

Primo Central – vesmír odborných článků na dosah

Martin Vojnar

MULTIDATA Praha s.r.o.

vojnar@multidata.cz

INFORUM 2010: 16. konference o profesionálních informačních zdrojích

Praha, 25. - 27. 5. 2010

Abstrakt:

Primo Central je souhrnný rejstřík sestávající z dat získaných od primárních a sekundárních vydavatelů a producentů databází. S použitím Primo Central uživatelé vyhledávají v lokálně dosažitelných sbírkách i ve světovém elektronickém obsahu. Z výsledků, které obdrží, je vytvořen jeden seznam, který zahrnuje všechny prameny seřazené podle relevance. Knihovny mohou upravit relevantní výsledky tak, aby zvýraznily své lokální sbírky.

Primo Central znamená významnou změnu ve zpřístupňování informačních zdrojů dostupných vědecké obci. Uživatelé mají nyní přístup k rozsáhlému bohatství elektronických zdrojů spolu s jejich lokálními sbírkami – a vše je snadno a rychle dosažitelné.

Nejenže Primo Central přitáhne pozornost k materiálům, které uživatelé dříve neznali (takže se více využije investice do elektronických zdrojů), ale také jasně ukazuje původ vyhledaných materiálů, čehož uživatelé rádi využívají.

Těšíme se na veřejné spuštění Primo Central, které je plánované na červenec 2010, a na to, až uvidíme tuto službu aktivní ve více než 270 knihovnách pracujících se systémem Primo po celém světě.

Úvodem

Během posledních let knihovny čelily (a dosud čelí) jasné výzvě – nabídnout svým uživatelům takový příjemný, jednoduchý a efektivní nástroj, který by byl schopný prohledávat všechny knihovní sbírky a nabídnout nejrychlejší cestu k hledanému primárnímu dokumentu. Za vzor takového nástroje z prostředí internetových vyhledávačů, bez kterých si dnes svůj pobyt na informační dálnici dokážeme jen těžko představit, obecně posloužil ve světě Google nebo v českém prostředí Seznam. Rychlý, přizpůsobivý, intuitivní – drtivá většina uživatelů je tak používá ke své spokojenosti.

V roce 2007 reagovala firma Ex Libris (<http://www.exlibrisgroup.com/>) na tuto poptávku vývojem nového softwarového nástroje, který byl pojmenován Primo. Jedná se vlastně o model knihovního „googlu“, který zastřešuje všechny knihovní dokumenty (tištěné, digitální, elektronické) bez ohledu na jejich formu zpřístupnění (lokálně, vzdáleně) a nad kterými nabízí koncovým uživatelům dvě základní funkce: vyhledávání a dodávání.

Podařilo se tak oddělit systémy, které knihovna provozovala pro vlastní evidenci dokumentů (typicky automatizované knihovní systémy), od jejich prezentační vrstvy, jež byla do té doby výrazně zatížena tradičním knihovnickým pojetím zpracování informací a standardy s tím související. Díky osamostatnění prezentační vrstvy se tak mohla knihovna soustředit výhradně na své uživatele a nabídnout jim výrazně lepší prostředí pro jejich práci. V systému Primo navíc má knihovna nad celým procesem plnění prezentační vrstvy, indexování, vyhledávání a dodávání absolutní kontrolu.

Aktuální verze 3 systému Primo tak především umožňuje:

- **hladkou integraci s automatizovaným knihovním systémem** v reálném čase díky adopci doporučení, které bylo výsledkem pracovní skupiny v rámci *Digital Library Federation* (<http://www.diglib.org/architectures/ilsdi/>) – jinými slovy uživatel pracuje se svým čtenářským kontem stejně, jako kdyby se pohyboval v knihovním systému
- **plně konfigurovat vlastnosti prezentační vrstvy**, které se týkají automatického získávání záznamů ze zdrojových systémů, indexování, zobrazování a dodávání dokumentů – jinými slovy, knihovna může začlenit do nabídky uživatelům libovolný dokument, k jehož metadatům má přístup (metadata mohou být i automaticky generována, např. z údajů relevantních www stránek)
- **rozlišit a stanovit speciální míry relevance** pro určité případy (např. shodné záznamy, podobné záznamy, oblíbené záznamy, nové záznamy, zásadní záznamy apod.) – jinými slovy, výsledky dotazů, které se nabízejí v první desítce, vyhovují drtivé většině uživatelů a nabízejí stejnou relevanci, jakou známe právě z populárních vyhledávačů

