

OPEN REPOSITORIES 2019

Zpráva ze zahraniční cesty
Mgr. Tereza Šorejsová, ÚK ČVUT

Datum a místo konání: 10. - 13. června 2019, Hamburk, Německo

Webová stránka: <http://or2019.net>
<https://www.openrepositories.org/>

Sociální síť: Facebook: <https://www.facebook.com/pg/ORConference/>
Twitter: <https://twitter.com/OpenRepos>
Hashtag: #OpenRepo2019

Videozáznamy: <https://lecture2go.uni-hamburg.de/l2go/-/get/l/5134>

Prezentace: <https://www.conftool.net/or2019/sessions.php>

Prezentace včetně odevzdaných abstraktů budou časem přístupné též v repozitáři Zenodo.

Bude vydáno speciální číslo OA časopisu *Publications* s vybranými články z konference.

Selected Papers from Open Repositories 2018 je již k dispozici u vydavatele MDPI¹.

Otevřené repozitáře je konference zaměřená na problematiku informačních systémů pro umístění, správu a dlouhodobou archivaci dokumentů v režimu otevřeného přístupu.

Hlavní téma 14. ročníku konference **All the user needs** se snažilo soustředit na uživatele, na zviditelnění obsahu repozitářů a v důsledku toho na jejich větší dosah na společnost.

Den před samotnou konferencí se tradičně konaly workshopy rozdělené podle jednotlivých platform, na kterých repozitáře běží, a které jsou vhodné především pro vývojáře a správce repozitářů. Já sama jsem se zúčastnila až hlavního programu.

Keynote

Outcomes over output: a user-centric approach to building successful systems

Úvodní přednáška Jeffa Gothelfa, hvězdy v oblasti UX, byla osvěžující, vtipná, plná obrázků.

User Experience design je na rozdíl od pouhého uživatelského rozhraní o komplexní zkušenosti a interakci uživatele se systémem, produktem nebo webovou stránkou. Jde o soubor technik a způsobů, jak řízeně ovlivňovat, jaký dojem např. návštěva určitého webu v lidech zanechá.

Zkušenosti řečníka s projekty, jež nemají problémy s financováním, se pak ale v reálu těžko přenášejí do světa otevřených repozitářů. Přesto na rozehrání před maratonem konferenčních příspěvků byl Jeff Gothelf asi dobrá volba.

Measuring use and reuse

The BitViews project: a new method for aggregating online usage data, a new route to universal Open Access

V této sekci mne zaujal ekonom Manfredi La Manna z univerzity St Andrews, jeden ze signatářů budapeštské OA iniciativy, který se po 15 letech, kdy problematiku otevřeného přístupu pozoroval jen zpovzdálí, vrátil a zjistil, že se za tu dobu vlastně nic nezměnilo.

¹ https://www.mdpi.com/journal/publications/special_issues/OR2018

„Pokud jde o množství nápadů a myšlenek, knihovnické prostředí je velmi plodné. Ale co se týče konkrétních řešení, je téměř sterilní. Jedním z důvodů je zvláštní postavení knihovníků, jejichž cílem je poskytovat svým akademikům a studentům cenné služby, zodpovídají se však správcům univerzit, kteří je kontrolují. Rozhodovací síla tak není v kompetenci knihovníků, ale spočívá na vedení univerzit a tvůrcích politik.

Komerční vydavatelé jsou malá skupina bohatých, odhodlaných a nemilosrdných lobbistů. Knihovníci jsou legie - rozptýlené, bez zdrojů, sloužící konfliktním zájmům a s velmi malým vlivem na tvůrce politik.“

La Manna nyní přichází s novým nápadem, jak myšlenky OA lépe prosadit, s projektem tzv. BitViews. Jde o technologii blockchain, distribuované decentralizované databáze chránící soukromí prohlížených dat a používané jako účetní kniha kryptoměn. Tato technologie by mohla lépe sledovat, sbírat a počítat jednotlivá shlédnutí (views) vědeckých článků. Data o online využitosti již existují, a však zpeněží je komerční vydavatelé.

Jde o absolutně inovativní bibliometrický nástroj a La Manna věří, že díky BitViews budou mít autoři mnohem silnější motivaci k ukládání svých recenzovaných článků do otevřených repozitářů.

Analyzing Aggregate IR Use Data from RAMP

RAMP The Repository Analytics & Metrics Portal je webová služba, která registrovaným repozitářům nabízí agregované analytiky o využitosti dat z repozitářů, včetně analytických nástrojů Google.

