Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project

Jiří Bernas, Michal Wanner

Long-term preservation of digital records in the Czech Republic and the National Digital Archives project

ABSTRACT

The paper deals with the currently most pressing problems of the Czech archives - digital archiving. In the last five years implementation of e-government in the country very advanced. Although the Czech archives paid attention to associated issues for many years, it is still necessary to solve issues in many areas. While so far the Czech archives have been concentrated on research, development strategies and policy approaches, and raising funds for the construction of the National Digital Archives (NDA), currently already working on the implementation of the NDA and other digital archives and legislative arrangements. This contribution outlines, on the background description of the e-government in the Czech Republic, a strategy to address the existing problems. The contribution submits the characteristics of the NDA and other planned digital archives in the Czech Republic, describes the legislative basis, standardization of electronic records management and new basic rules for the arrangement of archival records.

La conservazione a lungo termine di documenti digitali nella Repubblica Ceca ed il progetto Archivi Nazionali Digitali

SINTESI

L'articolo tratta dei più pressanti problemi relativi agli archivi ed all'archiviazione digitale nella Repubblica Ceca. Negli ultimi cinque anni, nel paese è stata molto aumentato l'amministrazione digitale. Sebbene gli archivi cechi abbiano prestato attenzione alle problematiche connesse per molti anni, risulta tuttora necessario risolvere alcune questioni in molte aree. Mentre finora gli archivi cechi si sono concentrati sulla ricerca, le strategie di sviluppo e la politica dell'accesso, nonché al reperimento di fondi per la costruzione dell'Archivio Nazionale Digitale (NDA, in originale), al momento si sta ancora lavorando sull'attuazione dell'NDA ed altri archivi digitali e sull'organizzazione legislativa. Questo contributo descrive, sulla base della descrizione dell'amministrazione digitale ceca, una strategia di approccio ai problemi esistenti. Il contributo fornisce le caratteristiche dell'NDA e di altri archivi digitali in programma nella Repubblica Ceca, ne descrive le basi legislative, la standardizzazione della gestione dei documenti digitali e le nuove regole base per l'organizzazione dei documenti d'archivio.

Dolgoročna hramba digitalnih zapisov v češki Republiki in projekt Nacionalni e-arhiv

IZVLEČEK

Prispevek obravnava trenutno najbolj pereč problem čeških arhivov - elektronsko arhiviranje. V zadnjih petih letih je uvajanje e-uprave v državi zelo napredovalo. Čeprav so češki arhivi temu vprašanju že v preteklosti posvečali pozornost, pa je še vedno potrebno rešiti težave na mnogih področjih. Medtem ko so bili češki arhivi osredotočeni na raziskovanje, pripravo razvojnih strategij in politike ter pridobivanje sredstev za izgradnjo Nacionalnega e-arhiva (NDA), pa se trenutno že ukvarjajo z uvajanjem NDA in drugega elektronskega gradiva ter pripravo ustrezne zakonodaje. Na osnovi opisa e-uprave v Češki Republiki, prispevek predstavlja strategijo reševanja obstoječe problematike. Na začetku podaja značilnosti NDA in ostalih načrtovanih e-arhivov v Češki Republiki, nato pa predstavlja zakonodajo, standardizacijo upravljanja z e-gradivom in nove osnovne principe urejanja arhivskega gradiva.

Dlouhodobé uchovávání digitálních dokumentů v České republice a projekt Národního digitálního archivu

ABSTRAKTNÍ

Článek se zabývá nejpalčivějším problémem českého archivnictví současnosti, totiž digitálním archivnictvím. Implementace e-governmentu v České republice velmi pokročila. Přestože české archivy věnovaly v minulých

letech mimořádnou pozornost problémům spojeným s tímto fenoménem, je stále nutné vyřešit řadu problémů. Zatímco doposud se české archivnictví soustřeďovalo na výzkum, vývoj strategických a koncepčních přístupů a získávání prostředků na výstavbu Národního digitálního archivu (dále jen NDA), v současnosti již pracuje na jeho výstavbě. Příspěvek naznačuje, na pozadí popisu e-governmentu v České republice, strategii řešení uvedených problémů. Článek podává charakteristiku NDA a dalších plánovaných archivů v České republice, popisuje legislativní základ, standardizaci elektronické spisové služby a nová Základní pravidla pro zpracování archiválií.

