Ladies and gentleman, dear colleagues, I would like to give you a brief information about **Manuscriptorium** (which is digital library making available elder written and documentary heritage provided by the National Library of the Czech Republic and technically maintained by AiP Beroun Ltd.) and about cooperation that is in base of the Manuscriptorium activities, about a wide network of partners-memory institutions that tak part in the Manuscriptorium network, and about the effort a goal of which is creating of the virtual research environment for the sphere of historical resources, especially medieval an early modern manuscripts, incunabula, early printed books and other prints, historical maps etc.

Main idea of Manuscriptorium is complex although at the first view relatively simple and quite clear. Since beginnings in 2003 it was aggregation of data so that National Library of the Czech Republic made available not only it own digital copies of manuscripts and other historical documents but also those of other Czech and foreign memory institutions, libraries, archives, musea, galleries, etc. Then an effort on integration of services increases that helps and all end users, written heritage professionals, scholars, and general public as well to navigate within the large data sets that are aggregated in Manuscriptorium. Also individual account is offered to end users in order to facilitate both scholarly work (and first especially heuristics) and general work. Last not least step is creating of virtual environment for work with the elder written and documentary heritage which should enable also digital publishingin the Internet and so to realise the turning point on path from the traditioinal printed information and comunication environment to the new networked digital one.

History of the digitization activities in the National Library of the Czech Republic is relatively long and it splits up into several phases by the very marked milestones. Firstly, initial attempts consisted in creastion of the pilot CD for the UNESCO Programme Memory of the World in 1991/2. In 1995/6 a systematic digitization of edieval and early modern manuscrupts started in the National Library of the Czech Republic. DOBM, i.e. standard for creation of the compound digital document (virtual book gasthering both image data and descriptive, structural, and technical metadata) was developed and in 1999 it became a recommendation for the UNESCO Programme Memory of the World. Thirdly, in 2003/4 Manuscriptorium resource (first as so called open catalogue of historical holdings and immediately after that as digital library of compound digital documents-virtual books-copies of historical documents) was created on base of MASTER+, i.e. extension of the newly created MASTER standard which enabled to put together both sophisticated descriptive metadata and the structural ones and consequently to develop an advanced standard for compound digital document. And fourthly, in 2007/9 a fundamental step to aggregation at the European, eventually global level was made through converting MASTER+ standard into the TEI P5 ENRICH Specification which is more general platform enablig interoperability of systems.

Projects are very important in building such huge and large resources of written and documentary heritage like Manuscriptorium. Work on projects must not be seen as a purely formal and organisational matter, it is rather progress in steps which are specific tasks and their solutions that relate more or less close to the aim of Manuscriptorium and its goals. All the projects focus especially on the basic tasks of resource building, i.e. on aggregation of data, on integration of services, on personalisation of the networked digital environment, as well as on virtualisation of scholarly work, at least on creating a base for these fundamental directions of building Manuscriptorium as resource that is estimated to be a nodal point in the whole of the cyberspace

European MASTER project (which is an acronym for the Manuscript Access through STandard for Electronic Records) was done since 1999 to 2001. There were six full partners in the project consortium: De Montfort University, Leicester, UK, Oxford University Computing Services, UK, Institut de Recherche et d´Histoire des Textes, France, Royal Library, The Hague, Netherlands, Institute of Nordic Philology at the Copenhagen University, Denmark, and National Library of the Czech Republic, Czechlands. The aim of the MASTER project was to create a standard for manuscript description, to test it, to disseminate it among both memory and academic institutions an last not least to start production of descriptive manuscript records using MASTER. National Library of the Czech Republic started a regular cataloguing of manuscripts using MASTER since 2001/2 and in 2002/3 developed MASTER+ extension which enabled to connect structural and technical metadata with the descriptive one as well as to connect image data with the matadata, i.e. to create a compound digital document-virtual book. Simultaneously NationalLibrary of the Czech Republic developed in 2002/3 so-called open catalogue of historical holdings and in 2003/4 the Manuscriptorium digital library.

