

Byli jsme první aneb Implementace discovery systému Primo na Univerzitě Pardubice

Blanka Jankovská

Univerzitní knihovna Univerzity Pardubice

blanka.jankovska@upce.cz

INFORUM 2013: 19. ročník konference o profesionálních informačních zdrojích
Praha, 21. - 22. 5. 2013

Abstrakt

První discovery systém nasazený v ČR do ostrého provozu byl financován z projektu "Univerzitní IT pro vzdělávání a výzkum - UNIT" (reg. č. CZ.1.05/4.1.00/04.0134) v rámci operačního programu Výzkum a vývoj pro inovace. Vítězem otevřeného výběrového řízení se stalo Primo společnosti ExLibris. Systém je na Univerzitě Pardubice (dále UPa) v plném provozu od 1. 10. 2012. Příspěvek mapuje osm měsíců od podpisu smlouvy, přes implementaci systému, testování až k současnému plnému využívání. První srpnová jednání s dodavatelem firmou MULTIDATA byla orientována hlavně na tvorbu lokálního indexu (indexace záznamů z knihovního katalogu a záznamů a plných textů z Digitální knihovny UPa), optimalizaci centrálního indexu pro prostředí UPa (aktivace elektronických informačních zdrojů předplacených UPa, napojení na linkovací server 360 Link), zapojení Prima do univerzitní infrastruktury (vizuál, autorizace a autentikace, Shibboleth, vazba na Konto čtenáře v knihovním katalogu). V průběhu testování pak byly přidány doplňky jako bX Hot Articles a Altmetric. Součástí příspěvku je popis jednotlivých kroků, překonaných překážek a dobrých zkušeností. V neposlední řadě jsou zmíněny i formy propagace (aktuality na zaměstnaneckém a studentském intranetu, komiksový strip, Facebook, články ve Zpravodaji UPa) a první reakce a odezvy ze školení uživatelů (knihovníků, akademiků, studentů). Šest měsíců oficiálního používání přineslo i zajímavá data v uživatelských statistikách. Závěrem je tedy provedena analýza těchto statistik s konkrétním zaměřením na přístupy celkem a v jednotlivých hodinách dne, nejvyhledávanější slova a dotazy s nulovým výsledkem.

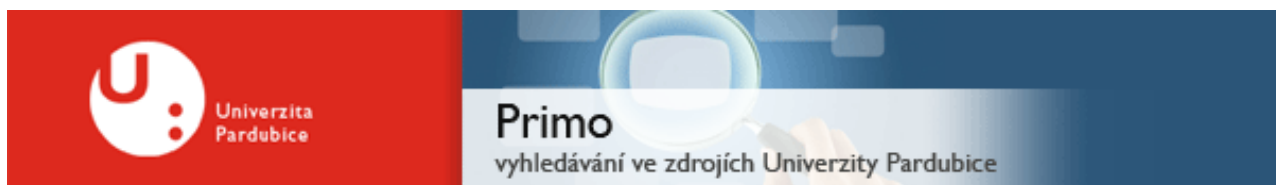
Jak to začalo a kde jsme na to vzali

V úterý 31. července 2012 byla podepsána smlouva mezi Univerzitou Pardubice (UPa) a firmou MULTIDATA Praha, s. r. o. a mohla tak začít implementace prvního discovery systému na půdě vysoké školy v ČR. Nákup systému byl financován v rámci operačního programu Výzkum a vývoj pro inovace, a to z projektu "Univerzitní IT pro vzdělávání a výzkum - UNIT" (reg. č. CZ.1.05/4.1.00/04.0134). Vítěz veřejné zakázky (nadlimitní otevřené řízení) byl jasný až napodruhé, protože při prvním výběrovém řízení se vyskytly v nabídkách nesrovnalosti a řízení bylo ukončeno bez vyhlášení vítězného dodavatele. Jako první hodnotící kritérium byla stanovena ekonomická výhodnost nabídky, tzn. cena. Druhé kritérium bylo rozděleno do dvou subkritérií – jak jsou v centrálním indexu pokryta metadata, resp. plné texty elektronických informačních zdrojů klíčových pro UPa. Ve středu 1. srpna 2012 tedy začal osmítýdenní implementační proces.

