Service Priorities and Best Practices for Digital Cultural Heritage

Edited by the DC-NET Working Group 3: New Services Priorities

Foreword

DC-NET is one of the results of ten years of cooperation among the Ministries of Culture in Europe in the field of digitisation of cultural heritage.

This project, co-financed by the European Commission, under the e-Infrastructure topic of the Capacities Programme of FP7 and leaded by the Italian Ministry for Cultural Heritage and Activities, coordinates the participation in-e-Infrastructures of the Ministries of Culture, their Agencies and other cultural bodies (museums, libraries, archives, audiovisual, archaeological sites, etc.), across Europe. Its main aim is to generate a powerful and comprehensive plan of joint activities for the implementation of a new data and service e-Infrastructure for the research in the field of the Digital Cultural Heritage.

A set of preparatory actions was carried out by the project in the areas of dissemination, training and definition of research priorities.

This document, edited by DC-NET WP3, intends to give an overview of Service Priorities and Best Practices for Digital Cultural Heritage. It is targeted to institutions from the cultural heritage domain involved in the creation and use of online data resources for DCH that could benefit from e-Infrastructures, as well as to e-Infrastructures service providers, which can learn more about cultural institutions needs in order to offer more empowered services.

Digital Cultural Heritage systems, both the more and the less complex ones, need to be maintained, constantly updated and enriched with new functionalities and contents, as well as with a wide range of end-to-end services and tools to facilitate the integration and to increase the research capacities in the sector.

The challenge is to guarantee the quality of content, performance, scalability and stability of these systems pursuing at the same time a cost reduction, not only as regards the costs connected to the digitising process in itself, but the costs for its sustainability.

These challenges may only be gained through the continuous collaboration among all stakeholders, and the constant relationship with the main European and international initiatives in the field of research on Digital Cultural Heritage.

Concerning this, DC-NET is also the reference point for the digital cultural heritage in the framework of JPI (Joint Programming Initiative), an initiative coordinated by MiBAC, the main goal of which is to coordinate the research programme on cultural heritage (tangible, intangible and digital) of EU member States and to define the European research agenda in this field.

Rossella Caffo
DC-NET Project Coordinator

Acknowledgements

Writing a handbook on the subject service priorities and best practices for digital cultural heritage is in itself a challenge. The branch of knowledge is still in the beginning of being explored, and good examples to follow on are few and not easily found.

Luckily, we are in the position to have at our disposal a number of high skilled professionals from all over Europe. First of all, we would like to mention the French and Swedish members of the DC-NET project Work Package 3 on New Services Priorities, Aurélie Audeval and Sanja Halling, who actually carried out the work on which this handbook is based.

The contribution of Rosette Vandenbroucke has been important to improve our understanding of the world of the e-infrastructures also with her final check of the content of this handbook.
Ciaran Clissmans impressive excellence in structuring information and organising it into understandable wording has – as always – been an invaluable contribution. Many thanks for that.

We will give special thanks to Rossella Caffo, Maria Teresa Natale and Marzia Piccininno in the DC-NET team who have provided important input and inspired us to move on in our work.

Special thanks also to Michael Hopwood from the international standards coordinating group EDItEUR, work package leader in the EU project Linked Heritage, who has helped us to shape the English language. Remaining peculiarities in the use of it must be blamed only on us.

Last but not least, we would like to express our gratitude to all the partners of the DC-NET project who have not only shared their experiences and great knowledge with us, but in fact produced some of the ground pillars for the work of Work Package 3 and, in the consequence of that, this handbook.

February 2012

Christophe Dessaux, Antonella Fresa and Börje Justrell

Table of Contents

Introduction

1




1.1 Document overview

This document is an outcome of the DC-NET project. It gives an overview of the results of the work carried out by the project in the field of new services for digital cultural heritage (DCH). Its aim is to serve as a practical handbook.

The first target audience of this handbook includes institutions from the cultural heritage domain that are involved in the creation and use of online data resources for DCH, such as content providers active in managing and adding value to the content itself, as well as consumers/end users of content, such as curators and researchers in the humanities.

1.2 Document structure

This document consists of four chapters:

Background – This chapter reviews relevant aspects of the DC-NET project such as aims of the initiative and key stakeholders, added value of the project, and the work carried out to date, in order to give the reader the context in which this document shall be considered.

New Services Priorities – This chapter presents a set of prioritised services for DCH together with some elements of the discussion carried out within the DC-NET project that led to the priority ordering proposed in this document. This ordering refers to the most important new or improved services for the DCH sector which can benefit from e-Infrastructures support.

Best Practices on New Services – Gives an overview of the most important practical guidelines and lessons learned collected by the DC-NET project in connection to prioritised services for DCH.

Conclusions – Summarise the overall results

Background



2.1 The DC-NET project: aims and key stakeholders

DC-NET is an ERA-NET Coordination action contributing to the European Research Area Network, and having e-infrastructures for the digital cultural heritage as its research field. The main activity of the DC-NET project is, according to its description of work:

- to bring people together;
- to explore how e-Infrastructures can add value to the research in the cultural heritage sector;
- to pool programmes and resources to sustain the envisaged research and development.

The cultural sector has a strong interest in using ICT to pursue its own missions and institutional objectives, but there are still constraints on the adoption of ICT by the cultural heritage. The DC-NET project seeks to step up digital culture and e-Infrastructures cooperation and coordination, across and within the Member States and Associated States. The key stakeholders in this process are

- cultural ministries and agencies;
- research organisations and activities in the field of digital heritage;
- national and international e-Infrastructures;
- cultural heritage institutions (in particular libraries, archives, museums and audiovisual archives).

The DC-NET project is closely aligned with the digital agenda for Europe[1] and with the Joint Programming Initiative (JPI) for Cultural Heritage[2]. DC-NET is an integral element of the strategic development of digital cultural heritage in Europe, particularly in how digital cultural heritage relates to e-Infrastructures.

1) Digital Agenda for Europe: http://ec.europa.eu/information_society/digital-agenda/index_en.htm
2) JPI for Cultural Heritage: summary at http://ec.europa.eu/research/pdf/citizens-summaries/jpi-cultural-heritage/summary_en.pdf#view=fitH&pagemode=none. See also OJ L 106/18-19 of 28/4/2010

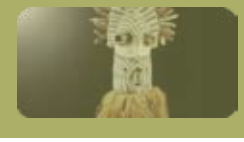
2.2 Added Value of the DC-NET project

While the use of e-Infrastructures for arts, humanities and social sciences has increased in recent years, the 'hard' sciences (notably High Energy Physics that is a dominant user of grid computing e-infrastructures) remain the dominant users of e-Infrastructures. The cultural heritage sector has the potential to be a significant user of e-Infrastructures capabilities, just as the e-Infrastructures sector can facilitate important research and services progress for DCH. For example: "digital humanities" programmes already exist in many universities, the cost of digitisation is high and demands for efficiency on low budgets, galleries, libraries, archives, museums (GLAM) are facing difficulties in implementing persistent identification of their digital content. All these are cases where the use of e-infrastructures can have impact.

The vision of the DC-NET project is a seamless data and services infrastructure for cultural heritage, which reliably provides key services such as preservation and backup, authentication and data integrity, collaborative research environments, advanced (cross-collection, multilingual and semantic) search and retrieval, while enabling intellectual property management and authorised use of DCH content. The project aims to bring this potential closer to realisation, by establishing networks of contacts between DCH and e-Infrastructures, by exploring priorities and capabilities, and by agreeing a common action plan for R&D into the future.

By doing this, the project will contribute to the establishment of an ongoing forum for coordination of European research in the application of e-Infrastructures to cultural heritage. This will contribute to improve the coherence and coordination of such research across Europe, via the development of a Joint Research Agenda. The ongoing coordination and Joint Research Agenda will support national actors to perform activities in cooperation, which they would not have been able to tackle individually.
The increased interaction between cultural institutions, ICT, digital cultural heritage research, and e-Infrastructures (GRID, NREN)





will stimulate the existing Virtual Research Community (in the humanities and in the other sciences that have liaisons with the DCH secotr), which will be a long-term benefit of a more effective coordination of research agendas, policies and programmes.