European VICODI project (which is an acoinym of VIsual COntextualisation of Digital content) was done since 2002 to 2004. There were seven full partners in the project consortium: SYSTRAN, s.a., France, RIDEMO, s.i.a., Latvia, Salzburg Research Forschungsgesellschaft m.b.H., Austria, Forschungszentrum Informatik an der Universität Karlsruhe, Germany, University of Newcastle, UK, National Library of the Czech Republic, Czechlands, and University of East Anglia, UK. The aim of the VICODI project was to enhance human comprehension of the digital content on the Internet. This is reached by introducing novel visualisation and contextualisation environment for digital content, i.e. geographical contextualisation of various historical topics based on so-called ontologies. On one hand a structure for historical ontologies was successfully developed in the VICODI project, on the other hand the contextalisation engine worked very slowly so that it was not user friendly. Thus, National Library of the Czech republic do not continue any development in this direction, only started a new step in use of the VICODI ontologies when analysing possibilities to implement it into the Manuscriptorium search engine.

European COMTOOCI project (which is an acoronym of **COM**putational **TOO**ls for the librarian and philological work in Cultural Institution) was done since 2004 to 2005. There were four full partners in the project consortium: Istituto di Linguistica Computazionale, Pisa, Italy, Institute of Mathematics and Informatics, Sofia, Bulgaria, National Library of the Czech Republic, Czechlands, and Institute for Bulgarian Language, Sofia, Bulgaria. The aim of the COMTOOCI project was to test a transcription tool developed by the Institute of Computational Linguistics in Pisa and try to disseminate it among memory institutions in Europe. It was an important experience with an advanced computational work with historical texts for the National Library of the Czech Republic, on the other hand workflow was too complicated to be used it practically in any common cultural institution. However, both sophisticated coding of text and correlation between part of the text and the corresponding digital image-copy of original document was a great inspiration for the implementation of full texts into the Manuscriptorium system.

European ECH:TOPICC project (which is an acoronym of **E**ndangered **C**ultural **H**eritage: **TO**ols for **P**reservation, **I**nvetsigation, and **C**oypright **C**learance) was done since 2003 to 2005. There were eleven full partners in the project consortium: Institute of Mathematics and Informatics, Vilnius, Lithuania, Library of the Lithuanian Academy of Sciences, Vilnius, Lithuania, Public Company Visoriai Information Technologies Park, Vilnius, Lithuania, National Library of the Czech Republic, Czechlands, iTEL, Informatics & Telematics Ltd/ R&D Unit, Athens, Greece, Cultural and Educational Technology Institute (CETI), Xanthi, Greece, Business Systems International SA, Athens, Greece, Narcea Producciones Multimedia S.L, Madrid, Spain, Videostudio of Rezekne Higher School, Latvia, Kolonasata museum of Fr. Trasuns, Latvia, and Graphical data processing centre, Rostock, Germany. The aim of the ECH:TOPICC project was to create the software for the management of multimedia data (text, images, sounds) and intellectual property rights as well as for the interactive internet-based presentation of results. The basic result ot ECH:TOPICC for the National Librarya of the Czech Republic and consequently for Manuscriptorium was implementation of hte on-line connectivity (Z39.50, MARC21 profile; OAI-PMH with profiles MARC21, DC Unqualified, MODS, OpenM. dtd) which was used for connection with the TEL, The CERL Portal and JIB (Czech National Gate). Is is a base for the wide Manuscriptorium network that was created in near future.

