

Discovery systémy

stručný přehled

Mgr. Petr Novák
Ústřední knihovna Univerzity Karlovy v Praze

Program

- proč discovery
- co je discovery
- komerční systémy
- opensource systémy
- aktuality
- budoucnost

Proč discovery?

- **zvýšení úrovně využitelnosti** elektronických informačních zdrojů pořízených z veřejných prostředků
- **konkurenceschopnost** knihovnických služeb v éře Googlu a dalších vyhledávacích služeb
- **šetření času** uživatelů knihoven
- **zjednodušení** vyhledávání
- **integrace informačních zdrojů** různých druhů a povahy

o to se snažily již paralelní vyhledávače jako je MetaLib

web scale discovery

- služba schopná **rychle** a **souvisle** vyhledávat ve velkém rozsahu předsklizeného a indexovaného obsahu, poskytovat výsledky a řadit je dle relevance a filtrovat podle velkého množství kritérií
- obsah centrálního indexu sestává z kombinace sklizených metadat a zaindexovaných plných textů
 - lokální katalog knihovny
 - repozitář či repozitáře
 - licencované EIZ
 - open access tituly

Dokumenty různé povahy

- elektronické i tištěné monografie a seriály
 - články z repozitářů
 - elektronická skripta
 - multimediální obsah (netextové zdroje)
 - ???
-
- trvalé stažení plného textu
 - zprostředkování výpůjčky dokumentu
 - zprostředkování digitální kopie

Centrální index

- odráží **licenční smlouvy** uzavřené mezi producentem discovery systému (a správcem indexu) a producentem licencovaného obsahu
- uspořádaná **kolekce sklizených metadat**, umožňující aktivování informačních zdrojů přístupných knihovně
- žádný centrální index bohužel nedisponuje metadaty všech existujících databází ani všech open access zdrojů – kámen úrazu celého konceptu
- existuje tlak ze strany producentů discovery systémů na sjednocení platformy discovery s dodavatelem obsahu
- částečné řešení:
 - srovnávací analýza PŘED pořízením discovery systému
 - nasazení linkovacího nástroje - umožní odkazovat na dodavatele obsahu nezávisle na nastavení discovery systému

Funkce rozhraní discovery

- jednotné **vyhledávací okénko**
- **detekce překlepů** v rámci vyhledávání
- **live search** (náповěda při psaní dotazu)
- **řazení** výsledků dle relevance
- nastavitelné **facety** pro navigaci (omezování pomocí filtrů)
- **obohacení** výsledků (obálky knih, náhledy, recenze)
- **doporučovací služby** (uživatelé hledali také ...)
- **sociální funkce** (komentáře, hodnocení, sdílení)
- mobilní **responsivní** rozhraní (nejen počítače, ale i mobilní telefony, tablety a televize...)

Rizika

- obavy knihovníků
 - investovali čas a prostředky do správy stávajících katalogů
- obavy producentů obsahu
 - obsah EIZ je prohledáván a zpřístupňován mimo nativní rozhraní jejich systému – nemají kontrolu nad procesem vyhledávání
 - nasazení discovery může mít negativní vliv na využívanost EIZ, pokud discovery upřednostňuje obsah či formu dodání od poskytovatele svázaného s producentem systému (není neutrální relevance)

Služby tradičních katalogů v rozhraní discovery systému

- Tradiční OPAC katalogy (knihovní systémy):
 - Rezervace, správa výpůjček, správa finančního konta, alertové služby, RSS kanály, upravený vyhledávací profil uživatele, funkce pro sdílení záznamů, kooperace ve skupinách, databáze uživatelů, služby dodávání dokumentů, generování citací ... atd
- Discovery systémy MOHOU disponovat službami které mohou výše uvedené funkce katalogů nahradit (ale nemusí)
- Discovery systémy disponují API rozhraními které umožňují rozšířit spolupráci s OPACy

Komerční discovery systémy a jejich indexy

- **EDS – EBSCO Discovery Service** od producenta databází EBSCOhost, služba hostována na serveru
- **Serials Solutions Summon** – také hostovaná služba na serveru, od producenta databází ProQuest
- **Ex Libris Primo** – běží na lokálním serveru i jako hostovaná služba
- **OCLC WorldCat Discovery Service**



Opensource discovery systémy

- **VuFind** – beta.knihovny.cz, cistbrno.cz...

vyvíjí Villanova Uni: <http://vufind.org>

- **Blacklight** – např. Stanford Uni, Yale...

