

Report on the activities and outputs of the Expert Advisory Board

Support Action Centre of Competence in Digitisation (Succeed)

December 2, 2014



Abstract

This report summarises the activities and recommendations made by the experts on the Succeed's Expert Advisory Board (EAB). According to the Description of Work, "the EAB will advise the Project Coordinator and the Coordination Board on those matters related to principal technical and strategic issues regarding the Centre of Competence in Digitisation. They will provide guidance on licensing, standardisation and good practices, in addition to strategic issues such as commercial and institutional relations, business models, etc., which will be compiled in this deliverable".

In order to achieve these goals, the EAB met in Madrid during Digitisation Days (May 19-20, 2014). At this venue, interviews with the members of the EAB were recorded and a debate was organised to discuss the future of digitisation in Europe. The members of the EAB also formed the panel responsible for choosing the winners of the Succeed awards, which were delivered in a ceremony, as part of Digitisation Days.



Succeed is supported by the European Union under FP7-ICT.

Document information

Deliverable number	D1.1	Start: M1	Due: M23	Actual: M23
Deliverable name (DoW)	Report on the activities and outputs of the Expert Advisory Board			
Internal/External	External			
Activity type	SUPP			
Participant	UA, BVC			
Estimated person months for this deliverable	2.00			
Dissemination level	PU (public)			

Document history

Revisions				
Version	Status	Author	Date	Changes
0.1	Recordings	Silvia Ponzoda	20/05/2014	
0.2	Transcription	Alicia Blaya	05/07/2014	
0.3	Draft document	Rafael C. Carrasco	07/11/2014	

Approvals				
Version	Date of approval	Name	Role	Signature
1.0	02/12/2014	Hildelies Balk (KB)	Internal supervisor	

Distribution (this document was send to)			
Version	Date of sending	Name	Role in project
0.3	12/11/2014	Alicia Blaya	T1.4 leader
0.3	12/11/2014	Isabel Martínez	Technical manager
1.0	02/12/2014	Cristina Maier	Project officer



About this document

This document is a public deliverable (D1.1) of the Succeed project (FP7-ICT-600555).

Copyright statement

This document can be distributed under the Creative Commons Attribution-Share Alike 3.0 license.¹

¹<http://creativecommons.org/licenses/by-sa/3.0/>



Contents

1	Introduction	1
2	Succeed awards	1
3	Digitisation days	2
	3.1 Expert Advisory Board meeting	3
	3.2 Panel discussion	4
	3.3 Summary of recommendations by experts	5
4	Conclusion	7
Appendix A Appendix: Interviews with the experts		8
1	Steven Krauwer	8
2	Milagros del Corral	11
3	Michael Keller	13
4	Jill Cousins	16
5	Frank Frischmuth	19

1 Introduction

According to Succeed's Description of Work, the Expert Advisory Board (EAB) "will advise the Project Coordinator and the Coordination Board on those matters related to the principal technical and strategic issues regarding the Centre of Competence in Digitisation. They will provide guidance on licensing, standardisation and good practices, in addition to strategic issues such as commercial and institutional relations, business models, etc., which will be compiled in this deliverable"

As decided by the project's General Assembly, on May 16, 2013, the members of the EAB would be the following highly acknowledged personalities in the fields of digitisation and digital humanities:

- Milagros del Corral (National Library of Spain, former director).
- Jill Cousins (Europeana, director)
- Frank Frischmuth (German Digital Library, director)
- Michael Keller (Stanford University Library, director)
- Steven Krauwer (Utrecht University, director of CLARIN)
- Andrew Prescott (King's College London, expert in digital humanities projects).

All of the aforementioned members agreed to collaborate with the project activities and to help in disseminating the outputs of Succeed. This collaboration was materialised, in particular, through the EAB's intense participation in the Digitisation Days (May 19-20, 2014, Madrid) organised by Succeed and the IMPACT Centre of Competence in Digitisation. At this venue, the EAB members were asked about key issues in digitisation and a debate was organised to discuss the future of the field in Europe. The members of the EAB also formed the panel for the Succeed awards, which were granted in a ceremony, as part of the Digitisation Days.