Zmíněné vlastnosti jsou jen střípkem z celkové funkcionality systému Primo. Mezi jeho další funkce bychom mohli zahrnout celou plejádu prvků, místo toho ale doporučíme si spíše vyzkoušet systém Primo tam, kde se používá. Nemusíte zkoumat všech cca 270 institucí, začít můžete např. v konsorciu rakouských knihoven Österreichischen Bibliothekenverbundes (<http://search.obvsg.at/OBV>) nebo na švýcarské Eidgenössische Technische Hochschule (ETH) v Curychu, která by měla své nové uživatelské rozhraní oficiálně zpřístupnit v polovině roku 2010 a jehož podoba bude určitě stát za pozornost.

Není žádným překvapením, že od začátku bylo klíčovou snahou systému Primo také zpřístupňování licencovaného elektronického obsahu. Jeho integrace typicky probíhala pomocí tzv. federativního vyhledávání – tj. v reálném čase byly paralelně dotazovány vybrané

externí zdroje, které s průměrnou dobou odpovědi 10-15 vteřin vracely výsledky. Pořadí záznamů ve výsledkových množinách a vrácený obsah byly determinovány každým externím zdrojem zvlášť, což neumožňovalo jejich rychlé a efektivní zpracování. V polovině roku 2009 se tak začal rýsovat velmi užitečný doplněk, který měl zásadně změnit přístup k vyhledávání na úrovni článků a který předznamenal, že Primo není pouze běžný nástroj pro vyhledávání a dodávání. Primo pro knihovnu představuje kompletní řešení, a proto systém v současnosti nabízí dva typy důležitých služeb:

- **indexovací služby** (Primo Central), které doplňují knihovní zdroje o externí souhrnný rejstřík skládající se ze záznamů, plných textů a dalších datových objektů pocházejících od největších světových provozovatelů informačních zdrojů a elektronických vydavatelů
- **referenční služby** (bX, BibTip), které „doporučují“ koncovým uživatelům ke každému nalezenému záznamu podobné dokumenty – jinými slovy dokumenty, o které se často zajímali jiní uživatelé (použitý mechanismus doporučování byl u obou služeb předmětem úspěšného vědeckého bádání a je založen výhradně na anonymním sledování uživatelských vzorců chování)

Lze předpokládat, že do budoucna se bude paleta dalších služeb systému Primo rozšiřovat. A právě na první uvedenou službu se soustřeďuje následující text.

Základní údaje o Primo Central

Máte-li cca 4 minuty času a jste-li připojeni k internetu, základní představu o Primo Central získáte nejlépe na YouTube (<http://www.youtube.com/watch?v=fDc6kJo6SX8>).

Primo Central je centrálně spravovaný index, který v sobě zahrnuje záznamy, plné texty a další datové objekty od největších světových provozovatelů informačních zdrojů a elektronických vydavatelů. Spolupráce s přispívajícími partnery a jejich výběr probíhá s ohledem na potenciál jejich odborného obsahu a na základě poptávky ze strany knihoven.

Aktuální obsah indexu činí více než 200 miliónů záznamů z odborných zdrojů, do kterého přispívá téměř 100 různých světových provozovatelů informačních zdrojů a vydavatelů. Tento počet dále roste, do konce roku 2010 by se měl minimálně zdvojnásobit.

Každá knihovna může získat před zahájením služby procentuální výpočet pokrytí obsahu svých elektronických licencí v indexu Primo Central.

Knihovny mohou Primo Central připojit jako externí zdroj do systému Primo nebo Metalib, čímž získají bezkonkurenčně nejrychlejší možnost prohledávání obsažených informačních zdrojů s vysoce přesnými výsledky.

Ačkoliv uživatelé mohou v indexu Primo Central vyhledávat i na úrovni plného textu, přístup k vlastnímu licencovanému obsahu jim Primo Central nezajišťuje. Zobrazení plného textu proběhne až prostřednictvím přímého odkazu v záznamu, je-li přítomen (např. u EBSCO), nebo pomocí linkovacího serveru (např. SFX), což je z hlediska určení adekvátní kopie pro konkrétního uživatele a jeho lokaci vhodnější a efektivnější.

Jak již bylo naznačeno, index Primo Central není v knihovně uložen lokálně. Jde o službu, ke které se přistupuje vzdáleně, a jejíž dostupnost, plnění a neustálou aktualizaci zajišťuje producent systému Primo, tj. Ex Libris.