While DC-NET, as an ERA-NET, focuses on bringing together the DCH and e-Infrastructures communities for the benefit of the researchers, the pan-European portal Europeana focuses on providing to the European citizens a gateway to European cultural heritage by bringing together metadata records and previews of digitised materials from cultural institutions across Europe and beyond. New services identified and specified by DC-NET for the DCH sector, which can be facilitated by collaboration with e-Infrastructures, will be valuable for Europeana too in the sense that they will enable DCH organisations to provide higher-quality content to Europeana[3].

2.3 Work carried out by the DC-NET project on the New Services

Work Package 3 New Services Priorities of the DC-NET project accomplished two main tasks:

1. to draw up a new set of priorities for digital cultural heritage research;

2. to establish a systematic exchange of information (such as national and regional programmes, funded projects, research priorities, evaluation practices, organization and management) and best practices on existing programmes that can be activated for the actual implementation of research.

In order to achieve these objectives, the activities of this Work Package were divided into two blocks:

3) www.europeana.eu

Service Priorities

Starting from a national level, three key inputs have been used: an internal project report on research priorities, the Handbook on e-Infrastructures (DC-NET deliverable D2.2), and an analysis of discussions held by all partners with their e-Infrastructure providers in order to combine cultural heritage research aims with the capacities of the e-Infrastructures on a national level.

Preceded by some preparatory meetings, a main workshop was then held in Paris in February 2011, where all cultural ministries from the partners' countries, stakeholders from Member States from beyond the consortium countries, and e-Infrastructures providers were invited. At this workshop, as well as at the earlier preparatory meetings, the research priorities of each partner were reviewed. The aim of this review exercise was to generate a common set of priorities for future work which is applicable to, in the first place, all partners' member states. The priorities discussed were then analysed and further investigated to form the basis for the content of the Digital Cultural Heritage Services Priorities Report (DC-NET deliverable D3.1).

Best practices

The overall goal for this work was to inspire by spreading good ideas, lessons learned and knowhow. Information and best practices have been gathered during the execution of the tasks of WP3. Targets have been defined in order to form the basis for a permanent exchange among the participants of the network of contacts that the DC-NET project has established.

New Services Priorities

3

A set of prioritised services is presented in this chapter. They shall be seen as the first step towards a common research action plan for DCH.

Two elements serve as a baseline for the discussion:

• Firstly, most DCH services can benefit significantly from access to e-Infrastructures. **Qualified performance, scalability and stability** are three ground pillars for all services which can be enhanced by access to faster networks, more CPU power and increased storage, facilities normally provided by e-Infrastructure.

• Secondly, **the concept of common services**, where a service is provided 'in the cloud' and used by multiple DCH websites, digital libraries, repositories, etc. is an appealing approach for the provision of services on an e-Infrastructures platform. By removing location as a factor to be taken into account, and by providing extensive processing and storage capabilities, e-Infrastructures can support services that remove replication, overlap and redundancy in the DCH community as a whole.

And naturally, allowing DCH to focus on priorities can contribute easing strain on stretched budgets.

The services priorities identified by DC-NET can be divided into three categories:

1. Services for content providers. These are the services related to the creation of online data resources for DCH. Content providers are the memory institutions which digitise and place content online.

2. Services for managing and adding value to the content itself. These services go beyond simply publishing digitised material online; they focus on the ways to enhance data, to make access to data more user-friendly and attractive and to facilitate re-use of data in different contexts like education, commercial ventures, collaborative projects, etc.

3. Services which enable support and enhance virtual research communities and the activities of content consumers. These services focus more on the users of the content and less on the content itself. Content consumers are those who consume content for research. They typically include academic researchers but also staff members of institutions in the cultural heritage domain involved in research. These institutions may, therefore, be both content providers and content consumers. Content consumers can also be referred to as 'end users'.

In the context of DC-NET, **content** refers to the digitised cultural heritage material which is held by content providers across Europe. This includes the digital holdings of libraries, museums and archives, as well as the digital representations of historic landscapes, buildings and towns (text, images, video, sound, 3D). Content can also include data resulting from the scientific investigation of cultural objects (such as chemical analysis of paintings and sculptures for restoration, physical characterisation of the materials, geological inspections, etc.). There are very large amounts of this material already available in digital form, and much more is being continuously generated. The path that the content will follow before reaching its user, can be described as in figure 1.

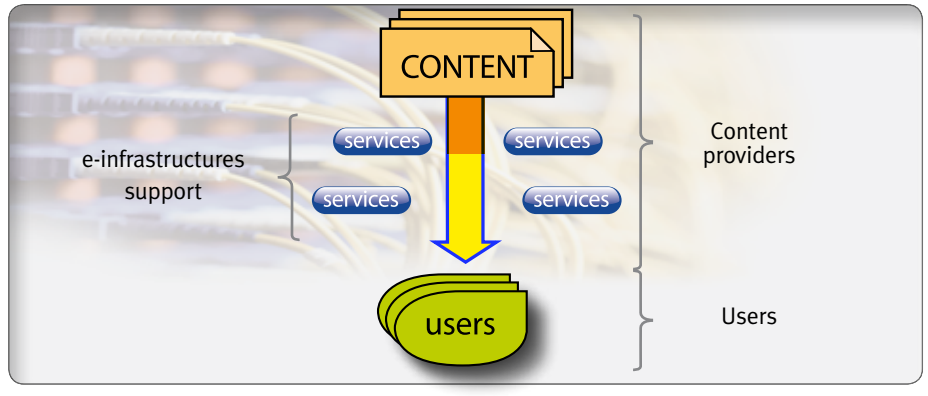
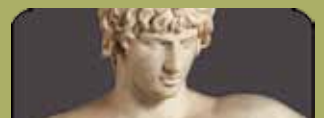


Figure 1: In the following sections each category of services is presented more in detail.

3.1 Services for content providers

3.1.1 Identification of areas of priorities

The creation of an online 'location' for the presentation of DCH materials online is a central part of any digital heritage initiative aiming at public outreach. The model applied will most commonly be a content management system, a portal system, a digital library or a digital repository which has been specifically designed and built for the purpose. Online resources have several challenges in common dealing with customised technical solution or blend of solutions: insularity in terms of searching, changes in location and consequent broken links, high cost of establishment, vulnerability to technical problems, limitations on server capacity and processing.

These challenges lead to the identification of the first set of areas of priorities for the DCH research community. These areas are:

Interoperation: these services are required to simplify the interoperation of online DCH resources. Such services may involve the transformation of metadata and database data sets to a particular common standard, or the creation of mappings between the metadata and data formats of different online resources.

Aggregation: these services can harvest and combine material from several DCH resources. They are needed to enable multiresource facilities to be delivered to the users.

Cross-Search: these services are needed to enable searching across multiple online DCH resources. This may involve the transformation of a centralised search into the local search calls for multiple DCH resources; alternatively it may build on the interoperation services mentioned above, or utilise linked data as described below.

Semantic Search: these services are needed to take advantage of semantic web technologies, such as linked data and ontologies. Such services can offer a new level of intelligence and relevance in meeting the needs of users and can support new models of searching navigating and experiencing complex and interlinked resources. Again, such searches must work across multiple DCH resources.

Persistent identification: these services are needed to simplify or automate the maintenance of persistent identifiers and their mapping to specific locations within DCH resources. Persistent Identification services which shield external users from internal reconfiguration are needed if portals and services which build on interoperability are to be stable in the medium to longer term. This is an interesting problem in relation to the semantic search question above. In fact, without PID, semantic search will be frustrating or useless.

Setup services: these facilities and tools are needed to simplify the construction of online digital culture resources. Templates and tools (ideally open-source, or available under unrestrictive licences) which can be customised to simply create digital libraries and repositories are a high priority. Such a service would reduce the cost of DCH initiatives; several data resources using the same tools would also have a high likelihood of interoperability and standards compliance.

