

Preservation of digital cultural heritage in Europe and in the Czech Republic

Bohdana Stoklasova
National Library of the Czech Republic
bohdana.stoklasova@nkp.cz

Jan Hutar
National Library of the Czech Republic
bohdana.stoklasova@nkp.cz

Pavel Krbec
Charles University, Prague
krbec@cuni.cz



Content

- **Preservation of digital cultural heritage in Europe – DPE survey**
- **Preservation of digital cultural heritage in the Czech Republic**
 - Introduction
 - National policy
 - National Library of the Czech Republic
 - Charles University, Prague
 - Standards
 - Sustainability
- **Conclusion – cooperation on national and international levels**



Objective 1. To create a coherent platform for proactive cooperation, collaboration, exchange and dissemination of research results and experience in the preservation of digital objects.

- **To identify and raise awareness of sources on the issues surrounding the curation and preservation of digital objects across the broad spectrum of national and regional cultural and scientific heritage activity in Europe**
- **To contribute to the elimination of the duplication of effort of research activities by researchers at different institutions and to enable identification, collection and sharing of knowledge and expertise.**
- **To create a conduit between the research community and practitioner community that will foster the collaborative approaches to preservation needs.**
- **To stimulate and co-ordinate further research on digital preservation in key areas and encourage the development of standards where gaps and opportunities have been identified. This will include promoting and developing research agendas.**

Objective 2: To increase prevalence of preservation services and their viability and accountability

- To support the development of a European-wide approach to the audit and certification of digital repositories as an essential stage in creating content management and delivery services and to repository federation.
- To stimulate ICT companies and software developers to incorporate some of the curation and preservation thinking into newer generations of software.
- To relate the digital preservation research agenda more directly to the development of exploitable product opportunities and to develop links with the industrial sectors.

Objective 3: To improve awareness, skills and available resources.

- To examine core issues that will deliver essential guidelines, methods and tools to enable preservation action with European public and private sectors.
- To examine core issues that will deliver essential guidelines, methods and tools to enable preservation action with European public and private sectors.
- To implement a suite of training seminars based on best practice and to identify where and what further practitioner training and staff development initiatives might be undertaken.

Market and technology trends analysis

Objectives of the Market and technology trends analysis

- ***Market analysis*** based on experience and knowledge of all the contributors and the consultation of main stakeholders on their needs and plans so that the outputs of the DPE project meet their present and future demands. examine core issues that will deliver essential guidelines, methods and tools to enable preservation action with European public and private sectors.
- ***Technology trends analysis*** providing main DPE target groups with information on technological solutions available for digital preservation.

Market and technology trends analysis

Survey

- **7 questions**
- **national libraries – relatively complex survey (34 from 54)**
- **archives and research institutions – snapshot, survey to be repeated**
- **results of the survey ->**

Market and technology trends analysis

Question 1: Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

- **82% of national libraries, 66% of national archives and 70% of research institutions indicated that long-term preservation is currently among their key strategic priorities.**
- **12% of nation libraries, 27% of national archives and 11% of research institutions indicated that long-term digital preservation will be a strategic priority in the next two to five years.**

Market and technology trends analysis

Question 2: Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?

- **30% of national libraries reported having a trusted digital repository, 61% indicated that they were working towards this goal. For national archives 32% reported having a trusted digital repository and only 27% indicated that they were working towards this.**
- **91% of national libraries and 59% of national archives recognise the importance of gaining trustworthy status and see this as both a priority and a target of nation libraries.**

Market and technology trends analysis

Question 3: Digital preservation is too big an issue for individual institutions to address independently. Who will your institution cooperate in this area with?

- Respondents were given the choice of memory institutions, research institutions, digital document producers and software developers.
- For all three categories of respondents memory institutions were the first choice for cooperation in these areas.
- While the national libraries and archives rated the other three choices equally, research institutions indicated a preference for cooperation with software producers and vendors as a second choice.

Market and technology trends analysis

Question 4: The building and operation of a trusted digital repository is a big and expensive business. Will you create and operate the repository only for your library or share it with others?

- While 85% of national libraries planned to cooperate with other institutions, only 53% of archives and 52% of research institutions planned to share in the creation and operation of their repository with other institutions.
- This difference could be explained by the existing experience national libraries have in the benefits of sharing their repositories of traditional documents and through coordinating activities in other areas such as web archiving.

Question 5: What system will your digital repository use?

- **The responses from national archives and research institutions were close with the majority, 38% and 53% respectively, favouring an open source system, and the rest split equally between commercial and in-house systems.**
- **53% of national libraries planned combined solutions: the use of in-house, commercial and open source systems.**

Market and technology trends analysis

Question 6: Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your institution?

- **Memory and research institutions alike indicated that the DPE website was the most relevant method of dissemination for them.**
- **The production of guidelines was also identified as relevant for both memory and research institutions.**
- **Conferences, seminars and workshops were considered more relevant by national libraries than by archives and research institutions.**

Market and technology trends analysis

Question 7: In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?