Jednou z podmínek účasti ve výběrovém řízení bylo dodání systému v režimu SaaS (Software as a Service, Software jako služba). Hardwarové nároky a povinnosti jako provoz, správa a údržba jsou na straně společnosti ExLibris a MULTIDATA. UPa má fungování Prima pouze předplaceno, a to na pět let.

Jak to pokračovalo...

První implementační schůzka proběhla v pátek 3. srpna 2012. V tu dobu již fungovala testovací instalace Prima vyhledávajícího nad centrálním indexem [Primo Central Index](#). Účastníci schůzky se shodli, že se pro Primo nebude vymýšlet nový název s vazbou k instituci, že se bude systém prezentovat pod původním názvem a jméno serveru bude zapamatovatelné [primo.upce.cz](#). Mezi první úkoly pro univerzitní knihovnu a univerzitní IT pracovníky patřilo zajištění kódu pro sledování Google Analytics, vytvoření logolinku odpovídajícího Jednotnému vizuálnímu stylu UPa (viz obr. 1), zaslání rozsahu IP adres a vygenerování SSL certifikátů.



Obrázek 1 Logolink v hlavičce stránky

Diskuze ohledně datových zdrojů měla tři části – vytvoření jedné části lokálního indexu z knihovních MARC záznamů, druhé části z metadat a plných textů Digitální knihovny UPa a propojení linkovacího serveru 360 Link s Primem tak, aby Primo ve výsledcích vracelo záznamy s propojením na dostupný plný text univerzitou předplaceného dokumentu.

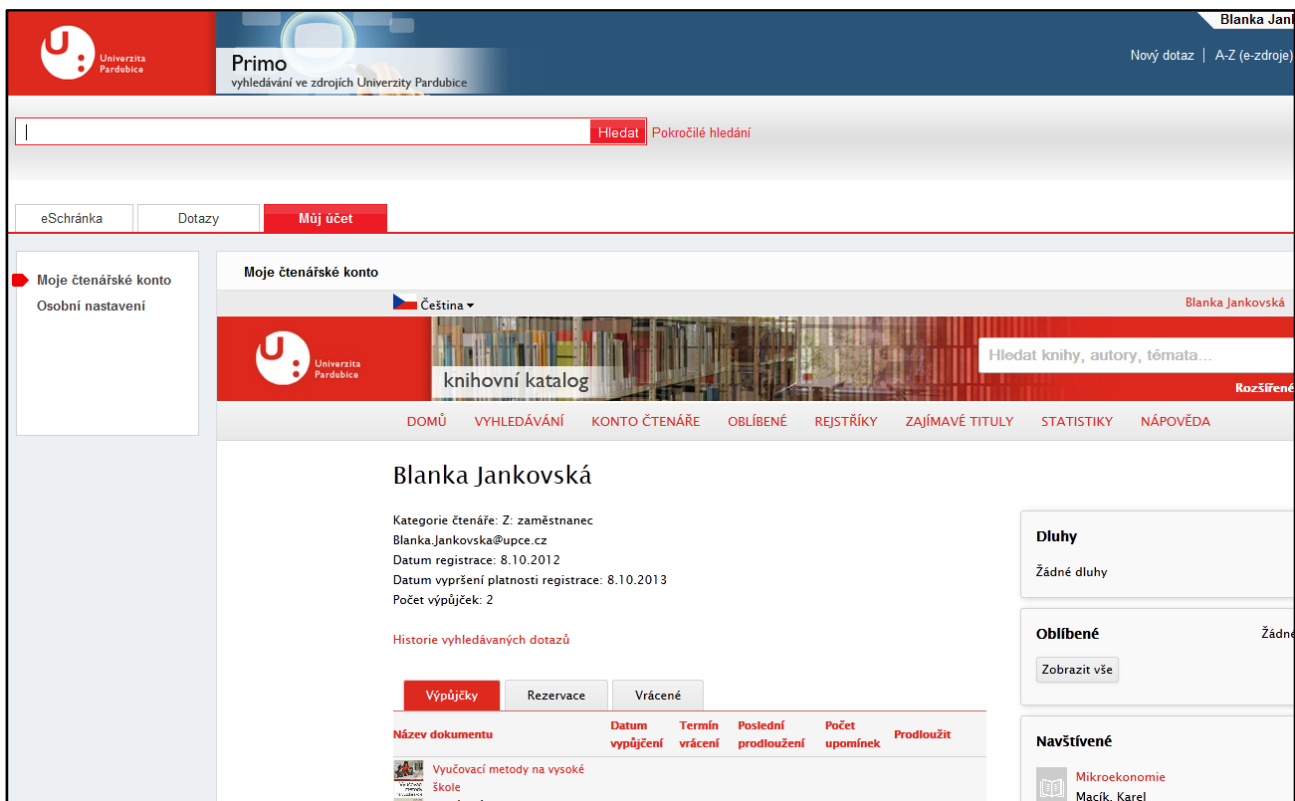
[Univerzitní knihovna UPa](#) používá knihovní systém Verbis a [online katalog](#) Portaro společnosti KP-sys. Pro první testování byl proveden jednorázový export všech MARC záznamů včetně exemplářových údajů. Z důvodu zautomatizování dodávky změn v záznamech a zajištění dalších kontinuálních exportů bylo rozhodnuto, že knihovna zakoupí u KP-sysu licenci na OAI-PMH provider. Aktualizace knihovních záznamů zobrazovaných katalogem Portaro tak od poloviny září probíhá v Primu díky protokolu OAI-PMH.