Stable platforms: these infrastructure services, which provide hosting, backup and availability of the DCH resources, are needed to raise confidence and to ensure that strategic assets and major investments are not jeopardised.

Scalability: these services are required, as the amount of digitised material grows and the usage levels increase. The last 3 problems, namely Setup services, Stable platforms and Scalability are strongly linked.





Intellectual Property Rights and Digital Rights Management: this is one of the top priorities for services needed. In general, all content providers are happy to see their content being made more widely available and being re-used, but only in a manner which complies with their own legitimate interests and policies. These policies vary from one content provider to another, reflecting different missions (public service versus profit) and business models, but all require that access to their content is controlled, recorded and acknowledged. To involve in the discussion the commercial sector perspectives can be considered here to help a more comprehensive understanding of the issues involved in these topics.

3.1.2 Technical services needed

The provision of DCH content online and the creation of data resources involve significant technical work. Apart from the act of digitisation itself, many issues arise around metadata, creation of the data resource, interoperability with other data resources, access control and user authentication, long-term preservation and more.

The technical services for content providers that are of huge importance for the provision of cultural heritage material online, in the sense that they can simplify or improve the quality of the content provider's technical work, are the following:

Data Resource Setup Services. They have been identified as a priority by the DCH community. The establishment of a new data resource, or the upgrading of an existing resource to a new platform, is an expensive and time-consuming process. The availability of services to simplify and support the creation of the data resource would reduce these costs; the use by many DCH organisations of the services would also increase interoperability and simplify aggregation. In order to deliver such services there is a requirement for the storage and rapid network access that

e-Infrastructures can deliver. Delivering such services also implies substantial software development after extensive consultation of DCH content providers, to ensure that any such facility will actually be used. However, a stable and scalable data resource environment, with tools which support setup and migration from current platforms would be attractive and valuable.

Interoperability. It is a perennial issue for DCH, due to the many in-house, proprietary and non-standard technologies which are deployed across different countries, sectors and even organisations. Much research effort has been expended on addressing interoperability, and good progress has been made in some areas (e.g. the widespread support for some form of Dublin Core for metadata, the increasing availability of RDF-based Linked Data as well as CIDOC-CRM and LIDO, the new proposal for standard). However, there remains a great deal of work to be done, if the data resources of organisations across Europe are to interoperate effectively. While several projects have attempted to address the interoperability issues in the past, the availability of e-Infrastructures provides facilities and an environment which can overcome previously critical issues such as scalability and complexity.

Aggregation. Effectively the same issue as interoperability, but here the aim is to combine heterogeneous data sources rather than enable them to share data. The same discussion applies as for interoperability. The case of aggregation includes also aspects of IPR, , as the combined data must embody some sort of licensing agreement.

Advanced Search Support (multi-site, multilingual, semantic): The ability to search across the entire DCH landscape, including several data resources, is a long-term strategic objective of DCH. The difficulties associated with such an objective are well established and have been addressed in research projects for many years. Some key issues include:



i. Similar concepts are described in different languages, depending on the country where the data resource resides or has been created;
ii. Similar concepts are represented in using different terminologies, even in the same language, by different data resources;
iii. Semantic similarity is difficult to capture, so that search results may omit relevant elements because of specific choices of query term.

Support for advanced search has been identified as a priority not only for content providers but also for end users (see below).

Reliable long-term preservation. It has been a major concern for digital cultural heritage for many years. The rapid evolution of technology has meant that both hardware and software used for storing precious cultural materials have become critical failure points. This is most critical for 'born digital' items where there is no option to 'redigitise' the original. There are many examples of this problem, such as the physical ageing of magnetic media with associated data loss, the dramatic drop in support for previously-ubiquitous storage media, the difficulties extracting data stored with obsolete file formats, etc. A number of solutions to such predictable obsolescence have been discussed, focusing on migration (from old media and software to more modern ones) and emulation (by new software of old software). New models, methods and tools are introduced on national level by memory institutions as well as on international level by the research community in general (including EU financed projects).

Persistent identifiers (PIDs). They are critical for the long-term access to DCH items, as they overcome the issues associated with internal restructuring, changes in website format, etc. Broken links are a particular issue for data aggregators and portals such as Europeana. PIDs have been addressed by several initiatives across different sectors (e.g. DOI, ARK, URI, URN, etc), with various schemes successfully in use. The requirement for PIDs is common

across almost all DCH data resources, so there is clear appeal to providing a single (possibly distributed) PID infrastructure for DCH which would enable the creation and maintenance of PID-to-URL mappings and their lookup on a regular basis. The requirement for rapid network access is well suited to an e-Infrastructures environment, while the need to reliably store the PID tables is also fulfilled by e-Infrastructures.

Linked Data Generation. The creation of linked data stores from digital cultural heritage resources is an attractive method to enhance interoperability and to link DCH resources to data sources from beyond the cultural sector. Benefits are also expected from incorporating commercial sector data where the use of PID is more advanced and established.

User Authentication & Access Control. User Authentication is a requirement for community-related services, such as annotation, discussion, conferencing and collaborative environments. It is also useful for controlling access to restricted resources, such as high-quality digital masters and high-resolution content which is not available to the public. User authentication and access control are identified as high-priority services because of their central role in digital rights management and the enforcement of intellectual property rights, a difficult technical area which requires extensive storage of access controls and permissions, of information about the usage of content, about tracking the sharing and re-use of material. Any solution will require substantial ICT resources, including rapid network access, and large scale data storage, services to be delivered in a robust and scalable manner. It should be noted, however, that technology is only part of the solution to IPR and rights management and that the legal basis and the agreements between content owners and those who enable access to the content, as well as with end users, are critically important.

User authentication and access control are services which are useful to both content providers and to content users.

3.2 Services for managing and adding value to the content itself

While the core of DCH is the digitised content itself, this content can be made much more useful and appealing if it can be fully appreciated by the content consumers/end users and if it can be re-used for serving new purposes such as education, combination with other technologies, etc.
The following services priorities focus on improving the usefulness of the content itself, by operating at the content level, as opposed to the online resource.

Geo-referencing associates one or more physical locations with a digitised item. It enables the item to be shown on a map, illustrating its relationships to other items and to other geo-referenced information. Maps-based interfaces are familiar and intuitive to many users, while a preference for local content is also very common. Services to identify the locations with which an item is associated (e.g. by analysis of the item metadata) and to convert this to latitude and longitude values are a high priority. They add a new perspective to the content itself and facilitate its re-use in 'mashups' and in powerful user interfaces.

3D visualisation and manipulation of complex digital items can make them much more attractive and useful. This is particularly the case for heritage landscapes, buildings, etc. (which can also benefit from geo-referencing), where virtual reality can provide unique immersive experiences. Services which create 3D visualisation from 2D images and/or geo-reference information are very attractive.

Annotation is a service aimed at the end users of the content. However, annotation also enriches the content which is annotated, increasing the understanding/appreciation of the item, linking it to similar or related content, etc.

Linked data generation is the generation of linked data is primarily a service which enables the linking of multiple data resources in combined or linked searches. As such, this service is most applicable to content providers. However, the availability of a linked data version of the content also enriches the actual content and makes it useful in a wider environment.

3.3 Services for content consumers

A research community adds value to the DCH material by annotating it, by identifying and validating links to other materials, by researching it and increasing the body of knowledge available about the material. Traditionally, the research around a corpus of material is geographically local to it; by enabling virtual research communities there are new opportunities for distributed researchers to contribute. Of course, issues such as security, accountability, intellectual property rights, etc. must be taken into account, so that online access and services do not impact negatively on the DCH resource or its content.

The content consumers/end-users in the DCH environment can be categorised into three types:

- A large number of relatively undemanding individuals, who are quite satisfied by the services already available, but remain ready to consider any new services that may emerge;
- A more modest number of individuals who have moderate requirements for change or new functionality;
- A few highly-demanding individuals with very advanced requirements. Requirements identified by these individuals may 'trickle down' to provide new services to the other user groups; however, some requirements may be infeasible.