Knihovna také nemusí vlastnit licence ke všem záznamům, obsaženým v Primo Central. Může se rozhodnout, zda bude nabízet vyhledávání pouze v záznamech, které má licencované, nebo i v těch ostatních. Pokud prohledává všechny záznamy, přístup k plným

textům je ošetřen na základě dohody s provozovateli (např. pomocí jednorázových poplatků za zobrazení plných textů – „pay per view“), nezávisle na Primo Central.

Koncový uživatel nijak nerozlišuje, zda záznamy přicházejí prostřednictvím lokálního indexu systému Primo nebo vzdáleného indexu služby Primo Central. Výsledky obou indexů jsou dynamicky a plnohodnotně sloučeny (jinými slovy: slučují se jablka a jablka, nikoliv jablka a hrušky, jak tomu často bývalo a bývá u federativního vyhledávání).

Podle výše popsaného rozhodování knihovny o zpřístupnění placených licencí dále záleží, zda je koncový uživatel přihlášen nebo ze kterého místa se systémem Primo pracuje. Pak obsahují jeho výsledky buď celý záběr indexu Primo Central nebo jen volně dostupnou část indexu Primo Central anebo jen část indexu Primo Central odpovídající jeho oprávněním (např. student fakulty A se dostane pouze ke zdrojům předpláceným fakultou A).

Služba Primo Central je po technické stránce založena na systému Primo, a díky tomu tak disponuje následujícími parametry:

- vícejazyčnost na všech úrovních (data, indexování, zobrazení) a nadstandardní jazyková podpora při vyhledávání
- konsorciální podpora
- otevřenost aplikační vrstvy (API) a dokumentované vývojové prostředí (SDK)
- rozšiřování funkčnosti systému Primo uživatelskou komunitou a její volné sdílení prostřednictvím portálu EL Commons (<http://www.exlibrisgroup.org/>)

Primo Central podrobněji

Při úvahách nad podobnou službou Vás možná napadne řada otázek. Na některé z nich se pokusím odpovědět níže.

Jak je zajištěna dostupnost služby Primo Central?

Poskytovatel služby Ex Libris provozuje službu Primo Central v prostředí výpočetního oblaku („cloud computing“). Díky tomu je plynule zajištěn dostatečný hardwarový výkon, dostatečná kapacita a propustnost internetového připojení a ideální prostředí pro průběžné plnění a aktualizaci obsahu, včetně zlepšování parametrů služby.

Jaká je cenová strategie služby Primo Central?

Poskytovatel služby Ex Libris se rozhodl za používání indexu Primo Central zavést jen minimální roční poplatek.

Jak je zaručena viditelnost lokálních sbírek knihovny v globálním centrálním indexu?

Při spojování výsledků z lokálního indexu systému Primo (tj. z lokálních sbírek) a vzdáleného indexu služby Primo Central se uplatňuje celá řada kritérií, mj. také významnost lokálních dokumentů, o jejichž důležitosti rozhoduje formou jednoduchého nastavení sama knihovna. Jinými slovy první desítky dokumentů rozhodně neobsahují pouze výsledky ze služby Primo Central, naopak obsahují ty nejvíce relevantní výsledky z obou indexů.

Při zapojení služby Primo Central se dramaticky rozšíří záběr pro vyhledávání, nezahltí to koncové uživatele?

Obdobně jako u internetových vyhledávačů jsou koncoví uživatelé zvyklí na výsledkové množiny většího rozsahu. Důležité je, aby relevantní výsledky byly ihned na začátku, o což se systém Primo včetně služby Primo Central snaží na prvním místě.

Má Ex Libris uzavřené smlouvy s poskytovateli informačních zdrojů ohledně přístupu k plnému textu prostřednictvím služby Primo Central?

Producent služby Primo Central neposkytuje licenci k elektronickému obsahu indexovaných dokumentů. Uzavřené smlouvy s poskytovateli informačních zdrojů opravňují uživatele pouze k vyhledávání dokumentů (buť na základě indexovaného plného textu). Vlastní zpřístupnění plného textu probíhá na základě licence mezi knihovnou a poskytovatelem informačního zdroje. Vyhledávání v licencovaném obsahu některých zdrojů v Primo Central může být dostupné pouze, pokud knihovna disponuje příslušnou licencí od daného poskytovatele.