2 Succeed awards

The members of the expert Advisory Board acted as the committee for the Succeed Awards recognising the best initiatives for the application of digitisation technology to digitisation programmes. After internal discussion and evaluation, two projects from the 19 submissions were selected as being the most innovative contributions:



- Hill Museum and Manuscript Library (<http://www.hmm1.org>). The Hill Museum and Manuscript Library is a remarkable, distributed and cooperative effort whose objective is to both preserve endangered content and provide online services that support full access to facsimilar documents.
- Centre d'Études Supérieures de la Renaissance (<http://cesr.univ-tours.fr>). The CESR is committed to producing both high-quality digital editions (in XML-TEI) and facsimile with reliable transcriptions, and this activity has led to the development of new open-source tools for both layout analysis and transcription.

In addition to these outstanding initiatives, it was considered that two more projects also deserved special recognition—in the form of Commendations of Merit:

- Tecnológica (<http://tecnologica.com>). Tecnológica maintains a constant innovative spirit, creating, for example, new technology for the digitisation of perforated hard cardboard disks and their conversion into MIDI files.
- London Metropolitan Archives - University College London (<http://www.greatparchmentbook.org>). The objective of the Great Parchment Book is to publish online both the images and transcript of this of fire-damaged book (165 parchment pages). As a result, new technology (planned to be open-sourced) has been created, specifically, algorithms that can be used to flatten pages from images obtained with multiple photographs.

The awards were delivered at a formal ceremony which was—together with the DATECH conference (www.datech2014.info) and joint the technology exhibition—, part of the Digitisation Days programme (www.succeed-project.eu/digitisation-days).

Additional information on the Succeed awards can be found in the deliverable D5.2 of the project, which deals specifically with all the awards and competitions organised by Succeed.

3 Digitisation days

During Digitisation Days (Madrid, May 19-20, 2014) Milagros del Corral, Jill Cousins, Michael Keller and Steven Krauwer met in a panel session together with Giuseppe Abbamonte (European Commission, Director of Media and Data, DG CONNECT) to discuss the European challenges in the digitisation domain. The members of the EAB also met the project coordinator and some partners of Succeed before the panel session.



Live interviews were also recorded and the full audiovisual content has been published on the Impact Centre of Competence video channels (www.youtube.com/user/theimpactproject and www.vimeo.com/impactcoc).

The principal conclusions reached and key topics discussed in these activities are summarised below.

3.1 Expert Advisory Board meeting

During the Expert Advisory Board's meeting with the Succeed representatives and coordinator, the following issues were raised:

- Digitisation must be more focused on users. In this respect:
 - There is still not sufficient thinking as regards who is using the content and how, a weakness which creates a gap between the content and its huge potential usage in, for example, research and education.
 - We therefore need more cooperative tools with which to annotate and share content.
 - It is necessary to promote advanced services such as transcription tools, measuring devices for maps, methods for the creation of mash-ups, utilities for easy connection with Geographic Information Systems (GIS) and Global Positioning Systems (GPS).
- Further development of the technology is required. In particular:
 - We still need improved support for scripts such as Hindi, Hebrew, and Arabic.
 - Organisations have little capacity to develop software and companies therefore need to be involved in the creation of new applications.
 - The promotion of higher interoperability will allow different platforms to work together.
 - Key lines for future development are the integration of machine translation working with automatic transcriptions, and new applications that will provide enhanced support for the semantic web and for linked data applications.
- The adoption of long-term preservation policies is urgent.
 - Culture promotes understanding, and the digitisation of cultural heritage must therefore remain a priority.
 - The storage of digitised content should be distributed for a safer preservation.



- The quality of digitisation must be a function of the expected usage of the content. For example, a manuscript, map, or document in which layout is relevant needs higher resolution than a printed source whose value is mainly the textual content.

3.2 Panel discussion

In the framework of the Digitisation Days and the DATeCH conference, a panel session to discuss hot topics about digitisation was chaired by Milagros del Corral (International Organisations Advisor and Chair of the Succeed Expert Advisory Board). The panel also involved some members of the Succeed EAB, Jill Cousins (Europeana), Steven Krawer (Utrecht University) and Michael Keller (Stanford University Library) in addition to Giuseppe Abbamonte, Director of Media and Data (DG CONNECT) of the European Commission. The discussion was structured in three key topics:

- Should the focus of digitisation move now into quality or into quantity?
- The business model for the digitisation of cultural heritage: new ways to fund digitisation and preservation.
- New usages of digitised content: requirements, expectations, and needs.