- **Memory institutions were identified by national libraries, national archives and research institutions as having the best backgrounds to become trusted national competence centres.**

Market and technology trends analysis

Technological solutions

- **Commercial**
- **Open source**
- **Combinations**





DP in the Czech Republic

Introduction

- Czech Republic - **long tradition in digitization and web harvesting.**
- National Library of the Czech Republic was awarded the first UNESCO/Jikji Memory of the World Prize for its contribution to the preservation and accessibility of its documentary heritage in 2005.
- When we look at digital preservation in all its complexity, we have to admit that **digital preservation has been underestimated and that it is only in its infancy.**
- Thanks to large national grant projects, our digitization projects including endangered books and periodicals and historical manuscripts started in the early '90s. We started with harvesting and archiving of Czech web resources in 2000. We have several million digitized pages and about 50 million archived web documents (about 30 TB together) covering the core of the national cultural heritage. These documents are accessible via the three national subsystems, **Manuscriptorium, Kramerius, and WebArchiv**, covered by the **Czech National Digital Library**. There are also many projects running in other libraries and universities that are expected to enrich the Czech National Digital Library in the future.
- The concept will be presented from the point of view of the **National Library** and also by one of the participating institutions – the **Charles University**.

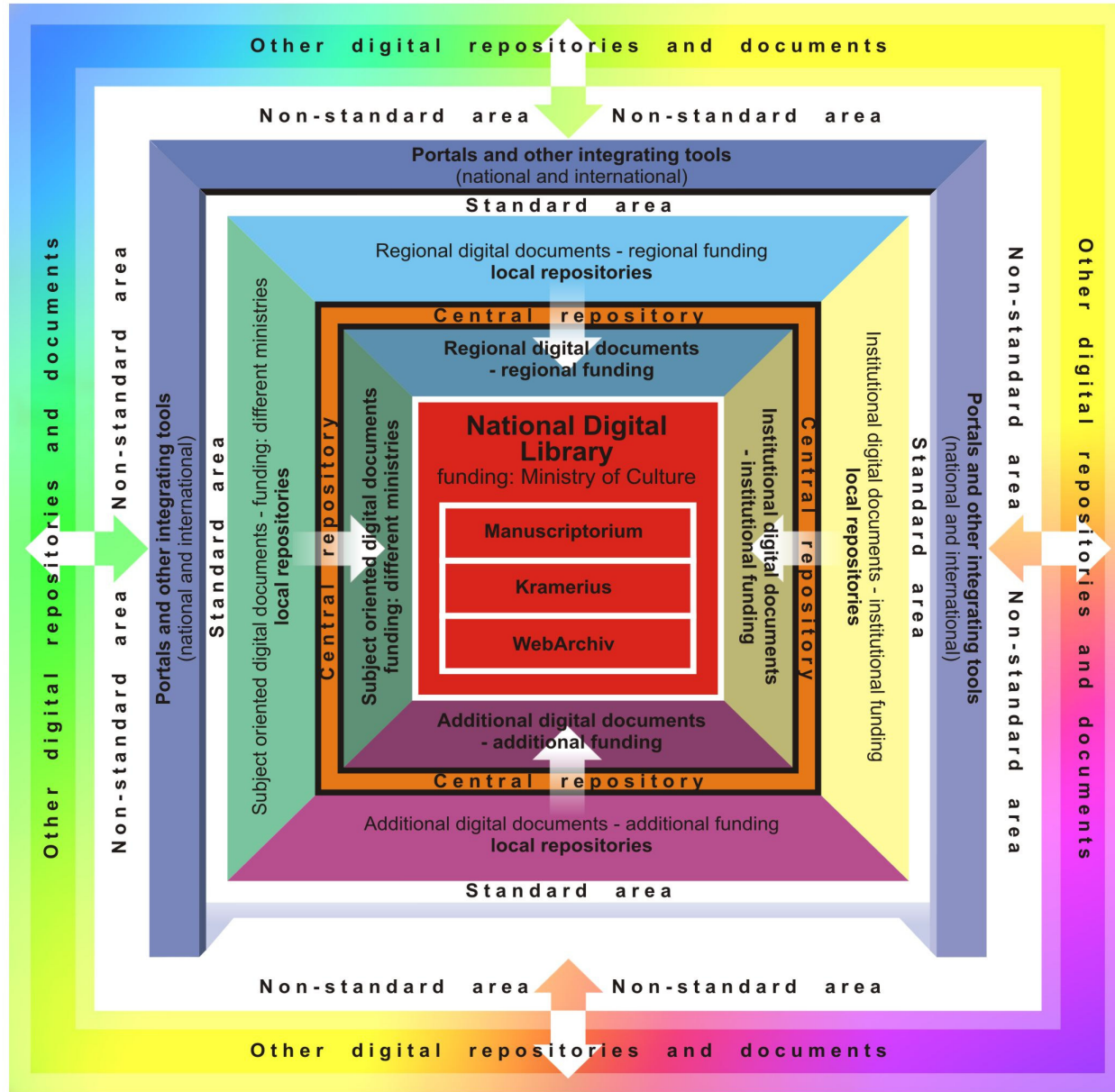


National policy

Broader context:

