

Open Content Alliance

www.opencontentalliance.org

Ladislav Chodil
N-AP, 4. semestr
6. prosince 2006

Stručná charakteristika projektu

Open Content Alliance (OCA) představuje úsilí skupiny kulturních, technologických, nekomerčních a vládních organizací po celém světě, které napomáhají vytvoření trvalého archivu mnohojazyčných digitalizovaných textů a multimediálního obsahu. OCA zastřešuje celosvětové společenství zaměřené na poskytování volného přístupu k obsahu archivu za současného respektování práv držitelů autorských práv. OCA je přímým konkurentem Google Book Search s tím rozdílem, že společnost Google chce digitalizovat dokumenty, aniž by získávala svolení od držitelů autorských práv.

Doba řešení, aktuální stav

Dne 3. října 2005 v San Franciscu bylo oznámeno formování Open Content Alliance (www.opencontentalliance.org) společnostmi The Internet Archive, Yahoo! Inc., Adobe Systems Inc., the European Archive, HP Labs, the National Archives (UK), O'Reilly Media Inc., Prelinger Archives, the University of California a the University of Toronto.

Milníkem pro OCA měl být říjen 2006 (podle článků z počátku roku, popř. z roku 2005), kdy byla plánována demonstrace toho, co přineslo otevřené a společné úsilí k vybudování jedinečné kolekce. Plně by se OCA měla rozjet v nadcházejícím roce.

O tom, jak to ve skutečnosti s vývojem projektu vypadá dnes nebo do budoucna, jsem nenašel žádné informace (samotné webové stránky www.opencontentalliance.org se odkazují na články, které jsou datovány do roku 2005).

Cíle projektu

OCA má nabízet široký rozsah materiálů zahrnujících kulturní, historický a technologický digitalizovaný tištěný a multimediální obsah z knihoven, archivů a nakladatelů. Obsah bude uložen v jedinečném, trvalém skladišti a souborem prací bude možnost vyhledávání a stahování dat zdarma pro každého.

Popis projektu

Projekt OCA je řízen přispěvateli a je navrhován tak, aby byl rostoucím sdružením, kde každý člen souhlasí s pravidly otevřenosti (tzv. *Call for Participation*). Počáteční obsah v archivu OCA budou poskytovat zakládající členové zahrnující organizace the University of California, the University of Toronto, the National Archives (UK), O'Reilly Media, Inc. a the European Archive. Adobe Systems Inc., HP Labs, and Yahoo! Inc., kteří jsou také zakládající

členové, budou vytvářet technické a v určitých případech finanční zázemí tak, aby byl možný přístup k archivu online každému. OCA bude zahrnovat obsah v rozmezí od historických prací přes dětské knihy po vysoce specializované technické zprávy.

Call for Participation

Přispívající organizace musejí podpořit následující pravidla otevřenosti (tzv. *Call for Participation*):

1. OCA bude podporovat v co největší možné míře přístup a znovupoužitelnost sbírek v archivu za respektování autorských práv vlastníků a přispěvatelů.
2. Přispěvatelé si budou určovat název a podmínky, pod kterými budou jejich kolekce distribuovány.
3. OCA nesmí být povinna přijmout veškerý obsah, který jí je nabídnut a může dát přednost obsahu, který může být přístupný širší skupině uživatelů.
4. OCA bude poskytovat metadata kolekcí v rozmanitých formátech.
5. OCA vítá snahu vytvořit a nabídnout nástroje (zahrnující poskytování pomoci, katalogů a indexů), které zvýší využitelnost materiálů v archivu.
6. Kopie OCA kolekcí budou uloženy mezinárodně v rozmanitých archivech pro zajištění jejich dlouhodobého uchování a dostupnosti všem.

Rozdělení rolí jednotlivých členů

Internet Archive bude poskytovat do archivu OCA materiál z jeho sbírek a bude také digitalizovat data od přispěvatelů. Yahoo! bude spravovat vyhledávací stroj na webových stránkách OCA a veškerý obsah z archivu OCA bude přístupný i přes Yahoo! Searcher (pro který je to další podkladový materiál do jeho indexů). Jakmile je jednou nějaký obsah dostupný přes archiv OCA, může být jinými vyhledávacími stroji zahrnut do jejich indexů. To významně přispívá ke snižování rozšiřování hlubokého webu (což jsou miliony webových stránek obsahujících data, na která nemají vyhledávací stroje typicky přístup), stejně tak k podporování celosvětového sdílení znalostí.

OCA má také rozděleny nástroje pro pomoc přispěvatelům při digitalizaci jejich dat. Adobe bude poskytovat online nástroje pomáhající jednotlivcům a institucím vytvářet PDF dokumenty pro archiv. Hewlett-Packard poskytne skenovací zařízení pro digitalizaci velkých objemů dat.

Respektování autorských práv

Zkušenosti ukazují, že nejtvrďší překážka pro digitalizaci leží v byrokratické politice a složitých zákonech. Open Content Alliance doufá, že bude fungovat i přes tyto problémy (je nutné ustanovit mechanismy pro sdílení obsahu).