<http://projectblacklight.org>

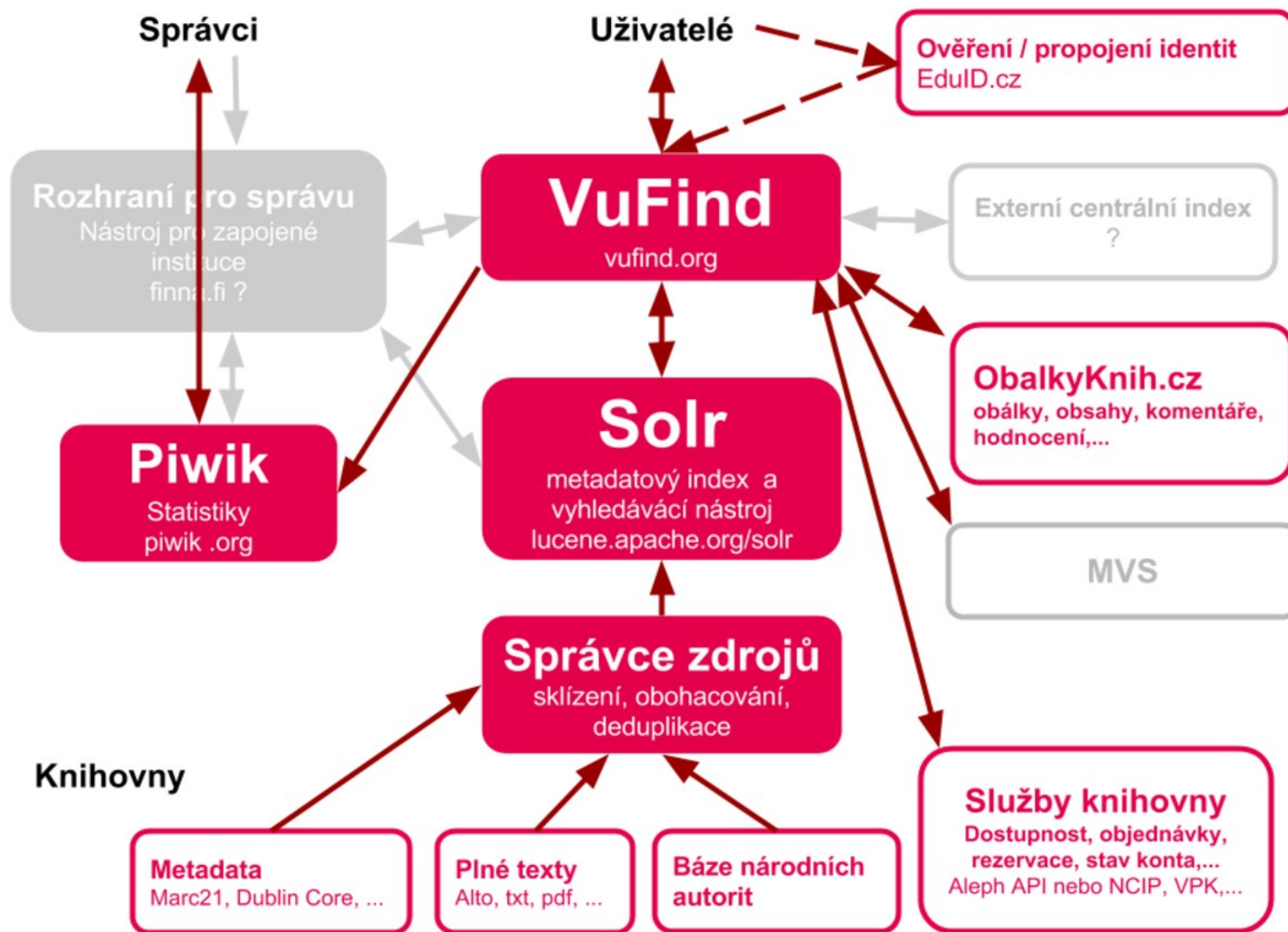
- **Xerxes** – např. UTB Zlín (index Summon)
<http://portal.k.utb.cz/>, vyvíjí California State University

- využívají **interní** index – výsledek vlastního sklizení

- **chybí** centrální index, ovšem lze napojit na komerční centrální indexy systémů v předchozím slajdu prostřednictvím API