- **Concept of long-term preservation of and access to national document cultural heritage** (both analogue and digital) – prepared by the Ministry of Culture, to be approved by the **Czech Government** (delayed as result of political situation)
- **Czech Digital Library** - conceptualizes a new national integration of the different digital libraries in the CR with digital repositories of other cultural heritage and research institutions

Czech Digital Library

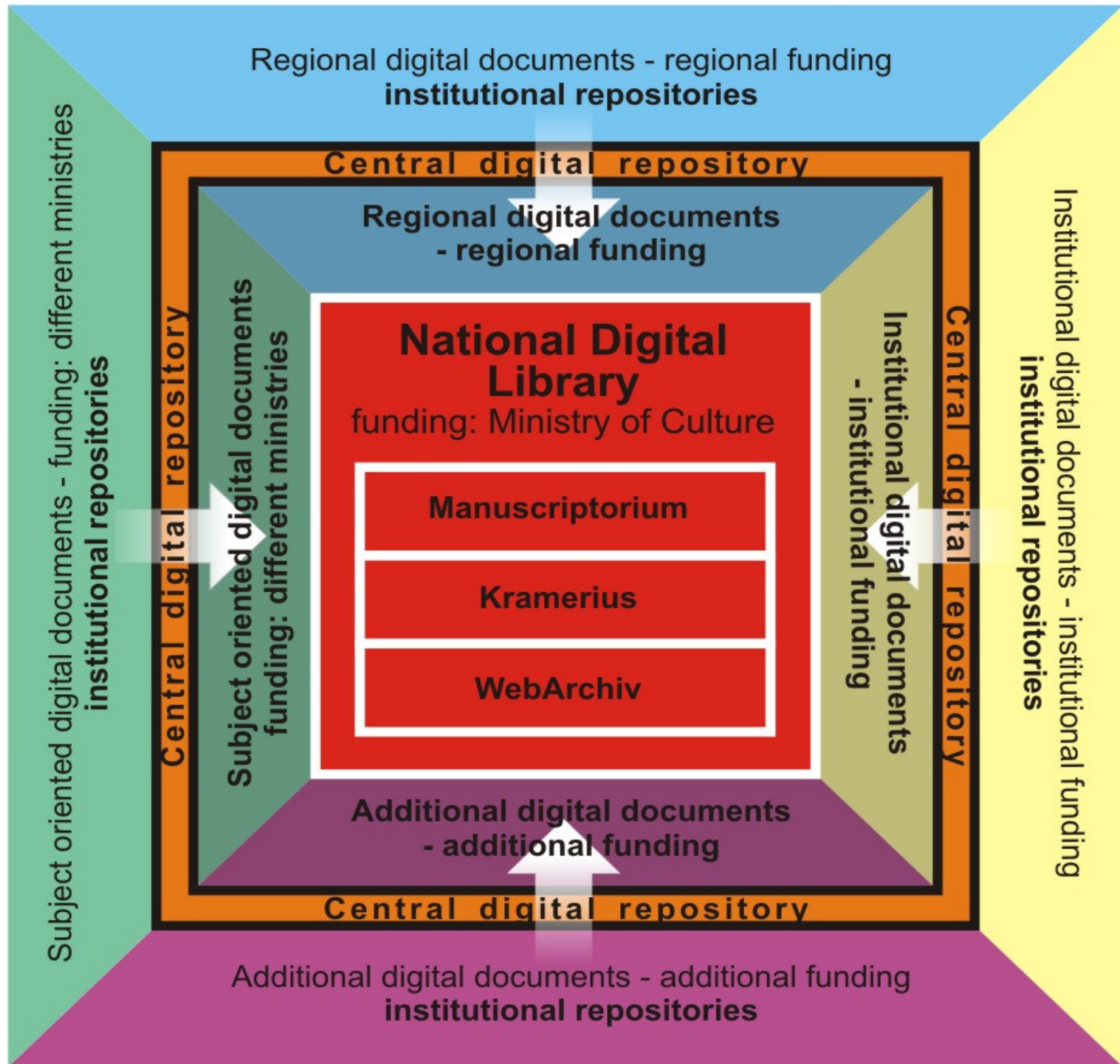




National policy

Digital preservation:

- **Central trusted repository** – national digital cultural heritage – funding Ministry of Culture (Manuscriptorium for digitised historical and rare documents, Kramerius for digitised books and periodicals, WebArchiv for archiving of web – all the projects have national framework), opportunity for others – funding from other resources
- **Institutional repositories** – subject-oriented, regional... inst. funding