Content consumers are of course free to pick and choose the services which they will use, and those which they identify as being of greatest importance. This model enables a broad range



of services to be made available, without the need to actually deliver them for all members of a (primarily undemanding) user community.

User authentication is a common requirement for many online services, not just in DCH. 'Single Sign-on' which provides authentication for several DCH resources (and perhaps other online facilities too) would be very useful, simplifying the use of several resources in a seamless manner and removing any requirement for the users to remember several logins and passwords. It is an important service for content providers and their data resources.

Group-based access control can further simplify the user's accessing several different data resources. It also offers a convenient way to give the user access to new resources, without having to do this on an individual-by-individual basis. It is an important service for content providers and their data resources.

Collaborative Environments can offer facilities for researchers and users to work together on their research. The environments can be private (i.e. with access restricted to certain groups and/or individuals) and provide services such as dedicated storage, document sharing, conferencing, discussion, annotation, etc.
Conferencing and Discussion services which enable researchers to meet virtually, to share and discuss images, text and other items, to record and replay such interactions would greatly 'tighten' virtual research communities and make them more closely resemble traditional co-located teams.

Annotation services which enable researchers to add their own contribution (text, images, links, etc.) to DCH materials are a high priority. Such services stimulate the users to engage with the material, add a sense of ownership and involvement and also lead to exciting and interesting material to be published. In addition, such services enable users to link items to other items, both in the same resource and beyond it. Such human-created linkages tend to be valid and relevant.

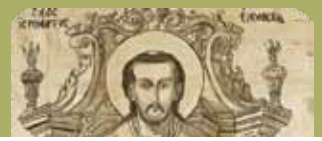
Advanced search support is a key service for content providers and for use with data resources. However, the advanced search service is primarily used by end users (content consumers), and has been identified as a priority from that perspective.

Visualisation services to enable improved visualisation of complex content entities have been identified as a priority. This includes services such as the creation of 3D model and virtual environments, immersive environments such as virtual tours of monuments and historic landscapes, manipulation tools to allow interaction with 3D items, etc.

3.4 Common services

An important potential value of e-Infrastructures to the DCH community is the possibility to act as platfoms from which services can be delivered which are used by several DCH projects, communities or bodies (so called 'common' services). Most DCH projects and initiatives have many requirements in common, and achieve these requirements in isolation; this leads to wasted time and investment, as well as poor potential for interoperability.

A common service is a technical facility which is useful to several independent DCH projects or initiatives. The service is delivered over the network, has clearly defined inputs and outputs and provides a distinct set of services. A simple example is permanent identifier (PID) maintenance – a single database of PIDs and URLs is maintained on the network, and any request to access a PID (regardless of the content provider or the data resource) is resolved into a URL by the service. PID-to-URL mappings are maintained via a web interface by the content provider, but the actual service is provided without any requirement for action by the content owner. This removes the need for individual data resource owners to maintain PID directories of their own.



The concept of common services relies on the availability of storage, CPU power and high speed data networks – precisely the facilities which e-Infrastructures can offer.

The following 'common services' are identified as being of high priority and also as being able to benefit significantly from the support of e-Infrastructures.

- **Multilingual** search and retrieval: relies on high-speed and accurate mapping of search queries in one language to metadata values in another. Higher-quality results are achieved where controlled terminologies and vocabularies are utilised. Such technical dictionaries and mappings should ideally be centralised for reasons of consistency and efficiency. In addition, the dictionaries and mappings are useful not just for one data resource, but for all such resources aiming to provide multilingual services – rather than have each site create its own dictionaries and mappings, it is much more efficient to provide such services centrally, on a platform which is responsive enough to support many users at once. These services require a stable platform (given their strategic value to the critical search function), rapid network access (so that search is not delayed due to its being multilingual) and substantial storage and CPU power (for rapid analysis and matching across large dictionaries and terminologies). A cloud-based solution is in many cases ideal.

- **Synonyms:** different terms to describe similar concepts (e.g. 'name' instead of 'title', 'author' instead of 'creator'), even within a single human language. Centralised, rapid-access tables of synonyms and alternatives are needed to enable this to happen seamlessly and without loss of quality of service. Here we are talking about 2 features of controlled vocabularies; the ability to specify multilingual labels, and authority control. Again, there are the same compelling reasons for centralising this process as a common service, rather than 're-inventing the wheel' for each data resource. These services require similar stable platforms, rapid network access and CPU/storage facilities as multilingual search.

- **Semantic similarity:** this is a 'hot' research topic both within and beyond DCH. Several European financed projects have a distinct focus on this area. Semantic analysis of both search queries and metadata (data resource descriptions) requires the use of ontologies and new ways of representing entities and relationships (e.g. using RDF triples for Linked Data). Once again, a centralised approach reduces the risk of overlap and repetition, as well as encouraging the identification of semantic similarities across metadata items from multiple sources. The large amount of data, the need to be able to search and retrieve from this data in a rapid manner and the expectation that it will be frequently used from multiple locations all suggest that an e-Infrastructures-supported approach would be very beneficial. It can also be added that standardisation makes semantic identification easier; it requires solid analysis to work, and serious buy-in from big players to gain traction though.

- **User Authentication** is a requirement for community-related services, such as annotation, discussion, conferencing and collaborative environments. The allocation of usernames, the generation, storage and protection of passwords and facilities for dealing with new members, with lost passwords/usernames etc. are all processes that do not differ significantly from one data resource or online service to another. This makes them excellent candidates for a centralised or cloud-based solution, or alternatively for an identical and interoperable solution distributed across several physical locations. From a user perspective, a single sign-on without the need for remembering multiple user names and passwords



is also very attractive, as is the ability to navigate from one data resource to another without the need for repeated login. This too argues for a common service, rather than several separate services. Finally, the technical expertise needed to ensure the security of authentication information means that such a service may best be delivered by the ICT experts of the e-Infrastructures, rather than the DCH community, where the main focus of expertise is elsewhere.

- **Access control:** Linked to user authentication is the concept of access control, which enables specific users or groups of users to have specific levels of access (e.g. view, annotate, print…) to specific items or groups of items. A common approach to access control will greatly facilitate the end user navigating from one collection (or data resource) to another; it will also simplify the creation of virtual collections and temporary exhibitions made up of items with similar access policies. The specifics of access control will vary from one data resource to another (it is infeasible to expect the access policies of all DCH collections to be similar), but the mechanisms (group creation, rights allocation, group membership management) could effectively and efficiently be centralised or at least harmonised across a distributed solution.

3.5 Priority ordering

The previous sections present the most important new or improved services for the DCH sector, which can benefit from e-Infrastructures support. The priority ordering (from the most important service, to the least important) must reflect both the input of the DCH community and also the impact which e-Infrastructures support can make. In addition, the amount of R&D needed to make the new service a reality must also be taken into account – 'low-hanging fruit' should be prioritised, so that tangible benefits are achieved in the short term. None of the services listed is 'low priority' – all of the services promise substantial benefit to the DCH community. The following ordered list takes these into account.

Priority 1: Long term preservation. This is given top priority because

- e-Infrastructures are in a position to offer substantial storage immediately.
- Preservation is an urgent issue, with hardware obsolescence and the need for media migration a looming problem for many digital culture collections.
- Relatively minor R&D is needed for 'raw' or 'simple' storage (clearly, digital repository frameworks and other middleware are another issue[4]).
- Moving critical DCH holdings to a cloud environment has the potential to place it on a new and stable platform for other new services.

Priority 2: Persistent identifiers. This is given second priority because

- Broken links and moving digital content due to website redevelopment, etc. are an urgent problem impacting on all portals and aggregators.
- There is excellent research and development already done in this area, and one or more of the existing schemes could be adopted with minimal adjustment.
- A PID registry is a service which can benefit significantly from a centralised or cloud-based implementation; e-Infrastructures have much to offer here.
- The PID requirements do not vary significantly from one DCH initiative to another – this represents a service useful to most DCH work, 'out of the box'.

