



europeana
think culture

Evropská digitální knihovna

<<http://www.europeana.eu>>

Tomáš Drusa

(256167@mail.muni.cz)

9. 12. 2009

Co je Europeana?

- Projekt Evropské unie
- Webový portál, který má
 - být přístupný v rodném jazyce všem občanům Unie
 - umožnit centrální přístup k multimediálním dokumentům
 - týkajícím se Evropského kulturního dědictví
 - uloženým ve stovkách repozitářů jednotlivých institucí v EU
 - umožnit v nich rychlé přirozené hledání a obsahovat souvislosti
 - uchovat kontrolu a správu repozitářů v rukách institucí

Zázemí

- Za projektem stojí Evropský spolek digitálních knihoven (EDLF)
 - sdružuje archivy, muzea, audiovizuální úložiště a knihovny v EU
 - má za cíl usnadňovat přístup občanů Unie ke kulturním hodnotám
 - největším prostředkem má být právě portál Europeana.eu



- Faktické sídlo (kanceláře, servery) v prostorách Holandské národní knihovny v Haagu

Zázemí

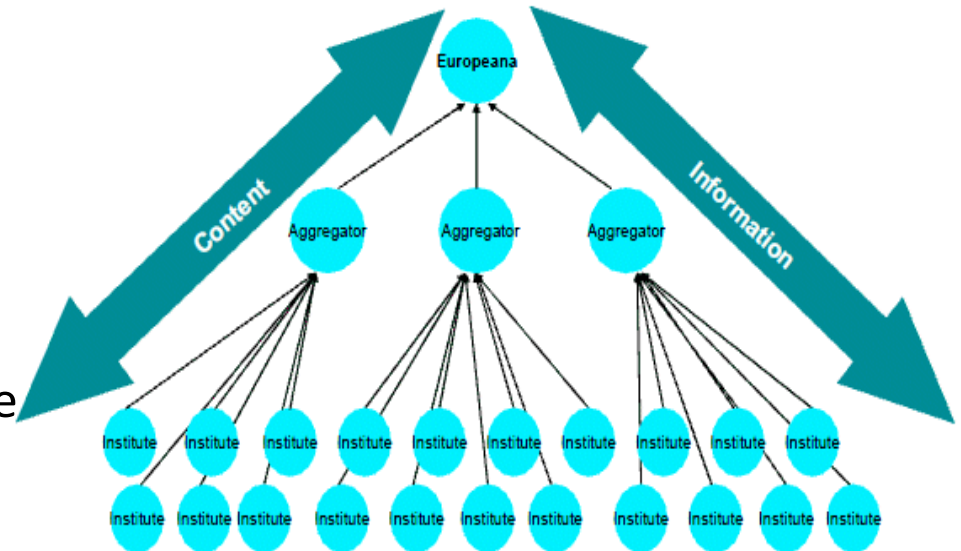
- Projekt je financován z programů EU
 - eContentPlus (do r. 2008)
 - ICT-PSP (Information and Communication Technologies Policy Support Programme)



- Ročně 2 miliony eur na projekt samotný
- Další miliony na podporu digitalizace ve státech EU a vznik agregačních center

Obsah

- Datové soubory zůstávají v repozitářích institucí zapojených do projektu, poskytují se pouze metadata
- Kontrola nad vlastními objekty, nižší náklady a složitost
- V konečné fázi uživatel přistupuje do jednotlivých repozitářů



Agregační centra

- mezičlánek mezi institucemi a Europeanou
- sama sklízí metadata z repozitářů, ověřují je, validují, doplňují, odstraňují duplikace
- zejména pro „menší hráče“, málo početné, hůře spravované sbírky

Protokoly

- Základní a povinný sklízecí mechanismus: OAI-PMH (tedy Dublin Core)
- Sklizení periodické, inkrementální, občasné plné obnovení
- Systém umožňuje použití takřka libovolného metadatového schématu
 - „programovatelná“ komponenta European SPM
 - pro nový formát metadat se nadefinuje algoritmus převodu do ESE (vnitřní metadatové schéma European)
 - SPM dostane metadata, převodní schéma a dle něj před uložením zpracuje