Currently the most urgent issue faced by the Czech archives is undoubtedly the digital archiving. In the last five years, the implementation of e-government in the CR advanced considerably. Although the Czech archives have paid major attention to the relevant issues for many years already, there are still a number of areas with issues outstanding. So far the efforts have been exerted especially in the field of research, strategy and policy development and fund raising for the construction of the National Digital Archives. At present, the implementation of the NDA project is already underway as well as the works on amendments to legislation allowing for the creation of other digital archives. This paper recaps the up-to-date development in addressing the matters of long-term preservation of digital records and against the background of e-government in the Czech Republic, outlines the existing strategy governing this area, characterizes the national digital archives and other scheduled digital archives in the CR, and it also describes the relevant legal framework. Thus, the paper interalia also points at substantial increase in matters to be addressed, quantity of knowledge in so far not investigated areas which the archivists shall focus on in the today's society. Often their role has considerably changed, but the ultimate goal always is to preserve the records and information for future generations.

Preservation of digital records in the CR

Matters associated with the long-term preservation of digital records emerged in the Czech archiving in the second half of 1990s. They were brought to attention by a group of experts focusing originally on the information systems in archiving. This small team composed of archivists as well as software and hardware specialist has established a close cooperation with the Faculty of Electric Engineering of the Czech Technical University in Prague and the National Archives in Prague. They formed the core of a research team which for several years (2000, 2001, 2003, 2004, 2007 and 2008), though intermittently, conducted research and generated a huge quantity of findings in multiple areas. The team mapped the fundamental methods of digital records preservation, explored the matters of authenticity of digital records, standards in the field of metadata and description of digital records, concentrated in detail on the preservation of individual types of digital records, and above all gathered vast knowledge through the visits to archives in Europe, the United States of America and Australia. Multiple foreign documents and standards have been upon agreement with the foreign colleagues translated into the Czech language[1]. Some results of this research were published in international journals, presented on the occasion of individual visits or conferences in the CR and abroad. Therefo-

---

1. Oskar MACEK, Michal WANNER, *Zkušenosti polských a nizozemských archivářů v oblasti využívání výpočetní techniky*, „Archivní časopis" 49(1999), n. 1, p. 1-9; *Co by se stalo kdyby? (problémy a možné důsledky vyplývající z využívání výpočetní techniky ve společnosti)*, „Archivní časopis" 50(2000), n. 1, p. 1-7; Oskar MACEK, Michal WANNER, *Řešení problematiky dlouhodobého ukládání digitálních dokumentů ve Velké Británii – poznatky ze služební cesty*, „Archivní časopis" 52(2002), n. 4, p. 219-225; Oskar MACEK, Michal WANNER, *Emulace nebo migrace aneb jak dlouhodobě ukládat dokumenty v digitální podobě*, „Archivní časopis" 53(2003), n. 2, p. 65-74; Oskar MACEK, Michal WANNER, *Zásady uchovávání dat v Datovém archivu Spojeného království*, „Sborník archivních prací" 2003, n. 1, p. 239-288; Oskar MACEK, Michal WANNER, *Strategie digitálního archivování Archeologického střediska English Heritage* (introduction + translation), „Archivní časopis" 2003, n. 4 – suplementum; Oskar MACEK, Michal WANNER, *Co je pravé a autentické? K problému autenticity a integrity dokumentu v digitální podobě*, „Archivní časopis" 54(2004), n. 1, p. 21-29; Oskar MACEK, Michal WANNER, *Evropské iniciativy v oblasti dlouhodobého uchovávání dokumentu v digitální podobě*, „Archivní časopis" 54(2004), n. 2, p. 81-94; Oskar MACEK, Michal WANNER, *Řešení problematiky uchovávání digitálních dokumentů v Národním archivu USA a v Kongresové knihovně ve Washingtonu*, „Sborník archivních prací" 2005, n. 2, p. 507-552; Oskar MACEK, Pavel MACH, Michal WANNER, *Metody a standardy pro přejímání dokumentů v digitální podobě v dánských archivech*, (together with Pavel Mach), „Archivní časopis" 55(2005) Zvláštní příloha 1, 48 p.; Oskar MACEK, Michal WANNER, *Skartační řízení, nebo automatizovaná sklizeň? K problematice dlouhodobého uchovávání webů*, „Archivní časopis" 55(2005), n. 4, p. 241-255; Oskar MACEK, Michal WANNER, *Metody, standardy a nástroje norského archivnictví využívané v péči o dokumenty v digitální podobě*, „Archivní časopis" 56(2006), N. 1, p. 40-50; Miroslav KUNT, Oskar MACEK, Michal WANNER, *Archivace digitálních dokumentů a zpřístupňování archiválií v archivech Austrálie a Singapuru*, „Archivní časopis" 58(2008), n. 3, p. 185-229,

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project, 101-111

re, we should most cordially thank all those who so willingly and selflessly offered their knowhow to the members of this research team[2].