Díky grafickému rozhraní Kibana jsou výstupy snadno dostupné a přehledné.

Díky tomu, že se do projektu registrovalo již přes 50 repozitářů a další přibývají, je možné zkoumat různé odchylky mezi repozitáři. Autoři projektu si například kladou otázku, proč v institucích, které rozdělují obsah repozitáře na kvalifikační práce a ostatní, jsou statistiky využitosti tak odlišné.

Collections and connections for research data

The Fast and the FRDR

Tento příspěvek podle mě výstižně ilustruje určitý posun ve vývoji repozitářů z čistě IT světa do světa knihovnického.

*„Navzdory technologickému pokroku jsou **kvalitní metadata** stále nejlepším způsobem, jak vyhledávání vědeckých výstupů podpořit.“*

FRDR (Federated Research Data Repository) je platforma pro ukládání, správu a vyhledávání výzkumných digitálních dat, za kterou stojí mimo jiné Kanadská asociace vědeckých knihoven. Hlavním cílem platformy je podpora tvůrců a kurátorů dat při plánování, tvorbě a správě popisných metadat, která umožní efektivní vyhledávání a opětovné použití výzkumných údajů v celé řadě oborů.

Výzvou pro pracovní skupinu Data Discovery se stala klíčová slova. Výsledky vyhledávání ve FRDR je možné zpřesnit pomocí faset. Metadata ve FRDR se scházejí z různých zdrojů a na klíčových slovech je patrné, že jde o nekonzistentní data. Použitá klíčová slova byla nejdříve analyzována pomocí nástroje Open Refine a následně namapována na heslář FAST (Faceted Application of Subject Terminology), který spravuje OCLC a je odvozen z Library of Congress Subject Headings.

Repositories and the future of scholarly publishing

A conceptual model for building publishing services on top of a distributed network of repositories

Jde o velmi ambiciózní návrh přeměnit architekturu vědecké komunikace na distribuovaný model o několika vrstvách.

Model PubFAIR se snaží nabídnout systém, který integruje možnosti publikování s možnostmi repozitářů. Tím se kombinuje síla obou světů. Repozitáře mohou v dalších nadstavbových vrstvách nabízet služby s přidanou hodnotou.

Zajímavá je především vrstva distribučních kanálů, autoři rozlišují kanály

- komunitní - analogické k časopisům current contents, tyto kanály jsou spravovány podle oboru samotnou vědeckou komunitou
- institucionální - instituce mohou zřídit a spravovat pro své vědce vlastní kanály s OA obsahem
- individuální - jednotliví uživatelé mohou spravovat obsah podle vlastních zájmů a potřeb, např. podle tématu výzkumu. Uživatelé se pak mohou navzájem sledovat a dostávat upozornění na nové „playlisty“.

Svět tradičních časopisů se posouvá ke službám typu Spotify.

Adding value to repositories through overlay journals

Dalším příspěvkem na téma přidaná hodnota repozitářů byl ohledně tzv. overlay journals, které mohou připomínat referátové časopisy.

Jde o časopisy, jejichž redakce si vybírá z OA článků, které jsou již dostupné online, např. v arXive. Články prochází recenzním řízením, jsou přístupné zdarma a nevyžadují žádné publikační poplatky. Jako praktická ukázka byl představen časopis *Advances in Combinatorics*², jehož první číslo má vyjít toto léto.

„Je potřeba najít nové udržitelné modely publikování. Overlay journals jsou skutečnou a životaschopnou možností, jak snížit náklady na publikování, zatímco repozitáře získají díky recenznímu řízení přidanou hodnotu.“

Open access, neoliberalism, compliance and the law

Do Authors Deposit on Time? Tracking Open Access Policy Compliance

Vzhledem k tomu, že se rapidně zvyšuje počet organizací, které se hlásí k politice otevřeného přístupu, se autoři snažili sledovat, zda přijetí politiky OA má vliv na zvýšení počtu uložených článků v repozitářích, a zda autoři ukládají články do repozitářů včas, tzn. bez zbytečné časové prodlevy.

Důkladná analýza potvrdila, že je skutečně vhodné zahrnout do politiky OA požadavek na časově omezené uložení článku do repozitáře.

Extending DSpace 7

Šlo již o druhou sekci na konferenci věnované dlouho očekávané verzi 7 systému DSpace, jež by měla vyjít do konce roku 2019 a přinese novou, podstatně odlišnou architekturu než předchozí verze.