European ENRICH project (which is an acoronym of **E**uropean **N**etworking **R**esources and **I**nformation Concerning **C**ultural **H**eritage) was done in 2007-2009. There were eighteen full partners in the project consortium: National Library of the Czech Republic, Prague, Czechlands, Cross Czech, a.s., Czechlands, AiP Beroun, s.r.o., Czechlands, Oxford University Computing Services, UK, Centro per la communicazione e l´integrazione dei media, Florence, Italy, SYSTRAN, s.a., Paris, France, Institute of Mathematics and Informatics, Vilnius, Lithuania, National Library of Spain, Madrid, Spain, Nordic Institute, University of Copenhagen, Denmark,Biblioteca nazionale centrale, Florence, Italy, University Library, Wrocław, Poland, Arne Magnusson Institute, Reyskjavík, Iceland, Computer Science fot the Humanities, Cologne, Germany, St. Pölten Diocese Archive, Austria, National and University Library of Iceland, Reykjavík, Iceland, University of Technology and Economics, Budapest, Hungary, Supercomputing and Networking Centre, Poznań, Poland, and University Library, Vilnius, Lithuania. The aim of the ENRICH project was to provide seamless access to distributed digital representations of old documentary heritage from various European cultural institutions in order to create a shared virtual research environment especially for study of manuscripts, but also incunabula, rare old printed books, and other historical documents. It builds on the Manuscriptorium Digital Library that has already managed to aggregate data from 46 collections from the Czech Republic and abroad. The aggregation of large resources, repositories and data sets was successful, TEI P5 ENRICH Specification, i.e. standard for distributed compound digital document was created, and some useful on-line tools were developed, such as MTool for creating distributed compound digital documents, MCan for checking documents in the simulated Manuscriptorium environment, My Library for personalisation of the Manuscriptorium use, especially for heuristics by means of both static and dynamic virtual collections and virtual documents, Gaiji Bank for inserting non-standard characters and glyphs into XML documents, EGE (ENRICH Garage Engine) for mutual conversions between TEI, MASTER and EAD. Thanks to ENRICH, a base for European digital manuscript library was created through the Manuscriptorium network.

European REDISCOVER project (which is an acronym of REunion of DISpersed CONtent: Virtual Evaluation and Reconstruction) was done in 2009-2010. There were four full partners in the project consortium: National Library of the Czech Republic, Prague, Czechlands, National Library of Lithuania, Vilnius, Lithuania, National Library of Poland, Warsaw, Poland, and National Library of Romania, Bucharest, Romania. The aim of the REDISCOVER project eas to gather, virtually reconstruct, reintegrate and make accessible manuscripts, old printed books, photographs and other documents dispersed during the Late Middle Ages and Early Modern Era. Digital data of all project partners representing medieval and early modern manuscripts as well as incunabula and early printed books were aggregated in Manuscriptorium and

Manuscriptorium networ was extended. However, the most imortant result of the REDISCOVER project is that a base for reconstruction of dispersed historical collections was made.

European EMBARK project (which is an acronym of **E**nhance **M**anuscriptorium through **BA**lkan **R**ecovered **K**nowledge) is progressing since September 2010 to April 2012. There was four full partners in the project consortium (National Library of the Czech Republic, Prague, Czechlands, Veria Central Public Library, Veria, Greece, Institute for Bulgarian Language, Sofia, Bulgaria, National Library of Serbia, Belgrade, Serbia) but one of them, i.e. Veria Central Public Library had to withdraw the work on the project because of catastrophic financial situation of Greece. The aim of the project is twofold, on the one hand professional mobility when experts from all partner countries will mutually transfer their most special and deepest knowledge regarding new digital technologies and methodologies in work with the written and documentary heritage and on the other hand Bulgarian and Serbian partners and memory institutions will aggregate their resources, repositories and data sets in Manuscriptorium. As for creation of distributed compound digital documents and preparation of full text editions of original historical texts as well as insert of non standard cyrillic characters and glyphs into XML documents and use of the lennatisation tool will be teached, eventually learned, especially Balkan Slavic countries and memoryinstitutions will profit taking part in the EMBARK project.