These activities had a positive impact particularly by setting up a special working team at the National Archives in Prague, identification of issues, their coverage by journals, establishment of cooperation with private sector and other memory institutions addressing similar issues, organisation of regular annual conferences with the topic of digital archiving and electronic records management on the agenda (What will be left after us?, Archives, libraries and museums in digital era) which were subsequently attended also by foreign experts. Eventually, a so far small in size, yet significant and extremely active community of persons interested in the given topic emerged in other Prague archives (Prague City Archives) as well as archives outside Prague (Moravian Provincial Archives in Brno, the Regional State Archives in Zámrsk). To a limited degree, the digital archiving matters were also incorporated in the archiving education (University of Hradec Králové). A major step ahead was also the joining of the Department of Archives Administration and Records Management of the Ministry of the Interior the DLM Forum in 2005, the involvement of Czech archivists in MoReq2 project, the Czech Republic accession to the European Union and the involvement of Czech archivists in cooperation initiatives at the level of EU[3].

All the referred to activities provided a broader framework for the development of a policy concerning the care of digital archival records and efforts to build the National Digital Archives. Approximately ever since 2000, the Czech archiving has been addressing the issues along two main lines, namely the financial and technical. The endeavours of the past years were thus targeted at raising funds for the National Digital Archives project, at drafting its concept, finding suitable premises for individual digital repositories and creation of an experimental workplace for the preservation of digital records. It was only in recent years when also the legislative and methodological issues were raised, i.e. those considered by the authors of the project ever since the beginning to be much more complex than the resolution of issues mentioned earlier.

Thus, at the beginning of the new millennium the Czech Republic, similarly to the majority of other countries, opted for the concept of a single well equipped central digital archives, though it has been foreseen ever since the very beginning that this facility would be made up of several remote archives repositories. For practical, staffing, financial and legislative reasons, the National Digital Archives has always been envisaged to be an integral part of the National Archives in Prague. Though, it has also been anticipated that it would play the role of the national repository.

NDA project

In the first stage, the archivists communicated the existence of this issue and potential solutions thereof to the government. It resulted in the issuance of several Government orders allowing for the development of the National Digital Archives project, namely the Order of the Government of the Czech Republic No 11 of 7 January 2004 on the long-term preservation and availability of digital records and Order No 500 of 10 May 2006 on electronic public procurement which facilitated the development of the National Digital Archives project. The NDA technological project was developed in

2. Oskar MACEK, Michal WANNER, *Czech Republic on the Road to Digital Archives: Research in the Field of Long-term Preservation of Documents in Digital Form*, Comma, 2004, n. 3-4, p. 185-190; Oskar MACEK, Michal WANNER, *Long-term Preservation of Digital Documents – Strategy for Solution in Czech Archives*, In *Archivy v mezinárodním kontextu-Archives in International context, Sborník příspěvků z mezinárodní konference 29. září - 1. října 2004 v Archivním areálu v Praze na Chodovci u příležitosti 50. výročí Státního ústředního archivu v Praze [Proceedings of the International Conference held on 29 September -1 October 2004 in the Archives premises in Prague - Chodovec on the occasion of the 50th anniversary of the State Central Archives in Prague]*, Praha 2004, p. 157-162,

3. Michal WANNER, *Fórum DLM o informačních technologiích*, "Archivní časopis" 49(1999), n. 1, p. 32-37; Michal WANNER, *Evropské iniciativy v oblasti dlouhodobého uchovávání dokumentu v digitální podobě*, "Archivní časopis" 54(2004), n. 2, p. 81-94; Michal WANNER, *DLM Forum Conference in Budapest* (for the first time with the Czech participation), "Archivní časopis" 56(2006), n. 2, p. 130-138; Oskar MACEK, Michal WANNER, *Nové cíle Fóra DLM: Závěry členského zasedání Fóra DLM ve Vídni 4. - 5. dubna 2006*, "Archivní časopis" 56(2006), n. 3, p. 203-208; Michal WANNER, Tomáš Dvořák, *Série archivních konferencí v Berlíně 23.-27. dubna 2007*, "Archivní časopis" 57(2007), n. 3, p. 179-195; Michal WANNER, *Mezi Lisabonem a Lublaní. Závěry dvou členských zasedání Fóra DLM o elektronických dokumentech*, "Archivní časopis" 58(2008), n. 2, p. 129-142; Oskar MACEK, Michal WANNER, *Konference Fóra DLM ve francouzském Toulouse*, "Archivní časopis" 59(2009), n. 1, p. 61-72; Michal WANNER, *Členské zasedání Fóra DLM ve švédském Härnösandu*, "AČ 60" 2010, n. 1, p. 74-85.