ESE a ESE XML Schema

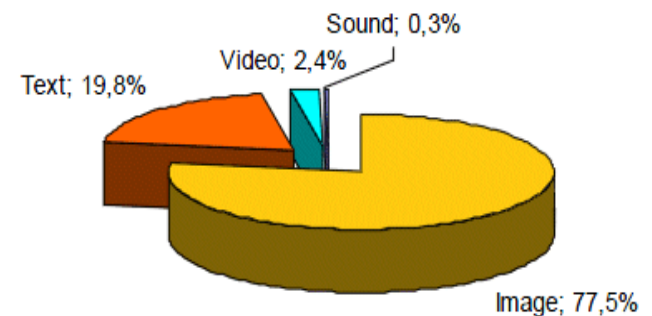
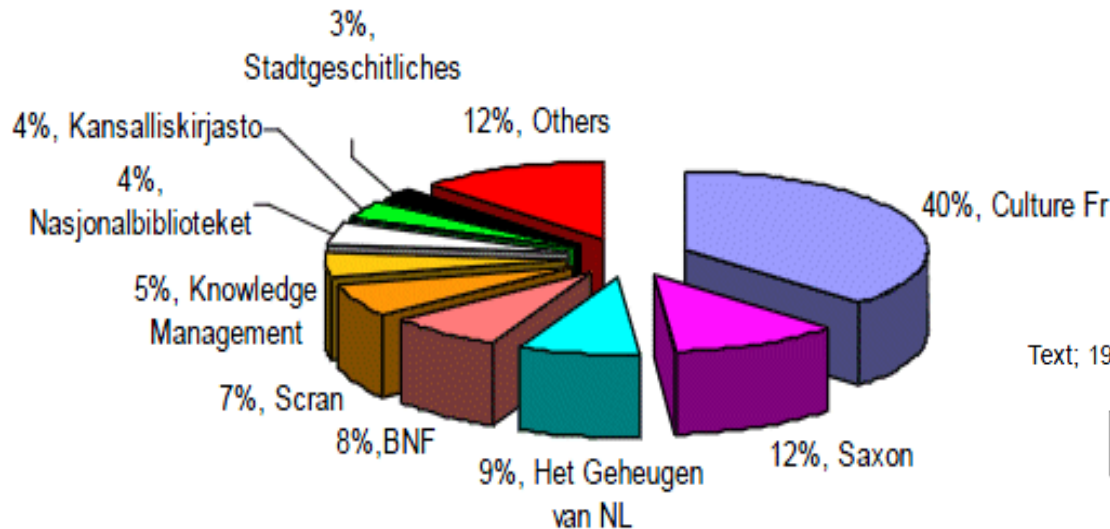
- Europeana Semantic Elements
- Vnitřní metadatové schéma Europeany, vše se na něj před uložením převádí
- Dublin Core (15+1 tag) a Europeana (11 tagů)
- Přidané značky pro
 - relace mezi objekty
 - odkazy na domovský repozitář, na datový soubor
 - informace o náhledech, informace přidané uživateli
 - typ objektu (text, obrázek, audio, video)
 - a další...

Sémantický web

- Metadata v prostoru Europeany obsahují sémantiku
- V souladu s Open Linked Data a Timem Berners-Leem
- Pomocí RDF:
 - časové a geografické zařazení
 - vztahy
 - podobnost
- Cíl:
 - přirozenější pohyb pro uživatele, otázky kdo, kde, kdy a co
 - jednotný dotazovací jazyk pro lidi i stroje (SPARQL)
 - krok směrem k machine-readable webu

Obsah

- V době spuštění přes 2 miliony DO, většina obrázků
- Hlavní přispěvatel Francie (50%), dále Nizozemí, UK, Finsko, Švédsko, ostatní kolem 1% či méně (ČR – NK, prakticky nic)



- V současnosti 4,6 milionu DO

Potíže se spuštěním

- Spuštění betaverze 20. 11. 2008
- Tři servery testované na 5 milionů přístupů za hodinu
- V den spuštění stabilně 10 mil/h, ve špičce 13 mil/h (3700 požadavků za vteřinu) – během 4 hodin překonána roční návštěvnost stránek EK
- Nouzové navýšení nedostatečné, odstavení a posílení hardware
- Opětovné spuštění koncem roku 2008, nyní čtyřnásobná kapacita (20 mil/h, 5500 hits/sec)

Budoucnost

- Projekt je stále ve stádiu betaverze, práce na systému + zapojování dalších poskytovatelů obsahu
- Start plné verze v roce 2010, má mít 10 milionů digitálních objektů
- Současnost: otřesná lokalizace, pokulhávající vyhledávání
- Dobré nápady, solidní zázemí
- Dodržení termínu, mnoho věcí najednou, velký rozsah

Děkuji za Váš čas, prezentace je u konce.

Pokud nám zbývá čas, můžete jej samozřejmě využít na dotazy.



Evropská digitální knihovna
<<http://www.europeana.eu>>

Tomáš Drusa
(256167@mail.muni.cz)