After a brief information about projects related to Manuscriptorium, I would like tell you some words about building of Manuscriptorium as a resurce and about its levels as well. When disregarding digitisation of historical sources, i.e. creation of digital copies, virtual books, compound digital documents of medieval and early modern manuscripts, there are four substantial steps or levels that concern building of resource that has aspirations towards an important node of the digital networked environment, and namely firstly aggregation of data, secondly integration of services, thirdly personalised tools and spheres consisting of catalogue records, images, full texts, music notation, comparison of full texts, and comparison of images, and fourthly virtual environment.

There are several crucial points concerning aggregation of data. Especially five features are significant. Firstly, it is desirable, even indispensible, for every resource in the network environment making mediaeval and early modern manuscripts accessible to be furnished with a database and information system relating to the end user and allowing, or facilitating, work for him/her as well as with an administrative system allowing and conditioning its sophisticated internal organisation. Secondly, such a resource then is not a mere digital replica of a physical library with its depository, catalogue, loan and basic information services but much rather a kind of elementary heuristic tool assuming not only the functions of unqualified as well as qualified service personnel but also replacing some basic, routine research activities. Thirdly, it is thus evident that the building of the individual resources making mediaeval and early modern manuscripts accessible already in itself in the indicated sense significantly surpasses the opportunities that were offered by the traditional scientific infrastructure until just recently, that hence already now the progress is substantial, in spite of the fact that the established scientific circles are usually not able – or willing – to see it. Fourthly, it is better, although it is materially, technically and financially more demanding, to aggregate the resources in such a way that their descriptive metadata, i.e. catalogue entries, are aggregated, and they are converted to a uniform format, whereas the data, the digital images and alternately full texts remain distributed in the original resources on the servers of individual partners of the aggregate resource. The advantage of this centralised-distributed conception of an aggregate resource is on the one hand the significantly faster as well as completely consistent searching, which provides considerably greater comparability and reliability, which is only minimally disturbed by the different information depth of the descriptive metadata, i.e. of the catalogue entries from the original resources. And fifthly, in this way, the joint search finds manuscripts which had never until now been found in it. To attain the same result with traditional heuristic methods is infinitely more exhausting and slower and in a number of cases is not possible at all.

Service in this sense is a new concept in relation to the codicological research not appearing in its full extent until in the digital network environment. In essence, it is what is called strategic services in the other cultural, social and economic sectors. They are hence neither the personal services which have traditionally been offered to end users by libraries and other memory institutions since time immemorial nor the type of services which has been replacing routine research work in the digital environment and is a component of the previous digitisation and aggregation levels. This type of services is in contrast much more sophisticated, because it already affects some elements of actual research work in that it makes it possible to perform various comparisons of the sources or their conversion to a more usable form, but it is necessary to point out that this level presupposes the transformation of the original physical analogue sources into the form of digital data, whether turnover or full/text or sound or multimodal/multimedia. In short, manuscripts are at this level usable solely in the form of digital image copies and their further processing exclusively in the form of electronic fulltexts, and so on, etc. It means that whatever exists only in traditional printed or written and other, simply analogue form does not exist for this way of utilisation and elaboration, because it is not machine-readable and consequently it is essentially ungraspable. It is in retrospect completely clear from this how colossal an importance the digitisation of the written and cultural as well as the scientific heritage has, because without it irreplaceable losses occur.

Personalisation of the environment does not mean having the possibility to create one’s own account in the formal sense only within the individual resources, i.e. the existence of an access password, on the basis of which the end user is recognised. It much rather means that each end user of this or that resource has his/her own digital space, which s/he can use independently. Having a personal account, i.e. a personal space, available entails first of all the possibility to accumulate the results of previous work and hence also to have one’s own tools for acquiring results as well as for their further processing and subsequent utilisation in the digital network environment there. And there is still another advantage here tied to this type of personal space. The user of the personal space may reserve all of his/her space for himself/herself and nobody else or make all of this space or some of its component parts or even individual documents and files accessible to other selected persons and also quite generally to all users of the digital network environment. If we look at it in terms of the method and approach of codicological work, the personal space allows both strictly individual work and team work as well as various combinations of them. Through general access for all of the digital network environment, however, also the base for electronic network publication is created, which does not have to be necessarily dependent on employer or publisher institutions, or such relations will be resolved in a different way than up to now.