Stejný protokol naplnil i druhou polovinu lokálního indexu, a to metadaty a posléze i plnými texty z [Digitální knihovny UPa](#). Zde proběhlo stažení metadat ve formátu OAI_DC úspěšně již napoprvé a postupně byly laděny jen detaily. Pokud šlo o indexaci plných textů kvalifikačních prací, bylo nutné uzavřít smlouvu o ochraně dat, protože ne všechny práce byly veřejně přístupné. V digitální knihovně se nacházejí i takové, ke kterým mají přístup pouze interní uživatelé z UPa. Smlouva byla v polovině září uzavřena (týkala se všech plných textů dokumentů uložených v digitální knihovně a online katalogu) a do lokálního indexu byly tyto plné texty zindexovány.

Napojení na Prima na linkovací server [360 Link](#) a zajištění pravidelné dostupnosti holdingových informací se proti všem očekáváním nakonec projevilo jako problematické. Firma Serials Solutions (poskytovatel 360 Linku) odmítla pravidelně generovat aktuální data o rozsahu předplacených informačních zdrojů. Překážka byla překonána díky vstřícnosti pracovníků MULTIDATA. Na serveru MULTIDATA je totiž momentálně provozována „neplnohodnotná“ instance linkovacího serveru SFX, která stínuje holdingové informace z 360 Linku. Prvním krokem tedy bylo vyexportování všech holdingů z 360 Linku a jejich import do SFX. Jakákoli změna v předplatném pak znamená její promítnutí jak do 360 Linku, tak nahlášení jejího rozsahu správci stínového SFX. Samozřejmě si všechny strany uvědomují dočasnost tohoto řešení.

