



NACRX004YOLQ

Dne: 21.4.2023 | Č. j.: NA- 1438-5/12-2023
Vaše č. j.:

| Vyřizuje/tel.: Rehtorik / 974 847 111

ZPRÁVA ZE ZAHRANIČNÍ PRACOVNÍ CESTY

MÍSTO:	Ljubljana, Slovinsko
ÚČEL CESTY:	3rd Geospatial Digital Records Preservation Conference , DLM FORUM Members´ Meeting
ÚČASTNÍCI CESTY:	Pavλίna Nimrichtrová, Martin Rehtorik, Zbyšek Stodůlka
ZPRÁVU PODÁVÁ:	Martin Rehtorik
NAVŠTÍVENÉ INSTITUCE:	N/A – pouze konference
TRVÁNÍ CESTY:	7. 5. – 11. 5. 2023
DATUM VYHOTOVENÍ:	22. 6. 2023

SCHVALUJE: **PhDr. Ing. Milan Vojáček, Ph.D.**, ředitel Národního archivu
podepsáno elektronicky

TEXT ZPRÁVY ZE ZAHRANIČNÍ PRACOVNÍ CESTY

ČÁST VŠEOBECNÁ:

- 7. 5. 2023** 06.24 hod – 10:49 hod., vlak Pha -> Vídeň
12.18 hod – 16:44 hod, vlak Vídeň -> Villach
17:23 hod – 19:02 hod vlak Villach -> Lublaň
Příjezd na ubytování v 19.45 hod
- 8. 5. 2023** 09.00–18.00, 3rd Geospatial Digital Records Preservation conference
- 9. 5. 2023** 09.00–17.00, 3rd Geospatial Digital Records Preservation conference
- 10. 5. 2023** 11.00–18:00, DLM FORUM Members´ Meeting
- 11. 05. 2023** 09:00 – 14:00, DLM FORUM Members´ Meeting
16.40 – 18:25 hod., odlet letadlem z Lublaně do Bruselu
21:00 – 22:25 hod., odlet letadlem z Bruselu do Prahy

ČÁST ODBORNÁ:

DLM Forum Lublaň + Geospatial Conference + EAG Archiving by Design

Ve dnech 8. 5. až 9. 5. 2023 proběhla již v pořadí 3. Geospatial Digital Records Preservation Conference. Z celkově velmi zajímavého a bohatého programu lze vyzdvihnout především příspěvky Ann-Kristin Egeland, který se detailněji zabýval formáty vhodnými pro trvalé uchování rasterových datasetů s geodaty a dále pak zejména prezentace o technologii BIM, které se zabývaly jak moderními stavbami od Görana Samuelssona, tak těmi historickými od Franco Niccolucciho. V rámci tohoto programu se konal workshop zaměřený na praktické zkušenosti s archivací geodat pod vedením již zmíněné Ann-Kristin Egeland. Do tohoto workshopu se aktivně zapojil kol. Rechterik se stručnými informacemi o prvních reálných výběrech archiválií z tohoto typu dat v České republice. Z tohoto workshopu vyplynulo, že jediným archivem, který se reálně a pravidelně pokouší trvale ukládat geodata je Dánský národní archiv, ale Národní archiv díky projektu z podpory TAČR a spolupráci s VÚMOP se rychle posouvá vpřed.

Kromě zmíněné konference proběhly ve stejných dnech dvě akce pracovní podskupiny EAG „Archiving by Design“. V pondělí dne 8. 5. workshop pro zájemce z řad původců k vyzkoušení připravené metodiky, za NA se účastnil kol. Rechterik. Následující den, úterý 9. 5., se konalo řádné zasedání pracovní skupiny „Archiving by Design“, které zapracovalo připomínky členských zemí před schválením konečného znění textu pro jednání EAG ve Stockholmu dne 25. 5. Dále byla dohodnuta komunikační strategie pro rok 2023, Z. Stodůlka se vedle finalizace připomínek dále bude podílet na přípravě školení pro školitele z členských zemí, které proběhne pod záštitou Evropské komise v listopadu v nizozemském Haagu.

Od 10. do 11. 5. se uskutečnilo řádné setkání členů skupiny DLM Forum, jehož program obsahoval jak společné části pro všechny účastníky, tak specializované bloky vč. několika praktických workshopů. V rámci sekce 2B během prvního dne prezentoval kol. Rechterik výsledky z proaktivního přístupu Národního Archivu ČR datům veřejnoprávních původců na úrovni celé České republiky ze sociální sítě Twitter v období od 1/2022 do 3/2023, kdy byly prováděny sklízň

vybraných účtů a připravené možnosti a nástroje pro další zpracování a zpřístupnění těchto datasetů.

Z bohatého programu konference lze vyzdvihnout např. příspěvky od Stephana Mackaye z Penwem Ltd. k digitálnímu archivu zdravotnické dokumentace všech obyvatel Norska, který je realizován ve spolupráci se společností Piql. Tento prozatím zcela výjimečný projekt v evropském měřítku využívá open source řešení digitálního archivu v podobě LTP Archivemata, podobně jako Národním archivem provozovaný Národní digitální archiv. K vlastní archivaci zdravotnické dokumentace se používá formát SIARD, který je sice vhodný především k archivaci relačních databází, ale jedná se standardizovaný formát vhodný pro trvalé uložení strukturovaných dat, a to včetně široké množiny nestrukturovaných dat, které jsou v souboru SIARD popsány. Formát, který má ve své podstatě podobu balíčku ZIP tak zajišťuje jejich integritu a autenticitu. Jako přínosné se jeví dále informace funkčních možnosti migrace datových balíčků SIP/AIP/DIP mezi nejčastěji využívanými standardy E-ARK a BagIt. Může to být zajímavá inspirace pro řešitele projektu výměnného AIP, který má podporu ze strany TAČR.

Velkou pozornost vzbudil také skvělý workshop pro účastníky, který byl připraven ze strany odborníků z Lublaňské univerzity. Tématem byly budoucí možnosti využití nástrojů strojového učení/umělé inteligence a data miningu v archivech, a to na příkladech automatizovaného popisu drobných grafik, vytvoření registů k rozsáhlým textům nebo také vytvoření tematických map na základě nejčastějších příjmení ve Slovinsku včetně migračního schématu obyvatel regionů. V následné panelové diskusi, které se vedle slovinských a finských inženýrů účastnil i John Sheridan z britského Národního archivu byl patrná jasná potřeba vytvoření evropské platformy, která by sdílela případy dobré praxe evropských archivů v oblasti využití AI/ML, případně vyhodnotila potřeby v oblasti regulace. Výkonný výbor DLM Forum se zavázal aktivně rozvinout aktivity v této oblasti. IT speciality prosazovaný přístup k neomezenému vytěžování dat uložených v archivech musí být v souladu s právním rámcem a vyhodnocením rizik masivního strojového zpracování archivních „velkých dat/big data“ (např. možná deanonymizace, zpracování osobních/citlivých údajů atd.). Díky v loňském roce ukončenému výzkumnému projektu bezpečnostního výzkumu ke zpracování osobních údajů v archiváliích, který pro vyhodnocení rizik při využívání osobních/citlivých údajů v archivech připravil mj. metodiku posouzení DPIA dle čl. 35 GDPR, má české archivnictví k dispozici moderní základ, na kterém lze dále stavět.