



Zpráva ze zahraniční cesty

13. – 16. 6. 2016, Dublin, Irsko

Helena Kovářiková

Díky podpoře AKVŠ a Ústřední knihovny ČVUT jsem měla možnost se zúčastnit 11. ročníku mezinárodní konference o otevřených repozitářích „OR2016“ <http://or2016.net/>, jejímž tématem bylo „*Illuminating the World*“.

Konference byla zahájena klíčovou přednáškou Laury Czierniewicz, ředitelky „Centre for Innovation in Learning and Teaching at the University of Cape Town (UCT), South Africa“. V přednášce s názvem "***Which knowledge online? Reflections on networked scholarly communication practices in a global age***" se zabývala nerovnostmi v produkci a šíření znalostí, přístupem k informacím a jejich publikování na severní a jižní polokouli. Dokladovala to na řadě příkladů (mapy, grafy) a porovnávala s prostředky, které jsou do publikačního procesu vkládány. V té souvislosti položila otázku, co je z tohoto obrovského množství a v jaké přesnosti vyhledatelné. Zmínila problematiku vyhledávacích nástrojů a jejich algoritmů vyhledávání, které nejsou neutrální, nestranné. Technologicky „digitální svět“ umožňuje otevřenost, ale digitální se ještě nerovná otevřený. Je zřejmý požadavek na přístupnost informací online – tady a teď. Zabývala se problematikou přístupu k informacím „třetího světa“. Přístup k informacím je v řadě případů ovlivněn omezenými prostředky (nemají elektřinu, nemají prostředky na stahování dat, provoz telefonu atp.).

Následně jsem se zúčastnila bloku přednášek „Repositories as tools for change“. Přednáška „***Implementing Orphan Works Legislation at the Digital Repository of Ireland***“ (Marta Bustillo¹, Rebecca Grant², Kathryn Cassidy¹ z ¹Trinity College Dublin, Ireland; ²Digital Repository of Ireland) se týkala zejména digitalizace a zpřístupňování „osiřelých děl“. Tedy děl, u kterých zatím neuplynula doba ochrany autorských práv a nelze autory nalézt. V důsledku toho pak nelze osiřelé dílo užívat, protože k tomu není možné získat povinný souhlas jeho autora. To je překážkou pro zpřístupnění některých děl prostřednictvím digitálních knihoven. Proto byla na evropské úrovni přijata směrnice 2012/28/EU o některých povolených způsobech užití osiřelých děl. Tato díla se musí ukládat do Orphan Works Database, kterou na evropské úrovni spravuje EUIPO. Příspěvek řešil vkládací formulář databáze a zkušenosti s ním.

Následovala prezentace „***Building a repository as an instrument for change***“ kolegyně Saskia van Bergen, Laurents Sesink z Leiden University Libraries, Netherlands, která od roku 2014 spravuje koloniální knihovnické sbírky Royal Tropical Institute (KIT) v Amsterdamu a také všechny sbírky Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV) v Leiden. Mnohé ze sbírek jsou v digitální podobě. Každá ze sbírek je přístupná z jiného místa, každou sbírku spravuje jiné oddělení, často ve spolupráci i s jinými institucemi (KIT, KITLV) nebo dokonce zeměmi (např. ostrovní země spadající pod Nizozemské království – Curaçao). Minulý rok se Univerzitní knihovna v Leidenu rozhodla pro zásadní redesign celé infrastruktury digitálních sbírek. Systémy, které dosud používala (DSpace, DigiTool), byly nahrazeny jiným softwarem – Islandorou.

Poslední přednáškou v tomto bloku byla „***Workflows and key messages to enable Open Research***“ (Jackie Proven, Federica Fina z University of St Andrews, United Kingdom), která informovala o tom,

jak knihovna reflektovala posun k „otevřené vědě“ a jak rozšířila oddělení Digital Research Division a snažila se spojit podporu Research data managementu s otevřeným přístupem. Příspěvek popisoval použití systému Pure (CRIS) pro nahrání a následné uložení všech vědeckých výstupů. Celé workflow zahrnuje import zdrojů, synchronizaci dat, propojení na repozitáře, webové služby nebo portály. Výsledkem je maximální viditelnost a dostupnost vědeckých dat a publikací.