Central repository

- **Data storage** As the central disk storage two IBM Systems Storage DS 4800 are installed, one in Klementinum and the second one in Hostivar data centre. These data centres are connected via dark fibre with CWDM modules and fibre channel switches SAN16B-2 built Storage Area Network (SAN) are used in each of above mentioned localities. For the backup and archiving services the Tivoli Storage Manager (TSM) together with an IBM tape library is used. The currently implemented solution for the Central Data Storage provides the possibility to store digitalized data on the safe platform with flexible capacity. The CDS offers also services concerning Disaster Recovery. Data are replicated between two data centres more than 20 km distant. Both this replication and distance between localities protect data against the physical destruction of the building, long power outage, etc. Together with CDS both backup and archiving systems are available to protect operating data against human or software error. Next steps beside appropriate CDS capacity for digitalization in year 2008 are to be finalization of backup and archiving strategy for selected applications to enhance data security.
- **DOMS** – to be selected in 2008-9 (commercial system – DIAS, DPS, other ???)
- Internal audit – **DRAMBORA** - recommendation



Introduction of the National Library of the CR and its key role in preservation in the national cultural heritage

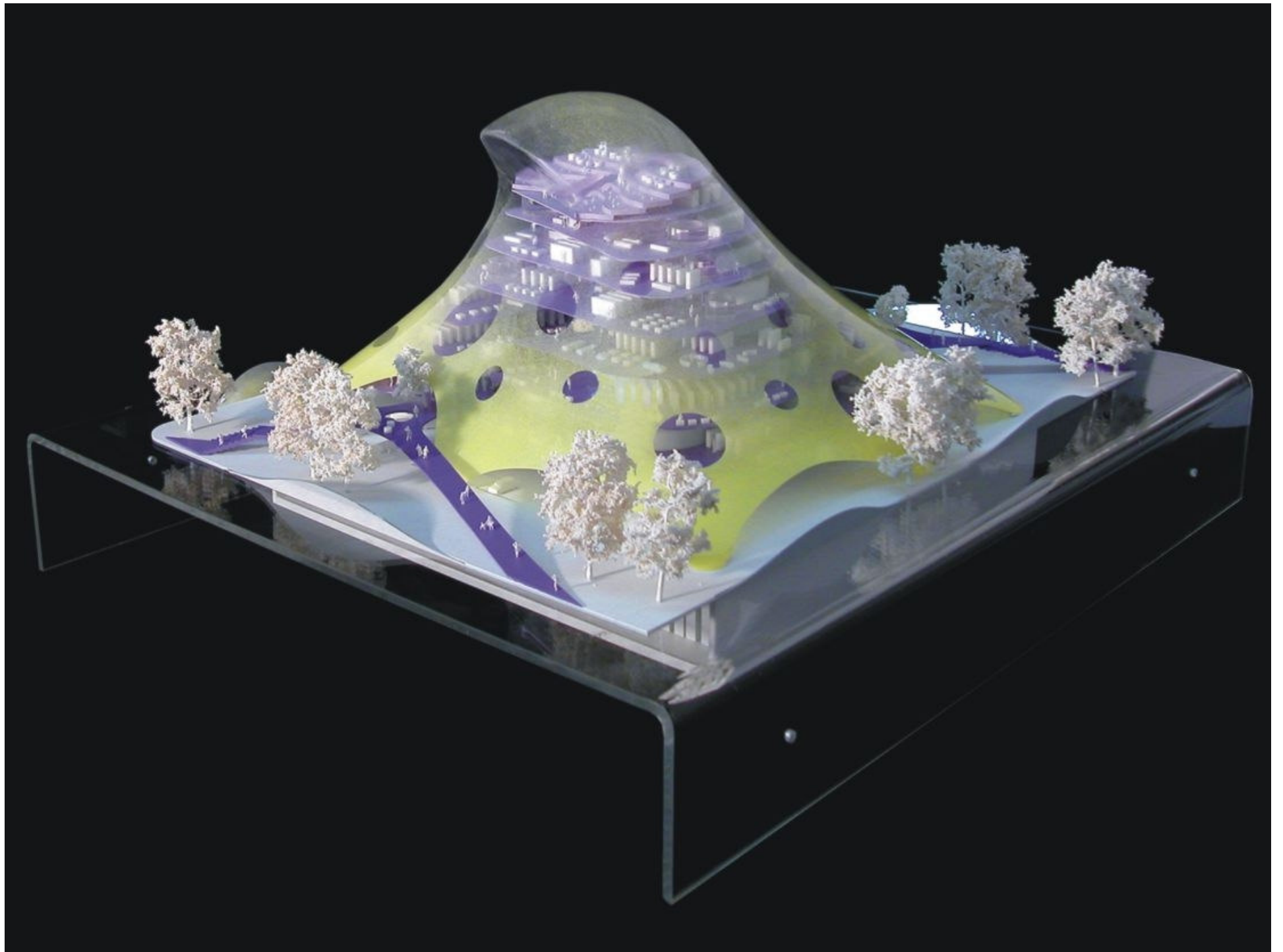
- **National library, more than 6 million volumes** (many of them candidates for digitization)
- **Digitisation since 1992** (3 million digitized pages, plan for 300 million digitized pages)
- **Web harvesting since 2000** (150 million files)
- **Negotiations with publishers** (legal deposit also in digital form – to avoid digitization of printed legal deposit in the future)
- **National coordination of digitization and digital preservation**
- **International cooperation**



