

The 7th International Conference on Open Repositories (OR2012)

9.-13. července 2012, Edinburgh, Velká Británie

Zpráva ze zahraniční cesty
Mgr. Pavla Rygelová (ÚK VŠB-TUO)

Letošní sedmý ročník konference nesl podtitul *Open Services for Open Content: Local in for Global out*, v němž se odráželo současné směřování k otevřenosti a obohacování obsahu repozitářů o přidané služby a budování potřebné infrastruktury.

Myslet v měřítku globální sítě

Zahajovací přednáška *Network Enabled Research: The possibilities, the path, and the role of repositories* Camerona Neylona velmi výstižně charakterizovala ideové směřování těch, kteří po celá léta konferenci pořádají a těch, kteří na ní aktivně vystupují. Nosnou myšlenkou konference je otevřenost, na níž jsou např. postaveny softwary pro otevřené repozitáře a všechny nadstavbové služby pro otevřené sdílení a opětovné využívání obsahu repozitářů. Cameron Neylon pojmenoval v podstatě to, co všichni intuitivně cítíme a v praxi i zažíváme. Shrnuje potenciál, jaký přinesl internet, ale který ještě nedokážeme naplno využívat. Jeho prezentací se neslo jedno spojující heslo: maximalizovat měřítko a konektivitu a minimalizovat překážky (maximise scale and connectivity, minimise friction). Na dvou příkladech z oblasti open science (Polymath Project¹ a Galaxy Zoo²) ukázal, jak síť kvalitativně mění naši kapacitu (výzkumnou, informační). Jednou z výzev dnešní vědecké komunikace je změna ověřování (filtrování) kvality vědeckých výstupů. Tradiční proces peer-review nemůže všem vyhovovat. Filtrování kvality by mělo být dynamické, personalizované, přizpůsobitelné podle potřeb uživatele. Využití potenciálu, který nabízí internet, závisí na naší schopnosti uvažovat v globálním síťovém měřítku, tj. na schopnosti sdílet a maximálně zpřístupnit informace všem bez překážek, neboť sami nemůžeme odhadnout všechny možnosti, jak může být sdílený obsah využitelný.

Sdílené služby repozitářů a infrastruktura

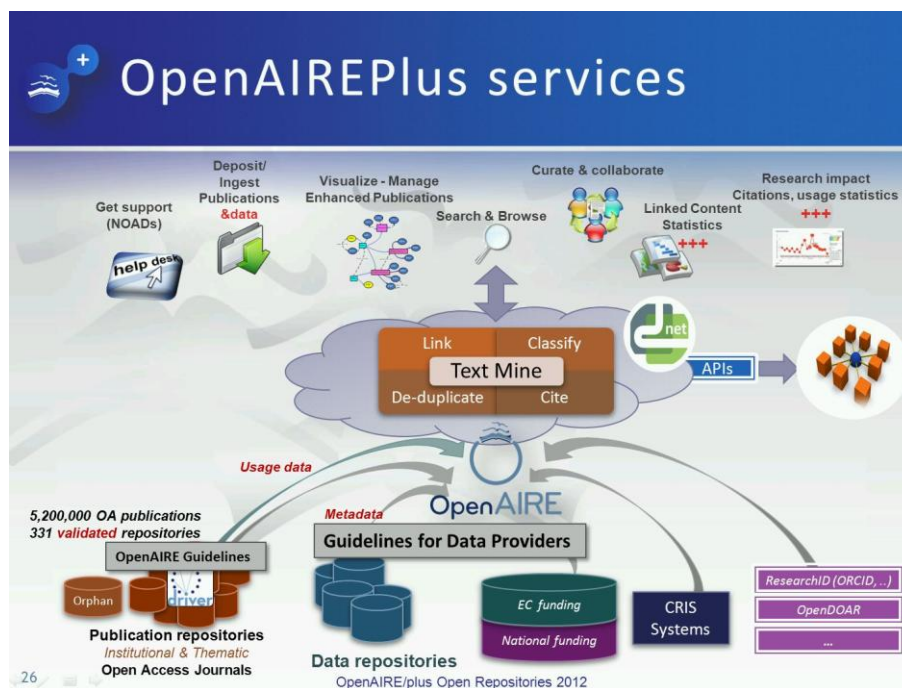
Hlavní téma konference - služby pro otevřený obsah - se nejvíce odráželo v sekci *Shared Repository Service and Infrastructure*.

