EUROPEAN COMMISSION
Information Society and Media Directorate-General

Digital Content & Cognitive Systems
Cultural Heritage &Technology Enhanced Learning

REPORT

WORKSHOP ON CENTRES OF COMPETENCE
FOR DIGITISATION AND DIGITAL PRESERVATION

LUXEMBOURG 14 NOVEMBER 2006

1. BACKGROUND

The context for the workshop was twofold:

- the Commissions' i2010 digital libraries strategy, built around the 3 pillars of digitisation, accessibility and long-term preservation and including the vision of the European Digital Library as a single entry point to a targeted 6 million digitised objects by 2010;

- and the research objectives for digital libraries and digital preservation, set out in the ICT Work Programme 2007-2008, under Challenge 4, Objective 4.1 – "digital libraries and technology-enhanced learning".

The research specifically targets scalability and cost-effective digitisation, as well as long-term preservation and accessibility, through the emergence of European centres of competence in these areas, able to support different organisational stakeholders and to provide over time access to a new generation of digitisation tools, services and skills.

As background, it was established that proposals around centres of competence need to bring together the high level expertise and capacity available from a limited number of organisations across Europe and make it available to support digitisation and digital preservation activities undertaken by other organisations (libraries, museums, archives) in the Member States. This approach should provide cultural institutions with an easier access to the best experts and to the technological infrastructure required to make their collections available online, increasing the availability of European digital content and making it accessible to citizens across and beyond Europe. The key issues are mobilisation of critical mass of resources, complementarity and reduction of fragmentation, and extending excellence.

Funding schemes will be implemented as appropriate using the shared-cost funding schemes available under FP7, namely Collaborative Projects and Networks of Excellence.

2. ORGANISATION AND OBJECTIVES OF THE MEETING

The workshop brought together around 40 stakeholders (cultural institutions, centres for digitisation, research, private sector – involving different types of content; see annex 1),

invited on the basis of their organisational involvement in digitisation projects and European activities so to bring proven experience to the discussions.

The workshop's objectives were threefold: to discuss and identify appropriate organisational models; to define the scope, boundaries and priorities for work, including the starting conditions of existing excellence; and to demonstrate the impacts at European level, in terms of added-value, leadership and sustainability.

The discussions were held on the basis of prior written inputs (see annex 2) from the invitees, providing information on existing experiences and organisational approaches; research issues to be tackled; and European and national impacts.

3. REPORT OF DISCUSSIONS

3.1. Existing experiences and organisational approaches

The workshop discussions identified that in most Member States digitisation competence seems to have strong institutional origins (often project based) – but there is also a growing tendency to national coordination and cooperation mechanisms. Private sector competences are mostly fragmented, offering niche market solutions and still showing a lack of industrialisation of processes. Audiovisual and text digitisation have followed different paths when it comes to competence organisation, whereas mass- and specialised digitisation tend to have developed in parallel. Outreach and training activities tend to be incidental, suffering from fragmentation, ad hoc funding and a lack of continuity.

Obvious trends are the development of strategies to better utilise capacity and cross-institutional cooperation, including support for standardisation. An increased sharing of experiences and involvement of private sector can also be identified. Digitisation is a maturing domain, moving from test applications to mass processes and from pilot projects to large scale implementation. PPPs in technology and funding are gradually developing.

Still lagging behind is the establishment of a real market offer for digitisation (and digital preservation) services, causing also slower convergence of standards and hampering the delivery of effective interoperability.

3.2. Research issues

The workshop discussed and identified both organisational and technical issues that were considered particularly critical for future research to be taken up by competence centres and that should ideally be dealt with in combinations:

- Organisational: workflow modelling, evaluation, benchmarking and quality management (of processes and services) for digitisation and digital preservation; cost and benefit modelling; efforts to create better interoperability; enhancements in access and new tools for usability and multilingualism; better availability of advice, best practice, exchange of know-how and training, pre-testing new tools, market monitoring. Moreover, competence building across institutions could contribute to a better outreach and sharing of training programmes, leveraging excellence ('best of breed') and sharing infrastructures in support of long-term digital preservation.

- Technical: new and improved mass digitisation techniques and technologies, covering various types of objects; further development of OCR technology (fonts, old paper, multiple languages, multiple documents layout); metadata (extraction, interoperability, preservation etc.); analysis of approaches to distribution of content; digital preservation (integrity and authenticity issues); multilingual search.

3.3. Impacts

In terms of impact, the workshop discussed the supportive approaches needed. In some areas of expertise a more centralised model was considered to be more suitable than in others. Discussion about the boundaries between generalist and specialised areas dealing with digitisation concluded that such boundaries were likely to be largely informed by the analysis of existing workflows. As an example, issues such as optical character recognition are likely to benefit from a concentration of capacity that could be applied to all sorts of published material. This was summarised in the expression 'horses for courses' – there is a need to find the most suitable arrangement for each particular situation, i.e. the type of activity being supported by the centres of competence.

Actual digitisation activities could be carried out by competence centres if they are geared towards the development of a capacity to deliver cost effective and more mass digitisation services. Attention would be needed for the consolidation and dissemination of work already started by various entities in this area. For digital preservation there was widespread consensus about the need to give priority to further structuring existing research capacity and promote more effective links with industry and the concerned user communities (not just limited to the cultural heritage sector).

