

- Follow up (including how to manage distributed services)
- Organisation (including definitions of roles and responsibilities)
- Interface (including forum for clients and service providers to meet)
- Working procedures
- Employees and competences.

The levels of governance each have different focus and perspectives:

- Strategic level: aiming at securing the long-term perspective; this is done from both an internal and an external perspective through, firstly, follow up and managing a consolidated service provider portfolio, and, secondly, establishing a forward-looking relation between the client and the service-provider;
- Tactical level: has a time middle-term perspective with focus on securing services and agreements at hand and that they are up to date;
- Operative level: focus is here on securing the follow up of the daily work and that problem and incidents that arise are handled in a proper way.

Depending on which type of service is involved the service providers can be classified as being strategic/non-strategic and providing services that are easily accessible/not easily accessible. For the institutions in the DCHH domain the results of such a classification will inform their approach to managing the situation.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Explore the issue of trust-building

The DCH-RP project outlined the design of a new trust model suitable for the use of e-Infrastructures digital preservation, including recommendations for user authentication and access control system. The Civic Epistemologies project will look into it to see if it is suitable also for services supporting citzen cience activities.

Establish a possible business model

A business model describes the rationale of how an organisation creates, delivers, and captures economic, social, cultural, or other forms of value. In both theory and practice, the term business model is used for a broad range of informal and formal descriptions to represent core aspects of a business, including purpose, target customers, offerings, strategies, infrastructure, organisational structures, trading practices, and operational processes and policies. There is also a clear connection between the business model used and trust-building.

Innovative funding models must, therefore, be investigated, for example models where the public and private sectors enters into new partnerships (e.g. the re-use of digital cultural content by creative industries, non-IPR based models for the exploitation of digital cultural resources in applications for educational and research, commodification of cultural heritage and cultural tourism exploitation, etc.;

Services supporting citizen science activities built on a distributed model also needs a business model suitable for the integration between the cultural heritage community and the e-Infrastructures. ITC managements have today started to implement new concepts for outsourcing, whether cloud-based or not. One of them is Vested Outsourcing. This is a hybrid business model, based on research conducted by the University of Tennessee Center for Executive Education and funded by the U.S. Air Force, In this model both clients and service providers in an outsourcing or business relationship focus on shared values and goals to create an arrangement that is mutually beneficial to each, in contrast to traditional outsourcing and businesses relationships that, according to Vested Outsourcing, focus on win-lose arrangements.[8]

The basic philosophy in the Vested model is "What´s in it for We", and it consists of five rules that have to be implemented in a relation-based contract, in this case for distributed services supporting citizen science:

Focus on results and not on transactions: conform to a business model that will give both parties unanimous interest with focus both on valuable results and on a joint vision for the partnership.

Focus on what to do instead of how to do it: this approach means to concentrate on what to achieve instead of how it shall be done. Traditional outsourcing contracts often have detailed

[8] http://en.wikipedia.org/wiki/Vested_outsourcing

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

texts on how a service provider shall provide a service. This, sometimes called the "outsourcing-paradox", can ends up in a situation where the client outsource a service to an expert organisation, but at the same time describe in detail how this expert organisation shall provide its expertise. The Vested model instead points out the need for both a definition of functions and a roadmap with strategic goals for how the service provider shall support the client in achieving his or hers objectives.

Agree on clearly defined and measurable goals and deliverables: traditional contracts on outsourcing often contain agreements about measuring different levels of services and how to compensate the client if the agreed levels are not reached. However, this is not the same as the client being satisfied with the results. In a result based business model, focusing on what to do, the goals and achievements must be clearly defined from the beginning.

Establish a pricing model with optimal incentives for the agreed partnership: the traditional price list is not used in the Vested model. Instead, the service provider shall be economically compensated depending on how the strategic goals are achieved. But the conditions for every pricing model are constantly changing, and both partners must, therefore, have a high degree of transparency regarding their actual costs and economical situation. Otherwise fruitful negotiations about changes of prices will not be possible.