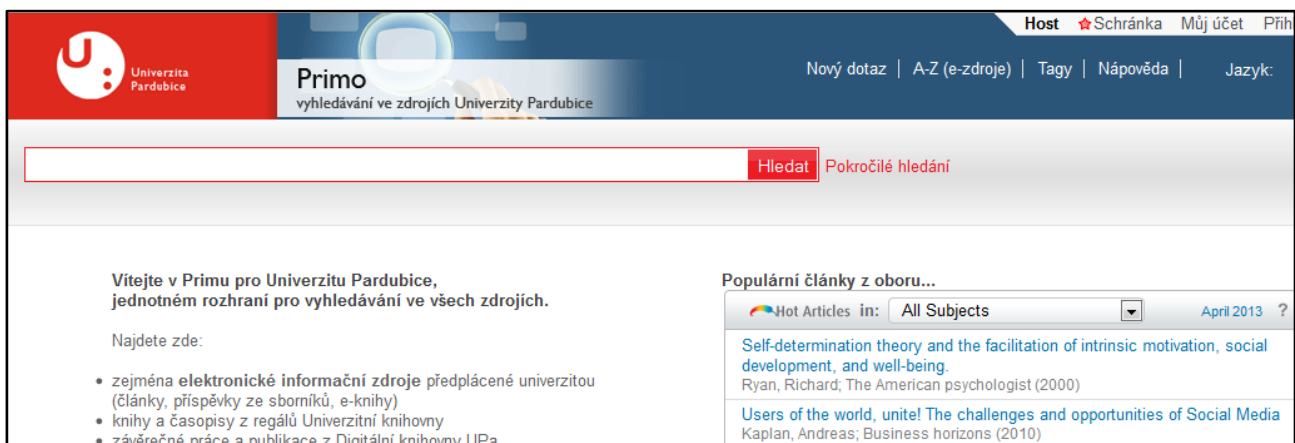
Na závěr první implementační schůzky bylo ještě diskutováno uživatelské přihlašování, zřízení e-mailové konference, konání úvodního školení pro knihovníky a předpokládané spuštění nejdříve do testovacího a poté do ostrého provozu.

Do 7. září proběhly ještě další tři schůzky. Postupně se realizovaly výše naznačené kroky. Pro přihlašování uživatelů byla využita technologie [Shibboleth](#) (IdP UPA) a došlo k propojení uživatelského účtu v Primu s Kontem čtenáře v knihovním katalogu (viz obr. 2). V Primu už šla provádět plnohodnotná vyhledávání, proto se práce pracovníků knihovny odpovědných za implementaci zaměřila hlavně na sestavení požadavků na změny v uživatelském zobrazení (barvy, fonty, ikony, návěští faset, mapování polí záznamů z digitální knihovny a další dílčí úpravy).



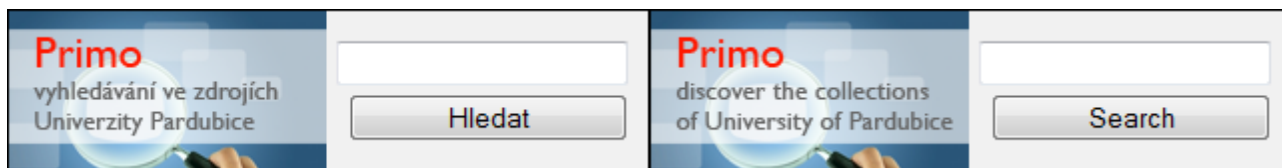
Obrázek 2 Propojení účtu v Primu s Kontem čtenáře v online katalogu

Dále byla na domovskou stránku Prima přidána služba [bX Hot Articles](#) (doporučování aktuálních článků z různých oborů), (viz obr. 3).



Obrázek 3 bX Hot Articles

Pro externí a intranetové stránky knihovny byly vytvořeny vyhledávací okénka (widgety) v české a anglické verzi, které umožňují přímé vyhledávání v Primu (viz obr. 4).



Obrázek 4 Vyhledávací okénka

V neposlední řadě bylo Primo propojeno s Altmetric.com, webovou službou, která díky DOI sleduje ohlas článků, jejich diskutovanost na sociálních sítích (Twitter, Facebook, Mendeley, Connotea ad.) (viz obr. 5).

Blueprint for antimicrobial hit discovery targeting metabolic networks
Shen, Y ; Liu, J ; Estiu, G ; Isin, B ; Ahn, Y-y ; Lee, D-s ; Barabási, A-I ; Kapatral, V ; Wiest, O ; Oltvai, Z N
Proceedings of the National Academy of Sciences of the United States of America, 2010, Vol.107(3), pp.1082-7 [Recenzovaný časopis]
Advances in genome analysis, network biology, and computational chemistry have the potential to revolutionize drug discovery by combining system-level identification of drug targets with the atomistic modeling of small molecules capable of modulating their activity. To demonstrate effectiveness of such a discovery pipeline, we deduced common antibiotic targets in Escherichia coli and Staphylococcus aureus by identifying tissue-specific or uniformly essential metabolic reactions in their metabolic networks. We then predicted through virtual screening dozens of inhibitors for several enzymes of these reactions and showed experimentally that a subset of these inhibited both enzyme activities in vitro bacterial cell viability. This blueprint is applicable for any sequenced organism with high-quality metabolic reconstruction and suggests a strategy for strain-specific anti-infective therapy.
[Je k dispozici plný text](#)

