



Zpráva ze zahraniční služební cesty

Jméno a příjmení účastníka cesty	Helena Dvořáková
Pracoviště – dle organizační struktury	1.6 - odd. retrospektivní konverze a technické podpory
Pracoviště – zařazení	1.6
Důvod cesty	Společné setkání pracovních skupin TEL (The European Library)
Místo – město	Haag
Místo – země	Nizozemí
Datum (od-do)	29.5.-31.5.2011
Podrobný časový harmonogram	Ne 29.5. přílet Po 30.5. Europeana Libraries, Full-text aggregation Út 31.5. Europeana – EDM/ESE, Data mining; odlet
Spolucestující z NK	PhDr. Bohdana Stoklasová
Finanční zajištění	NK – TEL
Cíle cesty	<ul style="list-style-type: none">• zastoupení NK jakožto člen pracovní skupiny TEL-TIWG - účast na techničtěji zaměřených prezentacích a v diskusních skupinách• seznámení s vývojem v oblasti TEL a Europeany
Plnění cílů cesty (konkrétně)	Získání informací důležitých pro další spolupráci s TEL a Europeanou.
Program a další podrobnější informace	Témata prezentací a diskusí: 1/ zapojení vědeckých a vysokoškolských knihoven (sdružených v LIBER) do TEL a jejím prostřednictvím do Europeany – projekt Europeana Libraries: <ul style="list-style-type: none">- zájem vědců z humanitních oborů o jednotný přístup k informacím (již je např. porátl Meril- knihovny LIBER mohou nabídnout cca 600 000 knih- programy řeší agregaci obsahu a infrastruktury a mapování knihovnických metadat na datový model Europeany EDM (WP3-WP5) – data by měla být integrována do konce r.2012 -

	<p>2/ Metadata pro Europeanu – nyní EDM (Europeana Data Model), výhledově model ESE (Europeana Semantic Elements)</p> <ul style="list-style-type: none"> - cílem je oddělit objekt od jeho digitální reprezentace - vlastnosti (properties) v EDM odpovídají prvkům (elements) v ESE, navíc jsou kontextuální informace jakožto třídy (classes): událost, původce, místo, předmět, pojem, časové úsek – pokud lze, tak mapovat na jiné systémy, např. VIAF <p>3/ zpracování fulltextů pro vyhledávání – výsledky programů TEL Plus WP1 a Impact</p> <ul style="list-style-type: none"> - harvest fulltextu s metadaty pro indexování SOLR, normalizace a obohacování metadat – vyhledávání v textu, zobrazení útržku (snipped), odkaz na dokument v lokálním umístění - zdokonalení OCR (Impact) – použití slovníků, moduly pro převod fraktury, využití výsledků z již provedených oprav atd. – použití ve FineReaderu i samostatně - potřeba persistentních identifikátorů, unifikace strukturálních metadat <p>4/ problémy s dodáváním fulltextů</p> <ul style="list-style-type: none"> - jak jsou schopny knihovny dodat data - perspektivou je ALTO - vyhledání musí vést až na konkrétní stránku - problém deduplikace (stejně dokumenty) <p>5/ možnosti řazení dle relevance, filtrování a sdružování zobrazených výsledků - jak využít data mining</p> <ul style="list-style-type: none"> - uživatelé chtějí řazení dle relevance, deduplikaci výsledků a persistentní linky - projekt ARROW využitelný pro řešení autorských práv (získání biografických dat autorů) i sdružování - využití „podobnosti“ záznamů pro FRBR - 40% záznamů bez věcného popisu - problém při věcných facetách
Přivezené materiály	Tištěné materiály nejsou k dispozici, bližší informace včetně některých prezentací jsou na interních stránkách TEL
Datum předložení zprávy	7.7.2011
Podpis předkladatele zprávy	
Podpis nadřízeného	
Vloženo na Intranet	
Přijato v mezinárodním oddělení	