Seat of the NL

- **Klementinum (+ Hostivar) – no space after 2010, long-term preservation of printed documents endangered)**
- **New building – Letna - ?2011?, international UIA competition, finished**



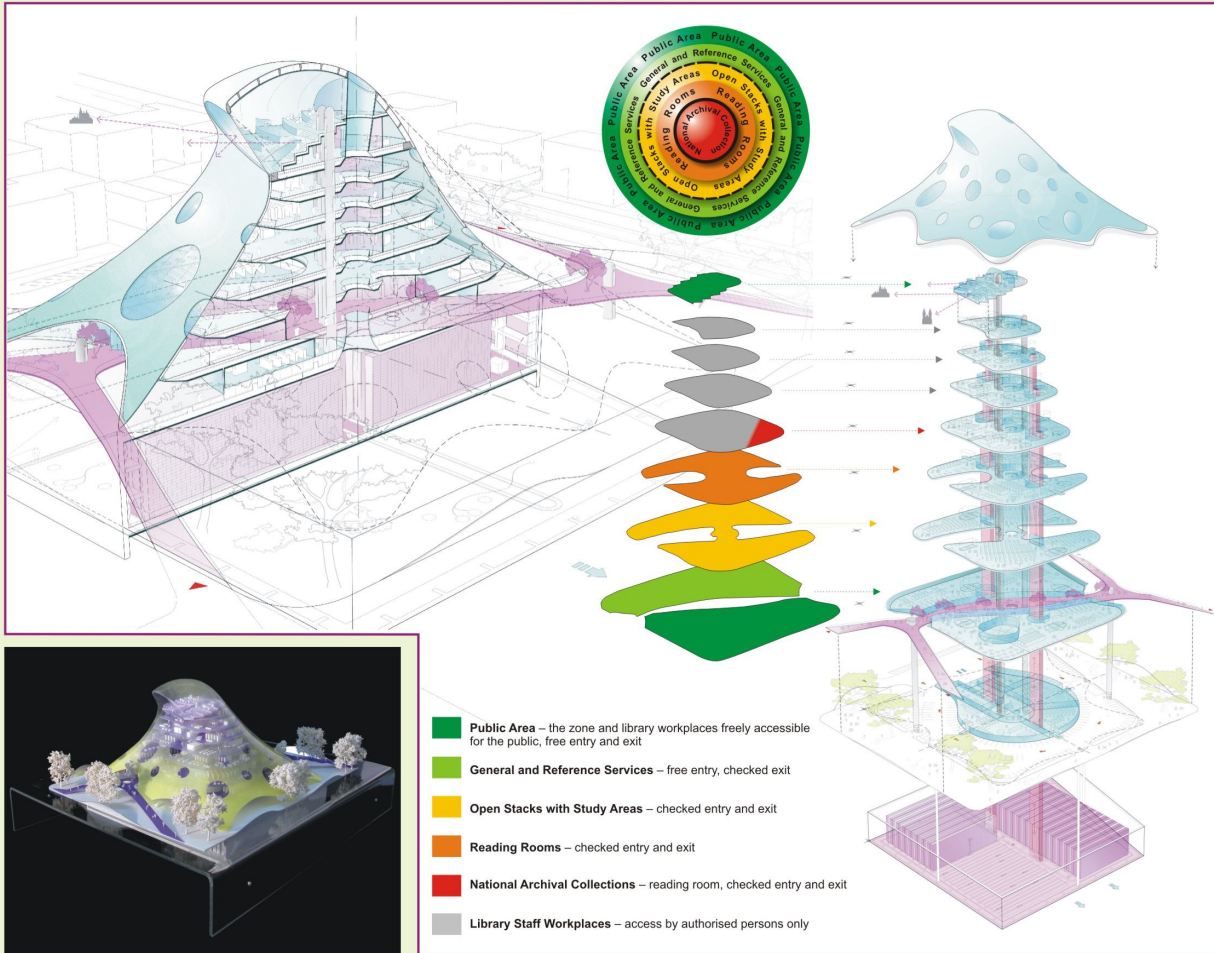


THE EYE



ABOVE PRAGUE

A LIBRARY FOR THE THIRD MILLENIUM



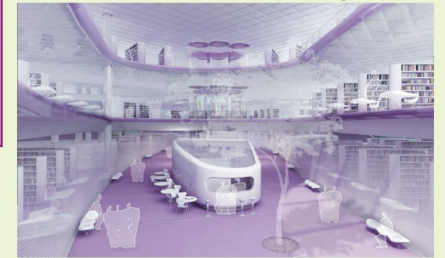
- **Public Area** – the zone and library workplaces freely accessible for the public, free entry and exit
- **General and Reference Services** – free entry, checked exit
- **Open Stacks with Study Areas** – checked entry and exit