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project, 101-111

2007. The winner of the tender was ICZ Company. The project budget amounted to roughly CZK 5 million, i.e. EUR 179,000, $ 227,000. Following the completion of project documentation in February 2008, the project was subject to extensive comments procedure in which a number of representatives of external institutions took part (The Czech Technical University, National Library, etc.). Based on their comments some aspects of the project were substantially supplemented and modified. The referred to record is available in the Czech language on the National Archives website - www.nacr.cz/zpravy/projekt-nda.aspx.
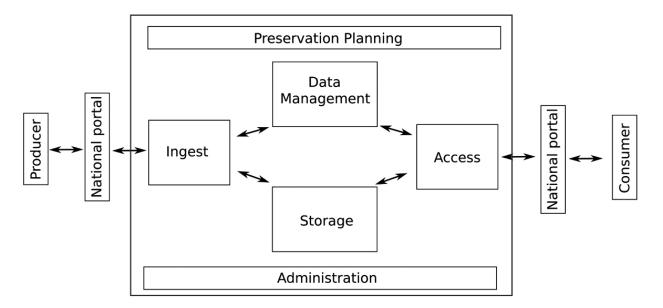
In general, the project described the life cycle of records, preservation technologies, required hardware and software, metadata model, management of archival records, methods for making the records available, testing and validation workplaces, methods for maintaining the authenticity of records. The project also comprises an overview of labour forces and their qualification structure. Finally, also some legislative amendments related to the project implementation were proposed.

Based on the technological project, the National Archives in cooperation with external companies conducted the Feasibility study and the CBA (Cost Benefit Analysis).

Characteristics of the project

It is appropriate to mention at least the key characteristics of this project. The Czech Digital Archives adopts the OAIS model. The project includes a preparatory workplace for the ingest of records, their testing for malicious codes, validity of metadata and output formats, etc. The archives should support a broad spectrum of digital archival records (ERMS and EDMS records, audio, video, databases, raster and vector graphic). Included in the project is the definition of SIP/AIP (based on METS, DC, MoReq2 metadata). Parallel storage in two geographically distant repositories is anticipated. Both the locations are already administered by the National Archives (one is somewhere in Central Bohemia, the other somewhere in South Bohemia). Both the repositories are located under the ground, in premises originally intended for state defence purposes. Parallel storage on two types of media (optical and magnetic) and automated operation are foreseen. The underlying principles of the technical solution are general enough to enable the application of several storage methods. Migration should become the primary maintenance method. A preservation strategy will be developed for each format received by the NDA. If not, such a format will not be received by the NDA. Where necessary, the visualisation and emulation are acceptable. Every 5-10 years an upgrade of technical equipment is scheduled. In mid-term horizon, the project reckoned with the data receipt of 15 TB annually, while now we believe this value will be much higher. The National Digital Archives will serve not only for storage of digital archival records, but it will also be prepared to store the digital copies of conventional archival records. It is assumed that the National Digital Archives will also play the role of the National Portal for making available the digital archival records and digital copies of conventional archival records (web interface). The project budget totals CZK 466 million (EUR 14.3 million, $ 18.2 million). 15 % of project financing should be covered from the state budget, 85% from the Integrated Operational Programme of the EU. The timetable of implementation has delayed as a consequence of organisational problems associated with financing which is why it will have to be revised. The rollout of its routine operation is still scheduled for 2014.

The design of the National Digital Archives builds on the concept of OAIS complemented by the Archives portal.

Archives portal is the only access point to NDA services. Originally, it was included in the concept especially for the reason of increased safety. Its purpose is to prevent the access of users directly to NDA modules. It shall also improve the user's comfort when working with the NDA. The portal will be accessible through the web interface or through web services. The access via the web interface is intended particularly for individual researchers. The information systems of archives and producers will, on the other hand, use the web services.

According to the purpose they serve, there are 3 basic types of operations with records: ingest of records, management of records and retrieval of records.

The process of ingest of records consists of the producer's proposal, decision of the archivist and ingest of the selected records for the NDA. Considering the circumstances, a decision was adopted not to develop a special software for the NDA to be distributed among the producers, but only a single web application. Horizontal introduction of ERMS among producers will allow for automation and alignment of the whole process of selection and receipt of records and for reducing its administrative burden. The list of records to be selected will be generated by ERMS (necessary functionalities are prescribed by the National standard for electronic records management systems).The archivist subsequently marks in the list the records which shall be transferred to the archives and the ERMS performs the transfer of selected records. It is anticipated that the ingest process will be applicable not only for "digital-born" records, but to a certain degree also for both the hybrid and purely analogue records. The chosen concept, however, results in certain complications. The work with the archives portal will necessitate reliable Internet access which might not be always in place at the producer, especially when conventional records are selected. It shall also be noted that the marking of records in the PC is less comfortable and much more time-consuming than marking in a paper list. These shortcomings, though, can be eliminated. For example by using the nowadays popular tablets with mobile Internet access.

