

Zpráva ze zahraniční služební cesty

Jméno účastníka cesty	PhDr. Ladislav Cubr
Pracoviště – instituce, adresa	Národní knihovna ČR, Klementinum 190, Praha 1, 110 00
Pracoviště – zařazení	Odbor digitální ochrany 2.3
Důvod cesty	-účast na tutoriálu v rámci konference iPRES 2010
Místo – město	Rakousko
Místo – země	Vídeň
Datum (od-do)	18.-20.září 2010
Podrobný časový harmonogram	18.9. cesta vlakem (Praha – Vídeň) 19.9. účast na tutoriálu 20.9. cesta zpět vlakem (Vídeň – Praha)
Spolucestující z NK	
Finanční zajištění	
Cíle cesty	-účast na tutoriálu -konzultace s přítomnými kolegy z jiných institucí
Plnění cílů cesty	-účast na tutoriálu -jednotlivé konzultace s přítomnými kolegy z jiných institucí
Program a další podrobnější informace	Viz níže
Přivezené materiály	
Datum předložení zprávy	
Podpis předkladatele zprávy	

V rámci sedmého ročníku mezinárodní konference „iPRES 2010 - The Seventh International Conference on Preservation of Digital Objects“ věnované problematice digitální archivace (dlouhodobá ochrana digitálních dokumentů), která se konala na Technické univerzitě ve Vídni (Die Technische Universität Wien), jsem navštívil speciální tutoriál věnovaný problematice trvalých identifikátorů digitálních dokumentů. Tento tutoriál nesl

název „Stability of digital resources on the Internet and strategies for persistent identifiers“ (Stabilita digitálních zdrojů na internetu a strategie pro trvalé identifikátory). Akci vedli tři přední evropští odborníci na danou problematiku – Jurgen Kett a Lars G. Svensson z Německé národní knihovny (Deutsche Nationalbibliothek) a Maurizio Lunghi z italské instituce Fondazione Rinascimento Digitale. Všichni tito odborníci se mj. podílejí na provozování národních resolverů (v Německu, resp. v Itálii) pro identifikátor URN:NBN. Identifikátor URN:NBN je v současnosti implementován ve zkušební verzi také v Národní knihovně ČR (včetně vývoje resolveru), tutoriál byl tedy vysoce relevantní s ohledem na aktivity Odboru digitální ochrany Národní knihovny ČR, který se této problematice podrobněji věnuje (podrobněji viz např. „CUBR, Ladislav. Budování důvěryhodného systému trvalé identifikace digitálních dokumentů. *Knihovna – knihovnická revue*. 2010, 21, 11, s. 23-31. ISSN 1801-3252.“).

Níže se blíže věnuji jednotlivým příspěvkům předneseným v rámci tutoriálu.

Jurgen Kett – Introduction on Persistent Identifiers / Use Cases

Jurgen Kett podal ve svém příspěvku obecný a dobře strukturovaný úvod do problematiky trvalých identifikátorů se zaměřením na identifikátory digitálních dokumentů. Zdůraznil základní teze systému trvalé identifikace – identifikátory jsou vždy závislé na kontextu a jejich implementace se odvíjí od požadavků konkrétní uživatelské komunity při současné zvážení jejich reálné využitelnosti v praxi. Pro důvěryhodný systém trvalé identifikace je nezbytné, aby identifikátory byly registrovány třetí stranou, která bývá někdy označována jako registrační autorita. Kett dále srovnal klasický pohled na URI (Uniform Resource Identifier) s pohledem novým (třídy vs. schémata). Od trvalých identifikátorů digitálních dokumentů se na rozdíl od tradičních identifikátorů navíc očekává, že budou využitelné nejen k jednoznačné identifikaci objektu, ale také k přímému získávání zdrojových dokumentů (problematika „resolving“ a „dereferencing“). Pro systém trvalé identifikace digitálních dokumentů je nutné v rámci dané komunity jasně vymezit, co se rozumí pod pojmem „trvalost“.