- **Reading Rooms** – checked entry and exit
- **National Archival Collections** – reading room, checked entry and exit
- **Library Staff Workplaces** – access by authorised persons only



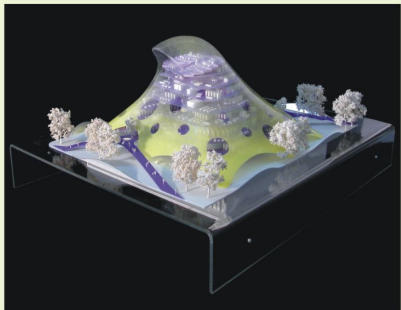
Café Eye



Reading rooms



Street



Automatic storage & retrieval system



Plasticity



Beauty



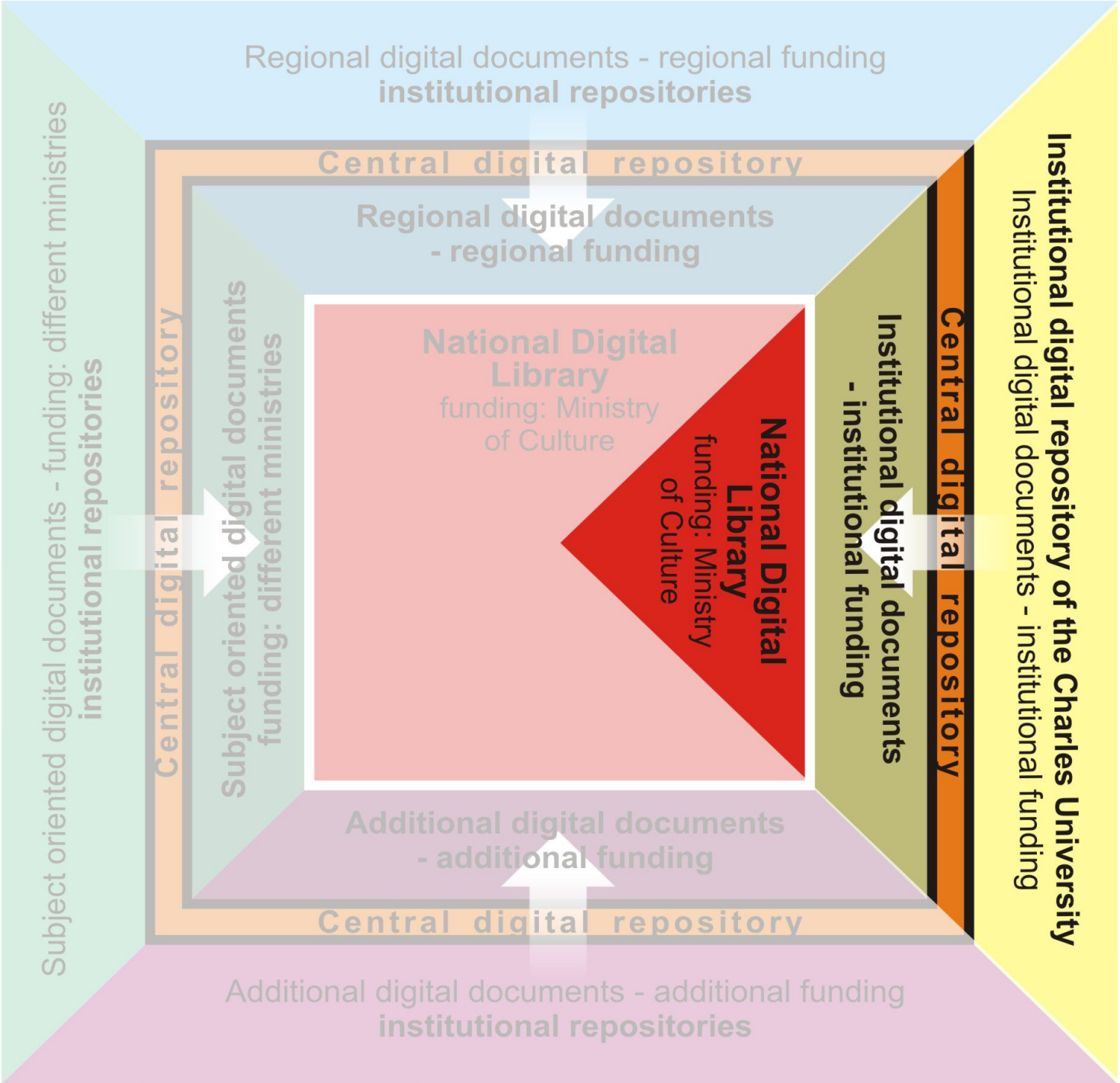
Colours





Institutional repositories

- **Data storage** – different solutions
- **Institutional repositories management systems** – research institutions – mainly open-source (Fedora, DSpace etc., some – commercial - DigiTool)