Establish a governance model that gives both parties both overview and insight: the important part in good governance is - according to the Vested model - to focus on the partnership as such and not on the partners. The partners work with a stratified structure, usually found in governance models (see above), but instead of just one interface for communication, with one responsible person per partner, several interfaces are used, one for each specific field in the contract.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

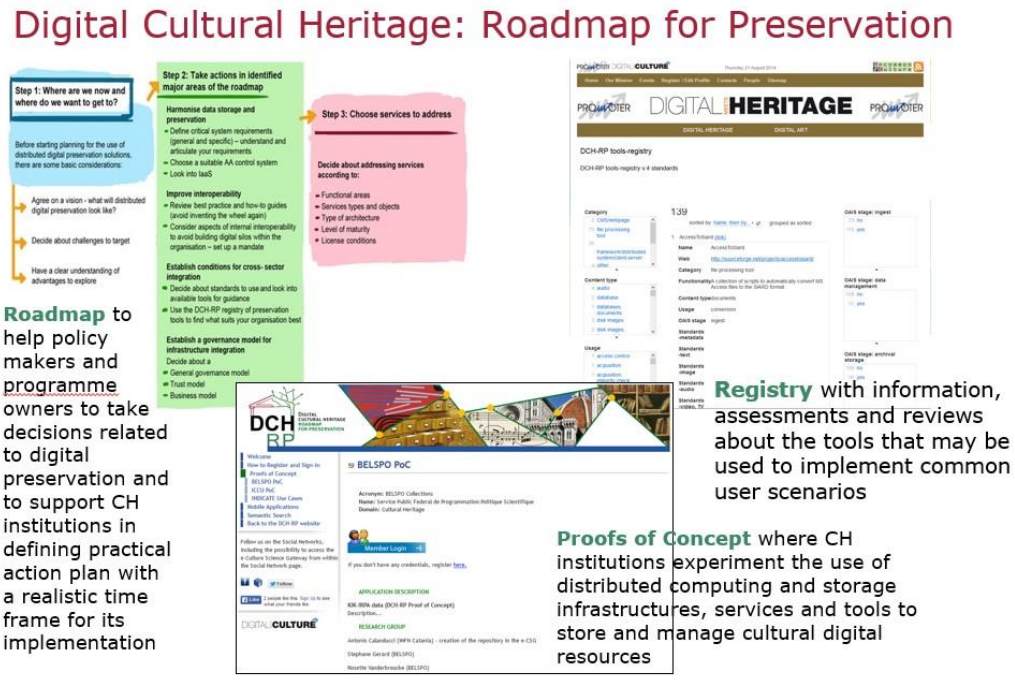
5 A WEB-SPACE FOR THE ROADMAP

The Roadmap for the implementation of an e-infrastructure to support citizen science and crowed sourcing in the area of cultural heritage and humanities represents the main outcome of the Civic Epistemologies project.

By definition, a Roadmap is not useful if it is not widely disseminated, validated and endorsed by the user groups that it aims to target. The Civic Epistemology project contributed substantially to the creation of a wide community of people coming from different sectors (policymakers, cultural heritage institutions, citizen scientists, activists, e-infrastructure providers, etc.) who demonstrated interest in the work done for the development of the Roadmap. Now it is important to keep alive and continue to nurture this community, creating awareness about the final version of the Roadmap and fostering its diffusion and implementation in Europe and worldwide.

Furthermore, a Roadmap cannot be considered as a final step. It has on the contrary to be considered as a living document that needs to be continuously maintained, updated and improved as time passes, technology changes, new requirements have to be taken into account, and so on.

It is for these reasons that the Civic Epistemologies project plan to create a dedicated web-space where it is possible to download the last version of the Roadmap, but also where it is possible for everyone to provide feedback and comments, a kind of Forum dedicated to the use of e-infrastructure services and facilities for citizen science and crowd sourcing targeting the DCHH domain.



Example of an overview of sections in the web-space for the Roadmap (taken from the DCH-RP project)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Apart from presenting and discussing the Roadmap, this web-space will link also to other relevant material, information and services that are linked to the Roadmap itself and that contribute to supplement it.

