

Digital Preservation Europe

Exchange Report

Award Holder	Hosting Institution
<p data-bbox="395 779 651 815">Gemma Tougher</p> <p data-bbox="316 891 735 1182">Former student on the Msc Information Management and Preservation course HATII George Service House 11 University Gardens University of Glasgow G12 8QH</p>	<p data-bbox="847 779 1294 853">National Library of the Czech Republic (NKP)</p> <p data-bbox="842 891 1302 1070">Digital Preservation Department Klementinum 190 110 00 Prague 1 Czech Republic</p>

17th December 2008

Contents

Introduction	3
Pre-Exchange	3
Exchange	5
DPE Starting Out Training Event	9
Post-Exchange	10
Bibliography	12

Introduction

This report records the activities, aims and outcomes of a DigitalPreservationEurope exchange trip at the National Library of the Czech Republic (NKP) from 6th to the 17th of October 2008.

The award holder has a background in archives and libraries and first became interested in digital preservation while studying the Information Management and Preservation (IMP) course at HATII, based at the University of Glasgow. In particular, the ‘Management, Curation and Preservation of Digital Materials’ module taught by Professor Seamus Ross, Director of DigitalPreservationEurope, offered a valuable introduction to the main principles and challenges involved in digital preservation. Furthermore, it highlighted how archivists and cultural heritage sector workers need to be aware of these issues and how to approach them if they are to ensure the accessibility and interoperability of complex digital materials for researchers and archive users of the future.

The DPE exchange programme offered an opportunity for the award holder to apply the theoretical knowledge gained from the IMP course to a practical digital preservation environment, and to engage with the methods, policies and standards adopted by the digital preservation community. The exchange also provided the award holder with a chance to analyse and evaluate the projects and activities undertaken by the Digital Preservation Department at the host institution.

Pre-Exchange

1. Institutional Context

The National Library of the Czech Republic (NKP) is an influential partner in the DigitalPreservationEurope project. Based in the Klementinum, an former Jesuit monastery in the heart of Prague, it serves as a national library, which acquires and preserves important Czech historical documents, and also strives to meet the demands of current researchers, such as students and academics from the nearby Charles

University.¹ It has been actively involved in digitisation projects and related research since the early 1990s² and over the years has gained ‘worldwide recognition for its long tradition and remarkable results in the area of digital archiving’.³ It has openly embraced digitisation as a means of disseminating and providing public access to a variety of manuscripts, books, periodicals, newspapers and other physical documents, and has tackled the problem of capturing and preserving online resources by creating and managing the WebArchiv, which will be discussed in more detail below.

2. Organisational Context-the Digital Libraries Department

The Digital Preservation Department, which is based at the Klementinum, is currently preparing, researching and conducting tests for the ongoing large-scale National Digital Library programme, which is funded by the Ministry of Culture. The team are involved in evaluating suitable preservation file formats and testing and implementing the software for persistent identifiers to ensure that the material preserved and disseminated in the national digital repository will be accessible over time.

The Digital Preservation Department at the NKP was chosen as a suitable host institution for this exchange mainly because it was expected that the award holder would obtain an awareness and understanding of the impact and aims of the digitisation projects currently in place at the NKP and also learn about the preservation planning and practical activities behind these projects. Members of the host institution were aware that the award holder had little practical digitisation and digital preservation experience, so the balance of the exchange was to provide an environment for learning and gaining knowledge and understanding through discussions, presentations and literature, in favour of practical hands on activities.

¹ ‘Library’s History’, *National Library of the Czech Republic*, [http://www.nkp.cz/_en/pages/page.php3?navez=Library%27s_History&submenu3=5], last accessed 30/11/08

² Brian Rosenblum ‘Digital access to cultural heritage and scholarship in the Czech Republic’, *University of Kansas Scholarworks*, [http://hdl.handle.net/1808/1715], pp. 1-24, last accessed 14/12/08 and Bohdana Stoklasova, Jan Hutar and Pavel Krbec, ‘Preservation of digital cultural heritage in Europe and in the Czech Republic’ (provided by staff at NKP)

³ Bohdana Stoklasova, ‘Czech Digital Library: paper presented at the 1st Colloquium of Library Information Employees of the V4+ Countries, Bystrica, Slovakia, May 11-13, 2006’, [http://www.svkbb.sk/colloquium/zbornik/obsah.htm], last accessed 7/12/08