V odpoledním bloku jsem se zúčastnila sekce „Metrics, Assessment, and Impact“, jehož první přednášku s názvem „**Exploring Semantometrics: full text-based research evaluation for open repositories**“ uvedl Petr Knoth ve spoluautorství s D. Herrmannovou z Milton Keynes, United Kingdom. Z důvodu potřeby nového hodnocení vědy a výzkumu vzrůstá poptávka po vývoji nových metrik a metod hodnocení. Příspěvek byl zaměřen na možné využití tzv. sémantometrie. Jedná se o metriku hodnocení výzkumu, která je postavena na vyhodnocování samotného plného textu (nejen toho, jak je dílo citováno). Autoři příspěvku se pokusili vytvořit první sémantometrické měření, založené na podílu, přínosu ve vědecké komunikaci. Metrika vyhodnocuje přenos znalostí z publikace citované, která prezentuje příslušnou práci, na které byl výzkum postaven, k publikaci citující, která představuje oblast použití znalostí. Experiment byl proveden nad plnými texty 12 mil. publikací.

Příspěvek „**Scaling Usage Statistics across Repositories as an OpenAIRE Analytics Service**“ (Dimitris Pierrakos¹, Jochen Schirrwagen², Pedro Príncipe³, Ricardo Saraiva³) prezentoval vytvoření služby, která monitoruje a analyzuje informace o využívání dat z OpenAIRE. Byla použita open source aplikace Piwik, která zobrazuje přehled o návštěvách webových stránek (zobrazení, stažení, kdo přistupuje), byly prezentovány zkušenosti ze zkušebního provozu. Ostrý provoz by měl být spuštěn v roce 2017. Pro větší rozšíření používání této služby bude projekt spolupracovat s NOADs - National Open Access Desks.

Příspěvek „**Gathering and visualizing statistical data at CERN Document Server**“ (Nikolaos Kasiousis, Joe MacMahon, Sebastian Witowski Department of Information Technology, CERN, Geneva, Switzerland) se týkal institucionálního úložiště CERN Document Server, které spravuje 1,5 milionu záznamů. Velkou pozornost tak věnuje statistikám využívání. Všechny údaje o využívání jsou samozřejmě zaznamenány jako logy, kdo je ale umí číst a hodnotit? V příspěvku byl popsán přechod od nástroje pro webovou analytiku – Piwik k nové službě Elasticsearch.

Ve středu dopoledne jsem se zúčastnila bloku přednášek „Identify yourself! Identifiers and attribution“. V příspěvku „**Giving Researchers Credit for their Data**“ (Neil Stephen Jefferies, (University of Oxford, United Kingdom) byl popsán způsob publikování výsledků vědecko-výzkumné práce (one-click process). Vložení dat, souborů, linků na soubory, jejich metadatový popis, přidělení DOI, ORCID ID, se v prvním kroku odehrává v rámci repozitáře. Poté v procesu následuje výběr časopisu, v jehož rámci chce autor publikovat, výběr odpovídající šablony, doplnění ORCID ID spoluautorů, výběr příslušné licence. Poté je celý „balíček“ prostřednictvím SWORD protokolu odeslán do odpovídající platformy vydavatele k publikování. Následně je ORCID profil autora (ů) aktualizován o nový článek. V příspěvku byly zdůrazněny především výhody pro všechny účastníky tohoto procesu.