Priority 3: Interoperability & Aggregation. This service receives third priority because

- Although very important for European DCH, it is technically

4) Consider, for example, the substantial work being done by the DRIVER initiative in e-Infrastructure-enabled digital repositories.



challenging and will require substantial effort and investment. The effort increases linearly with the number of data resources made interoperable – a solution for one resource will not often be immediately applicable to another.
- Once achieved, interoperability can underpin aggregation, some advanced search services and some elements of digital rights management. As such, it is strategically important, even if difficult to achieve.

Priority 4: Advanced Search. This service receives fourth priority because
- It is technically challenging and requires a great deal of effort. Several initiatives are underway, others have been completed and more will be needed before this realises its full potential.
- It can benefit significantly from interoperability and from PIDs; thus, these should be addressed first. Because it includes topics such as multilingual support, it can be argued that this should be completed before interoperability; it may emerge that both of these services will be addressed hand-in-hand.
- Of course, this is perhaps the single most important user-facing new service.

Priority 5: Data Resource Setup. This service receives fifth priority because
- It is effectively an 'add-on' for persistent storage and preservation, rather than a key issue in its own right.
- It will require customisation and adjustment for each DCH organisation which sets up a data resource – thus as a research proposition it is not a neat package like PIDs or multilingual terminologies, for example.
- However

  ◦ It is identified as valuable by many DCH organisations
  ◦ It can aid in interoperability and in advanced search
  ◦ It can facilitate the migration of data to the e-Infrastructures environment, with associated benefits in stability, scalability and network access.

Priority 6: User Authentication and Access Control. This service receives sixth priority because
- Persistent storage/preservation and advanced search are higher priority.
- The R&D required will be significant, and customisation (especially for access control) for each data resource means that the research itself will not be 'neat package'.
- None the less.
  ◦ This research is essential for community-focused, collaborative and interactive/annotation services .
  ◦ This research is an important element of the enforcing of IPR and DRM policies.
  ◦ User authentication is an ideal centralised and scalable services for e-Infrastructure implementation.

Priority 7: IPR and Digital Rights Management. This service receives seventh priority because
- Other services are also of very high priority, and benefit more from e-Infrastructures.
- However
  ◦ IPR is an urgent issue for all collaborative and portal initiatives
  ◦ DRM must be addressed effectively if a content industry is to be nurtured in the EU.
  ◦ and IPR is and aspect linked with all the other priorities.

In summary, all the services identified here are important and valuable. Ordering them in terms of priority is both difficult and somewhat subjective – the ordering presented here combines short-term feasibility of implementation with e-Infrastructures added-value and DCH strategic importance, but even the last items are critically important for digital cultural heritage in Europe.

Best Practices on New Services

4



4.1 Introduction

Work Package 3 in the DC-NET project has as one of its main objectives to systematically gather best practices. These best practices are a number of practical rules, so called 'rules of thumb', which are recommended to be considered before implementing the services priorities defined in previous chapter.

In this chapter we will, firstly, define the targets, procedures and gathering tools for the collection of good practices. The aim is that it will form the basis for a permanent exchange among the participants in a future network. Secondly, we will give examples of best practices gathered by the DC-NET project.

4.2 Targets, procedures and gathering tools for the collection of best practices

4.2.1 Targets

In the previous chapter, the services priorities identified by the DC-NET project have been divided into categories, namely

1. Services for content providers and related to the creation of online data resources for DCH;
2. Services for managing and adding value to the content itself;
3. Services which enable, support, and enhance virtual research communities and the activities of content consumers/end users.

These categories have also been used as targets for the collection of good practices.

5) Workshops in Tallinn January 2010, Best Practice meeting in Rome in July 2010, DC-NET Conferences in Brussels in October 2010 and in Budapest in June 2011

4.2.2 Procedures and gathering tools

To begin with, the procedure for collecting best practices was carried out as follows:

- Out of the descriptions, given by the partners in the DC-Net project, how e-Infrastructures are combined with DCH research priorities in their respective Member States, some basic best practices were extracted.

- During meetings, conferences, seminars and workshops organised by the DC-Net partners, collection of best practices has been one major issue . In particular, a workshop in Rome in July 2010 was dedicated to best practices.

- At the new services workshop in Paris in February 2011, best practices captured so far were validated and complemented through prepared questions for discussion.

From now on, best practices will be built on the results presented in this document and on complementary collections of updated information for which the DC-NET website will be used as the main channel.

4.3 Selection of best practices

4.3.1 Services for content providers

The following specific activities have been identified by DC-NET as important for content providers to have in mind:

Creating data resources. Content providers are the memory institutions which digitise and place their content online, either by themselves, through outsourcing or as a mix of these two alternatives. Their work includes, in general terms, besides selecting and digitising content and preparing metadata, also building data resources:

- for public outreach (websites, portals, digital libraries etc.);
- for keeping their data accessible and usable on long, mid- or short term (back up, storage, digital repositories).

The creation of these data resources is identified by the DC-NET project as a core task for any DCH initiative and here e-infrastructure can fulfil an important role. Every DCH initiative needs some type of data resource as a structure within which digital content can be placed and managed, preferably online. Typically, issues to be included are interoperability across resources, tools for searching and navigating the resources, the preservation of the resources (long, mid- or short term), etc.

Networking. Of great importance to content providers is effective networking as a base for sharing information across DCH initiatives and sectors including best practice. While DCH research has been an active field for some decades, there is still an urgent need for greater integration, consensus and collaboration. The research field has tended to fragment along national and sectoral (museums, libraries, archives, monuments, etc.) lines; additional fragmentation has been caused by the involvement of multiple ministries in some countries, by different funding models, etc. Enabling and developing a Europe-wide virtual research community dedicated to digital cultural heritage should, therefore, be of high priority for content providers who seek to develop their content on line for a broader audience.

Knowledge and documentation of user needs. The content consumers/end users of DCH work tend to be DCH researchers and/or members of the general public. In addition, commercial end-users can also be considered (e.g. picture researchers who want to license images from museums and galleries). As in other areas of ICT, there is a temptation to deliver solutions which take advantage of new technologies, without first ensuring that such a solution is actually required. Developments in the broader Web, such as social networking, semantic web technologies, adaptive systems etc. all offer potential for new DCH services – however, it is essential that end user appetite for such services is actually present. Research into the interests and requirements of the content consumers/end user communities is, therefore, needed. This research must be ongoing over time, as the content consumers/end user priorities and expectations will themselves evolve. Their priorities must be documented and validated so that DCH providers, technology partners and the users share a common vision.

Who is who. DCH research is typically fragmented along national and sectoral lines. This can lead to repetition of research, to the re-invention of common solutions and to organisations failing to learn from the experiences of others. A key solution to this issue is to be able to identify the organisations and the individuals who are active in DCH and to summarise their particular areas of expertise and experience. A ‘directory’ or ‘who is who’ guide to Europe’s DCH community would be a valuable addition to the current situation, where contacts tend grow organically, serendipitously and by word-of-mouth. As with user needs documentation, such a directory would need to be updated on a regular basis, to make sure that levels of activity and types of experience are truly reflected.

Policies and programmes. A further perspective on the DCH landscape is that of policies and programmes. Each country, each sector and often each organisation will have policies and guidelines for accessing, sharing and processing the content under its control. Such policies range from digitisation guidelines to access control, commercial re-use to integration with national and international portals. Where DCH research is being envisaged, it is important for all concerned to have a clear picture of the policies that will impact on the research. An awareness of the programmes already in place, and how new research initiatives can impact on those programmes, is also important. The impact or take-up of a new service will be affected if a similar programme already exists in a target ‘market’.
Research into this domain must be kept up to date and must cover




as many countries, sectors and organisations as possible. It is to be hoped that ongoing efforts at consensus (including DC-NET) may help to simplify the policy landscape over time; however, it remains a critical element of any successful collaborate DCH research initiative.