The session was opened with an overview, during which Jill Cousins presented the Europeana strategic plan for 2014-2020. She recalled the still minor fraction of cultural heritage that has already been digitised in Europe —about 10% of the total. She also stressed the importance of making this content reusable, and recommended focusing on quality —and not only on quantity—, for better reusability. As a good example of such practices, Jill mentioned the Rijksmuseum.

Milagros del Corral asked the participants the first question: “Should we move the focus towards quality or quantity?” Steven Krawer explained that, from a researcher’s point of view, “we need to resist the temptation to go for volume”, while Michael Keller distinguished between “documents that make it possible to understand the national entity, culture, etc., which need more digitisation with a sufficient quality”, and documents digitised by private institutions, in which “the high quality is terribly important.”

Milagros del Corral moved then the debate to a new question: “The business model: How to fund all these exercises, including preservation”. With regard to this issue, the participants had slightly varying points of view. According to Giuseppe Abbamonte, the hardware needed for digitisation is already available on the market at low prices. In this respect, companies, and, of course, also the Member States, can then play a very important role by sponsoring digitisation.

Steven Krawer stated that public funding is needed for digitisation and he claimed that the European Union must view digitisation as an opportunity,



since there is a commercial potential in the digitisation of cultural heritage. There should therefore be incentives for companies and universities to work together and to invest further in digitisation.

Jill Cousins suggested a model in which people using the content pay for its digitisation. In this respect, Michael Keller gave the example of the Mathew Parker library at Stanford University, at which the content is freely available in low resolution and users can pay for a higher quality version. These fees allow the library to improve content. Michael Keller stressed that this model has received several criticisms, but no alternative business model has been proposed.

Milagros del Corral stated that the model of universities in the United States is quite different to that of the public universities in Europe, where people believe that once they have paid taxes which sustain these universities, they should not need to pay for extra services.

At this point, Giuseppe Abbamonte highlighted copyright as one of the most important barriers to digitisation in Europe.

After the discussion on the business models, a final topic was raised: “New usages: requirements and needs”. Steven Krauwer pointed out that History or Social Science researchers are not aware of their needs with regards to the usage of digital content, and he noted that we need to make an effort to educate them as regards the potential usages of digital content. Michael Keller mentioned, as examples, several projects that the University of Stanford is undertaking in this field.

Milagros del Corral then ceded the floor to some of the attendees in order to obtain new points of view on the issues discussed. Some attendees raised questions such as the need to provide open access to open data, the fact that the EU is moving the target from creating content to reusing content and the need for a central policy establishing which documents have priority in the digitisation of the 90% of cultural heritage that is still to be digitised.

The session was closed by Giuseppe Abbamonte, who presented the EC point of view concerning the situation of the digitisation of cultural heritage in Europe, focusing on museums. He stated that there is no reason for not providing high quality content since the hardware needed is available on the market at a low cost. He also remarked that the EC provides member countries with regional funds. In Spain, for example, about 1% of these funds are spent on culture.

The full video of this session can be viewed at <http://vimeo.com/97923641>.

3.3 Summary of recommendations by experts

This section summarises the main opinions and recommendations —classified according to the topic— put forward by the EAB members during the personalised interviews performed. The full transcription of these interviews can be found in the appendix.



Promoting the digitisation of cultural heritage

The experts reminded conference attendees of the Parliamentary adoption of new EU Open Data rules and stressed their relevance as regards unlocking cultural heritage potential.

The EC has done a lot and has attempted to encourage member states to digitise, but there are also other priorities to which the structural funds are often applied. The EC can promote the use of part of the structural funds to improve the roads of the future, which are digital.

The stress should now move towards quality and reusability of the digitised content. The paradox is that even if about one tenth of the cultural heritage has been digitised, only a small fraction of the digitised content is available in aggregator platforms —like Europeana— and an even smaller fraction of this open content is ready to be used on mobile devices or to develop added-value applications for education, tourism, etc. Enhanced availability of the content is crucial.

Pilot schemes demonstrating the potential of digitised content could trigger new usages for new collections, languages or domains. It is therefore highly recommendable to invest in prototypes and observe new techniques and identify excellence. Furthermore, the use of digitised content should be promoted in the education system: education can play an essential role in allowing younger people to access and exploit digitised cultural heritage.

Funding digitisation

Cultural heritage institutions should increase their efforts in digitisation: there must still be some public funding for this purpose since there is not sufficient economic interest in cultural assets such as old texts or small languages.

In this respect, libraries need to play a role as places for work and co-creation, promoting new usages and easing the search for and discover of content. The digitisation of masterpieces that help to understand history, culture or international relations should be prioritised.