Příkladem pro budování infrastruktury byl příspěvek o evropském projektu OpenAIRE (přednesla Nejla Rettberg a Birgit Schmidt z univerzity v Göttingenu). Přednášející shrnuly základní informace o projektu OpenAIRE a o navazujícím projektu OpenAIREPlus. Vyzdvihly hlavní aktivity obou projektů a zastavily se blíže u výzev, které si řešitelé obou projektů vymezili. Hlavními aktivitami projektu OpenAIRE je vybudování technické infrastruktury, partnerské sítě všech zapojených států a zprovoznění služeb pro koncové uživatele. Výsledky projektu OpenAIRE stručně v číslech: 27 tisíc publikací z 5ti tisíc projektů (10 tisíc v režimu

¹ <http://polymathprojects.org/>

² <http://www.galaxyzoo.org/>

open access) harvestovaných z 94 repozitářů a časopisů. Na tomto poměrně velkém úspěchu bylo možné postavit navazující projekt OpenAIREPlus, který se zaměřuje na propojení otevřených publikací s datovými sadami a na propojení s informacemi o financování projektů.



Obr. 1 Schéma služeb projektu OpenAIREPlus

Schéma na obrázku 1 názorně ukazuje služby, které by měl integrovat projekt OpenAIREPlus. Vedle služeb, které již fungují (ukládání publikací do repozitářů nebo do sirotčího repozitáře, vyhledávání publikací na portálu OpenAIRE nebo helpdesk prostřednictvím kontaktních osob - tzv. NOADs), budou poskytnuta pravidla pro poskytovatele dat pro sklizení metadat datových sad a informací ze systémů financování Evropské komise a národních systémů, apod. Technická infrastruktura OpenAIRE nabídne deduplikaci a klasifikaci záznamů, textové vytěžování a statistiky využívání.

(Projekt OpenAIRE dostal prostor také v příspěvku *How to make your DSpace repository OpenAIRE compliant*, který přednesl Eloy Rodrigues v druhé části konference, v sekci určené uživatelům DSpace. Z dotazu moderátora vyplynulo, že z přítomných účastníků této sekce (kterých, nutno podotknout, po čtyřdenním přednáškovém maratonu značně ubylo) má repozitář přizpůsobený pravidlům OpenAIRE pouze VŠB-TUO.)

Příspěvek o portugalském projektu RCAAP (*Enhancing repositories and their value: RCAAP repository services*) byl ukázkou přidáných služeb.

Projekt RCCAP (Repositório Científico de Acesso Aberto de Portugal = Open Access Scientific Repository from Portugal)³ začal v roce 2008 a měl dva základní směry. První byl zaměřen na

³ <http://www.rcaap.pt/>

propagaci a rozvoj repozitářů, druhý se orientoval na budování portálu RCCAP pro sběr, indexování a agregování obsahu portugalských otevřených repozitářů.

V další fázi se projekt věnuje rozvoji přidaných služeb pro portugalskou vědeckou komunitu. Portál RCCAP od roku 2010 sklízí a indexuje také brazilské repozitáře. Prohledávat lze přes 40 tisíc dokumentů. Výsledky vyhledávání lze zpřesňovat pomocí filtrů. Každý záznam může uživatel obohatit o související dokumenty nebo jej sdílet na sociálních sítích.

Portál RCCAP implementoval v roce 2008 pro zajištění kvalitních metadat pravidla DRIVER, která jsou povinná pro všechny agregované zdroje. Portál nabízí administrátorům repozitářů validátor, který upozorní na chyby v metadatech, pokud se nějaké najdou.

Druhým významným okruhem služeb projektu RCCAP je SARI (Institutional Repositories Hosting Service). SARI je služba postavená na systému DSpace a je volně dostupná pro instituce, které chtějí provozovat repozitář. SARI momentálně využívá 26 institucí. Projektový tým RCAAP zajišťuje řízení infrastruktury SARI (hardware, hostování, operační systém, aplikace, bezpečnostní požadavky), administraci repozitáře si zajišťují správci jednotlivých repozitářů. RCCAP pořádá školení pro administrátory. SARI poskytuje několik addons. Ke klíčovému rozšířením patří:

- OAextended: pro přizpůsobení repozitáře pravidlům DRIVER a OpenAIRE
- Request Copy: pro zaslání požadavku autorovi o dokument, který není volně dostupný
- Minho Stats: nástroj pro administrátory ke sledování statistik využívání dokumentů
- Degóis addon: propojení s národním systémem vědeckých profilů (životopisů)
- OpenAIRE authority control addon: umožňuje vkládání ID projektů do vkládacích formulářů
- Sharing bar addon: umožňuje sdílení dokumentů na sociálních sítích a v bibliografických manažerech (Endnote, Mandelley nebo Bibtex) a umožňuje export záznamu do CV vědce do systému Degóis Curricula.