Architektura prototypu CPK založeného na VuFind

z prezentace B. Stoklasové
CPK, mojeID/eduID aneb Jak to celé funguje (knihovny.cz)



Instalace discovery služeb

<http://librarytechnology.org/repository/item.pl?id=21096>

System	2007	2008	2009	2010	2011	2012	2013	2014	Installed
EBSCO EDS							1774	2634	8246
Primo	12	37	53	506	111	101	98	88	1528
AquaBrowser	55	339	64	69	74	58	81	6	89
Encore	72	72	109	56	72		36		346
BiblioCommons							41		~200
Summon			50	164	214	158	238	195	697 *)
Enterprise		16		75	100	102	123	150	538
Infor Iguana								18	74
Axiell Arena			61	57	33		35	95	404

*) dle Vladimíra Karena (AiP) v roce 2015 bylo v provozu 1950 instalací Summonu

Spojení producentů discovery systémů

- 6.10.2015 bylo oznámena akvizice společnosti Ex Libris (Primo Discovery) společností ProQuest (Summon Discovery).
- Není jasné zda bude pokračovat vývoj obou systémů (a indexů) paralelně.
- počty zaměstnanců obou společností:

	vývoj	podpora	prodej	admini strativa	ostatní	celkem
ProQuest	224	107	75	63	11	480
Ex Libris	208	234	65	47	11	565
PQ a ExL součet	432	341	140	110	22	1045

Open Discovery Initiative

- projekt pod záštitou NISO
- cílem je prozkoumat a **zlepšit prostředí discovery služeb, vyzvat poskytovatele obsahu** k zapojení, větší transparentnosti a zavést doporučené předpisy, které povedou ke spokojenosti všech zúčastněných stran
- spolupracují knihovny, vydavatelé i producenti discovery systémů
- výstup:
whitepaper **The Future of Library Resource Discovery**
(autor Marshall Breeding) – únor 2015



Výzvy do budoucna

- mnoho zdrojů stále **není zahrnuto v indexech**
- schází lepší pokrytí **open access** obsahu
- nedostatky v podpoře obsahu a rozhraní v **jazycích** nepoužívajících latinku
- nutno zlepšit **transparentnost** v algoritmech stanovování **relevance** a řazení výsledků
- nedostatky v získávání **netextového obsahu**
- záhodno vylepšit **integraci s LMS** (moodle aj).

K zamyšlení I

- je třeba, abychom si neustále připomínali, že **uživateli** je zcela jedno, zda knihovna užívá ten či onen systém, zda data ukládá v DC či MARC, zda jsou plné texty uloženy v knihovně či zda se k nim přistupuje vzdáleně, že jsou zdarma či licencované, vytvořené v digitální podobě či digitalizované - a je zcela **škodlivé** jej těmito informacemi **zatěžovat**.
- Tyto **rozdíly** by mu měly být zcela neviditelné a nezřejmé. Pouze pokud mají rozdíl vliv např. na **komfort** jeho práce či **dostupnost** dané služby, pouze pak by měl být uživatel informován o tom, že dokument A je dostupný až za x dní (díky nutnosti dovozu z deponitáře), avšak již nyní existuje podobný dokument B, který dost možná jeho potřeby uspokojí.

K zamyšlení II

- Knihovny by měly těžit z faktu, že uživatelé **nevyhledávají informace pro samotný akt vyhledávání**, ale že informaci skutečně potřebují. Vyhledávání na Internetu může často vést k frustraci, neboť fyzické získání vyhledaného i přes všechny možné archivy není zaručeno. Pomocná ruka virtuálního knihovníka by měla být k dispozici.
- Knihovny mají tu **schopnost efektivně vyhledávat** existující zdroje, skryté běžným uživatelům i běžným robotům vyhledávačů. Je nutno tuto schopnost efektivně nabídnout jako svoji přednost.
- Jasně vyjádření **dostupnosti** nabízeného dokumentu či zdroje je přímo povinností knihovny, která chce v konkurenci uspět.

K diskusi

- Může existovat velká akademická knihovna programově bez discovery služby?

Ano: **Utrecht University Library**

- Je Google Scholar discovery systém?

Děkuji za pozornost

Petr.Novak@ruk.cuni.cz

čerpá jsem z:

<http://librarytechnology.org/repository/item.pl?id=21096>