Introduction of the Charles University in Prague of the CR and its role in digital preservation

- **founded in 1348**
- **17 faculties**
- **40000 students, 7000 employees**



Charles University and its role in digital preservation

Institutional digital repository of the Charles University
Institutional digital documents - institutional funding

NDL – Charles University documents

Institutional digital documents
- institutional funding

National Digital Library
funding: Ministry of Culture

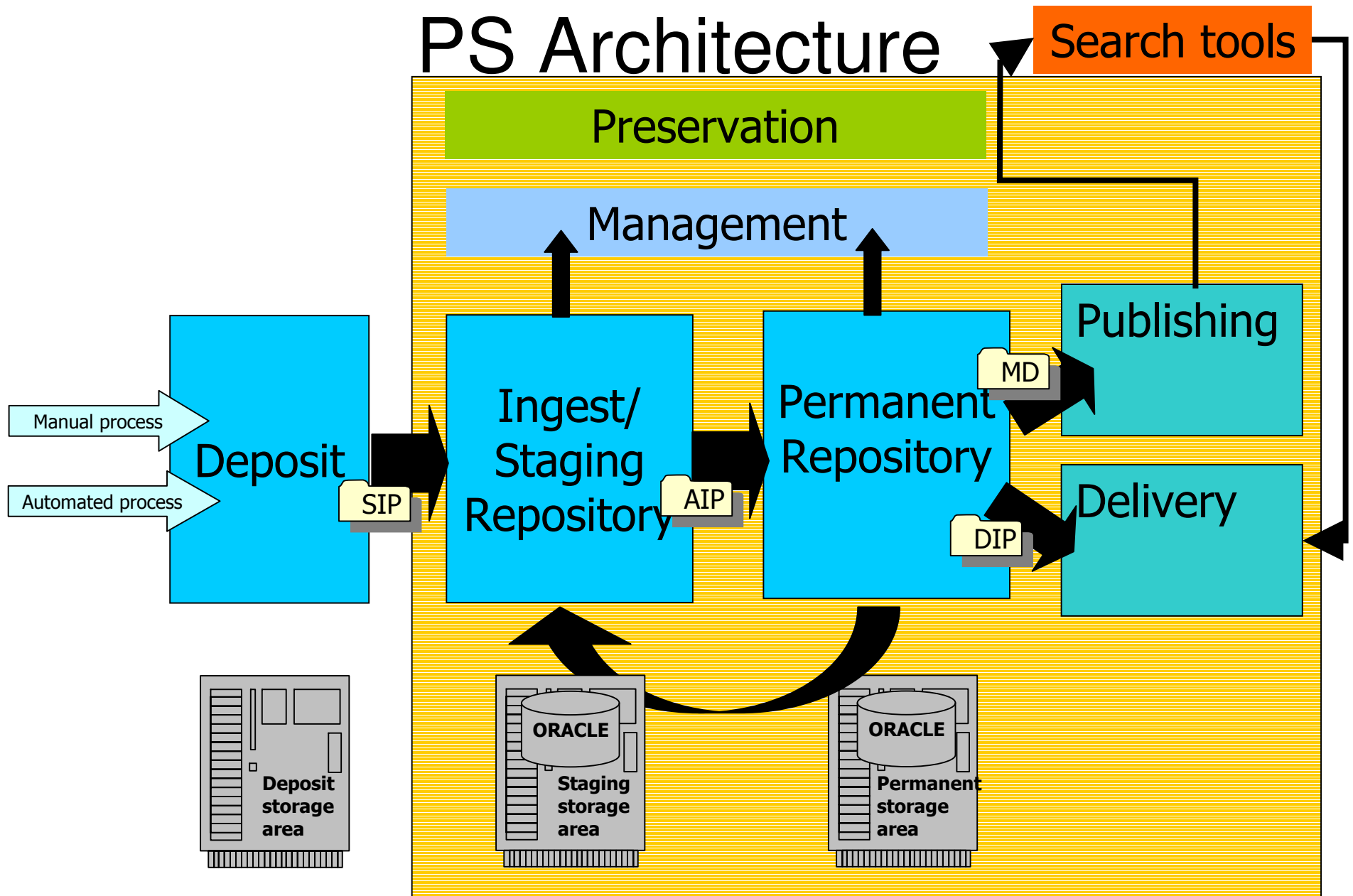


Charles University and National Digital Library SW

**Charles University - Digitool, stress
on Presentation**

NDL- Preservation

PS Architecture





Charles University digital documents

- **Thesis, Papers, e-learning support,...**
- **Rare historical manuscripts**

Maps

Administration



Charles University and its role in digital preservation

Institutional digital repository of the Charles University
Institutional digital documents - institutional funding

