

MASARYKOVA UNIVERZITA

FAKULTA INFORMATIKY



Projekt DRIVER - Networking European Scientific Repositories

<http://www.driver-repository.eu/>

Karol Kubanda

N-IN BIT [sem 3, roč 2]

2. December 2009

Úvod

S rastúcou cenou recenzovaných vedeckých časopisov a zároveň s rastúcim počtom samotných vedeckých článkov sa narušila rovnováha v oblasti vedeckého publikovania, čo spôsobilo nespokojnosť najmä vo vedeckej komunite a následné hľadanie nových možností ako flexibilne sprístupniť výsledky výskumov a umožniť tak lepšiu komunikáciu odborníkov vo vede a výskume. V roku 2002 vzišla „Budapest Open Access Initiative“ z podnetu Open Society Institute, ktorá začala presadzovať myšlienky voľne dostupnej, kvalitnej vedeckej literatúry s neobmedzeným využitím. Takzvanou „Zelenou cestou“ k dosiahnutiu Open Access je auto-archivovanie v otvorených repozitároch, čo viedlo k zakladaniu takýchto digitálnych repozitárov na jednotlivých vedeckých inštitúciách. Podľa projektu OpenDOAR, ktorý registruje Open Access repozitáre, sa ich počet v roku 2007 vyšplhal celosvetovo na tisíc, z toho približne polovica sa nachádzala v Európe.^[4] Vznikla tak potreba spojiť tieto izolované repozitáre do jedného logického celku a uľahčiť tak prístup k ich obsahu. Európska komisia si uvedomila prínos takéhoto logického celku pre vedu, výskum a vzdelávanie v celej Európe a finančne podporila projekt DRIVER.

Projekt DRIVER

Digital Repositories Infrastructure Vision for European Research, skrátene tiež označovaný ako DRIVER, je projekt snažiaci sa vo viacerých na seba nadväzujúcich fázach vytvoriť kompaktnú infraštruktúru vedeckých digitálnych repozitárov z celej Európy a poskytnúť tak prepracovanú funkcionálnu ako pre samotných vedcov tak aj pre bežnú verejnosť. Vytvorenie progresívnej infraštruktúry DRIVER pre uloženie a sprístupnenie budúcich znalostí v oblasti európskej vedy a výskumu sa má stať doplnkom k sieti GÉANT2, ktorá je už dnes úspešnou infraštruktúrou pre zdieľanie výpočtových zdrojov, úložísk dát ako aj samotný prenos dát v akademickú a vedeckú obci. Na rozdiel od siete GÉANT2 má však DRIVER poskytovať prístup k vedeckým výstupom, akými sú vedecké a technické správy, pracovné podklady, pre-printy, vedecké články ako aj pôvodné výskumné dáta. Víziou projektu je zabezpečiť úspešnú spoluprácu medzi dátovou sieťou a úložiskom znalostí v podobe digitálnych repozitárov a vytvoriť tak celistvú platformu pre vedu, výskum a vzdelávanie v Európe. V dobe písania tejto práce sa uskutočnili prvé dve fázy projektu DRIVER.

Prvá fáza projektu DRIVER

Prvá fáza projektu DRIVER trvala osemnásť mesiacov, začala v júni roku 2006 a bola financovaná najmä grantom z kapitoly „Research Networking Testbeds“ agendy 6. rámcového programu Európskej Únie. Na projekte spolupracovalo desať partnerov z ôsmich krajín Európskej Únie, z ktorých mnohí sú dobre známi účasťou na budovaní národných i nadnárodných repozitárových riešení (CNRS, SHERPA, SURF, UKOLN, University of Ghent) a zvyšní majú skúsenosti z budovaním potrebnej technickej infraštruktúry a služieb (University of Athens, University of Bielefeld, CNR, ICM, University of Goettingen). Prvá fáza sa zamerala na vývoj testovacej verzie s dôrazom na päť hlavných cieľov:^[3]

1. Zostaviť a vytvoriť virtuálnu sieť európskeho rozsahu z existujúcich repozitárov jednotlivých inštitúcií Holandska, Veľkej Británie, Nemecka, Francúzska a Belgicka.
2. Určiť a implementovať technickú infraštruktúru vzhľadom na aktuálny stav technológie, ktorá by spravovala fyzicky distribuované repozitáre ako jeden rozsiahly celistvý virtuálny zdroj obsahu.
3. Určiť a implementovať niekoľko základných služieb pre užívateľov.