Handbooks and technical reference. The experience of completed and ongoing DCH initiatives needs to be shared with new and planned projects. While all projects produce reports and deliverables, a greater emphasis should be placed on the generation of widely-applicable guidelines and handbooks which formalise and integrate the lessons learnt. The loss of knowledge when projects end is an important source of inefficiency in DCH research; measures to address this are a priority. While this may not represent a research domain in its own right, it is an important aspect of the activities of any new research project.

Consensus Building – Seminars, workshops, dissemination. While sharing knowledge and experience is very beneficial, international and cross-sectoral impact can best be achieved by face to face discussion and consensus building. Best Practice Networks, ERA-NETs and bodies such as the MSEG all contribute to this process, but such efforts need to be ongoing, need to involve as many Member States and organisations as possible and need to address as many issues relevant to DCH as possible[6].
The mechanisms used for consensus building are typically meetings, seminars, conferences, etc. The results can be usefully validated through dissemination events where they are exposed to, and receive feedback from, the broader DCH community. The results of the consensus should, as noted above, be documented as handbooks and technical reference, so that they are useful for other initiatives.

Training is a particular example of knowledge transfer which is of particular urgency for digital cultural heritage. As noted above, expertise in various DCH areas tends to be distributed around Europe, with different organisations having particular areas of expertise, often held by a small number of individuals.

6) This, for example, is the driving force behind the DC-NET ‘follow-on’ project INDICATE (www.indicate-project.eu), which involves many Mediterranean-bordering countries including several outside the EU.

‘Traditional’ cultural heritage skills, such as curatorial and preservation expertise, are widely distributed in the cultural heritage domain. However, in order to deliver digital cultural heritage, there is a requirement for individuals whose skills include both these ‘traditional’ ones and also the variety of technical and ICT skills which impact on DCH.

An important research topic is the development and agreement of a training syllabus specifically for DCH, and its implementation across Europe. The use of such a syllabus would deliver a new generation of expertise, with a common level of knowledge on the many topics relevant to DCH. Clearly, such a syllabus could be delivered both to new students and as professional development for existing cultural heritage practitioners.

The following topics have been identified by DC-NET as being a minimum set of modules for such a syllabus:

- End user interaction and target analysis
- Collection and digital repository management
- Selection of material to be digitised
- Digitisation and creation of digital resources
- Management and curation of digital objects
- Long term preservation
- Multilingual information access
- Semantic web technologies
- Collaborative environments
- IPR
- Usability

Risks to handle

There are some broader aspects on information technology that are more critical than other when handling digital data. They can easily become failure points or at least holdbacks if not treated as


potential risks when digitising and placing content online.

Hardware obsolescence. The use of 'cloud' computing and e-Infrastructures removes one of the main issues with long term preservation. Such platforms entirely hide the physical computing infrastructure from the end user, so that the user is no longer concerned with disk formats or sizes, tape drives, magnetic media age or other hardware issues. The cloud provider continually upgrades his hardware as the technology evolves and migrate the content stored there on to the new platforms.

Software ageing. Data stored in old file formats relies on the availability of software to decode the files. The availability of the software in turn relies on the demand for such software and the distribution of the software. Again, a centralised or 'cloud-based' approach offers advantages – large libraries of rarely-used codecs, media software and emulators can be maintained centrally, so that finding an appropriate player is simplified. This addresses an important and common issue – requirements for archaic media players are relatively infrequent, but are high-priority when they occur. Where large files are involved, centralised or 'cloud' computing offers large-scale storage, ample memory and extensive CPU resources.

Older software to run on operating systems which are no longer common or even available. Thus not only is older software needed, but the operating systems which they rely upon must also be available (or at least emulated). This requires substantial computing resources which, again, may not be used very regularly but are essential when needed.

Content growth. The amount of material which needs to be preserved has grown explosively, following the trend of overall growth of digital and online data. As the creation and publication of digital information has become a mass-market phenomenon, the amount of information that requires preservation (both within memory institutions and by the broader public) has exploded

and can be expected to continue to expand rapidly. A scalable management and storage solution is a priority. Where storage is available from the e-Infrastructures providers, such a service can offers an excellent solution to digital preservation (short-, mid- and sometimes even long term) because

- It is hardware-agnostic and avoids issues with ageing magnetic media, changes in hardware configuration, etc;
- It has the capacity to host and run (or to emulate) a wide range of operating systems and application software needed to access and render older file formats;
- It moves the responsibility for ICT support and maintenance to a central, ICT-focused environment, enabling DCH researchers to focus on their own area of expertise;
- It is scalable to accommodate arbitrarily-large amounts of data without any requirement to re-engineer the technology.

Use of common services

Common services are often good candidates for funding at a national or even international level, because of their broad applicability and usefulness, and because they are not tied to any single data resource, memory institution or organisation.

But – common services are normally only a useful concept where very similar services are required by multiple data providers and their users. A service which requires a good deal of customisation is not a good candidate – where a 'common' service serves only a small user base then it is typically more efficient to deliver it locally.

4.3.2 Services for managing and adding value to the content itself

Managing and adding value to the content itself goes beyond just publishing digitised material online. There are many attractive ways to enhance the data, make it more user-friendly and to facilitate its re-use in education, in commercial ventures, in collaborative projects, etc. Key issues here include visualisation, geo-referencing, 3D representation and manipulation, IPR management, annotation, statistical analysis and other activities which add value to the actual content and/or make it easier to manage.

However, basic for both managing and adding value to content is the selection of an appropriate meta-data model. This is crucial for enabling search and retrieval, especially when the case is searching across multiple collections and holdings stored in different locations. There exists already an impressive collection of standards for meta-data covering various aspects, and also a large number of specific meta-data models (institutional and sectoral) that do not meet cross-domain demands properly. The key to success is in this case to choose a meta-data model based on the organisations own goals and on its general conditions, and at the same time avoid creating a new model or, more likely, one with too many exceptions to conform to standards.

Another type of services in this class are those that enable organisations to adhere to the standards that exist; for example, training, best practice guides, consultancy to set up the standard in the first place, automatic validation or “scoring” mechanisms for test data, etc.

4.3.3 Services which enable, support, and enhance virtual research communities and the activities of content consumers/end users

Services for virtual research communities focus more on the users of the content and less on the content itself. The DC-NET project is not supposed to focus on individual end-user, and we, therefore, have broadened the definition of the term ‘user’ to cover the research community as such, regardless if the research is conducted by a university or by a cultural heritage institution. A major priority for the holder of any DCH corpus is to add value to it by nurturing a research community around the material, by enabling annotation, discussion and user-contributed content. Underpinning such new facilities are requirements for user management (authentication, auditing, accountability), for tools and environments which stimulate collaboration across virtual.
The best practices collected up to now are in this section structured to reflect the average researchers (content consumers/end users) work process.

Accessing existing resources (databases, content etc.). There are mainly external database providers (over the Internet) that can offer services in this area. However, researchers are often using a mix of external and internal database service providers, where databases consulted through the Internet are in the first place provided by public services. Within the public sector, the NREN and e-Infrastructure are not used broadly for database access, at least not for the moment. But an e-Infrastructure can normally offer a more reliable data base service in terms of high capacity and sustainable connection to the Internet.

It goes without saying, that the offered services have to be defined in accordance with identified needs and avoid ‘over-equipment’. Provided database services have to be wide, but also accurate. Attention has to be paid to specific needs, especially in the arts and humanities domain. It is important to consider, that ‘new services’, in sense of accessing new external resources, can result in


positive enlargements of the researcher's models, methods and tools, and also in increased multi-disciplinary involvement.

Tools for scientific analysis on content. Researchers sometimes insist on the necessity to improve and build databases analysing tools especially for DCH. This demand mainly concerns database e-book access, better database software/tools environments for monitoring and analysing (such as visualization, processing, System-wide 3D image quality analysis tools, etc). Some researchers are also asking for aggregation and semantic research facilities in their specific domains of research.