NDL – Charles University documents

Institutional digital documents
- institutional funding

National Digital Library
funding: Ministry of Culture



Charles University and its role in digital preservation

NDL – Charles University documents

**Institutional digital documents
- institutional funding**



Title: Sign. I/7

Creator: Karel IV.

Description: Karel IV., římský císař a český král, zakládá kolej pražské univerzity, jíž dává jméno kolej Karlova, a dává jí dům na Starém Městě pražském, jenž se říká Lazarův. Dále koleji dává vsi Horní Počernice, Drahelčice, Uněbuzy, Nenačovice a Holonohy (Pocznicz, Drahelczicz, Vneboz, Nenaczowicz, Holonohy) a pět kop grošů ročního platu ve vsi Čertousích (Trčúsích, "Truss") se vším příslušenstvím, jak bylo zapsáno i do zemských desek. Kolej je zřízena pro dvanáct mistrů svobodných umění, z nichž jeden má vykládat bibli, jeden knihu Sentencií a zbylí kolegiati svobodná umění, přičemž mají zároveň studovat teologii. Za porušení těchto ustanovení je určena pokuta tisíce hřiven zlata, z nichž polovina má připadnout královské komoře, druhá polovina poškozeným. Z přítomných svědků je v listině dvacet pět jmenovitě vypočteno. Na horním okraji listiny je poznámka rukou 14. století o zanesení jejího obsahu do zemských desek: "Hoc privilegium habetur in tabulis regni, in 2o Daczicz G 2o". Ohlášeno je ověření listiny císařskou majestátní pečetí, která byla přivěšena na zlatých hedvábných nitích, avšak dnes není dochována. Součástí textu listiny je monogram Karla IV. Na plíse kancelářská poznámka "Per dominum imperatorem cancellarius", na dorzu listiny registrační poznámka "R. Johannes Saxo". Na dorzu dále poškozený regist rukou 14./15. století "Littera fundacionis collegii imperialis et est duplex" (?), záznam rukou 16. století o vložení obsahu listiny do zemských desek a signatura a regist rukou 16./17. století.

Publisher: Archiv Univerzity Karlovy v Praze

Date: 1366-07-30

Type: Still Image

Format: Pergamen, 42,7 x 55 cm, výška plíky 9,9 cm

Identifier: Sign. I/7

Source: Codex diplomaticus almae Carolo-Ferdinandae Universitatis Pragensis originem, incrementa, privilegia, jura, beneficia aliaque fata, tum prospera, tum adversa illustrans : Monumenta historica Universitatis Carolo-Ferdinandae Pragensis. Tomus II. Pragae, 1834. Č. 55, s. 413-416.

Source: K. Kučera, M. Truc. Archiv Univerzity Karlovy. Průvodce po archivních fondech. Praha : Univerzita Karlova, 1961. S. 83, sign. I/7





Charles University and its role in digital preservation

NDL – Charles University documents

**National Digital
Library**
funding: Ministry
of Culture



National cooperation

- National Digital Library website
- National Board for Digital Repositories (representatives from different libraries, observers from archives)
- National IT Library Board – meetings concerning digital preservation, coordination and cooperation



Sustainability

- Digital preservation can not be covered from regular budgets of memory or research institutions
- Based on grants – limited for several years, sometimes even for one year, problem of continuity and sustainability
- National concept of long-term preservation – sustainability for the central repository (core of the national digital cultural heritage)
- ??? Institutional repositories ???



Standardisation

- **Container: METS - new**
- **Descriptive metadata – coordination on national level, long tradition, based on international standards (DC, MARCXML, MODS)**
- **Administrative metadata - new:**
 - **Technical metadata: PREMIS-object, MIX**
 - **Rights metadata: PREMIS-rights for administrative rights, METS-rights for intellectual property rights**
 - **Digital provenance metadata: PREMIS-events**



Standardisation

Present situation

- **Proposal – mounted on National Digital Library website for comments of the National Board for Digital Repositories and the National IT Library Board**
- **Tested on Kramerius**
- **Consulted abroad**



Conclusions

- Importance of **cooperation** in DP
- **National level:** concepts, standards, repositories, different working groups, conferences, trainings, guidelines, cooperation of memory institutions and research institutions
- **International level:** extremely helpful especially in the early stage – learning from more advanced countries. International conferences, projects etc, audit tolls, external auditors etc.