The Exchange

The award holder spent the two week exchange in the Digital Preservation Department with Jan Hutar, Marek Melichar and Ladislav Cubr, who provided a very warm and friendly welcome. The first day consisted of a brief tour of the Klementinum and an introductory discussion with Jan about the work of the Digital Libraries department, an overview of the three main digitisation projects, and the projected aims and outcomes of the DPE exchange. The award holder was supplied with a desk, a plentiful supply of tea and a basic plan of action. It was agreed that the first week would be spent analysing and evaluating the three main digitisation projects through reading literature and meeting with those involved, and the second week would focus more on the National Digital Library programme and how the Digital Preservation Department are involved in preparing and planning for the different stages of this programme.

1. Manuscriptorium and Kramerius

Through discussions with Jan and Marek, and analysing the portals on the NKP website, the award holder gained a good grounding in the background and aims of these two ongoing digitisation projects. The award holder first spent some time becoming familiar with the Manuscriptorium project, which is a ‘virtual research environment’⁴ for old books and manuscripts, and was impressed with the quality of the images and the browsing capabilities on this portal. The project allows the NKP to provide access to its rare physical documents in a stable digital environment which supports the contribution of digital material from other institutions in the Czech Republic and other countries, such as Poland, Slovakia and Turkey.⁵

The aim of the Kramerius project on the other hand is to digitise fragile and deteriorating modern physical documents as a means of preserving them. Newspapers, periodicals and other documents suffering from acid erosion will soon be too fragile to be handled by researchers and the project attempts to solve this problem by providing access to digitised versions. Therefore, it is crucial that the digital material is stable, reliable, and stored on a suitable preservation file format to ensure that the

⁴ Stoklasova, ‘Czech Digital Library’

⁵ Rosenblum, ‘Digital Access to Cultural Heritage’, p. 4

intellectual content of these records will survive into the future, despite the continuing deterioration in the condition and quality of the originals.

Through analysing both projects the award holder became aware of many digital preservation issues affecting the NKP, such as the importance of providing metadata and suitable preservation file formats, and of making sure that the digital information is trustworthy and reliable. The project analysis also highlighted the prevalence of copyright issues which can affect access to certain digitised material. Most of the online digitised images on Kramerius can only be accessed within the National Library and the award holder agrees with Bohdana Stoklasova, Director of Library Collection and Services at NKP, who argues that current copyright legislation may have to change in order to accommodate access to the growing number of published digital resources.⁶

2. WebArchiv

The award holder was given the opportunity to meet with Libor Coufal, who manages a small team of staff working on the WebArchiv. The discussion was based around the aims of the WebArchiv, the criteria used to select web resources; the practicalities of selecting and harvesting material; the frequency of these activities; and significant problems regarding copyright and access. The NKP seem to have used their expertise to establish a very successful solution to the problem of securing and preserving web resources and other online publications. The award holder was interested to note that the approach to acquiring and preserving web archive material is similar to that of traditional manuscript and paper archives: selection guidelines are firmly established, the integrity and completeness of the web resource is valued, and automatic metadata generator tools, which are based on an International standard (Dublin Core), are in place to ensure that there is contextual and descriptive information about the digital material.

⁶ Stoklasova, 'Czech Digital Library'. She claims that 'authorizing legislation, terminology, and responsibilities for preservation of and access to traditional documents are relatively stable and satisfactory. However, this is not the case for digital (both digitized and digitally originated) documents. Legal deposit legislation does not include Internet resources and must be updated'. See also Rosenblum 'Czech Digital Access to Cultural Heritage', p. 6

However, just like the Manuscriptorium and Kramerius projects there are also ongoing copyright problems. According to Libor only around 40% of website creators respond to requests asking permission for their sites to be added and freely accessed by users of the WebArchiv.⁷ Therefore, the staff working on the WebArchiv are fully behind initiatives such as Creative Commons, which allow authors to ‘easily mark their creative work with the freedoms they want it to carry’ in order to promote the sharing and dissemination of digital resources.⁸

The discussion with Libor provided the award holder with the opportunity to ask questions in order to supplement knowledge gained from the analysing the project portal. One particular point of interest was approaching how these websites will be preserved once they have been archived. A significant amount of time and resources have been spent on developing tools for selection and automatic harvesting and the success of the WebArchiv can partly be attributed to these efficient and innovative tools. But the real challenge will be in planning the preservation strategy behind this web archive. According to Libor the working group of the International Internet Preservation Consortium,⁹ of which he is a member, is still in the early stages of preservation planning and no one is entirely sure of the correct way to approach the challenge of preserving websites. The award holder found this conversation interesting and thought provoking and is now keen to explore in more detail how UK web archives, such as the UK Web Archive Consortium, and web archives of other countries and institutions are operated and controlled.