4. Identifikovať, vytvoriť a uviesť do života relevantnú množinu štandardov.
5. Pripraviť nasledovné rozšírenie a zlepšovanie infraštruktúry digitálnych repozitárov naprieč Európou a zaistiť najširšiu možnú angažovanosť užívateľov a využitie infraštruktúry.

Počas samotného projektu DRIVER sa vytvorilo množstvo strategických a koordinovaných štúdií na tému digitálnych repozitárov a s nimi súvisiacou problematikou. Jednou z nich bola aj publikácia „DRIVER's Guide to Repositories“, ktorej úlohou je motivovať a propagovať ďalšie vytváranie a vývoj repozitárov, ako aj ich následné spájanie pomocou sietí. Boli vytvorené smernice projektu DRIVER, ktoré umožnili harmonizáciu a validáciu dát s cieľovou interoperabilitou na sémantickej aj syntaktickej úrovni. Prvá fáza skončila úspešne v novembri 2007.

Infraštruktúra projektu DRIVER

Systém DRIVER je implementovaný na otvorenej servisne orientovanej softvérovej architektúre logicky organizovanej vo funkčných oblastiach, ako sú zobrazené na obrázku 1. Táto architektúra zaisťuje rozšíriteľnosť a interoperabilitu služieb, rozšíriteľnosť samotného systému a zároveň autonómiu lokálnych repozitárov. Samotné rozmiestnenie systému DRIVER môže obsahovať viaceré inštancie jednotlivých služieb aby tak zaisťovalo lepšiu kvalitu poskytovaných služieb (dostupnosť, výkon) alebo podporu rozmanitej funkcionality (rozličné dotazovacie jazyky). Tieto inštancie, ktoré môžu mať špecifické konfigurácie, sú distribuované na rôznych fyzických miestach. Voľba kam umiestniť jednotlivé inštancie služieb je závislá ako na organizačných, tak aj technických parametroch systému (nezávislosť, bezpečnosť, dostupnosť, výkon atď.).

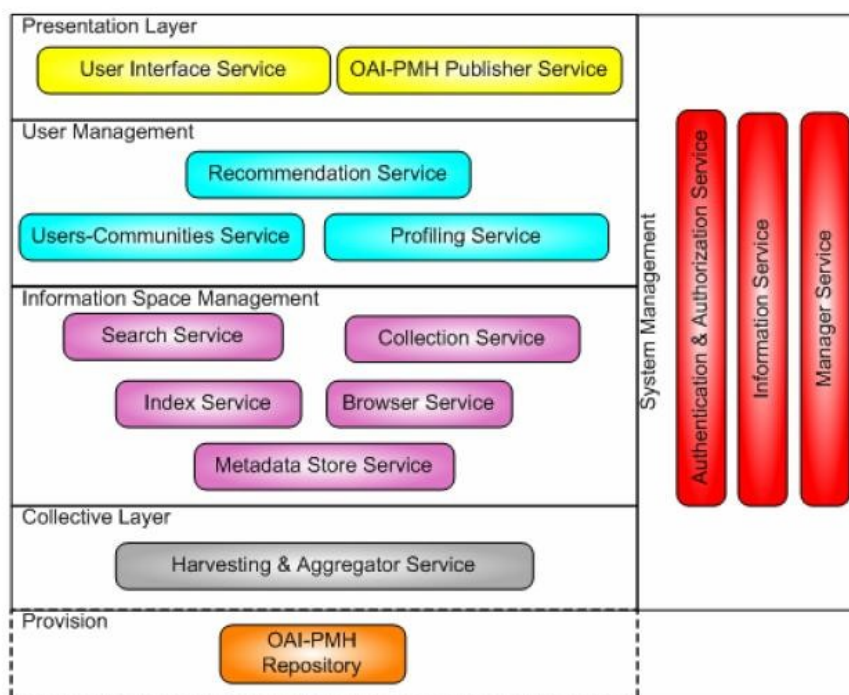


Figure 1 — DRIVER Functional Areas

Ďalším dôležitým výstupom projektu DRIVER je pomocné programové vybavenie nazvané D-Net, ktoré umožňuje poskytovateľom dát a služieb agregovať repozitáre kompatibilné z OAI-PMH do zdieľaného informačného priestoru a aj budovať a upravovať svoje aplikácie nad digitálnymi knižnicami tak, aby dokázali pracovať nad takýmto informačným priestorom. Táto technológia podporuje servisne orientovanú elektronickú infraštruktúru, v ktorej sú distribuované a zdieľané zdroje implementované ako štandardné webové služby a samotné aplikácie pozostávajú z množiny vzájomne komunikujúcich služieb. D-Net v súčasnosti ponúka služby vyžadované pre budovanie