The first sphere that has its role in the work in the personal and personalised environment is catalogue record. It does not mean any basic and/or simple descriptionthe working of the Manuscriptorium technical and daabase and information system is based on, but it means diverse variants of descriptive records both from the subject specific point of view like e.g. form point of view of musicology, history of arts, etc. and many types of analytical descriptions like in-depth description of manuscript content, composition and/or structure of intelectual units of higher level and/or individual intellectual units of lower levels etc.

The second sphere that has its role in the work in the personal and personalised environment is sphere of images, i.e. digital copies of original historical documents. Now only gallery thumbnails, previews, lower and higher are provided in Manuscriptorium, in some cases also black and white optimisations that due to the bigger contrast improve possibility of reading difficult legible texts. In future, black and white optimisation should became default and details of selected pages as well as UV scanns (for palimpsests) should be available. That means, both images and descriptive catalogue records will be provided in variants.

Not only distributed compound digital documents-virtual books-digital copies of manuscripts have to be provided in Manuscriptorium as well as in other aggregated resources providing written and documentary heritage. Full texts have to be published in networked digital resources. In digital environment an idea of fluid text arises. Fluid text is a very controversial concept, which completely contravenes the existing conceptualisations of the humanities disciplines. The text in semiotic delimitation is a set of features with a meaning and associated on the basis of syntactic rules. In this way, the text can be anything, perhaps even a culture in the sociological sense or a manuscript in the context of with its external and internal features. This could be close to social and cultural anthropologists, less however to historians or philologists. I will then speak here of fluid text as of a text in the common sense, i.e. as a record of a language expression with letters. Nevertheless, not even here can we avoid certain collisions and misunderstandings, namely because historians and philologists see something else in such a text: whereas for historians the text is merely a direct or indirect trace of an external reality, which is given by human activity and its material and ideal results, for philologists the text is a part of an external reality itself (and for some it even replaces this external reality). The text and as a consequence of this also the work, the artefact therefore has a different position in the structure of the world as understood on the one hand by historians and on the other by philologists. The text in a flowing or fluid state two things: the first is the methodological approach, which arises from the fact that a number of the records of the texts have not been preserved, which concerns particularly sermonic, university and didactic paraliterature, and hence in a number of cases it is impossible to find any fixedly given text, which would be a clear starting point for the further history of the text; the second is a conceptualisation of mediaeval paraliterature, which lacked a sufficiently clear and strong awareness of authorship and originality and where the individual works were much less different from one another than in the case of so-called high literature and than we are accustomed today. And since particularly historians work rather with this paraliterature, the concept of fluid text has begun to be close to them, because it places the more fluidly given text among the other paraliterary expressions while seeing in this context not only the text itself but already its individual records so it is possible to consider that the individual manuscript record is a component state of the fluid text, which spreads not only to all preserved but also to all of the formerly created but no longer preserved records and texts. In the traditional printed environment, such an idea was conceivable only purely abstractly and potentially, without any possibility to be realised. In a digital network environment, which can be characterised as infinity closed in borders, it is on the other hand possible to realise albeit not in the absolute sense. A contextual edition, i.e. one of possible representations of the fluid text, is focused on the representation of the text and its record in connections with other texts and records. Already this basic starting point predetermines that the text does not necessarily have to be understood as fixedly given, but that it is characterised by a quite significant level of variability. This must be reflected also in the contextual edition.

While we are able to think on fluid text and contextual editions in the case of full texts, regarding music notation we did not go so far. Of course, we understand that music notation is a written record of music as well as that i is a text in the semiotic meaning, i.e. that it is a set of meaningful signs related each to other according to syntactic rules. Consequently, we understand that digital representation of the music notation is edition is similar sense like usual textual edition. On other hand, there are difficulties with coding and displaying of different systems of music notation as well as with mapping them and consequently converting them, i.e. converting one notation systém into other and vice versa. Thus, first attempts are beeing done concerning music encoding, its XML representations and correlation of music edition with images representig original historical document now.