Special tools for analysis often bring about special requirements for running them, which can cause problems when used in a standardised technical platform meant for systems in multiple every-day use. The knowledge of these tools is also normally limited inside the organisation, and special attention has, therefore, to be put on activities like upgrades and maintenance.

Publishing the results of the research. A publishing platform is not commonly used by institutions practising research (even if some begins to be available), but the need of such a tool seems more obvious in the art and humanities area. If a publishing platform is in use, it often relies on institutional solutions like NREN, other available e-infrastructures, electronic libraries etc., and sometimes also on internal solutions

The main enhancement proposed, when asking the individual researchers, often concerns the possibility to have a collaborative environment system. Here, it is possible to distinguish between different kinds of needs:

- Publishing research data for a broader audience (the researcher being a content provider);

- Collaboration between researchers in:

  - closed but wide networks,
  - in smaller and limited groups.

A standard Content Management System could in many cases be a smart solution for dedicated groups of researchers. These systems normally contain functions like

- Controlled access to data;
- Easy storage and retrieval of data;
- Reduction of duplicate input;
- Easy writing of reports;
- Internal communication between the users of the system;
- Automatic control and update of document versions.

An alternative could also be Software As A Service (SAAS), sometimes referred to as 'software on demand'.

If a publishing platform is to be implemented, evaluate the conditions carefully, and choose a service provider (internal or external) which can guarantee the security of data as well as the adaptability to researchers' needs.

Collaborative environment. There are normally a few commercial national service providers of conference system available, and researchers mostly use solutions already in place in their organisation or/and E-infrastructures. Services offered as 'conference services' normally contains a possibility to share screen and documents, to chat, to use video functions and to send files.

Using existing solutions offered by private companies like Skype, Alfresco, Adobe Connect or SharePoint is normally a first alternative, but it could be limiting and not specifically adapted to the researchers needs. Developing modified versions for specific usage of tools in collaboration with such companies could be a solution, but is normally expensive and require specialised software competence to be locally built up and maintained.




Requests for an adapted multiplatform solution have, therefore, to be carefully balanced towards needs for a better centralised service with high capacity level for all-round usage.

Regardless of which solution is used, it is important ensure that the local technical platform can handle it. Images, videos, sound system sharing etc. are capacity-consuming.

Store and preserve data. The researchers normally express three kinds of need:

- **i.** to have a back up security service to avoid computers crashes,
- **ii.** to have some individual possibilities to archive their own files short or mid-term for their own use,
- **iii.** and to have long term preservation and possibilities to share the data over time.

Backup services are broadly used by researchers. It's a central need, not only on individual level but for the whole research community, guaranteeing the immediate permanence of the research. The simple backup is a basic service for the individual/ team to handle during a research in progress. Data should normally be backed up following a fixed schedule and if of high value stored on two kinds of media separately.

The organised forms of preservation, mid-term (records) or long term (archives), are more problematic and need special expertise, but are also necessary to permit continuous access to the research process as well as to its results.

Already in the beginning of a research project, considerations have to be taken about the need of 'long term preservation' of data and the establishment of a digital preservation routine. Such a routine has to consider at least the following aspects:

1. File format. Before deciding about the file format, the relevant standards in place have to be considered and also to what degree file formats are supported by software in use by the organisation. The global user base can also be an indicator of the expected future support for a particular file format, and indicate the likelihood of sustainable possibilities for migration, when the file formats are changing. Even if a proprietary or national format may appear attractive from a technical standpoint, not using standard formats can be an obstacle to international exchange of data and use of networked resources.

2. Media choice. Regardless of choice of media, it is a fact that the media in use will become obsolete. Normally, migration to new storage media is likely to need to be made within a five years cycle. In order to avoid end user concerns about hardware issues like disk formats and media age, the physical computing infrastructure could be out sourced to an e-infrastructure or centralised 'cloud' solution. Clear agreements about level of services and responsibilities are such cases of huge importance.

3. Migration strategy. A clear strategy is needed for moving data to new file formats and different storage solutions/ storage medium in the foreseeable future (possibly less than five years, probably less than ten years). An important part is to examine the relevant standards, as compliance with standards is an indicator that a format or medium is expected to have some support into the future.

Finding models, methods and tools for mediating knowledge from digital preservation experts to those who currently have limited access to specialist know-how in this field, is an often forgotten aspect. It is also important to not underestimate the needs for mid-term archiving on an individual basis. If not properly taken care of these needs easily grow into a jungle of person-based solutions that can put whole research projects at risk.



Security Services. Most researchers have access to security services, but not to all services. Security is often an issue that researchers are not engaged in; instead they trust/delegate it to the ICT service provider. There is often a knowledge gap between the service level that researchers ask for and the level of security needed not to jeopardise the researchers' own computer environment. The simplest way to handle this gap is to intensify the communication between security service providers and end-users and to plan for actions to raise the awareness about the importance of a solid security service. Otherwise implemented services risk not to be in line with actual needs, or users will try to sidestep the security routines they think prevent them from acting as freely as they would like to do (not realising the risks).

Conclusions

Conclusions

In this document the key outcomes of the DC-NET project in new services for cultural heritage (DCH) are presented.

- Firstly, a list of service priorities has been identified that can benefit from e-Infrastructures support and facilities. This list is based on three identified categories of services:

  a. Services for content providers and online data resources;
  b. Services to enhance the content itself;
  c. Services for content consumers and end users.

  The priority listed services are Long term preservation (priority 1), Persistent identifiers (priority 2), Interoperability & Aggregation (priority 3), Advanced Search (priority 4), Resource Setup (priority 5), User Authentication and Access Control (priority 6), and IPR and Digital Rights Management (priority 7). The ordering combines short-term feasibility of implementation with e-Infrastructures added-value and DCH strategic importance. But looking at them individually all these services are important of significant value for digital cultural heritage in Europe, even the last items. This list, therefore, requires an on-going discussion and more in-depth technical validation but points towards a common research action plan.

- Secondly, the document gives an overview of practical guidelines and lessons learned collected by the DC-NET project in connection to prioritised services for DCH. These best practices consists of a number of practical rules, so called 'rules of thumb', which are recommended to be considered before implementing the defined services priorities.

- Thirdly, a general conclusion is that most DCH services can benefit significantly from access to e-Infrastructures . Qualified performance, scalability and stability are three ground pillars for all services, and they can be enhanced by access to faster networks, more CPU power and increased storage, facilities normally provided by e-Infrastructure.

- Fourthly, the concept of common services, where a service is provided centralised or "in the cloud" and used by multiple DCH websites, digital libraries, repositories, etc. is an appealing approach for the provision of services on an e-Infrastructures platform. By removing location as a factor to be taken into account, and by providing extensive processing and storage capabilities, e-Infrastructures can support services that remove replication, overlap and redundancy in the DCH community as a whole. However, common services are normally only useful as a concept where similar services are required by multiple data providers and their users. Services that require a lot of customisation are not a good choice and services with a small user base are normally more efficient to deliver locally.

Traditionally, the research around a corpus of material is geographically local to it. Removing location as a factor will enable virtual research communities and open up for distributed researchers to contribute and take active part in scientific work processes.