V prezentaci příspěvku **Don't get irritated, get automated! Accelerating repository workflow with ORCID, Crossref and Datacite** (Elizabeth Krznarich, ORCID, United States of America) byly představeny novinky ohledně identifikátoru a profilu ORCID, automatizace různých procesů a služeb, které s identifikátorem souvisí. Jednou z nich je schopnost automatického updatu ORCID profilu, kterou na podzim 2015 implementovali v systémech Crossref a Datacite. Jakmile je vytvořeno DOI, publikovaná metadata jsou přímo z registrační agentury posunuta do autora profilu ORCID. ORCID tak nyní naplňuje svůj potenciál spolehlivého datového zdroje pro repozitáře. Prozatím bylo díky těmto novinkám do profilů ORCID automaticky dotaženo 800 tis. DOI z Crossref a 150 tis. DOI z Datacite.

V příspěvku **Identify to simplify: integrating identifiers into national infrastructures with the ORCID HUB** (*Andrea Bollini¹, Josh Brown², Susanna Mornati¹, Michele Mennielli¹¹Cineca, Italy; ORCID, US*) byla představena integrace identifikátoru ORCID na národní úrovni v rámci italského konsorcia Cineca. Konsorcium Cineca bylo založeno již v roce 1969 a jeho členy je 70 univerzit, 4 výzkumné instituce a ministerstvo školství. Pro integraci identifikátoru ORCID byla využita open source infrastruktura ORCID HUB, která zjednodušuje integraci více aplikací ORCID. HUB byl původně určen pro národní iniciativy, ale může být využit také institucemi, které potřebují ORCID integrovat do několika různých a komplikovaných aplikací. HUB poskytuje každému vědci místo, ze kterého může autorizovat a monitorovat všechny aplikace, které komunikují s jeho ORCID profilem. Administrátoři a manažeři systémů a aplikací pak mohou masově updatovat data, získávat reporty a mají tak jednoduchý a aktuální přehled o všech ORCID datech.

V dopoledním bloku navazovala sekce „What's in here anyway? Metadata and identifiers“. První přednáška s názvem „**RIOXX: a Modern Metadata Application Profile**“ (*Paul Walk z EDINA, University of Edinburgh, United Kingdom*) prezentovala požadavek Research Councils UK, aby publikace vzniklé za finanční podpory grantové agentury byly k dispozici v rámci institucionálních repozitářů. Tím vzrostly požadavky na dodávání metadat o těchto publikacích v jasné a jednoznačné podobě. Byla provedena analýza stávajících metadatových profilů a rozhodnuto vytvořit profil nový (využití struktury metadat OpenAIRE se jevila jako nevhodná). Jako základ byly použity prvky Dublin Core a NISO OA Metadata and Indicators. Byl tak vytvořen metadatový profil a návod RIOXX. Profil RIOXX byl vyvinut během 2-3 měsíců. V současné chvíli je k dispozici verze RIOXX 2.0. a je implementován v 56 repozitářích ve Velké Británii.

„**COAR Resource Types - a SKOSified Vocabulary for Open Repositories**“ (*Jochen Schirrwagen¹, Imma Subirats², Kathleen Shearer³Bielefeld University Library, Germany; ²Food and Agriculture Organization (FAO) of the United Nations, Italy; ³Confederation of Open Access Repositories (COAR)*). Příspěvek prezentoval činnost pracovní skupiny Controlled Vocabularies for Repository Assets, založené v roce 2014 v rámci COAR, která má za cíl vyvinout sadu řízených slovníků pro metadata použitá v záznamech popisujících výsledky výzkumu. V prvotní fázi byl jako základ použit aplikační profil info: eu-repo, byly provedeny analýzy, statistiky a revize použitých termínů v metadatových záznamech v repozitářích, dále byly porovnány s jinými slovníky. Výsledný řízený slovník odpovídá standardům SKOS, byl implementován v pluginu EPrints, v DSpace University of Minho. Bude též podporován v nadcházejícím Guidelinu OpenAIRE. Další řízené slovníky, které bude pracovní skupina vyvíjet, jsou přístupová práva, verze dokumentů, druhy datace (včetně např. embargo).