K dalším službám projektu RCCAP patří „common repository“ – společný repozitář pro organizace bez vlastního repozitáře; SCEUR (Centralized Repository Usage Statistics) – centralizovaná služba pro administrátory repozitářů ke sledování statistik využívání repozitářů; SARC (Scientific Journal Hosting Service) – služba pro hostování a řízení vědeckého časopisu postavená na platformě Open Journal System.

Úspěšnost projektu RCCAP dokazují čísla: v roce 2007 bylo v Portugalsku 10 repozitářů,

v roce 2011 již 34; výrazně narůstá počet dokumentů indexovaných portálem RCCAP. Stoupl také počet politik povinného otevřeného přístupu ze dvou v roce 2007 na 12 v roce 2012.

K úspěchům patří funkční interoperabilita díky dodržování pravidel DRIVER a OpenAIRE, vybudování silné komunity nabízející podpůrný helpdesk, školení a podporu myšlenkám OA a v neposlední řadě integrace repozitářů jako pevné součásti výzkumného systému.

Evolution of Repositories



Universidade do Minho
Serviços de Documentação

Evolution of Repositories in Portugal



Universidade do Minho Serviços de Documentação

Obr. 2 Nárůst repozitářů v Portugalsku

Rozšířený obsah

V sekci Augmented content nejvýrazněji vystihl pojem rozšířený obsah příspěvek *Beyond Bibliographic Metadata: Augmenting the HKU IR* Davida Palmera z univerzity v Hong Kongu. Univerzita v Hong Kongu se rozhodla použít institucionální repozitář HKU Scholars Hub⁴ pro zviditelnění výsledků své výzkumu a svých vědců. Od výrazného zviditelnění výsledků očekávala zvýšení možnosti k získání kontaktů a spolupráce na výzkumných úkolech s jinými univerzitami.

Nejdůležitějším úkolem bylo rozšířit a zkvalitnit bibliografické údaje, získaná z univerzitní databáze *Research Output System*, která obsahovala nekvalitní a sporá metadata. Pro zkvalitnění metadat bylo použito několik postupů – např. doplnění dat pomocí DOI a ISBN z databází CrossRef a OCLC nebo sjednocení jmenných autorit.

V repozitáři byly vytvořeny profilové stránky přibližně 1400 vědců. V profilech se zobrazuje: seznam publikací s odkazy na plné texty; další výsledky (získaná ocenění, vedení závěrečných prací apod.); řešené granty (v profilu se zobrazují ID projektů a publikované výsledky vzniklé z projektu); externí metriky (data ze Scopusu) a interní metriky (interní statistiky systému DSpace).

⁴ <http://hub.hku.hk/>

Pro jednoznačnou identifikaci vědců je v profilech využito ResearcherID. (Thomson Reuters nabízí API pro automatické vytvoření ResearcherID pro každého autora z univerzity) Repozitář shromažďuje kromě publikovaných výsledků také závěrečné práce, granty a patenty. Obohacený obsah (obohacená metadata) rozšiřuje funkčnost repozitáře tak, že se stává skutečným komunikačním uzlem univerzity. Repozitář nabízí opětovné využití uloženého obsahu např. hypertextovými vizualizacemi osobních vazeb na základě autorství, stejných klíčových slov v publikacích, výzkumných zájmů apod. Repozitář nabízí webovou službu pro vytvoření profilů vědců na webových stránkách jejich pracovišť s využitím metadat uložených v repozitáři. Z interních statistik je zřejmé, že zviditelnění výzkumu a vědců hongkongské univerzity se podařilo.