The records management in terms of archiving theory and practice comprises particularly the supplementing and modifying of records' metadata, or changes in their position within the archive fond (archive group). It shall be highlighted that both the selection and management will not be performed by the NDA staff but by archivists in the relevant archives competent by subject-matter or venue. It will be possible to provide for the management through the web interface or through the information system of the respective archives. At the time of launch of the NDA operation, the management will most likely be executed via the web interface and only later via the information systems.

The access to records stored in the NDA will again be possible through the web interface or the information system. The web interface will be used primarily by individual researchers. The interface will offer the option to define a query and subsequently will give the results, including the preview of records. The access through information systems will be used by applicants availing of ERMS. The access to records will be subject to access rights defined by the archivists in keeping with the legislation.

Apart from operations related to the ingest, management and provision of access the Archives portal will serve as a single access point to information from the archives. It will collect information on archives, archive fond (archive group), archival records, producers etc from individual archives. It will also offer unified code lists of authorities.

Ingest is a module safeguarding the processing of records submitted to the digital archives. It is the responsibility of the module to ensure that records with malicious code, records lacking metadata and records in inappropriate formats are not submitted in the archives.

The records will be scanned for malicious codes at two intervals with a certain time gap between them. The time gap could be defined for individual formats in line with the potential risk of presence of a malicious code.

Each record submitted to digital archives shall contain at least the basic set of metadata. The list of obligatory metadata will be further specified once the NDA is operational. However, it shall always be defined to what archive fond (archive group) the record belongs, which archives will manage the record and who decided on its storage.

Apart from the requirements mentioned above, the submitted record shall be stored in a suitable format. The initial set of supported formats is in compliance with the legislation and comprises PDF/A, TIFF, JPG, PNG, GIF, MPEG-1, MPEG-2, MP2, MP3 and WAV formats. In the near future, this set should be extended by the format of an e-mail message and XML schema for storage of databases. For each format stored in the NDA, a special storage strategy shall be developed. The strategy includes the method of format identification, methods for checking its validity, format documentation, previewing method and forecast lifecycle of the format. The record which is not in a matching format, or its format is not valid, is rejected by the digital archives. When this occurs, there are three solutions at hand. The first solution suggests that the producer removes the detected deficiencies and re-submits the record to the digital archives. The second solution consists in the suspension of the process of ingest until the storage strategy for the new format is developed. The third solution consists in storing the record in the form of a bit stream. The last alternative, however, is only an extreme and is resorted to only if the above referred to procedures cannot be applied and the archivist insists on storing the record. Nonetheless, in such a case the digital archives provides only for the record storage, not for its readability.

The repository serves for secure storage of AIP. The digital records contrary to the paper ones are more susceptible to damage. For this reason the AIP will be stored multiple times in two repositories located at different sites. It is requested for any change in the record in one repository not to be transferred to the second repository. Therefore, no data replication may occur between repositories.

Major attention is paid to storage media. Despite the claims of manufacturers, the storage life of media cannot be guaranteed since it has not as yet been tried and tested in practice. Even higher risk is inherent in the lifetime of readers. That is why it appears more suitable to store the records on media with tried and tested, though time-limited lifetime, and to gradually replace these media. The mentioned approach will also facilitate more flexible response to the dynamic media development. Moreover, it is also better not to rely on a single type of media only. Each type has its weaknesses which can, however, be eliminated by using multiple types of media. Thus, in the framework of the NDA, two types of media with a different method of information recording will be used in each repository. The first type based on a magnetic recording (disk arrays), the second type based on an optical recording (UDO, or Bluray).

Data management is a data basis of the digital archives. It safeguards support for long-term and safe preservation of records. It preserves information used by the other NDA modules (except for the

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project, 101-111

Archives portal) and descriptive, preservation and structural information on all the AIP packages stored in the NDA. Moreover, it contains information for checking the access rights to data packages, all system code lists and data for the support of process management and other operations.

Access ensures the processing of requirements for the access to stored records and metadata placed through the Archives portal, their evaluation and provision of metadata or records. The required records and metadata are transformed into a DIP package which is disseminated to the user through the Archives portal.

The particular hardware and software used in the National Digital Archives are subject to a tender and therefore are not known as yet. It is required that the supplied system is modular and as much as possible independent of technologies of a single manufacturer. It is apparent that in some segments this requirement will not be easy to meet. Especially the basic information system for records management will have to be supplied and administered by a single supplier. Nevertheless, this system too will have to be modular to a certain degree and allow for the connection of third party's applications. E.g. tests of formats and their validity during the ingest, development of previews, applications for record migration during the access or tests elaborated by the National Digital Archives team. Fully independent of any particular manufacturer should be all the hardware (disk arrays, servers, active network components, etc.).