4. CONCLUSIONS

The workshop concluded that competence centres should generate equal access to excellence from anywhere – but not duplicating excellence unnecessarily (support for 'best of breed'). They should improve digitisation capacity (quality, coherence, efficiency, widespread accessibility), provide services of guaranteed quality (directly or brokerage access) and should have a clear impact at European level (leadership). They should link EU added value and national initiatives (leveraging effect) and must be able to facilitate training and spreading knowledge, competences and skills to all types of institutions. As such, competence centres will have a structuring effect starting from existing excellence and move towards longer term issues.

Therefore, potential proposers for competence centres will need to bring in expertise, experience and competence, such that at European level these will deliver the right complementarity and increase the expertise and knowledge available. Funding schemes should allow core significant players to come together to provide the complementary and baseline competence and also allow the options of including new partners through the life of a project. Consortia should not try to include everyone from all across the EU but to focus on excellence at the start. Links should be created between different funded projects to further consolidate the picture at European level.

ANNEX 1 — WORKSHOP PARTICIPANTS

PARTICIPANT	ORGANISATION	COUNTRY
Adam Horvath	National Szechenyi Library	HU
Aly Conteh	The British Library	UK
Birte Christensen-Dalsgaard	State and University Library	DK
Börje Justrell	National Archives of Sweden	SE
Catherine Lupovici	BNF	FR
Daniel Terrugi	PRESTOSPACE / INA	PROJECT
David Arnold	EPOCH	PROJECT
Markus Brantl	München Digitalisierungszentrum	DE
Paulo Costa	Portuguese Institute of Museums	PT
Edwin van Huis	Netherlands Institute for sound and vision	NL
Emmanuel Amasio	Infotechnique	FR
Erlend Kolding Nielsen	Royal Library Denmark	DK
Frank Poireau	Infotechnique	FR
Guibert Vanhoof	Infotechnique	FR
Hans Jansen	National Library NL	NL
Hans Petschar	Austrian National Library	AT
Istvan Monok	National Szechenyi Library	HU
Katarzyna Slazka	National library of Poland	PL
Lucien Scotti	BNF	FR
Majlis Bremer-Laamanen	Centre for Microfilming and Conservation	FI
Marc Pinter	Medea Services	HU
Marek Jindrich	MEMORIA digitisation centre	CZ
Marianne Backes	CVCE	LU
Mathias Hemmje	NESTOR	DE
Minna Kaukonen	Centre for Microfilming and Conservation	FI
Monique Kieffer	National Library of Luxembourg	LU
Patrick Peiffer	National Library of Luxembourg	LU
Rafael Carrasco	Biblioteca Miguel Cervantes	ES
Ralf Stockmann	Göttinger DigitalisierungsZentrum GDZ	DE
Rosella Caffo	Biblioteca di Storia Moderna	IT
Seamus Ross	DPE/HATII	PROJECT
Stefan Gradmann	Göttinger DigitalisierungsZentrum / NESTOR	DE
Stefanos Kollias	Image, Video and Multimedia Systems Lab	HE
Stephan Cernohorsky	MEMORIA digitisation centre	CZ
Teresa Malo de Molina	National Library of Spain	ES
Ute Schwens	Deutsche Bibilothek	DE

How are competences organised in your countries?

- In most cases, institutional origins (often project based) – but growing tendency to national coordination & cooperation mechanisms
- Private sector – fragmented, niche market solutions, lack of industrialisation
- AV and text digitisation – followed different paths
- Mass- and specialised digitisation – in some cases co-resident, but in most tend to have developed in parallel
- Outreach and training activities tend to be incidental: suffer from fragmentation, ad hoc funding, lack of continuity

How are they developing / would you want to see them further develop?

- Trends:
  - strategies to better utilise capacity and cross-institutional cooperation, including support for standardisation
  - better sharing of experiences and involvement of private sector
  - from test to mass – maturing domain, move from pilot projects to large scale
  - development of PPPs in technology and funding
- Wanted:
  - Structuring of the industry (esp. reference to audio-visual)
  - Interoperability and standardisation

What are the needs for digitisation and digital preservation competence in your country that could be serviced at EU level?

- Evaluation, benchmarking and quality management (of processes and services)
- Workflow models (digitisation and digital preservation)
- Approaches to cost-reduction (semi-automation)
- Coordination - standards & their deployment, interoperability, registries
- Consolidation of expertise at national level
- Enhancements in access, new tools
- Usability, multilingualism
- Advice, best practice, exchange of know-how and training (the skills problem)
- Also: political and strategic issues (i.e. IPR, orphan works, legal deposit, web archiving)

How and where should we organise competence at EU level to meet these needs?

- Shared activities:
  - Creating registries to support producers by identifying what exists, who owns
  - Registries of service providers, supported by common quality criteria
  - Anticipating the potential use of content
  - Technology watch – foresight activities, pre-testing new tools, market monitoring
  - Promotion of standards & interoperability
- Competence building across institutions
  - Outreach – sharing of training programmes, research programmes
- Leveraging excellence
  - Best of breed excellence – trans-nationally?
  - Specialised expertise – audio-visual, manuscripts, webarchiving, preservation
- Archiving infrastructure in support of long-term preservation