Coordination of initiatives

It could be said that a world wide digital library already exists. However, it is not coordinated and it will constantly become larger and more linked. It is therefore recommended that communication be reinforced between active digitisation initiatives in Europe and America. These conversations must involve practitioners and, most importantly, also policy makers.

More intensive transnational research is needed since research is still highly fragmented in Europe. An initiative to set up a coordination board (perhaps a centre of competence or a network of excellence) should be started. A website on which to register digitisation initiatives could help to advance coordination.



The role of centres of competence

Centres of competence can coordinate and guide local, regional and national initiatives and function as an information provider on all subjects concerning digitisation (including legal problems).

Centres of competence can also assist in the integration of complementary technologies. For example, effective access to content may require image processing and OCR to be linked to machine translation technology. Bridging the gap between technologies is a key role of the centres.

The centres should also:

- study user requirements (for example, an observatory could facilitate the understanding of users' needs);
- lead innovation;
- establish guidelines for practice and standards;
- create networks of experts who share their experience, thus helping others to advance faster;
- promote the transfer of knowledge.

New lines of research

Access should be simplified in the future, particularly through the use of linked data and semantic web technologies. The integration of multilingual support for information retrieval is, for example, widely demanded. Improving users' understanding of what will of course help as regards not creating false expectations.

4 Conclusion

Succeed's Expert Advisory Board (EAB) met in Madrid during Digitisation Days (May 19-20, 2014). At this venue, interviews with the members of the EAB were recorded and a debate was organised to discuss the future of digitisation in Europe. The members of the EAB also formed the panel responsible for choosing the winners of the Succeed awards, which were delivered in a ceremony, as part of Digitisation Days. From this experience a number of recommendations for the advance of digitisation in Europe were proposed and they will feed into the roadmap which will be produced by Succeed's WP7.



A Appendix: Interviews with the experts

This appendix contains the full transcriptions of the interviews with the members of the Succeed EAB.

1 Steven Krauwer

To see the video, please visit <http://vimeo.com/97209469>

1. What challenges does multilinguality create for digital libraries? What technology is available or necessary to address such challenges?

I've spent fifteen years of my career in machine translation, which is one effort to try to solve this problem because if we had had perfect machine translation, all language barriers would've been taken away and we could easily get access to any document in a library and we would just have it translated by a wonderful translation system and we could all read it. But unfortunately, if you look around you, you see that the quality, the present quality of machine translation is awful. It's embarrassing, it's bad, it's usable in certain, very restricted contexts and, for normal use it's still very unusable. So libraries do not have access to such an instrument to bridge language barriers and basically I think people have to accept and realise that. But you also have to realise that even so, in the absence of perfect machine translation systems, there are lots of things that people can do, and libraries can do across languages. Because some people go to a library to read a book, and of course if I don't read Russian and they give me a Russian book, I can't read it. But many people also go to libraries to find information; they want to get answers to specific questions and sometimes the answer can be found in books or in documents, digital documents, in a different language. But, fortunately, even if machine translation is really still very poor, trans-lingual, cross-lingual information retrieval or information extraction is flourishing. There, I think the researchers are doing quite well. So if I have with a query for some specific piece of information I'm sure the digital libraries can easily, using existing methodology, give me the answer I want to have. So you have to explain to people what sort of things they can expect in terms of crossing language barriers. So you really have to manage people's expectations. Don't expect a perfect translation of a novel, but you can expect getting information across language barriers.



2. What initiatives are still needed to improve European coordination in the creation of digital content?

I think that what you see is that in every single country in Europe, every single European country, people are doing their own digitisation according to their own national priorities and national research programmes, and that's good. But, at the same time, it's a pity that since countries are doing different things, there's very little opportunity for, let's say, transnational research. So it would be very interesting to see a European effort and, let's say, the simultaneous creation of similar data in different countries so the researchers could really do joint research across national borders, transnational research, because that would really mean a big difference, I think, in research. So that you start a sort of defragmentation exercise in the whole area of research.