Legislative process

Essential for successful implementation of the project is also the legislative process. Of major importance is primarily the compliance with basic legislation governing the e-government implementation in the Czech Republic. Aside from the Act on basic registers, it is first and foremost the Act No 300/2008 Coll., on data boxes and authorised conversion. The Act listed the entities that were obliged to establish a data box and starting from 1 July 2009 communicate by means of this tool which guarantees the delivery of a data message. It has been proven by practice that the vision of purely electronic communication will not be feasible for many legislative, technical, economic as well as organisational reasons, the referred to legislation has introduced a hybrid environment for the Czech authorities. Nonetheless, the digital records are a fact of life and have to be reckoned with for long-term preservation.

Significant changes have been introduced by Act No 190/2009 Coll., amending Act No 499/2004 Coll., on archives and records management and amendments to some other laws. The amendment modified primarily the position of the National Archives vis-à-vis the National Digital Archives. To this Act two implementing regulations were adopted – amendment to Decree No 645/2004 Coll., on the application of certain provisions of the Archiving Act (relevant for the description of digital archival records) and a new Decree on the performance of records management (relevant among other things for the definition of output formats from ERMS system).

The Act authorised the National Archives to store and provide access to digital archival records from all the public archives in the CR. This provision does not apply only to the archival records selected by security archives of intelligence services of the Czech Republic. The National Archives was entrusted with the methodological and consultancy function in the field of pre-archival care for digital archival records and digitisation of archival records.

Whereas the preceding legislation preferred the keeping the records in conventional (paper) form, the amendment, on the contrary, in compliance with the implementation of the data boxes system preferred the electronic records management. In this context, a number of provisions were added to the Act concerning e.g. the conversion of records from paper to digital form, but also, rather unusually, validating the authenticity of digital records, or to be more accurate, the records sent via the data boxes system. This was solved by legal presumption of authenticity of electronic records and shifting the burden of proof onto the person questioning the authenticity of the record: "Unless proven otherwise, the digital record is considered to be authentic if signed by a valid advanced electronic signature or furnished with a valid electronic mark of the person who at the time of signature or marking was authorised thereto, of the person responsible for conversion of an analogue record to digital record, or changing the format of a digital record, or the person responsible for the conduct of authori-

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project, 101-111

sed conversion of records, and furnished with a qualified timestamp"[4]. This definition allows for preserving the record authenticity in mid-term horizon only (tens of years), however, it fails to meet the current conditions of e-government.

The changes in records management were so extensive that they called for the replacement of the existing Decree No 646/2004 Coll. by Decree No 191/2009 Coll., on details concerning the performance of records management[5]. This Decree laid down not only the details governing the ingest of records at agencies, their marking, recording, distribution, circulation, execution and signing, use of official stamps, but also dissemination and storage of records, appraisal process or separation of files. The Decree newly stipulated also the output data formats of records for the outputs from the electronic records management system, storage of records in electronic registry and for transfer of digital archival records to digital archives. They are PDF/A formats for text records, TIFF, JPG and PNG formats for static images, MPEG-1 and MPEG-2 formats for video and MP3, MP2 and WAV formats for audio.

Of key importance is also the authorisation issued by the Ministry of the Interior to adopt the national standard for the electronic records management systems.

National Standard for Electronic Records Management Systems

The standard stipulates basic requirements for the functions of these systems and creates unifying parameters for the performance of records management services in relation to the digital records. It ensures that the output records from electronic records management system have standardised metadata enabling further work with the respective records. The national standard was published in the Bulletin of the Ministry of the Interior[6].

The standard follows from the European specification "Model Requirements for the Management of Electronic Records - MoReq2"[7] which was supplemented by the Czech specific requirements. It concerned primarily the requirements for the records register as the most typical form of records management in the Czech Republic.

As to the text of the standard was simplified and clarified. Some requirements were specified and made more stringent in order to keep the Czech national standard for ERMS in line with MoReq2 requirements. In two cases, another entity than stipulated by the MoReq2 was authorised (in compliance with the legislation). The Czech national standard does not contain the chapter Non-functional requirements since the majority opinion is that it very quickly becomes out of date. The MoReq2 standard forced the Czech archivists to translate some of the terms used even in the CR, add them to the terms taken over from MoReq2 and explain them in Chapter 0. The Czech national standard on the other hand accepts certain re-defined modules and elements, not included in MoReq2 standard, the application of which is so common in the Czech Republic that their omission would cause difficulties (records register, reference number, some specific feature of the workflow). Important are the links between the standard and the ingest of digital archival records in the digital archives. Crucial, here, is metadata. Based on the so far gained experience the Czech national standard has two own metadata models with two corresponding XML schemas – a schema for transferring the records and their metadata to the archives (SIP) and XML schema for exchange of records and their metadata between ERMS.