Abbreviations

AHSS	Arts, Humanities and Social Sciences
ARK	Archival Resource Key, a scheme for the persistent identification of information objects
CH	Cultural Heritage
CIDOC-CRM	CIDOC Conceptual Reference Model (CRM), www.cidco-crm.org
CPU	Central Processing Unit
DCH	Digital Cultural Heritage
DC-NET	Digital Cultural Heritage Network, the ERA-NET project supported by the European Commission in the frame of FP7 e-infrastructures programme.
DOI©	Digital Object Identifier, system for identifying content objects in the digital environment
DoW	Description of Work (Annex 1 to the DC-NET Grant Agreement)
DRM	Digital Rights Management
EC	European Commission
ERA-NET	European Research Area Network, a type of project supported by the European Commission in the frame of FP7 to contribute to the joint programming with Member States on strategic areas of the research, to contribute to the establishment of the European research area.
FP7	The seventh Framework Programme for research and technological development of the European Commission.
GLAM	Galleries, libraries, archives and museums
ICOM	International Council of Museums, icom.museum
ICT	Information and Communications Technologies
IPR	Intellectual Property Rights
JPI	Joint Programming Initiative summary at http://ec.europa.eu/research/pdf/citizens-summaries/jpi-cultural-heritage/summary_en.pdf#view=fitH&pagemode=none. See also OJ L 106/18-19 of 28/4/2010
LAM	laboratoire d'acoustique musicale http://www.culture.gouv.fr/culture/conservation/fr/laborato/labo_acoustique.htm
MSEG	Member States Experts Group
NREN	National Research and Education Network
PID	Persistent Identifier
RDF	Resource Description Framework, a general language for conceptual description and modelling of information that is implemented within web resources
R&D	Research and Development
URI	Uniform resource identifier, a string of characters used to identify a name or a resource
URL	Uniform (or universal) resource locator; it is a specific character string that constitutes a reference to an Internet resource
URN	Uniform resource name; it is a type of uniform resource identifier (URI)
WG	Working Group
WP	Work-package




Glossary

- *Best practice*
Best practice is, as the DC-NET project defines it, a number of practical rules and examples of successful implementations where these rules are applied, so called 'rules of thumb', which researchers in research organisations and others using e-Infrastructure are recommended to consider before implementing services in this field. This document contains examples of best practice for content providers, those who are managing and adding value to the content itself as well as consumers/end users of content.

- *CIDOC*
CIDOC is the International Committee for Documentation of ICOM. It provides the museum community with advice on good practice and developments in museum documentation.

- *Cultural Heritage Domain*
The Numeric Study , ordered by the Commission and performed 2007 – 2009, used the following list of types of institutions as belonging to the cultural heritage domain:

Archive/records office
Audio-visual or film institute
Broadcasting institute
Museum of art, archæology, history
Museum of science, technology, ethnography/ethnology?
Other type of museum
National library
Higher education library
Public library
Special or other type of library
Other type of organisation

- *e-Infrastructure*
The new generation of integrated ICT-based infrastructure. E-Infrastructures exploit and also interconnect several separate components and layers, like networks, grids, data centres and collaborative environments, and can include supporting operation centres, service registries, certificate authorities, help-desk services, storage and other remote resources.

- *GRID*
A set of services over the network that allows geographical dispersed users to share computer power, data storage capacity and remote instrument documentation.

- *LIDO*
harvesting schema for delivering metadata for use in the service environment of an organisation's online collections database, portals and aggregations

- *New services*
In the context of the WP3 of DC-NET, new services are those services that have not been implemented before. In some cases new services exist but are not known and therefore not used by content providers and/or the research community.

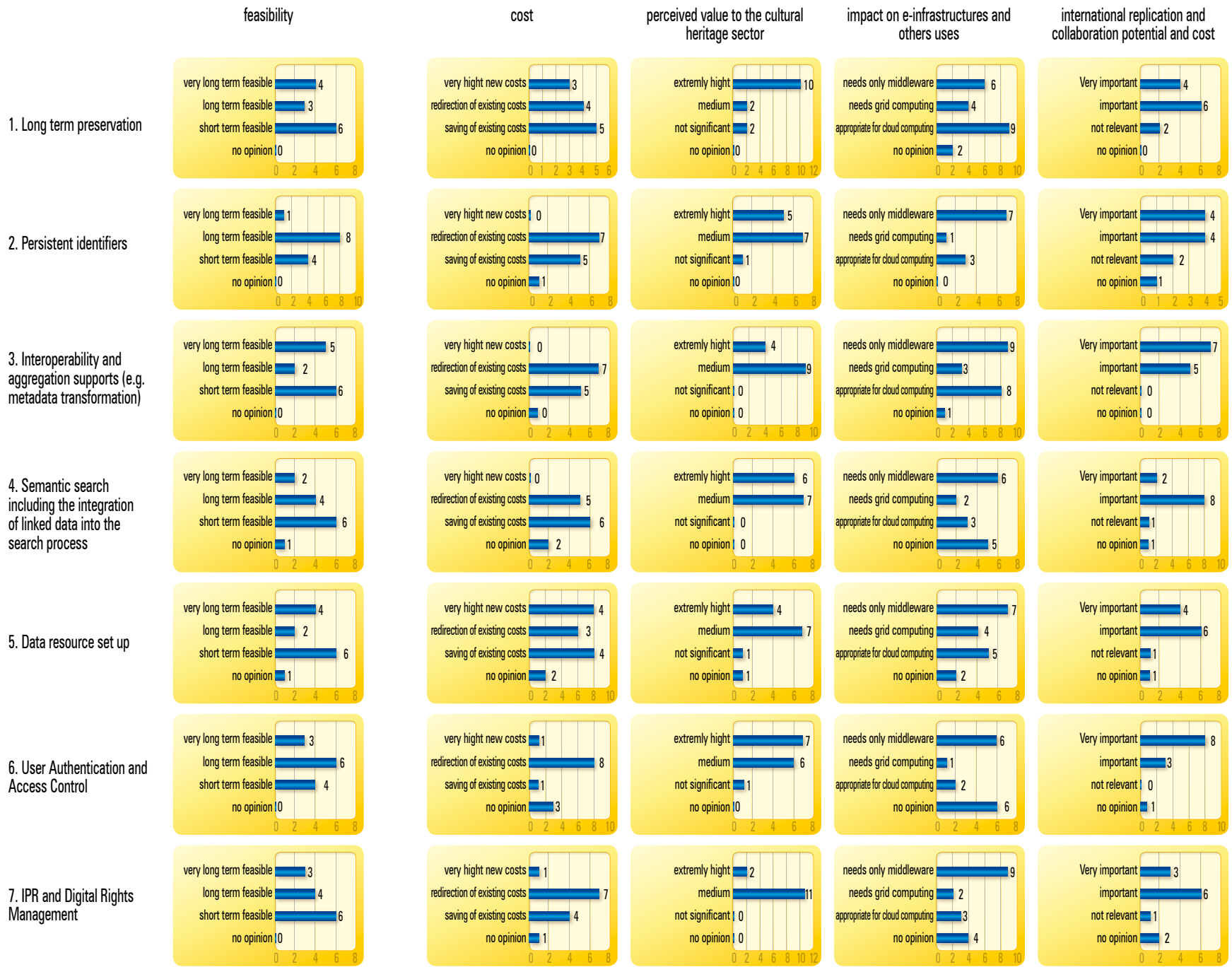
- *Research community*
The research community is, in this document, the community of researchers working in the field of DCH. Our definition is broad and includes any actor researching in the field of DCH. We take into account research centres, universities, and also cultural institutions with research activities in ICT applied to cultural heritage and research activities in the humanities where ICT is used to support such research.

- *Services*
Any e-Infrastructure that can be seen as an answer to needs of the research community and/or of the content provider is a service in the context of this document. It can be defined as a general service or as a specific service. General services mean technical capacity and power. Specific services mean special services which answer to specific needs.

- *Users*
Users are in the context of this document the research community as such and also the individual researchers in the humanities.

Contribute to the work in progress

An online survey has been opened to validate the service priorities described in this Handbook.
The picture below illustrates the results of the survey so far, but this is totally a work in progress.
If you are interested, you can contribute to the survey with your opinions.
The online questionnaire is available at dcnet.situla.org



Contacts

This publication is intended as a first step towards the implementation of a federated e-Infrastructure dedicated to the Digital Cultural Heritage able to valorise and integrate the many efforts made by the whole sector in the last years and many people are working in this direction.

For further contacts:

Rossella Caffo, DC-NET Project Manager
rosa.caffo@beniculturali.it

Antonella Fresa, DC-NET Technical Coordinator
fresa@promoter.it

Work Package 3 Leaders:

Borje Justrell
borje.justrell@riksarkivet.se

Christophe Dessaux
christophe.dessaux@culture.gouv.fr