[Online zdroj](#) [Podrobnosti](#) [Recenze a tagy](#) [Další služby](#) [Odeslat e-mailem](#) [Ohlas](#)

Altmetric.com

7

- Blogged by 1
- Subject of a Research Highlight
- 61 readers on Mendeley
- 1 reader on Connotea
- 12 readers on CiteULike

Obrázek 5 Metriky Altmetric v Primu

Školení knihovníků se uskutečnilo v pátek 21. září 2012. Během posledního týdne implementace se poté ještě doladřovaly drobné nedostatky týkající se denní aktualizace dat online katalogu, zobrazování aktuální dostupnosti exemplářů z katalogu, přihlašování uživatelů, sdílení záznamů a zaslání záznamu e-mailem ze stručného zobrazení.

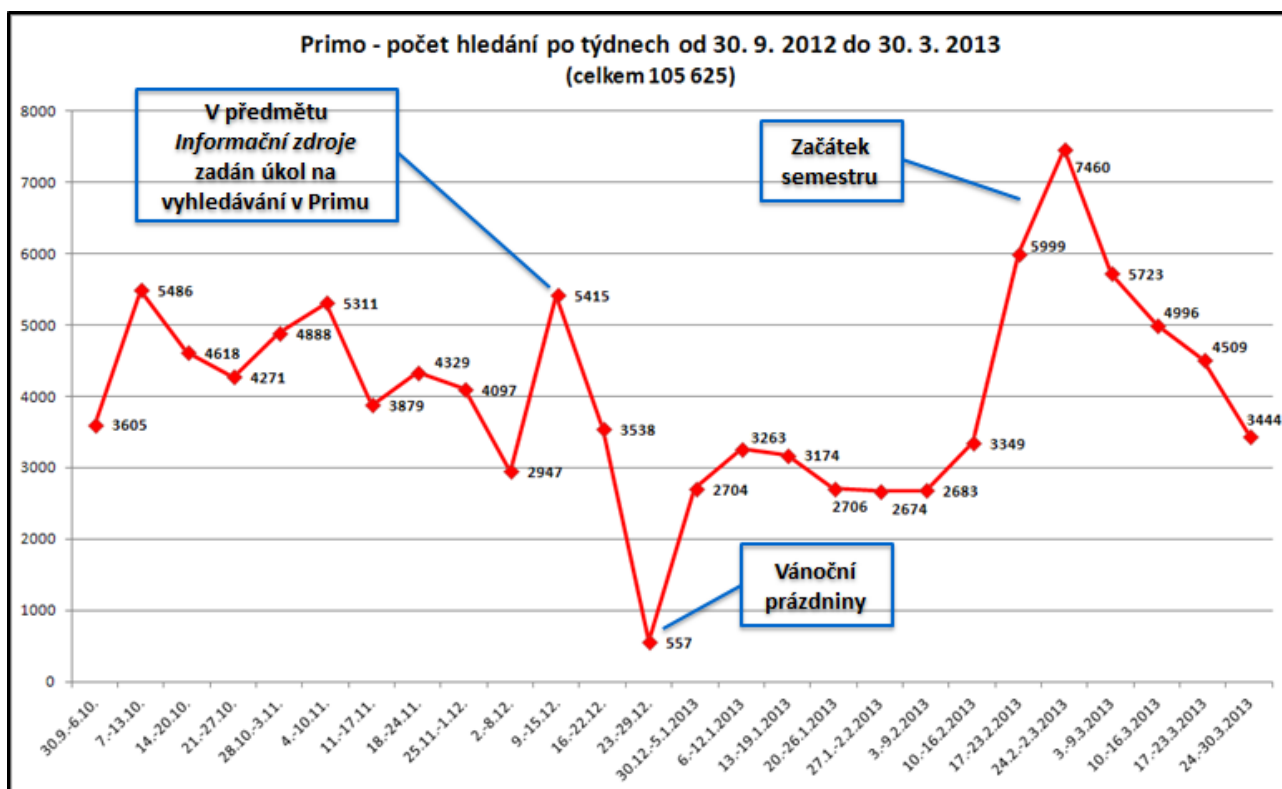
Ve středu 26. září 2012 byla oficiálně ukončena implementace, potvrzeno poskytnutí přístupu, které vyžadovala smlouva a zahájen 30denní testovací provoz. Během něho se knihovníci soustředili především na pátrání po případných nesrovnalostech v detekci dostupnosti plných textů a exemplářů z knihovního katalogu. Všechny nalezené chyby nebyly nijak zásadní a byly průběžně, často i obratem odstraněny. V pondělí 1. října 2012 tak mohl být zahájen bezproblémový ostrý provoz.

