

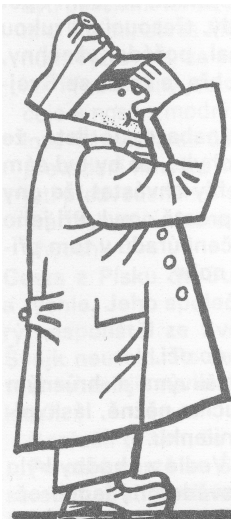
Vlastíkova ženevská anabase (OA18)

Vlastimil Krejčíř, krejcir@ics.muni.cz

Ústav výpočetní techniky, Masarykova univerzita, Brno

30. – 31. 10. 2012

Aby nedošlo k nějaký mejlce ...



CERN Workshop on Innovations in Scholarly Communication (OAI8)

(CERN/Ženeva) Workshop zaměřený na nové směry a technologie v oblasti webu, sdílení informací, repozitářů, otevřeného přístupu etc.

V rámci workshopu proběhl i jednodenní DSpace User Group Meeting.

Má účast:

- příspěvek na DSUG Meeting (FFdigi a DML-CZ),
- dva postery (DML-CZ a EuDML).

Letošní OAI8:

- protokol ResourceSync
- sémantický web
- metriky a věda
- data ve výzkumu
- Open Access
 - OA v humanitních vědách
 - Gold OA (diskuse)

- zahřívací kolo pro OR 2013
- cca 25 účastníků

Z programu:

- ORCID identifikátor
- implementace protokolu SWORDv2
- ResourceSync v DSpace
- novinky v DSpace 3.x
- *use case* přednášky
- diskuse k budoucnosti DSpace

- synchronizace zdrojů na internetu
- nástupce OAI-ORE
- autorem skupina kolem Herberta van de Sompela
- obecný protokol umožňující synchronizaci jakýchkoli zdrojů
- modularita - není nutné použít všechny funkce
- poslední draft protokolu ze srpna 2013
- předpokládaná finální verze na přelomu roku

- možnost popsat obsah protokolu
- „zipování“ dat
- rozdílové aktualizace, historie změn
- mirroring
- linkování na starší verze obsahu
- discovery vlastností zdroje
- selektivní synchronizace
- posílání notifikací o změnách
- technicky XML na bázi Sitemap a Siteindex rozšířené o vlastní elementy

- původní koncept sémantického webu zastaral
- dodávání sémantických informací automatizovaně
 - v současnosti už lze s poměrně zajímavými výsledky
 - např. projekt VIVO
- OpenAnnotations
- zajímavá přednáška o klasickém problému webu - neplatných odkazech
 - přednášející prof. Thompson (spolupracovník T. Bernes-Lee)

- impakt faktor a jeho problémy
- zajímavý koncept měření úspěšnosti vědce z dat na sociálních sítích
 - projekt *altmetrics.org*
 - ne každý je na Facebooku/Twitteru
- peer-review v OA
 - neprůhlednost procesu u predátorů (publikují hodně ⇒ klesá kvalita peer-review)
 - řešení = plná transparentnost
 - *OA tool*: měření transparentnosti - korelace s kvalitou časopisu?

- sémantické hledání v PubMedu
 - „heart attack“ = „myocardical disfunction“
 - před dotaz přidat [Mesh]
 - detekce klíčových slov v abstraktech
- detekce klíčových informací a ideí z vědeckého článku
 - cílem je získat výsledky výzkumu
 - detekce na základě „typických“ frází

- problematika uchovávání vědeckých dat a jejich analýza
- doporučené politiky pro uchovávání
- velká data v CERNu
 - data z LHC
 - obrovské datové toky (10 PB/s)
 - po filtraci cca 6 GB/s
 - po 3 letech provozu cca 100 PB
 - disková pole, redundance
 - přes lokální síť asi 3 EB dat
 - spolupráce manažerů, IT a knihovníků

- standard pro vytváření sémantických vazeb mezi objekty na webu
- obohacování informací – strojově zpracovatelné
- technicky JSON (JavaScript) vytvářející vazby RDF
- tagging, poznámky na okraji apod.
- podle diskuse se využití stále hledá

Open Access v oblasti HSS

- zdálo se, že trendem je spíše Gold OA
- kritika současného majoritního modelu publikování (zejména nevhodný pro oblast HSS)
- britský model (RCUK) - podpora Gold OA, povinnost po 6 měsících (12 pro HSS) dát dílo volně k dispozici
- striktní vyžadování Gold OA – nedostatek peněz
- vydavatelé v oblasti HSS odmítají
- model *OpenBooks*
 - včetně udržitelného modelu financování
 - SpringerOpen
 - má někdo zájem?
- open data ve World Bank

Gold OA (panelová diskuse)

Co je potřeba udělat pro podporu Gold OA?

- získat uživatele (ukázat jim proč je GOA lepší)
- problém financování (velká propast)
- hromadné smlouvy s vydavateli (levnější)
- chybí infrastruktura (SW, znalosti, podpora) – nebudovat znovu, spíše využít stávající
- netrvat na konkrétních licencích, nestavět překážky
- podporovat GOA vydavatele
- jít do toho naplno