distribúovaných agregáčnych systémov a aplikácií pre koncového užívateľa. Agregáčne systémy umožňujú aby boli uniformné informačné priestory budované pomocou záznamov zozbieraných z heterogénnych repozitárov inštitúcií. V tomto prípade služby uloženia, indexácie a agregácie predstavujú vyspelé nástroje pre zbieranie, triedenie a integrovanie záznamov metadát pomocou protokolu OAI-PMH. Výsledný informačný priestor môže byť následne prístupný pre aplikácie digitálnych knižníc, ktoré sú vytvorené pomocou služieb D-Net akými sú „Recommendation“, „Collection“, „Browsing“, „User Interfaces“ a podobné (porovnaj s obrázkom 1).

Druhá fáza projektu DRIVER

Schválená žiadosť o grant Európskej Komisie v rámci 7. rámcového programu naštartovala druhú etapu projektu DRIVER, ktorá trvala dva roky od 1. decembra 2007 do 30. novembra 2009. Druhá fáza projektu nazývaná tiež DRIVER-II predstavila kľúčové inovácie vzhľadom na prvú fázu pričom vychádzala z jej výsledkov. Prechod z testovacej verzie na plne funkčnú infraštruktúru sa stal prioritou projektu. Ustanovenie Európskej Konfederácie Digitálnych Repozitárov, zahrnutie repozitárov s iným než textovým či publikovaným obsahom (obrázky, prezentácie, pôvodné výskumné dáta) či tvorba rozšírených publikácií, ktoré kombinujú spolu súvisiace objekty informácií do jedného logického celku (publikácia s prezentáciou a dátami) sú významné ciele druhej fázy. Hlavné ciele druhej fázy zahŕňajú aj snahu rozšíriť, obohatiť a posilniť už dosiahnuté výsledky v rôznych oblastiach a dali by sa zhrnúť v nasledujúcich ôsmich bodoch:^[1]

1. Strategicky rozšíriť geografické aj spoločenské pokrytie pomocou DRIVER konfederácie.
2. Ustanoviť robustnú a dobre škálovateľnú infraštruktúru repozitárov spolu s Open Source pomocným programovým vybavením D-Net.
3. Pokryť širšiu oblasť obsahu a typov informácií s využitím rozšírených publikácií.
4. Vytvoriť vyspelú funkcionálnu podporu pre koncových užívateľov a vyvinúť nástroje pre podporu vedeckého bádania s použitím zložitých digitálnych objektov
5. Vyvinúť bohatšie výukové a propagačné programy pre všetky typy komunit.
6. Nepretržite podporovať a podnecovať prácu manažérov repozitárov.
7. Vytvoriť smernice pre zaistenie interoperability v rozsiahlej európskej komunite okolo digitálnych knižníc
8. Podporovať a obhajovať otvorený prístup k európskym výskumným materiálom.

V druhej fáze sa rozšírila aj skupina inštitúcií priamo zapojených do projektu, a tak projekt DRIVER-II má trinásť hlavných partnerov z prostredia vedy a výskumu (University of Athens, University of Bielefeld, Consiglio Nazionale Delle Ricerche, SURF foundation, University of Nottingham, Centre National de la Recherche Scientifique DIS, University of Bath, University of Warsaw, University of Gent, Goettingen State and University Library, Danmarks Tekniske Universitet, Universidade do Minho a Narodna in univerzitetna knjiznica). Vznikla aj „DRIVER Confederation of European Digital Repositories“ majúca za úlohu podporu a propagáciu rozvoja digitálnych repozitárov vedeckých inštitúcií a ich začlenenie do infraštruktúry DRIVER. Konali sa aj dve konferencie, prvá v januári 2007 v nemeckom Göttingene a druhá v októbri 2008 na univerzite v Ghente. Keďže životnosť infraštruktúry DRIVER má byť dlhšia ako grant Európskej Únie, bolo potrebné vytvoriť organizáciu schopnú jej dlhodobej údržby a riadenia. Tak vznikla organizácia COAR (Confederation of Open Access Repositories), ktorá je medzinárodná nezisková organizácia s víziou propagovať väčšie sprístupnenie a využitie vedeckých výsledkov na základe globálnej siete digitálnych repozitárov s otvoreným obsahom.