At present the last tool that leads from the traditional codicology to the digital one is a tool for comparison of full texts. Originally and principially, it is a tool based on the frame of the computational linguistics for philological work, but it can be used also outside this philological area. It is a tool that facilitates work with internal attributes or content features which is a crucial step to both enlarging traditional codicological themes and leading to interdisciplinarity. Comparison of full texts is based on vectoral statistics, i.e. it compares not simply words, but word strings or text strings. A big problem is that in the past stages of language evolution, language systems not only in Latin but especially in vernacular languages lacked any conception of strict prescriptions, i.e. languages in the past used neither a prescriptive grammar nor a normative orthography. Thus, a more sophisticated technique and methodology like e.g. graphical variants (for texts that use no normative orthography) must be used for the comparison to be effective and successful. It works with such standards like XML, HTML, and TXT. Now, full text comparator is only an off-line prototype, not yet implemented into the Manuscriptorium platform. Representation of results is only numerical-statistical, but after implementing it as regular on-line service (supposed in end of 2011 or early 2012), it will be enriched with the graphical representation of results, too. A tool for comparison of full texts is an important step into the virtual research environment and into the area of the digital codicology.

Tools for comparison of images accomplish usually distinguish human faces, trade marks etc. However it has a little sense in the sphere of written and documentary heritage, in manuscript work, digital codicology and likewise. Thus, the first step is distinguishing and eventually selecting typycal groups of images, especially manuscript pages with text, music notation, illuminations, borders, diagramms and tables. As for algorithms are already known, a convenient tool can be developed and programmed, so that Manuscriptorium team tests such a tool at present. As for we have a long-run cooperation with the department of cybernetics at the Technical University in Prague, we suppose to solve also further tasks, e.g. finding the same or very similar illumination, finding the same or very similar detail of illumination like e.g. coat of arm, attribute of saint and the like, finding similar composition of figurative and/or non-figurative illumination, etc. There can be also many other ideas, but development of these sort tools is not trivial so it will také some time to realize it.

Knowledge is the ability to treat and elaborate information, to complement it with metainformation, i.e. additional information, and on that basis to create transinformation, hence new information. At the same time, it is necessary to be aware that information is identical with data, that it is not a form of content but communicated particular which has an influence on a change in thought and subsequently also behaviour. Knowledge this is not and cannot be only the representation of information, because in this regard there is nothing to represent generally, but it is the ability to work with information, to create metainformation and transformation and eventually also new knowledge *ad libitum ad infinitum*. Virtualisation is a path of how to uncover implicit, not only explicit, features of the objects of the natural and artificial worlds. It means that virtual reality is such a reality that is not evident at first sight and is difficult to capture with traditional analytical methods found between induction and deduction and can be captured only with the additional use of abduction, i.e. an approach that is beyond the set of the existing methodologies, methods and techniques used in the field concerned and that is adopted from another field or even is based on mere common sense. We thus come to a field which seems to be entirely new and specific for digital codicology as against traditional codicology; nevertheless, we must again draw attention to the fact that the first steps towards this were made by culturomics, which had emerged based on the massive digitisation of the later as well as earlier written and documentary heritage.

In future Manuscriptorium will continue the same way. Aggregation of data and metadata means to cooperate with more and more partners in more and more countries and continents, cultures and civilisations in order to achieve not only European but just the global level. Integration of services on the Manuscriptorium site means further steps to more closely network of memorys institutions holding and providing written and documentary heritage, especially medieval and early modern manuscripts. Further virtualising the networked digital environment means to work on further development of personal space, of external tools and to go through the path of interoperability and specialised webservices that will enable cooperation both between individual resources like Manuscriptorium and over them.

Zdeněk Uhlĩř

National Library of the Czech Republic