In particular, a section will be dedicated to the Registry of Services and Tools that was developed in the Civic Epistemologies project as a practical instrument to help different stakeholders.

By the end of the project, the web-space will be hosted as a section in Digital meets Culture (http://www.digitalmeetsculture.net/heritage-showcases/dch-rp/). The partners are committed to continue the work on the Roadmap even after the end of the project period and in this framework they are discussing about creating a URL dedicated to the Roadmap to be maintained on a longer period.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Example of an entrance page to the Roadmap in Digital meets Culture (taken from the DCH-RP project)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

6 CONCLUSIONS

This first draft version of the Roadmap has the ambition

- to provide a description of what a Roadmap for the use of e-infrastructure to support citizen research could look like;
- be used as a basic document for discussions.

Therefore, the report raises more questions than it can answer. It sketches the profile of a possible Roadmap, coving the most important issues to be handled by Civic Epistemologies project.

A ground breaking part of the concept that the project is aiming to introduce, is the possibilities to customise the citizen science focus services provided by e-Infrastructure, i.e. tailoring the service portfolio and characteristics to the actual tasks and requirements. However, even if the e-Infrastructure resources seems to be allocated in ways that could support citizen science activities quite well, the general conclusion must be that the market for those distributed services is still in its infancy, even if this market is developing quiet rapidly with a focus on the reach domain.

We now that ICT are powerful drivers of creativity in a number of areas, but technical know-how is still often lacking. An important issue is, therefore, the level of maturity in the DCHH domain to handle distributed services for citizen science. E-Infrastructures can reach their maximum potential in serving the DCHH domain in practice only if the domain is prepared to exploit the opportunities offered by using e-infrastructures.

From contacts with different stakeholders it is clear that parts of the DCHH domain is not yet taking full advantage of technologies to engage with wider audiences.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

ANNEX 1 INITIAL SET OF CRITICAL SYSTEM REQUIREMENTS

General needs and requirements

Examples (listed regardless of priority):

Miscellaneous issues

- Reliability and robustness
- Assurance of valid licensing procedures, commercial conditions, and transactions
- Open, scalable, and flexible solutions (built on open industry standards like J2EE and XML)
- Ease of use (for example, user-friendly interfaces)
- Multilingualism

Content/information issues and metadata issues

- Mechanisms for integration and automation of appraisal and ingestion of digital material
- Automatic metadata capture and extraction
- Separation of content (information) and metadata
- Various content formats (from print-based documents to digitized images)
- Ontologies for both visual and textual concepts
- Annotation services

Performance issues

- Scalability (up to hundred terabytes or more)
- Performance for hundreds of thousands of electronic documents

Trust issues and security issues

- Authenticity and integrity of data
- Continuity (which means the handling of information, both data and metadata, for at least the next 100 years)
- Identification of digital objects which are in danger of becoming inaccessible due to changes in technology
- Security during transmissions of files between countries
- Validation (certification) of software and hardware environments required to render the digital objects
-

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

Infrastructure-related issues

- Distributed systems
- Virtualisation

Hardware-related issues

- Support of many storage media and devices
- Backup and restore

Specific requirements

Need for simplicity

Integrating research data workflows with e-Infrastructures normally requires significant levels of computing and IT expertise, not always available in DCHH institutions. The solutions developed need, therefore, to be tested for their simplicity of installation, management and use.

Metadata

The metadata connected to a digital object is crucial for the possibilities to retrieve it and to preserve it for future use. It has to include basic descriptive information about the file as well as information about the file format of the object. The metadata collected about a digital object helps to place it in context, as well as give specific information, which is essential for making sure the object in mind is authentic (hasn't been added to or modified in any way). This is especially important for digital files, which in contrast to print media can be easily changed in ways that may not be easily apparent. Metadata can be linked to the digital object or encapsulated with the digital object itself. Encapsulating the metadata with the object ensures that the information stays with the file, no matter where it goes. Linking the metadata but storing it separately ensures that the information about the file can be recovered even if the object itself is lost. Depending on the actual situation, a decision about metadata has to taken before a institution in the DCHH domain enters into distributed services to support citizen science.