3. What role can Centres of Competence play in the digitisation field?

Well, the Centres of Competence, I think, they are extremely important because, you know, in my daily life I am the director of a research infrastructure and we are dealing with language. We want to give our research community access to digital language material wherever it's located in Europe, from all of Europe. But the problem is that if you start digitising language material, especially older material, it exists as a picture, and the community we are serving is not very much helped with having access to pictures. What they want to have is, they may want to see a picture of the document, but they also may want to have access to the linguistic content. Because if they are linguists they want to study linguistic features of a text or they might be interested in the content in a particular document. So they have to be able to extract the content from a document, and that means that between the picture and, let's say, the usable variant of the same document, there's a very wide gap that needs to be bridged. So that you start from the picture and, thanks to the expertise of the Centre of Competence, the gap between the picture and the usable format of the text or document will be bridged and that's really essential, I think. So I think it's very good that the Centres of Competence exist and I hope that they will continue to exist and that someone will be prepared to continue to fund them. Because, you know, they can't live on air. They really have to have some financial support, you know, to be active.



4. How can Europe promote private-public partnerships for technology development in this area? What type of good practices can be recommended?

Europe can do various things because it's important to keep in mind that for digitisation of old texts, there's no big market. So I think it's very difficult for a private enterprise to embark on, let's say, development of such technologies because there is, there may be little chance that they can earn back their investment. So it would be very good if at the European level there would be funding to encourage private players to invest in the development of these technologies that are necessary. So, that's one thing, and especially, and that's very important, if you think of the smaller European languages, the Commission, the European Commission, should support the development of such technologies not just for the big languages, but might be a potential market also for the very small languages that are spoken by just a handful of people. I mean, there are few very big languages in Europe, but most of them are small and they don't, they cannot really offer big markets, market opportunities, to private enterprises. So I think there should be some European funding to get this off the ground.

5. There is a growing number of documents online, often in full text format. How can we foster new usages of this content?

I think there are different ways to do it. I think one way is simply to try to think of interesting demonstrator projects that show the potential of using this material. Just investing in small projects so that other people, who might be researchers or school teachers or whatever, when they see it say: "Wow, this is something that I would like to be able to do as well. On my own language, on my own document collection." So I think these pilot demonstrators could be a very interesting instrument and, secondly, I think it would be very interesting to promote the use of this material in the whole education system. So that people from a very young age are used to working with the digital material and learning what the potential is to use digital methods in their research or in their future career.

6. Do you think they should focus on Education?

I think there is an enormous digital divide if you look at the ages, that people belong to different ages. People of my age



category, they are basically, I think, lost for digital life, you know, they're too old to change, they want to stick to their old methods. You have to capture the young people, and that means that you have to start in the education system and make using digital methods as natural as using a mobile phone.

2 Milagros del Corral

To see the video, please visit <http://vimeo.com/97209471>

1. Do you think that the European Commission's support for the digitisation of the European cultural heritage has proved sufficient? What additional measures/steps should be taken?

No, it hasn't been enough. I mean, it has been good that they've at least funded Europeana, which is a common purpose of all European libraries. But the cost of digitisation in a large scale, the one we need, and in the way we need it for users' needs, for users' requirements nowadays, goes far beyond that. And it would be needed that, if they consider, if European states really consider that digitisation is a pillar for the new society, knowledge society in knowledge economy they want to be the one for Europe, really we need... it is necessary that they put aside some more funds for this topic.

2. Should the main libraries increase the fraction of their resources invested in digitisation?

It is difficult to answer so generally. There are many libraries that are really doing their best in terms of according a good budget to digitisation issues. Some of them maybe not, it depends on the institution, also on their leaders, but in general my impression is that, taking into account the present situation, economic and financial situation around Europe, or in most countries of Europe, they are doing their best. So we should mainly reassure them in what they are doing and applaud their efforts more than thinking that they need to be pushed. They are completely aware that this is a need, this is not a luxury or a caprice somebody had, and I think they are doing quite well.

3. What initiatives or measures can help to further the coordination of European digitisation initiatives?



There is an area where coordination or a better coordination will be needed; it is probably the battle for quality digitisation. We have, when it all started, it's a lot... already some years ago, and we didn't know how quickly technologies advance, and develop and change and so on, and at that time, in many libraries, not in this one, but in many libraries, the idea was massive digitisation. Quality was not the first priority, and in the case of the National Library of Spain, it is true that we favoured quality, of course with the techniques that were there at that time, but we favoured this rather than quantity.