OCA může do svého archivu přidávat materiály veřejně dostupné (otevřené) nebo byla možnost přidání materiálů do archivu povolena držiteli autorských práv. Podle přání držitele autorských práv může být obsah chráněný autorským zákonem šířen i pod licencí *Creative Commons license*. *Creative Commons* je nekomerční organizace, která podporuje licencování pro osobní využití digitálního obsahu. Data, která mají být přístupná přes webové stránky OCA budou dostupná v PDF a dalších široce používaných formátech (v roce 2005 byl dalším

podporovaným formátem například formát Déjà Vu, ale OCA se nebrání změnám a vývoji dalších formátů). Tento přístup umožňuje sdělovacím prostředkům a nezávislým vydavatelům rozšířit pole působnosti, jelikož stačí předložit obsah v jednom podporovaném formátu (do ostatních je obsah překonvertován automaticky). Tím překlenuje kategorie, formáty souborů a jazyky, zatímco si data ponechávají svá autorská práva.

Financování OCA

OCA bude financována jejími přispěvateli a bude také přijímat dary z celosvětových institucí zahrnujících vládní, komerční a dobročinné organizace.

Interní dělení organizace

Jádro funkcí OCA poskytují jednotlivé pracovní skupiny, které radí v provozních záležitostech a pomáhají vybudovat nezbytné politiky a postupy. Tyto skupiny jsou navrženy tak, aby byly pracovníě orientované a nezávislé.

Pracovní skupiny v OCA

Metadata and collaborative collection development working group

Tato skupina pracuje s knihovnami a ostatními poskytovateli, doporučuje a pomáhá ohodnotit přístupy k popisování digitalizovaných materiálů, které budou přispěny. Aby byl zachován směr v tom, co je digitalizováno, pomáhá informovat vybrané osoby o tom, co by mělo být digitalizováno (včetně stavu autorských práv). Skupina navrhuje a vyvíjí přístupy k vedení výběru. Také vyvíjí cenově efektivní služby pro umožnění koordinace přístupů k velmi rozloženému vývoji OCA kolekcí.

Digital Preservation Working Group

Doporučuje minimální požadavky nutné pro vytvoření důvěry široké komunity v OCA kolekce jako trvalé, kulturní, digitalizované zdroje. Očekává se, že základní přehled minimálních požadavků bude rozsáhlá práce právě vedená mnohými knihovnami, archivy a vědeckými organizacemi.

Contribution Working Group

Vybízí přispěvatele o příspěvky digitálních materiálů, zařízení, nástrojů a finanční podpory od organizací (podepsanými) pod principy Open Content Alliance. Tato skupina bude vyhodnocovat a stanovovat prioritu příspěvků nabízených do OCA právě tak jako rozšiřovat a vysvětlovat principy OCA. Tato skupina bude mít také povinnost k formálnímu (oficiálnímu) uznávání příspěvků, které čekají na uznání. Uznání může být uvedeno na webových stránkách OpenContentAlliance.org, v dopisech podpory, v uvolněném tisku atd.

Book Format Working Group

Tento výbor doporučuje několik formátů, které mohou být použity ve skenovaných projektech nebo v přispívaných digitálních knihách.

Scanning Protocol

Vytváří dokument, který popisuje detailně průběh práce a příslušné náklady pro knihovny a ostatní zainteresované strany ve skenování tištěných materiálů udržovaných v jejich sbírkách.

Data Transfer Protocol

Vytváří dokument, který popisuje protokol (zahrnující minimální standardy) spojený s přenosem digitálně přeformátovaných dat pro Internet Archive pro zahrnutí jejich dat mezi sbírky v OCA archivu.

Vlastní zhodnocení projektu a jeho přínosu

Projekt Open Content Alliance je (tedy zatím bude) zajímavá alternativa digitalizace a zpřístupňování digitalizovaných dat oproti projektu Google Book Search. Otázkou zůstává, jestli se OCA v budoucnu prosadí mezi uživateli hledajícími nějaké informace na webu. Jelikož budou v archivu OCA přístupné pouze otevřené materiály nebo ty, ke kterým dostane OCA oprávnění pro zveřejnění, přikloní se pravděpodobně větší počet uživatelů ke Google Book Search (další část uživatelů bude produkty Googlu používat už jen proto, že jsou na něj navyklí), jelikož tam bude digitalizované větší množství dokumentů a knih. Zároveň si Google může zaindexovat veškerý obsah dostupný přes archiv OCA, takže Google bude vždy nabízet širší možnosti vyhledávání. Na druhou stranu se uživatelé nerozhodují jen podle prohledávaného množství dat, ale také například podle uživatelského prostředí nebo dalších poskytovaných služeb, takže se teprve uvidí po úplném spuštění projektu OCA, zda ho uživatelé budou využívat přímo nebo dají přednost konkurenci.

Seznam literatury/zdrojů

- Open Content Alliance (<http://www.opencontentalliance.org>)
- Wikipedie (http://en.wikipedia.org/wiki/Open_Content_Alliance)
- InfoToday – Open Content Alliance Rises to the Challenge of Google Print (<http://www.infotoday.com/newsbreaks/nb051003-2.shtml>)

Metadata v DC

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
<meta name="DC.title" lang="cs" content="Open Content Alliance" />
<meta name="DC.creator" content="Ladislav Chodil" />
<meta name="DC.description" lang="cs" content="Záv re ná ese j do pedm tu
PV070 Digitální knihovny" />
<meta name="DC.date" scheme="DCTERMS.W3CDTF" content="2006-12-07" />
<meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text" />
<meta name="DC.format" content="application/pdf" />
<meta name="DC.format" content="473202 bytes" />
<meta name="DC.identifier" scheme="DCTERMS.URI"
content="http://www.fi.muni.cz/~xchodil/pv070/esej/OCA.pdf" />
<meta name="DC.language" scheme="DCTERMS.RFC1766" content="cs" />
```