The law imposes upon the public producers, regions and the City of Prague to adjust the electronic records management systems to meet the requirements of the national standard no later than by 1 July 2012. It is obvious from what was said above that the national standard, through the producers,

4. Section 69a para 8 of Act No 499/2004 Coll. on archives and records management and on amendments to some laws, as amended.

5. Decree No 191/2009 Coll., on details governing the performance of records management. It was published in 57/2009 Coll, on page No 1773. The decree and its English translation are accessible on http://www.mvcr.cz/clanek/archivni-legislativa.aspx.

6. Bulletin of the Ministry of the Interior No 76/2010. The National standard for electronic records management systems

7. At the time of elaboration of this paper the Czech translation of MoReq2 specification was available on the website of the Ministry of the Interior of the CR - http://www.mvcr.cz/clanek/archivni-standardy.aspx?q=Y2hudW09Mw%3d%3d.

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the National Digital Archives Project, 101-111

is also intended for business companies active in the development and application of relevant softwares ensuring the conditions for the performance of the electronic records management.

The one year of experience with the application of the standard indicated that a number of requirements included in the MoReq2 standard has proven to be in the Czech Republic not always necessary, or not for all types of producers, and that the number of obligatory requirements for the records management of public authorities (PA) was thus over and above the requirements laid down by the Czech legislation which made the purchases of new or upgrades of the existing ERMS far too expensive. That resulted in a risk of non-compliance with or circumvention of the provisions of the standard. This threat led to the optimisation of the standard thanks to which the risk was mitigated without disrupting the underlying principles of the standard. The optimisation consisted in labelling some of the so far obligatory requirements as recommended. The modifications also covered certain specifications and interpretation of some terms. Several requirements were modified with respect to the establishment of the National Digital Archives. The optimisation of the National Standard of the Electronic Records Management System resulted in considerable reduction of costs of the acquisition of the electronic records management systems born by public authorities. The metadata models and the above referred to XML schemas have not been strongly impacted by the optimisation[8].

New Basic rules for the arrangement of archival records

The standardization of records management has a substantial impact on the method of arrangement the archival records and the structure of finding aids. It is obvious that the so far prevailing classification of archival records by diplomatic category will have to be replaced by models better reflecting the structure of the original registry. The implementation of the national standard for electronic records management systems together with the increased need of exchange of information with abroad (data compatibility, application of internationally recognised EAD and EAC data exchange formats) thus supported and encouraged the existing activities performed promoting the international descriptive standards of ISAD (G), ISAAR (CPF) and ISDIAH that constituted the starting point for wording the new Basic rules for the arrangement of archival records. The specific wording of the rules is, of course, influenced by other factors too - e.g. the necessity to interlink the archives with related memory institutions, links to the state administration and global information systems, reflecting the new information environment created by modern technologies (Internet) with the new application of the principle of provenance, creation of new types of archival records to which the traditional processing method can hardly be applied (digital records, periodical records, various types of technical documentation), the need to align the description (apart from the hierarchical description, especially the contextual description), change in the design of registries and the systems of registries and creation of new types of non-conventional agencies for the definition of which the traditional criteria are not relevant (international companies, multi-level agencies). Of certain importance was also the public pressure insisting on broad availability (of both the digital descriptions and data) on-line or within a network. At the time this article was written, the team of archivists, recruited from a number of Czech archives and headed by dr. Wanner, commenced the works on the new Basic rules.

The implementation of the new Basic rules will call for the development of new software solutions. The basic functionality will be provided for by the Archives portal. The Archives portal is an information system within the NDA enabling remote selection and ingest of digital archival records and their metadata from agencies to the NDA, but it will also facilitate keeping and making available the register of the National Archival Heritage, the data on those archives that take care of the respective archival records, description of archives and cultural and/or scientific institutions, ingest and presentation of digital finding aids, access to digital archival records and digital records created by conversion from analogous archival records.

Development in the NDA-related legislation

In general, the amendment to the Act on archives No 499/2004 Coll. laid down the fundamental prerequisites for the functioning of archival records agencies and archives networks in the digital

---

8. The Bulletin of the Ministry of Interior Chapter 101/2010 (The National Standard for Electronic Records Management Systems). It is downloadable, including the metadata models and XML schemas, from the website of the Ministry of the Interior of the CR - http://www.mvcr.cz/clanek/narodni-standard-pro-elektronicke-systemy-spisove-sluzby.aspx a http://www.mvcr.cz/clanek/vestnik-ministerstva-vnitra-vestnik-ministerstva-vnitra.aspx .