Ajedeme naostro...

Během prvního říjnového týdne (1. – 7. 10. 2012) navštívili stránky primo.upce.cz 623 unikátní návštěvníci (77 % z nich tvořili noví příchozí, 23 % se opakovaně vracelo), celkem uskutečnili 798 návštěv, zobrazili si 4 281 stránku a průměrná doba trvání jejich návštěvy byla 4 minuty a 54 vteřin (zdroj: Google Analytics). Tito uživatelé pak zadali 3 474 vyhledávací dotazy, 1 053x klikli na vyhledávání pomocí jednoduchého vyhledávacího rozhraní, 102x klikli na vyhledávání pomocí pokročilého vyhledávacího rozhraní, 454x klikli na další stránku, poslali 10 záznamů e-mailem, 56x se přihlásili do svého účtu a 52x přidali záznam do eSchránky (zdroj: Primo Click Through Statistics).

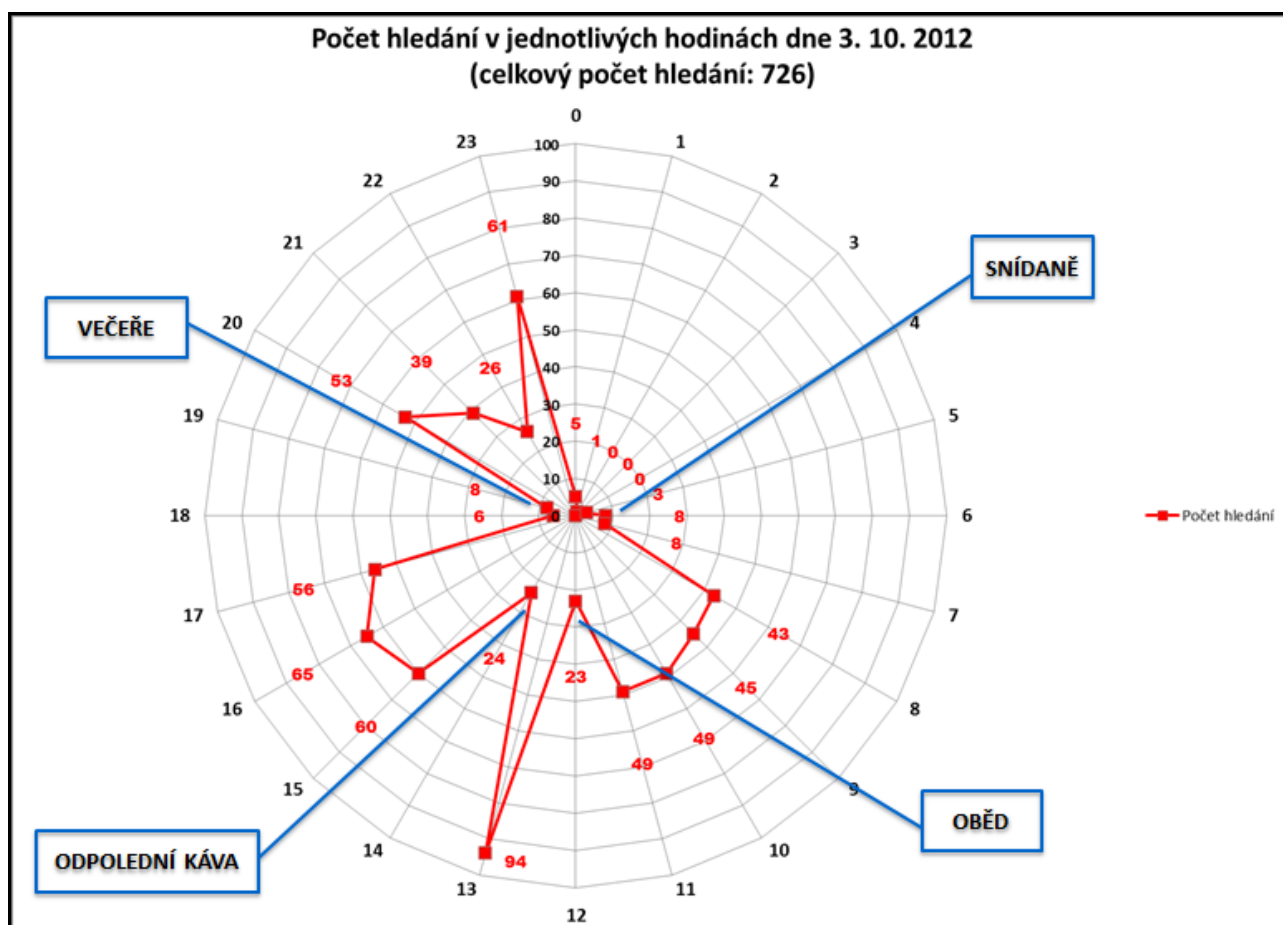
Po šesti měsících (1. 10. 2012 – 31. 3. 2013) používání se statistiky změnilo takto: 6 695 unikátních návštěvníků (53 % nově příchozích, 47 % vracějících se), 12 551 návštěva, 74 953 zobrazení stránek a průměrná doba trvání návštěvy se zvýšila na 6 minut a 27 vteřin (zdroj: Google Analytics); 105 625 vyhledávacích dotazů, 21 057x klik na jednoduché vyhledávání, 3 010x klik na pokročilé vyhledávací rozhraní, 9 900x kliknuto na další stránku, 273 záznamů e-mailem, 565 přihlášení, 295x přidat do eSchránky (zdroj: Primo Click Through Statistics).