DSpace 7: open for integrations

Potřeba poskytnout plnohodnotnou Rest-API vrstvu, nad kterou bylo postaveno nové uživatelské rozhraní za použití frameworku Angular (nahradí XMLUI a JSPUI) poskytla příležitost se posunout

² <https://advances-in-combinatorics.scholasticahq.com/>

k novějším technologiím, standardům a osvědčeným postupům, jež REST-API nabízí, a kterých se dřívější verze systému DSpace jen lehce dotkly.

DSpace 7 - The Power of Configurable Entities

Novinkou ve verzi 7 jsou konfigurovatelné entity, volitelné nové typy položek, např. časopisy (vytváří a propojuje název časopisu, ročník, číslo, článek, autora) nebo organizační jednotky (jednoduchá možnost, jak prezentovat, fakulty, katedry, projekty, publikace).

Tyto nové entity byly převzaty z rozšířené verze DSpace, projektu DSpace-CRIS, o němž byl další příspěvek.

Extending DSpace 7: DSpace-CRIS and DSpace-GLAM for empowered repositories and digital libraries

Systém pro správu a ukládání vědecko-výzkumných výsledků byl poprvé představen v roce 2014, ve verzi DSpace 4 a stojí za ním italské konsorcium CINECA a University of Hong Kong. Na rozdíl od jiných (komerčních) systémů pro správu a ukládání vědecko-výzkumných výsledků má DSpace-CRIS jako hlavní součást vlastní repozitář.

Spolu s novou verzí DSpace 7, tak dozná změn i rozšíření DSpace-CRIS, tzn. nové REST-API a uživatelské rozhraní Angular. Díky novým funkcionalitám bude možné dodržet OpenAire guidelines pro systémy CRIS, přidat publikace z externích zdrojů (ORCID, Scopus, WOS, PubMed, ArXiv, ADS atd.), obousměrně integrovat ORCID, nebo synchronizovat DSpace-CRIS s dalšími systémy (účetní, personální apod.).

Dále bylo zmíněno též rozšíření DSpace-GLAM, platformy vyvinuté pro potřeby institucí spravující kulturní dědictví. DSpace-GLAM má rozšířený datový model, který je možné snadno integrovat s prohlížeči obrázků, média servery nebo moduly OCR.

ORCID in repositories

Working together to make ORCID work for repositories: ORCID in repositories task force

Překvapila mě nízká účast zástupců ORCID. Do Hamburgu přijela pouze Liz Krznarich a představila výsledky práce pracovní skupiny pro repozitáře.

Od samého začátku měl ORCID spoustu nápadů a myšlenek, jak by se měl identifikátor ORCID využívat a integrovat. Několik institucí přišlo s vlastními in-house řešeními, nicméně pro velkou většinu repozitářů je integrace obtížná, protože funkce ‚out of the box‘ nejsou k dispozici.

S tím, jak významně narůstá počet založených identifikátorů a počet členských organizací ORCID (včetně konsorcií), ale přichází doba, kdy se hotová řešení, např. plugin v EPrints již objevují. Pomoci tomu má právě i pracovní skupina, která je velmi rozmanitá a tvoří ji 15 členů z 12 zemí. V únoru tohoto roku vydala skupina sadu doporučení³, jaké funkce mají repozitáře podporovat. Doporučení jsou celkem předvídatelná. Jde o sbírání a zobrazování identifikátorů ORCID, výměnu dat s ORCID, zahrnutí ID do strojově čitelných výstupů, možnosti konfigurování a vytváření nástrojů pro reporty.

Na základě těchto doporučení se bude nyní ORCID snažit více zapojit prodejce a vývojáře repozitářů. A pro vývojáře, kteří doporučení implementují, bude ORCID vytvářet dokumentaci a podporu. Díky tomu, že v nadcházející verzi 7 systému DSpace budou implementovány nové typy položek - entity, se bude moct integrace ORCID do tohoto systému rozjet mnohem rychlejším tempem.

³ <https://doi.org/10.23640/07243.7777274>

V sekci byl dále popsán vznik konsorcia ORCID ve Spojeném království (UK ORCID Consortium).
Cultivating ORCID - growing a sustainable national consortium

A byla představena integrace ORCID v Národním digitálním repozitáři Irska, kde pracují s různými metadatovými standardy (Dubline Core, MARC, MODS, EAD) a tak řeší, nerudovskou otázku, kam v těchto formátech s ním. **Using ORCID in the Digital Repository of Ireland**

Tools for measuring and monitoring

V sekci Měření a nástroje pro monitorování bych zmínila 2 přednášky týkající se nástrojů využívající data dvou agregátorů – BASE a CORE.