Every library has made its own policy out of this, according to their needs, their users... also, you know, the sensibility of government. Were governmental leaders more in favour of digitising all newspapers or this is a massive thing? Or were they particularly proud of seeing masterpieces of the library being digitised and accessible to everybody? Is this the correct policy? You can also reflect, we can also reflect, about this... For us, yes. For us, librarians, it seems to be our, you know... But at the end of the day, is this usable and highly interesting for the common people in the Internet? You know, those are issues. So, your question, how should we better coordinate? You know, for me coordination comes together with some budget backing; if we want to be better coordinated, the stimuli can be economic and financial, and I think this is also very much the issue of the European Commission I guess.

4. How can international organisations (UNESCO, UN, Europeana...) help to define the priorities in the digitisation of cultural artefacts?

Yes, priorities yes, but as long as this is recommendations, recommendations have the weight they have. I mean it depends on your... yes. So, in relation to Europeana, this is what they are doing, in fact, and this is a very good thing. It's standardising the standard, the technical standards, in all areas: digitisation, transcription, translation... all these things. We need to use the same or compatible, because otherwise the whole work doesn't function at the European level, and the very objective of Europeana is to be compatible and accessible for everybody, so this, yes. For the rest, recommendations in terms of policy, I mean, of promoting first seeing masterpieces or massive things, this will be very difficult to coordinate. And, well, I happen to know UNESCO quite well, and so I can tell you they will always remain in the area of recommendations, they will not go further. And they are also



involved in this digitisation process. You know, they held from very much earlier than Europe started, the so-called 'Memory of the World Programme'. Of course, the objectives are not necessarily the same, but it's digitisation and it is, yes, masterpieces of the whole world. Not in terms of beauty; for UNESCO this is not the basic interest, but in terms of having changed history; something that is really key material, a key piece to understand in certain history or international relations between this country and this other and all this. Or for the origins of a country, all this kind of things; so this is gained in other policies. But this is a voluntary issue for the states to propose the inclusion of a given piece of their own collections to be included in the programme. And then, there are committees and things that really apply their criteria and decide if this can or cannot be included in the list. It's a bit like a world heritage. In fact it's like a world heritage but in documents, for documents, not for monuments.

5. What role can Centres of Competence play in the digitisation field?

This is a very personal opinion, maybe my colleagues, I don't know what they may have said, but this is the interesting thing, that everybody has his own. For me where research is absolutely needed and it is very, almost impossible, for a single library or institution to obtain these data, this could be a very interesting task for Centres of Competence; it is to follow very, very closely the behaviours of users. What is what they really need and to do what? You know, because we may assume that they need this or don't need that and be completely wrong. So this is, happily enough, I guess, that the users' behaviours are more and more homogeneous through the net, the social networks. People are getting used to working in a given way and patterns of behaviour are maybe probably quite similar in different parts of Europe. But to study this, to have the result of such a study or observatorium or something like that, would be interesting because this could lead and help a lot the policy makers, at the library network or system, who are involved in the digitisation. Otherwise, some of us, or some of them, may invest money and time in things that in a few months or before they are really in exploitation, they are already over.

3 Michael Keller

To see the video, please visit <http://vimeo.com/97209470>



1. What can be done to enhance the cooperation in digitisation between Europe and America?

It seems to me that there are many possibilities. I think one of them has to do with increasing the conversation among those who are involved in digitisation already. So it would be useful, I think, to gather not the practitioners, per se, but the policy makers above them who help to define what might be done and then apply the staff and the assets to a topic. So, for instance, we have done work with European collaborators in various areas. At Stanford we worked with our very fine colleagues at the World Trade Organisation to digitise the archive, most of it, of the General Agreement on Tariff and Trade. It's an incredibly important archive to understand the decisions and the processes that have shaped the global trading market as we know it now. That's an example of collaboration that occurred because scholars drove the need for such a project. We've also collaborated with our friends and colleagues at the Corpus Christi College, Cambridge University, in the Matthew Parker Library Medieval Collection – Medieval Manuscripts. I think that that also provides an opportunity for those who are interested in ancient historic documents, images, manuscript images, could be art images, to get together and talk about what might be important, what might be useful, what might drive new scholarship and new learning. The important thing though, as I said, is to try to get the practitioners who are already at work together so it's not an educational enterprise, it becomes a policy and a kind of decision-making business having to do with the assignment of assets to tasks. And those tasks should be cultural ones it seems to me.