Po příspěvku následovala debata, ve které zazněly z řad účastníků tutoriálu představy o tom, jaké různé funkce by měl trvalý identifikátor splňovat pro potřeby instituce, která jej implementuje (požadavky čitelnosti identifikátoru člověkem, škálovatelnosti apod.).

Maurizio Lunghi – Persistence is a matter of trustworthiness: a benchmarking model

Italský odborník Maurizio Lunghi se svém příspěvku zaměřil na uvedení do technologického kontextu trvalých identifikátorů. Základními kategoriemi pro posuzování trvalých identifikátorů jsou podle zjištění italských výzkumníků z instituce Fondazione Rinascimento Digitale (Emanuele Bellini, Chiara Cirinna, Maurizio Lunghi) rozsah jejich užití, míra standardizace, možnost využití v resolversch, spolehlivost a trvalá udržitelnost. Italští experti řeší problém, jak srovnávat různé typy trvalých identifikátorů. Pro tyto potřeby vyvíjejí model, na základě kterého by bylo možné trvalé identifikátory v jisté míře objektivně srovnat. Model navrhuje vzájemně poměřovat identifikátory v rámci pěti hlavních kategorií: technologický aspekt, funkcionalita, pravidla, komunitní užití a ekonomické aspekty. Dále italští experti srovnali nejznámější systémy trvalé identifikace užívané v oblasti digitálního dědictví –DOI (Digital Object Identifier), ARK (Archival Resource Key) a URN:NBN (Uniform Resource Name : National Bibliography Number). Dospěli k závěru, že žádný systém nelze označit za nejlepší, každý je třeba hodnotit z hlediska aktuálního kontextu užití a míry konkretizace institucionálních požadavků.

John Kunze - EZID

Tvůrce velmi sofistikovaného systému trvalé identifikace ARK John Kunze z Kalifornské digitální knihovny (California Digital Library) představil zbrusu novou internetovou službu pro trvalé identifikátory nazvanou EZID (ee-zee-eye-dee). Tato služba byla právě v době konání tutoriálu spuštěna na internetové adrese „<http://n2t.net/ezid>“. EZID má sloužit jako nástroj pro vytváření a správu trvalých identifikátorů nejen pro digitální dokumenty, ale i pro jakékoliv jiné entity (fyzické, abstraktní apod.). Nástroj mohou využívat jakékoliv paměťové instituce i další subjekty. EZID slouží i jako resolver a registr trvalých identifikátorů a je provozován Kalifornskou digitální knihovnou. Budoucnost ukáže její reálnou využitelnost a užitečnost této služby.

Lars G. Svensson - Unified Access: Resolving PIs the Europeana Way

Lars G. Svensson představil některé výstupy evropského projektu Europeana Connect. Jedním z nich je vývoj opensourcového resolveru, který má sloužit k vyhledávání jiných resolverů pro různé typy trvalých identifikátorů. V rámci téhož projektu bude provozována také celoevropská služba ERDS (Europeana Resolution Discovery Service), která umožní vyhledat konkrétní resolver pro daný typ trvalého identifikátoru, ať již to bude URN:NBN, DOI, Handle nebo jiný identifikátor. ERDS je nástroj, který pomůže dalšímu rozvoji digitální knihovny Europeana a rovněž bude sloužit libovolné evropské paměťové nebo podobné

instituci jako pomůcka pro identifikaci a vyhledávání digitálních dokumentů. Vývoj softwaru pro resolver se momentálně nachází ve stádiu betaverze (<http://sourceforge.net/projects/metaresolver/>). Služba ERDS by měla být v plném provozu do několika měsíců.