Storage in different locations

Archival data (master files) can often be stored offline, since they are infrequently accessed. It is best practice in many cultural heritage institutions to write digital archival data to more than one type of media and then store these in different locations.

Digital resources in continual use (surrogate delivery files) will typically be stored online. Online storage is often mirrored across multiple disks using redundant disk arrays (RAID).

Today clustered (data center) and distributed storage systems are normally used for distributed storage. A storage cluster consists of at least two independent storage nodes, running under the control of relevant software. When one of the nodes fails, the other immediately takes over all of its duties.

A data center is a facility housing computer systems and associated components like telecommunications and storage systems. It generally includes services such as redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression) and security devices. The concept Dynamic Infrastructure is a design of data centers making it possible for the underlying hardware and

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

software to respond dynamically to changing levels of demand in more fundamental and efficient ways. This concept is also known as *Infrastructure 2.0* and *Next Generation Data Center*.

Cloud storage is often implemented with complex, multi-layered distributed systems built on top of clusters of servers and disk drives. Sophisticated management, load balancing and recovery techniques are needed to achieve high performance and availability. While there is a relative wealth of failure studies of individual components of storage systems, such as disk drives, relatively little can be found reported, so far, on the overall availability behavior of large cloud-based services connected to Citizen science (including storage). Special care has therefore to be devoted to this issue before entering into a solution based on distributed services.

Migration of data and metadata

A routine error-checking schedule should be implemented and a strategy drawn up for migrating data and metadata to suitable formats as necessary. If a file format is becoming obsolete and a migration is planned, archival master files should be migrated to new formats that are non-proprietary. Quality control checks should follow any migration or refreshment so that any loss of data integrity can be identified and quickly addressed.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

ANNEX 2 LICENSE AGREEMENTS AND TERMS OF USAGE

When using distributed services for the support of citizen science activities is it important to understand and communicate the license agreements and terms of usage that are associated with digital resources, "born digital" ones as well as digitised representations of other cultural heritage artefacts. The Linked Heritage project investigated this topic and reported seven overall license types relevant here and broke these out further, for example describing at least four variations of the Creative Commons (CC) licenses in routine use.

The following table briefly summarises the licenses mentioned.[9] The table also mentions a highly structured method for license expression, namely ONIX-PL; this is not a license in itself but rather a machine-readable framework for conveying licensing and usage terms, conditions and prohibitions.

License	Description/purpose	More information
BSD Berkeley Software Distribution	One of a group of permissive software licenses, imposing minimal restrictions on the redistribution of the software covered by the license	http://en.wikipedia.org/wiki/BSD_licenses
CC Creative Commons	A series of public copyright licenses. Currently seven such license types exist	http://creativecommons.org/licenses/ See the website for more information on each license type: CC BY, CC BY-SA, CC BY-NC, CC BY-ND, CC BY-NC-SA , CC BY-NC-ND and CC0
GNU FDL GNU Free Documentation License	A "copyleft" licence designed for the free documentation of software, but which can be used for other text works	http://www.gnu.org/copyleft/fdl.html
GNU GPL GNU General Public	A free software licence granting the licensee the	http://www.gnu.org/copyleft/gpl.html

[9] More details can be found in Linked Heritage deliverables.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 632694

License	Description/purpose	More information
License	right to change and redistribute the software free of the prohibitions of copyright law	
ODbL Open Database License	A license covering data in databases and allowing licensees, under certain conditions, to share create or adapt the database or its content	http://opendatacommons.org/licenses/odbl/
ODC PDDL Open Data Commons Public Domain Dedication and Licence	A license covering data in databases and allowing licensees, without attribution, to share create or adapt the database or its content	http://opendatacommons.org/licenses/pddl/1-0/
ONIX-PL ONIX for Publication Licenses	An XML format for the communication of license terms for digital publications in a structured and substantially encoded form	http://www.editeur.org/21/ONIX-PL/