Jak probíhá slučování výsledků s ohledem na fakt, že stejný článek se může vyskytovat ve více zdrojích?

Knihovna se může rozhodnout, zda se mají zobrazovat všechny výskyty takového článku s uvedením každého zdroje, odkud článek pochází, nebo zda má docházet k tzv. seskupování a slučování jednotlivých záznamů s výběrem nejlepšího reprezentanta (podle rozsahu metadat a dostupnosti plného textu).

Plný text konkrétního článku může být typicky získán od různých poskytovatelů. Jak je zajištěna optimalizace přístupu pro koncového uživatele?

Služba Primo Central plně podporuje standard OpenURL, který umožňuje s pomocí linkovacího serveru (SFX nebo jiný) efektivně určit tu nejlepší variantu dodání plného textu. Pokud je součástí záznamu také přímý odkaz na plný text, je uživateli rovněž nabídnut, což může být výhodné zejména tam, kde poskytovatelé informačních zdrojů nepodporují přímé linkování na úrovni článku.

Jaká je přidaná hodnota služby Primo Central při porovnání s alternativními řešeními?

S ohledem na dostupné analýzy chování uživatelů a zpětnou vazbu knihoven z beta testování služby Primo Central není rozhodujícím kritériem kvantita záznamů, tj. velikost rejstříku. Na druhé straně, za důležité lze považovat:

- role Ex Libris coby producenta služby Primo Central je vzhledem k indexovanému obsahu neutrální; cílem je maximální spokojenost uživatelů knihovny bez ohledu na vlastní indexovaný obsah, neboť Ex Libris není poskytovatelem informačních zdrojů
- služba Primo Central byla vyvíjena v souladu se systémem Primo – řada jejích parametrů je konfigurovatelná a volitelná a záleží na rozhodnutí knihovny, zda a v jakém rozsahu je aktivuje a nastaví; cílem je, aby si knihovna sama mohla určovat a průběžně své rozhodnutí měnit, jak má služba Primo Central fungovat
- služba Primo Central je poskytována v kontextu systému Primo ve vazbě na jeho další aspekty a služby (např. referenční služba bX), využívá tak synergií jak dalších

technologií producenta služby, tak funkčních rozšíření, na jejichž vývoji se formou otevřené platformy podílí samy knihovny

Lze upravit záběr vyhledávání článků např. jen na recenzované (peer-reviewed) tituly?

Ano, možnost filtrování výsledků nabízí nejen volbu pro články pocházející z recenzovaných titulů, ale je možné výsledky filtrovat také podle dostupnosti plného textu.

Proč je v souvislosti se službou Primo Central důležitá i referenční služba bX?

Referenční služba bX nabízí doporučení vedoucí k podobným článkům na základě aktuálního chování informačních spotřebitelů (tj. uživatelé, kteří se zajímali o tento článek, se zajímali často také o tyto články). Výrazným způsobem tak napomáhá objevování informačních zdrojů, které jsou důležité, ale které by jinak při běžných technikách vyhledávání na základě klíčových slov nebyly součástí výsledkové množiny a uživatel by se o nich nedozvěděl.

Jak uživatel pozná, z jakého informačního zdroje záznam v indexu Primo Central pochází?

Každý záznam indexovaný službou Primo Central obsahuje informaci o svém původu, tj. z jakého informačního zdroje pochází a také v jakém titulu byl článek publikován. Jednoduchým způsobem tak může uživatel filtrovat výsledky nejen v závislosti na zdroji, ale také v závislosti na konkrétním titulu.

Jaký je aktuální stav služby Primo Central?

Vybrané knihovny (např. Vanderbilt University) se v lednu a únoru letošního roku zúčastnily úspěšného beta testování služby Primo Central. Na základě velmi kladné reakce ze strany uživatelů se některé z těchto knihoven nezávisle na plánovaném termínu oficiálního uvolnění rozhodli provoz služby v rámci systému Primo již zahájit. V současnosti se služba Primo Central nachází před vlastním oficiálním spuštěním, které proběhne na začátku července 2010.

A nejlepší na konec

A nejlepší zpráva na konec – do června 2011 je služba Primo Central poskytována knihovnám, které provozují systém Primo, bez jakýchkoliv dalších poplatků, tj. zdarma.

Služba Primo Central bude oficiálně uvolněna producentem Ex Libris na začátku července 2010. Sledujte <http://www.exlibrisgroup.com/category/PrimoCentral>.