Maurizio Lunghi – Persistence is a matter of trustworthiness: ongoing initiatives

Ve své druhém příspěvku v rámci tutoriálu představil Maurizio Lunghi tři evropské projekty věnované problematice trvalé identifikace digitálních dokumentů a s tím souvisejícím tématům. Prvním projektem je italský projekt jNBN. Jde o projekt zaměřený na implementaci identifikátoru URN:NBN v italském národním kontextu. Architektura italského systému trvalé identifikace digitálních dokumentů navržená v projektu jNBN je založena na distribuované správě trvalých identifikátorů v rámci podřízených národních jmenných prostorů. Důvodem této volby je zajistit vyšší dostupnost digitálních dokumentů v Itálii.

Druhým představeným projektem je PersID, mezinárodní projekt zaměřený na interoperabilitu různých systémů trvalé identifikace digitálních dokumentů. Projekt vychází z předpokladu, že na světě nevznikne žádný jediný globálně dominantní systém trvalé identifikace digitálních dokumentů, a tudíž je důležité přesunout se od otázek hledání ideálního systému k problematice interoperability mezi již existujícími systémy trvalé identifikace. Projektu PersID se účastní několik významných evropských národních knihoven a dalších evropských institucí, včetně národních knihoven Finska, Německa a Švédska. Jedním z cílů projektu PersID je vytvořit evropskou infrastrukturu spojující jednotlivé národní resolvers založené na trvalém identifikačním systému URN:NBN. Prototyp celoevropského resolveru pro tuto infrastrukturu by měl být podle Lunghiho k dispozici do konce roku 2010. Jedním z dalších úkolů projektu je také revize internetových dokumentů RFC 2141 (syntax pro URN) a 3188 (URN:NBN). Projektu se také aktivně účastní známý finský odborník na problematiku trvalých identifikátorů Juha Hakala (původce systému URN:NBN).

Posledním představeným projektem je APARSEN. Jde o projekt, který byl v době konání tutoriálu ve fázi schvalování (měl by být financován Evropskou unií). Na projektu se bude podílet Sdružení pro trvalý přístup k vědeckým záznamům evropských sítí (Alliance For Permanent Access To The Records Of Science In European Networks). Toto sdružení se mj. zaměřuje na problematiku využití trvalých identifikátorů pro označování vědeckých digitálních dokumentů v celoevropském měřítku. Cílem projektu APARSEN bude spojit velmi odlišné typy výzkumných organizací za účelem zajištění trvalé udržitelnosti přístupu k digitálním informacím vytvořeným v různorodých oblastech vědecké činnosti. Dalším cílem

projektu APARSEN je vytvořit Virtuální centrum pro nejlepší odborné znalosti z oblasti digitálních archivace (Virtual Centre of Digital Preservation Excellence).

Závěrem tutoriálu Jurgen Kett shrnul připomínky a zkušenosti účastníků. Podle Ketta nemá velký význam přidělovat trvalé identifikátory digitálním dokumentů v soukromých archivech, protože vytvoření systému trvalé identifikace je velmi náročná činnost a digitální dokumenty v těchto archivech většinou nejsou ukládány s úmyslem dlouhodobé archivace. Kett zdůraznil, že existuje úzká vazba mezi systémy trvalé identifikace digitálních dokumentů a systémy pro trvalou správu a ochranu těchto dokumentů (problematika důvěryhodných digitálních repozitářů). Digitální dokumenty, které jsou označeny trvalým identifikátorem, by měly být nejen dlouhodobě archivovány v digitálních repozitářích, ale rovněž trvale zpřístupňovány v příslušných digitálních knihovnách nebo na jiných místech s využitím internetové sítě. Od trvalého identifikátoru digitálního dokumentu se v komunitě uživatelů totiž očekává, že tento identifikátor bude v rámci internetových služeb dlouhodobě fungovat jako prostředník k přímému zpřístupnění dokumentu, nebo přinejmenším k zpřístupnění metadat k tomuto dokumentu se vážících. Zajištění dlouhodobé funkčnosti identifikátoru digitálního dokumentu je však organizačně a finančně náročná činnost a vyžaduje značné zapojení relevantních institucí, proto je nezbytná opatrnost institucí při jeho zavádění. Jinak hrozí ztráta důvěry uživatelů.