Súčasný stav projektu DRIVER

V súčasnosti, na konci druhej fázy projektu, je vytvorená plnohodnotná infraštruktúra spájajúca viac ako 200 inštitucionálnych alebo tematických digitálnych repozitárov a tvoriaca tak informačný priestor vedeckých publikácií obsahujúci približne milión dokumentov z dvadsiatich troch európskych krajín v dvadsiatich piatich jazykoch. Tento informačný priestor je pre bežnú aj vedeckú verejnosť sprístupnený cez vyhľadávací portál s možnosťou vytvárať komunity a kolekcie tematicky príbuzných dokumentov. Z technického hľadiska sú súčasné služby systému DRIVER distribuované v štyroch krajinách Európskej Únie (Taliansko, Poľsko, Grécko, Nemecko) pomocou už spomínaného konceptu servisne orientovanej architektúry. Pravidelne sa zbierajú metadáta vo formáte OAI_DC z registrovaných repozitárov pomocou protokolu OAI_PMH, spracovávajú, indexujú a následne sprístupňujú. Samotný dokument môže byť sprístupnený hneď po vyhľadaní v metadátach, a to priamo cez odkaz vo výsledku vyhľadávania, bez zbytočnej registrácie, prihlásenia či prechodu viacerými stránkami. Toto správanie je zaistené pomocou direktív kladených na registrované repozitáre, čím je zaistená čistá forma otvoreného prístupu k vedeckým informáciám. O ďalší vývoj aj samotný manažment infraštruktúry DRIVER sa po skončení grantu bude starať organizácia COAR.

Zhodnotenie projektu DRIVER

Z môjho pohľadu projekt DRIVER v prvých dvoch fázach splnil ciele, ktoré si predsavzal, a tak čiastočne prispel k naplneniu vízie o infraštruktúre vedeckých digitálnych repozitárov v rámci Európy. Dokázal prekonať problémy plynúce jazykovej a kultúrnej diverzity cieľovej oblasti a vytvoril funkčný a perspektívny nástroj pre zdieľanie výsledkov a záverov vedeckých výskumov. Táto dobre škálovateľná, ľahko rozširiteľná, robustná a flexibilná infraštruktúra má potenciál prerásť zamýšľané hranice a stať sa globálnou záležitosťou. Dôležitá je aj poradenská iniciatíva, ktorú projekt DRIVER poskytuje záujemcom o vytvorenie infraštruktúry digitálnych repozitárov na medzinárodnej, národnej či inštitucionálnej úrovni. Len čas ďalej ukáže, či sa výsledky tohto projektu stanú štandardnou súčasťou vedeckej a výskumnej práce, alebo ich vystrieda iná technológia. Dúfam, že krajiny Európy pochopia strategickú výhodu, ktorú im projekt DRIVER ponúka v oblasti vedy, výskumu a vzdelávania a budú ju podporovať, využívať a ďalej vylepšovať.

Zoznam použitej literatúry a zdrojov

- [1] Webová stránka projektu DRIVER
<http://www.driver-repository.eu/>
- [2] Webová stránka podpory pre projekt DRIVER
<http://www.driver-support.eu/>
- [3] Článok o projekte DRIVER na stránke projektu SHERPA
<http://www.sherpa.ac.uk/projects/driver.html>
- [4] Článok o projekte DRIVER vo webovom magazíne Ariadne.
<http://www.ariadne.ac.uk/issue53/feijen-et-al/>
- [5] Článok z portálu Belief – Global Access Point To e-Infrastructure
<http://www.beliefproject.org/zero-in/zero-in-second-edition-emagazine/driver-the-digital-repository-infrastructure-vision-for-european-research/?searchterm=driver>

Metadata v Dublin Core

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Title" content="Projekt DRIVER - Networking European Scientific Repositories" />
<meta name="DC.Creator" content="Kubanda, Karol" />
<meta name="DC.Subject" content="Projekt DRIVER - Networking European Scientific Repositories" />
<meta name="DC.Description" content="Projekt DRIVER vytvára infraštruktúru vedeckých digitálnych repozitárov z celej Európy a sprístupňuje tak vedecké výstupy vedcom i bežnej verejnosti." />
<meta name="DC.Date" content="2.12.2009" />
<meta name="DC.Type" content="Text" />
<meta name="DC.Type" content="Esej" />
<meta name="DC.Format" content="application/pdf" />
<meta name="DC.Format" content="computerFile" />
<meta name="DC.Identifier"
content="http://www.fi.muni.cz/~xkubanda/PV070/ProjektDRIVER_esej.pdf" />
<meta name="DC.Source" content="http://www.driver-repository.eu/" />
<meta name="DC.Language" content="sk" />
```