2. Is a world-wide digital library possible in the next future or utopia?

It exists now. As all libraries are fragmentary in their collection of documents and genres, and publications, and manuscripts, there's no complete library anywhere. So the world digital library, such as it is, is a statement of how far we've come with materials that we've digitised, we've converted from analogue form to digital form and materials born digital, it's not complete; it may never be complete. But we have to understand that this is part of the grand tradition, we none of us collect everything. We have a big library collections programme at Stanford and I think we collect, I would say, maybe two per cent of what's published every year. Still a pretty good library, people get work done there. And we know how to reach out



and find materials that our scholars require, although we try to anticipate their need, we never can do so perfectly. So the world digital library exists, it's not very coordinated, but we work on the basis of the themes and processes and basic rules of the World Wide Web and the Internet. It will get larger, it will get more linked, it will get more relevant, the discovery will become better as new techniques become possible, but I think to say that it doesn't exist is a mistake, it exists.

3. How can the community help to advance the state-of-the-art in digitisation? (e.g., open-source software developers, open-content creators, crowd-sourcing projects,...)

I think the answer there lies in understanding behaviours. So, if we now have, if you can accept that we now have sufficient material in digital form to help teachers teach the young, to help students in higher education learn and to create new knowledge, to help scholars create new knowledge, if we have enough material now to show some examples, then we should survey those examples and extract from them principles, examples and so forth, which would help us understand what more we might do and how we might do things differently. So one would look therefore to what we might regard as common practice and at the same time we would look at what we might call steeples of excellence, or steeples of innovation, to see what's going on and to see which among them we might believe to offer some great promise for the next steps. The community therefore is a kind of hypothetical creation, it has to be looked at across borders because, as we said, the world's digital library exists and the world's digital library promotes collaboration among people of varying backgrounds, varying locations, varying circumstances, various disciplines. We need to look at those steeples of innovation, steeples of excellence, as we've also got to look at common practice and understand how we can cause the tide to lift all the boats, on the one hand, and on the other hand, to invest in a few more prototypes that show what's possible with the new techniques, and we're always developing new techniques, and we're always observing new techniques. So the work in the digital humanities, the computational social sciences, the great literary labs, based on digitised materials or born digital materials should inspire us to think how we might go further and farther to present new interpretations, new research possibilities and thus new knowledge.



4. What role can Centres of Competence play in the digitisation field?

Well, always these Centres of Competence, I would... let's call them centres of innovation, let's call them centres that provoke thought on standards and on possibilities, all these contribute to the sense of motion, the sense of development, the sense of taking advantage of new possibilities. I think it's also important for the Centres of Competence to see themselves in a larger environment, a larger workflow. So that the Centres of Competence and digitisation should see themselves in the context of the world digital library. What are those requirements? Well, certainly ease of access for the general user who may not have any technical apparatus except a computer screen. Secondly we need to pay attention to how we provide intellectual access. Right now we do this on the basis of metadata that is extracted and sometimes interpreted, we will soon be able to do linked data for these objects and even semantic web techniques in some years. Third is the importance of having real honest-to-goodness digital libraries that are meant to store the original scans and then the revised and curated scans. The linkage among objects is another category. How do we take advantage of all these possibilities? In principle, to inform how the Centres of Competence provide suggestions, guidelines for practice. But those practices always have to be seen in this larger environment of behaviours, which allow people on their own, people in courses, people in their individual studies to make good use of the documents of civilisation.

4 Jill Cousins

To see the video, please visit <http://vimeo.com/97209466>

1. Do you think the European Commission's support for the digitisation of the European cultural heritage has proved to be sufficient? What additional measures/steps should be taken?

One always wants more, especially when you start on something as big as digitising cultural resources of Europe. I think they've done a lot and they've tried to encourage member states to digitise. I think unfortunately we then hit an economic crisis in Europe and different places reacted in different ways. They have made structural funds available for, which can be used for, digitisation, but it's not always what different



states use it for. And personally I think that's a shame because I think you can use the structural funds to improve the roads, but you can also use the structural funds to improve the roads of the future, which are probably digital ones, and the access to this material digitally will give us new economies, new jobs and real new feature.

I think the Commission has tried, and I think they continue to try, and I don't think with the money that is available they are going to be able to fund more digitisation programmes but I'm sure they can help us make what we have got better available, so that we improve the quality of what we have done. And the programmes that they've now got in place are very much about how to make that better used than it currently is. So, yes, they can do more, but with the money available I think we're probably better off at the moment concentrating on making what we have got truly reusable.