Data Science Tools for Monitoring the Global Repository Eco-System and its Lines of Evolution

Zástupci univerzity v Bielefeldu, která provozuje vyhledávač BASE, představily možnosti, které BASE bude moci nabídnout ohledně statistik repozitářů, jež pohrabuje. Spojením fasetového vyhledávání, kdy je možné získat statistiky podle země, druhu dokumentu nebo typu licence a externích zdrojů dat (díky API) bude možné vypočítat různé korelace. Externí zdroje dat mohou obsahovat například informace o velikosti populace v zemi, počtu vědců, studentů, publikačních výstupů nebo o financování výzkumu. Vrstva API, která interní a externí data propojí, je sice ve fázi vývoje, ale BASE pak bude moci nabídnout různé výstupy typu: počty diplomových prací v repozitářích podle země, procento OA dokumentů podle repozitáře/země/kontinentu/globálně, procentuální nárůst vědeckých výsledků podle země za rok, seznam typů licencí podle repozitáře/země a mnohé další.

Výstupy je pak možné vizualizovat např. jako mapu světa (pomocí Google Charts) nebo evaluace konkrétního repozitáře či země v podobě grafů (pomocí javascriptové knihovny D3.js).

CORE Analytics Dashboard

CORE Analytics Dashboard je nástroj pro analytické hodnocení výsledků výzkumu univerzit a jejich porovnávání. Využívá data agregovaná v systému CORE a různé metriky zachycené v otevřených zdrojích, včetně čítačů citací (Microsoft Academic Graph) nebo metrik sociálních sítí (Crossref Event Data, Mendeley). Díky API je možné CORE také integrovat s interními systémy univerzit. Data z různých systémů jsou propojována díky DOI. Nástroj je sice zatím ve fázi prototypu⁴, nicméně má ambice expandovat do celého světa. Jde o odpověď na komerční řešení typu SciVal a InCites. Na rozdíl od nich se zaměřuje na ukazatele výkonosti z otevřeně dostupných zdrojů a snaží se tak k hodnocení výzkumu přidat na transparentnosti.

How to be discovered?

Analysing the performance of open access papers discovery tools

Třetí z příspěvků stálice a hvězdy konference Petra Knotha se zabýval analýzou discovery nástrojů pro vyhledávání dokumentů OA. Jak už u Petra a jeho spolupracovníků bývá zvykem, šlo o velice důkladnou a propracovanou analýzu systémů Unpaywall, OA Button, CORE OA Discovery a Kopernio. Analýza byla provedena na náhodném vzorku 1 000 DOI z Crossrefu.

Na základě zjištění, které analýza přinesla, pak v CORE Discovery⁵ vyvinuli možnost na jedno kliknutí získat přístup k bezplatným kopiím výzkumných prací (pokud existují), kdykoliv uživatel narazí na nutnost platby. Jde o rozšíření, které je možné si instalovat do internetového prohlížeče a o plugin,

⁴ <http://bit.ly/core-analytics-dashboard>

⁵ <https://core.ac.uk/services/discovery/>

který se instaluje do repozitáře. V repozitáři je tak záznam bez plného textu obohacen o link na volně dostupnou kopii.

Závěr

Na závěr bych ještě zmínila, že velká část příspěvků se snaží naplňovat doporučení COAR Next Generation Repositories recommendations⁶. Jde o doporučení, která v roce 2017 vypracovala pracovní skupina při Konfederaci repozitářů s otevřeným přístupem a mimo jiné zahrnují doporučení přijmout nové technologie, standardy a protokoly, které pomohou se repozitářům lépe integrovat ve webovém prostředí. Umožní jim tak hrát důležitější roli v prostředí vědecké komunikace.

Líbilo se mi, že na konferenci se člověk dozvěděl spoustu praktických informací, o jejichž implementaci v rámci své univerzity je možné ihned začít uvažovat. Zmínila bych především registraci v portálu RAMP, nebo plugin a rozšíření webového prohlížeče CORE Discovery.

Děkuji Asociaci knihoven vysokých škol České republiky, že díky její finanční podpoře jsem se mohla konference OR2019 zúčastnit.

⁶ <http://ngr.coar-repositories.org/>