3. National Digital Library

During the second week of the exchange the award holder learned about the impressive Digital Library programme which is currently in place at the NKP and the digital preservation challenges and issues it raises. The aim is to use digitisation to provide one portal where all digital resources relating to Czech cultural heritage can be accessed. With backing from the Czech Ministry of Culture and the Czech Government, the NKP and the Moravian Library in Brno now have the funding and

⁷ Conversation with Libor Coufal at NKP on October 2008. For more information see Rosenblum, ‘Czech Digital Access to Cultural Heritage’, p. 7

⁸ *Creative Commons*, [<http://creativecommons.org/>], last accessed 14/12/08.

⁹ ‘Mission’, *International Internet Preservation Consortium*, [<http://netpreserve.org/about/mission.php>], last accessed 14/12/08

resources to embark on a large scale digitisation project and to provide ‘long-term preservation of and access to digital documents’.¹⁰ At the ‘heart of the whole system’ are the three National Digital Library projects described above which are thought to contain the ‘core’ documents of Czech national cultural heritage’.¹¹ The overall aim of the Czech Digital library is to provide the facilities to allow Czech cultural heritage organisations, such as archives, museums or libraries to submit documents for preservation in the Czech Digital Library and also to use ‘integrating tools’ to provide links to ‘digital data stored in local repositories’.¹²

The analysis of the project raised a number of issues. Firstly, the literature discussing the project proposals and project aims highlighted the need for the use of standards across the entire digital repository from the system model itself (which is based on the OAIS reference model) to metadata and persistent identification standards.¹³ Considering persistent identification, after testing a number of different options the Digital Preservation Department have decided to implement the Italian solution for Uniform Resource Name National Bibliographic Number (URN NBN) as a way to ensure that every digital resource on the Czech Digital Library can be easily identified and retrieved with a unique number and that this system will cope with new unique identifier methods in the future.¹⁴ The award holder spent a bit of time speaking to Jan and Marek and researching persistent identifiers to gain a better understanding of their significance and how they are implemented. The URN NBN is a global identifying system which works on an international level (rather than being a method specific to a particular institution) and will aid the preservation of digital resources by ensuring that digital documents can be easily retrieved and accessed through their unique URN NBN number, regardless of changes in their location. At the moment the Digital Preservation Department are setting up the software (resolver) phase.

¹⁰National Library of the Czech Republic, ‘IOP Project Proposal: Broader Context-Digital Preservation of National Digital Cultural Heritage’ (provided by staff at the NKP)

¹¹ Ibid. and Stoklasova , ‘Czech Digital Library’

¹² Ibid. and Stoklasova , ‘Czech Digital Library’

¹³ National Library of the Czech Republic, ‘Request for documentation for a project prepared within EU Structural Funds - Smart Administration (IOP, line 1.1)’ August 2008, draft version, p5 and pp. 9-10 (provided by staff at the NKP)

¹⁴ Ibid. p. 13

Secondly, the literature demonstrated that the NKP are keen to ensure that the Czech Digital Library will become a certified trusted repository and that, through this, users will have the guarantee that the digital records are reliable and authentic. As an archivist working with original and authentic physical documents the award holder believes that the implementation of certified trusted repositories and of self-audit tool kits such as DRAMBORA will be crucial for archives and digital libraries of the future. At the moment many library and archive buildings stand as a testament of authority and trustworthiness and these attributes are often transferred to the records deemed worthy of being stored inside them. But virtual libraries and archives will need to secure a means of establishing their reputation and expertise in the management and preservation of digital material so that users of these records are convinced that they are genuine and have not been altered. The award holder intends to find out more about trusted repositories, which will hopefully lead to fruitful discussions with fellow archivists and the award holders own institution gaining trusted status through such means as DRAMBORA in the future.

Overall the award holder felt that Czech Digital Library was an innovative and challenging project which once again establishes the NKP as an organisation committed to leading research and development in the management, dissemination and preservation of digital materials. The discussions with Jan and Marek and the analysis of the project literature encouraged the award holder to think about the planning issues that have to be dealt with during such a large scale project and how its success is largely dependant on the generosity of funding bodies.