Přednáška **Enhancing Citation-only Repositories with Digital Object Identifiers, OpenURL and Memento Links** (*Mark Timothy Hakkarinen z Georgetown University, United States of America*) informovala, jak knihovna buduje index citací z oblasti bioetiky (1974 – 2009), který přidala do repozitáře a v jeho rámci obohatila citace o další doplňující informace DOI, OpenURL a Memento links. To zlepšilo viditelnost této literatury, zvýšily se požadavky na ni a tato kolekce reprezentuje nejvyužívanější zdroj v repozitáři Digital Georgetown.

V odpoledním středním programu jsem se zúčastnila bloku „24x7 session“, v jehož rámci ČVUT prezentovalo své zkušenosti se zaváděním identifikátoru ORCID do prostředí univerzity „**Solving a problem of authority control in DSpace during ORCID implementation**“ (*Helena Kovarikova, Tereza Sorejsova, Lenka Nemeckova, Ivo Prajer, Petr Karel, Rostislav Novak*). Příspěvek navázal na řešení propojení IS ČVUT s repozitářem ČVUT, které bylo prezentované na konferenci OR2014 v Helsinkách. Autoři se v příspěvku nezabývali detailně vlastním workflow přidělování identifikátoru ORCID, ale soustředili se zejména na přínos, který se může projevit v prostředí platformy DSpace, a to je rozšíření autoritního modulu DSpace o možnost vložení dalšího externího identifikátoru (např.

ORCID) k internímu identifikátoru (UUID) jména autora a jeho propojení na univerzitní informační systém. Řešení bylo pozitivně oceněno ze strany představitelů ORCID.

Ve čtvrtek jsem se zúčastnila panelu „Repositories of the Future“. (*Eloy Rodrigues¹, Paul Walk², Kathleen Shearer³* panelisté z ¹*University of Minho, Portugal*; ²*EDINA, University of Edinburgh, UK*; ³*Confederation of Open Access Repositories*). Východiskem panelu byla přednáška „**Next generation repositories: building the repository of the future**“ věnovaná institucionálním repozitářům budoucnosti. Před deseti lety se začaly objevovat první repozitáře, za tu dobu prošly určitým vývojem. Hlavně v prostředí vysokých škol vznikla infrastruktura celé sítě repozitářů. Nerealizují však ještě plně svůj potenciál a jejich funkce jsou hlavně pasivní. Vzniká potřeba diskuse ohledně rolí a funkcí úložišť v příštím desetiletí, aby byly lépe webově orientované a příjemné (web-friendly), aby se v rámci repozitářů vyvíjely funkcionality pro měření kvality obsahu, rozšiřovalo se workflow a funkcionality pro lepší podporu životního cyklu výzkumu, definovaly se a aplikovaly spolehlivé a interoperabilní metriky, dále aby se věnovala pozornost architektuře a technologiím pro příští generace repozitářů. Tyto cíle byly definovány v rámci pracovní skupiny COAR pro rok 2016.

V poslední klíčové přednášce **Making an Open Information Age: Subtitle: Power, freedom and inequality in an age of bits** se *Dr. Rufus Pollock, Founder and President of Open Knowledge* zamýšlel nad otázkami, kde a kdo dnes produkuje informace a komu patří. Nepřístupnost informací znamená moc a bohatství těch, kdo je vlastní. Je zastáncem „otevřeného světa“, volně přístupných informací, hudby, softwaru atd. Propaguje myšlenky platit autorům, ne prostředníkům. Lidé jsou ochotni i v prostředí internetu zaplatit „rozumnou cenu“ za informace. Zabýval se otázkami, jak tuto otevřenost implementovat (ekonomicky – platit za první kopii, politicky – překonat partikulární zájmy). Naznačil možnosti pro „open business model“. To, na čem se dá vydělat, je „omáčka nad tím - ketchup and fries“.