ATLANTI • 21 • 2011

Jiří BERNAS, Michal WANNER: Long-Term Preservation of Digital Records in the Czech Republic and the
National Digital Archives Project, 101-111

era. Unfortunately, at the time of adoption of the amendment to the Act, neither the main features of the National Digital Archives (hereinafter referred to as the NDA) project nor of the technical solution of real-time data repositories etc. were clear. These matters were supposed to be provided for by a separate law on the NDA. The discussions on the subject-matter of this law revealed that another amendment to the Act on archives would be more fitting[9].

This brings us to the last piece of legislation which was only debated at the time of elaboration of this article. Its drafting is closely related to the implementation of the NDA project at the National Archives in Prague. The drafted amendment to Act on archives endeavours to embed in law consistent and comprehensive care of digital archival records and to create conditions for their massive transfer to archives and to define the role of individual archives and the NDA. Thus, the amendment and the subsequent implementing legislation stipulate also the rights and obligations of the digital records′ agencies, specify the evaluation and selection of digital records, details concerning their registration and management of archival records metadata, preservation of documentation on the digital archival records, methods of transfers of digital records to the NDA or other digital archives, new types of finding aids, etc.

Pivotal is the definition of a digital archival record. In the event an archival record is defined as a digital record, it is registered as an archival record and its replica (copy) is stored in the National Archives or the digital archives; the digital record from which this digital archival record is created is not considered to be an archival record. A replica for the purpose of storing the digital archival records means a chain of signs (characteristics) identical with the digital records from which it was created.

The most notable change was made in the concept of digital preservation. While the original concept reckoned only with the existence of the National Digital Archives, the amendment newly includes the establishment of several digital archives. Namely because not only the TV and radio archives, but also by the National Film Archives, the Military History Archives, the Archives of the Academy of Sciences, security archives and others expressed interest in setting up their own workplaces capable of long-term preservation of digital records. That is why the Act lays down the general conditions for the storage of digital archival records. The archives which intends to store this type of records will have to satisfy the given requirements, based on which it will be granted an authorisation to store the digital archival records. If archives are not granted such authorisation, they will have to store the digital records in the National Digital Archives at the National Archives or in any other authorised digital archives.

The NDA is designed to act as a service workplace catering for the receipt and long-term preservation of digital archival records of the National Archives in Prague and provides technology for the storage, permanent readability and integrity of digital archival records selected by the regional state archives and specialised archives. It is anticipated that the NDA will simultaneously manage the digital copies of conventional records created by the archives and will also serve as a national portal for making accessible the digital and digitised archival records through the remote access via the Internet network. The national portal shall also operate the register of agencies, archives, archival fonds and collections and finding aids (see the subchapter on new Basic rules).

The archives within the network of conventional archives will retain the functions related to the evaluation and selection of records, validation of output formats and validity of metadata in keeping with the national standard for records management performed in the electronic form and other standards (validation in line with XML schema), their organisation, development of finding aids, or assisting in making them accessible. These tasks should be implemented by means of specialised softwares through remote access.

Last but not least we must not forget the methodological and consultancy function of the National Digital Archives in the pre-archival care of digital records and in the field of digitisation of conventional archival records. The National Digital Archives conducts scientific and research activities with respect to life cycle of records and digital archival records. In addition, the amendment also mo-

---

9. The document called "Subject-matter of the Act on the National Digital Archives and digital archival records" was discussed by the Government of the CR on 1 March 2010. The Government tasked the Minister of the Interior not to prepare an articulated wording of a new law, but rather to concentrate on the amendment to the existing Act on archives.

difies all the provisions and functions, the implementation of which is impacted by the specific features of the digital environment.

SUMMARY

The article deals with the currently most pressing issues of the Czech archives - long-term preservation of digital records. In the last five years implementation of e-government in the country very advanced. Although the Czech archives paid attention to associated issues for many years, it is still necessary to solve problems in many areas. While so far the Czech archives have been concentrated on research, development strategies and policy approaches, and raising funds for the construction of the National Digital Archives (NDA), currently they work on the implementation of the NDA and other digital archives and legislative arrangements. A strategy to address the existing problems, of the characteristics of the national digital archives and other planned digital archives in the Czech Republic are outlined on the background of description of the e-government in the Czech Republic. Contribution describes also the legislative basis, standardisation of electronic records management and crutial changes in arrangement of archival records.