DSpace a CRIS (Current Research Information System)

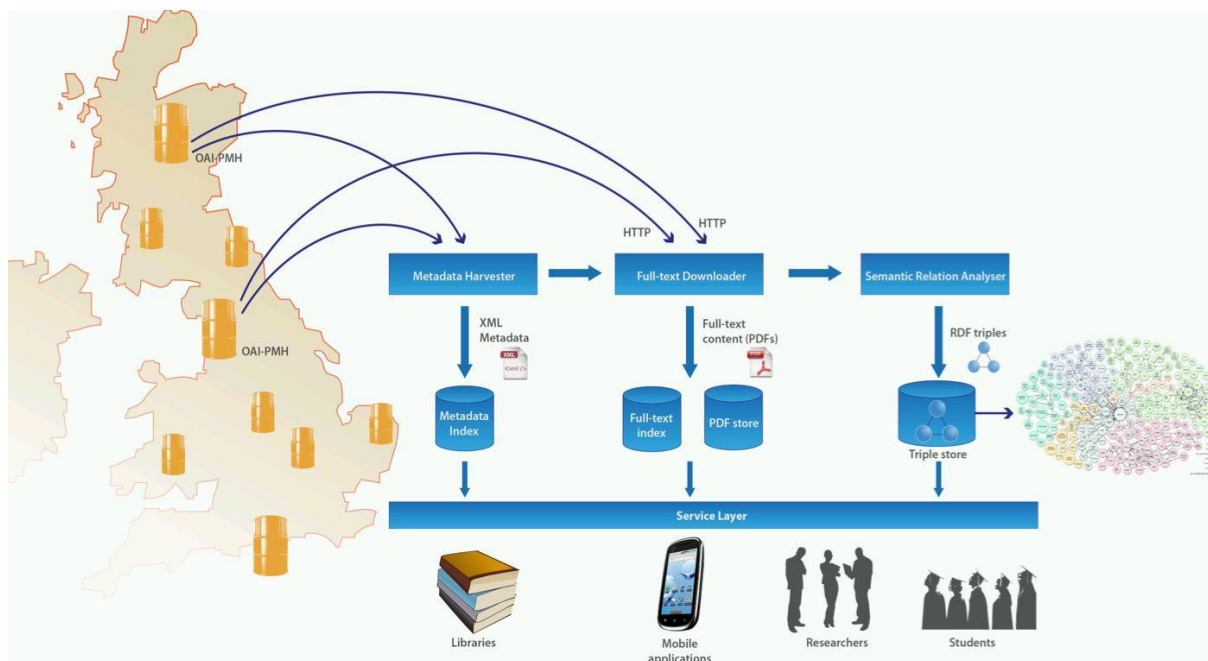
Na rozšíření repozitáře HKU Hub se podílelo italské konsorcium univerzit CILEA. Na konferenci se prezentovali s posterem *Moving DSpace to a fully-featured CRIS System*. CILEA nabízí modul DSpace-CRIS⁵ jako open source. Také další postery představovaly různá řešení integrace CRIS a institucionálního repozitáře, např. *Integrating an institutional CRIS with an Open Access Institutional Repository* z University of London nebo projekt španělské rady pro výzkum (Spanish National Research Council) *CSIC Bridge: Linking DIGITAL.CSIC to institutional CRIS*.

Projekt CORE

Developerské soutěže (Developer Challenge) a také posterové sekce se účastnil projekt CORE⁶ z Knowledge Media Institute, Open University (KMI). Projekt prezentoval Petr Knoth, doktorand z ČR, zaměstnaný v KMI. Projekt CORE (COncnecting REpositories) sklízí metadata a plné texty z více než 200 repozitářů. Nad sklizeným obsahem buduje další služby. Základním navigačním prostředím projektu je webový portál včetně základní funkce vyhledávání dokumentů. Portál nabízí sémantické vyhledávání odkazující na obsahově podobné dokumenty. Další službou je vystavování informací podle cílových skupin uživatelů – knihovnám, vědcům a studentům, mobilním zařízením. Základními aplikacemi projektu CORE je zmíněný portál, dále CORE API pro obohacení informací obsažených ve vlastním repozitáři, CORE Mobile, Repository Analytics – nástroj pro administrátory repozitářů, CORE Plugin nabízející doporučení souvisejících dokumentů uložených v jiných repozitářích na základě obsahové shody. Portál sklízí již také repozitář DSpace VŠB-TUO.

⁵ <https://github.com/CILEA/dspace-cris>

⁶ <http://core.kmi.open.ac.uk/search>



Obr. 3 Schéma služeb projektu CORE

Závěr

Konference tradičně přinesla opravdu velmi nabitý program a mnoho příležitostí i k neformálnímu setkání s ostatními účastníky konference. Účast na konferenci je užitečná pro sledování vývoje repozitářů (inspirace pro rozvoj služeb vlastního repozitáře) a byla také užitečná jako zdroj nápadů pro pořádání příštího ročníku setkání českých uživatelů systému DSpace.

V Ostravě 8. srpna 2012