DPE Starting Out Training Event

The DPE Starting Out training event was taking place at the same time as the exchange trip and the award holder was given a fantastic opportunity of being able to sit in on some of the lectures given. This was a very worthwhile experience as it complimented some of the subjects and issues being discussed with members of the Digital Preservation Department. In particular issues such as preservation file formats, the OAIS reference model and trusted repositories. It also provided the award holder with an opportunity for learning about collaborative digital preservation projects such

as CASPAR, PLANETS and PLATO, and to engage in discussions with participants from a variety of different backgrounds. This training event allowed the award holder to think about how digital preservation is approached collectively by European institutions and where the NKP fits in to all of this.

Post Exchange: Impact and Evaluation

The award holder thoroughly enjoyed the exchange and the issues and discussions it raised. Spending two weeks working in a department that is actively engaged in managing and planning digitisation projects and the preservation of these digital documents demonstrated the practicalities of approaching digital preservation, such as thinking about the selection of material and how to retain its intellectual content as well as its reliability and authenticity, how to store it on suitable file formats, how to implement standards and unique identifiers, how to add metadata and to provide access to users, and how to continually manage the material and make decisions on which preservation methods to adopt.

As mentioned at the start of the report the award holder had little knowledge and no practical experience of digital preservation prior to the exchange but now has confidence in discussing preservation planning and how to approach it, as well as an understanding of some collaborative European projects and current digital preservation research trends. Furthermore, the award holder learned about the mission and aims of the NKP, how it uses digitisation to promote access to Czech cultural heritage as well as to preserve fragile and rare physical documents, and its plans and ambitions for the future.

The award holder has just begun a career as an archivist and is well aware that over the next 40 or so years there will be massive changes in the ways in which information is transferred, disseminated and stored. Furthermore, there will be huge increases in born digital material and it will be necessary for archivists to work with IT specialists and other information workers to ensure that the digital records worthy of permanent retention can be adequately preserved for future generations, whilst remain authentic and reliable. This exchange has given the award holder the

knowledge and confidence to discuss the principles and challenges connected with digital preservation and to think about how it will affect archives. The exchange also allowed the award holder to establish relationships with members of the European digital preservation community and it is hoped that these relationships will help the award holder to keep up-to-date with new research and developments in this field.

Overall then, the exchange has given the award holder the chance to develop an interest in digital preservation which can be taken forward and disseminated with future employers in the cultural heritage sector. In addition, DigitalPreservationEurope benefits from sharing knowledge and expertise with a young practitioner and ensuring that the importance of digital preservation is recognised at the earliest possible opportunity and passed on to other young practitioners, thus leading to the expansion and growth of the digital preservation community. This may, in turn, lead to the development of new ideas and approaches to research.

Finally, the award holder would like to thank Professor Seamus Ross, Emily Nimmo, Jan Hutar, Marek Melichar, Ladislav Cubr, Libor Coufal, Bodhana Stoklasova, and all who made this exchange possible and enjoyable.

Bibliography

Creative Commons, [<http://creativecommons.org/>], last accessed 14/12/08.

International Internet Preservation Consortium,
[<http://netpreserve.org/about/mission.php>], last accessed 14/12/08

National Library of the Czech Republic, 'IOP Project Proposal: Broader Context-Digital Preservation of National Digital Cultural Heritage' (Provided by staff at the NKP)

National Library of the Czech Republic, 'Request for documentation for a project prepared within EU Structural Funds - Smart Administration (IOP, line 1.1)' August 2008, draft version, pp.1-13 (Provided by staff at the NKP)

National Library of the Czech Republic, [<http://www.nkp.cz>], last accessed 14/12/08

Rosenblum, Brian 'Digital access to cultural heritage and scholarship in the Czech Republic', *University of Kansas Scholarworks*, [<http://hdl.handle.net/1808/1715>], pp. 1-24, last accessed 14/12/08

Stoklasova, Bohdana 'Czech Digital Library: paper presented at the 1st Colloquium of Library Information Employees of the V4+ Countries, Bystrica, Slovakia, May 11-13, 2006', [<http://www.svkbb.sk/colloquium/zbornik/obsah.htm>], last accessed 7/12/08

Stoklasova, Bohdana Hutar, Jan and Krbec, Pavel, 'Preservation of digital cultural heritage in Europe and in the Czech Republic' (slides provided by staff at NKP)