Ten per cent. According to the enumerate survey we have done ten per cent of the total amount that should be digitised. (I have got the statistics somewhere but I don't have them off the top of my head. . . I can look them up for you.) So ten per cent of what has been digitised of which in Europeana, which I represent, we have got twelve per cent of that ten per cent, and of that twelve per cent only 8 per cent is really reusable. And by reusable I mean that it will work on a tablet, you know, so you can do this or it works on a mobile phone. . . yes, all that kind of thing. There is a direct link to the objects that is under open licence so you know what you can do with it, so then it is actually reusable in new apps, education apps, in tourism apps, in new ways of doing things.

2. Should the main libraries increase the fraction of their resources invested in digitisation?

I mean I think it is, where they can, I think it's good to do so, because it's making sure that the stuff is relevant for the future, so that fundamentally if it's not going to be available online then I think a lot of younger users will struggle to find the tool to make use of it. That doesn't negate the role of libraries in being spaces for work and co-creation and seeing the original object. I mean I can (offer) a very personal story. . . I did my PhD in maps of the 16th-17th century, Arabic maps. I needed to see the original because I actually needed to see where they put the pinpricks in the vellum and really how that worked out on the whole map. So I couldn't make do



with the facsimile or whatever. But it would have been amazing if that stuff had being digitised because then the place names would have been blown up; I would have been able to read them properly, I could have done much, much more with it; and in fact when I first started doing this job I found (there were eight of these Arabic maps), and I found another one because it had actually been digitised and was sitting in Bulgaria, and I would never have known about it otherwise. So I think there are all sorts of reasons why you do it, one because you can make use of the material in new ways; the other is that you can find stuff that you would not otherwise be able to find.

3. What initiatives or measures can help to further the coordination of European digitisation initiatives?

I think actually also more and better use of what we have already done. So if you look at the work in Succeed, they have tried very much to make sure that the tools are available or the different ways of digitising to make it more cost-effective or better digitisation. And we're not making that yet sufficiently widely known, so as I think that's the idea behind the Centre of Competence, that you have somewhere to go to say "Ok, I've got this problem: Who's done it this way in the past? Can I make use of that? So that, I can do it in a better way or in the same way"... but improve on what we've got rather than continually looking for new things. I think we can make use, much better use, of what we have already done in order to make available in better ways what is already digitised.

4. How can international organisations (UNESCO, UN, Europeana...) help to define the priorities in the digitisation of cultural artefacts?

I think potentially they are extremely useful in terms of transfer of knowledge, don't reinvent the wheel and sharing the tools on the applications that have already been developed. And creating a network of experts who can then improve on each other's work; say you build on what's there, and in fact the whole principle of Europeana is now moving towards... we have a platform where we don't expect to do, to show everything that's in there, but we do think other people can make use of the tools and the data to build something that they need for their purposes, whether that is for education or for tourism or a thematic site on shoes or on food or on maps because they have the curatorial capabilities which we don't



have, so we need to make sure that the core is there and available for them to build on; then another way is not Centres of Competence but it's a platform, and I think Centres of Competence and platforms have quite symbiotic relations.

5 Frank Frischmuth

Owing to restrictions in Mr. Frischmuth's agenda, this interview took place via electronic mail.

1. Do you think that the European Commission's support for the digitisation of the European cultural heritage has proved sufficient? What additional measures/steps should be taken?

Digitisation will be a major challenge for all cultural heritage institutions and national governments for many years from now. Beside the activities of institutions in digitising content and the support of their governments in such projects I am positive that the European Commission must also do its share in supporting such projects or set up its own projects as was done in the past. As there should be national strategies in digitising cultural heritage, a strategy should be built by the European Commission with those tasks and objectives to be achieved. Bearing in mind the background of the new EU Open Data rules "that public sector information represents an enormous economic potential, reaching tens of billions of euros per year across the EU", but at the same "studies conducted on behalf of the European Commission showed that industry and citizens still had difficulties in finding and re-using such data" the EU should be obliged to make more efforts in digitisation projects.

2. Should the main libraries increase the fraction of their resources invested in digitisation?

I cannot speak for the libraries. In general all cultural heritage institutions should increase their efforts and therefore budgets have to be increased.

3. What initiatives or measures can help to further the coordination of European digitisation initiatives?

An initiative to set up a coordination board (resp. the centre of competence or a network of competence) should be started. May be a website to register digitisation initiatives could help for the moment.



4. What role can Centres of Competence play in the digitisation field?

They can coordinate and guide local/regional/national initiatives and function as an information provider on all subjects concerning digitisation including legal problems