Na grafu (obr. 6) je pak vidět, jak různé události ovlivnily počty vyhledávacích dotazů.



Obrázek 6 Počet vyhledávacích dotazů po týdnech

Na rozložení dotazů v průběhu jednoho dne hned na počátku představení Prima uživatelům (3. října 2012) je dobře rozpoznatelný jejich denní režim (obr. 6).



Obrázek 7 Rozložení vyhledávacích dotazů během jednoho dne

V administrátorském rozhraní Prima a jeho sekci s uživatelskými statistikami lze nalézt i další velmi zajímavá data jako jsou nejčastěji hledané výrazy dle počátečních písmen a přehled dotazů s nulovým výsledkem. Písmeno P dlouhodobě vede „hitparádu“ s výrazy *php*, *print*, *prostitute*, *park and ride...* (3 536 hledání). Následuje ho písmeno M a výrazy jako *marketing*, *management*, *montessori* (2 870 hledání). Na 3. místě žebříčku je S a *semtin*, *stres*, *social marketing* (2 637 hledání). Na 4. a 5. místě D a B s variantami obligátních pojmů diplomová a bakalářská práce (2 024 a 1 839 hledání). Na dalších místech se pak střídají písmena A (*analysis of whisky*), T (*tanec*), C (*carbonic anhydrase*) a K (*kouření*).

Dotazy s nulovým výsledkem jsou výrazy, po jejichž zadání do vyhledávacího pole nenabídne Primo žádný záznam. Na počátku provozu Prima patřilo k problematickým formulacím obecně psaní s překlepy, zadávání URL článku nebo dokonce celé jeho bibliografické citace a v podstatě i česká klíčová slova. Ale tady se nejedná o chybu zadání, pouze není mezi českými zdroji (zastoupenými v indexu menšinově oproti těm cizojazyčným) relevantní záznam odpovídající dotazu. Bylo tedy na knihovně, aby osvětou mezi uživateli vysvětlila funkci discovery systému, jeho výhody i omezení. Postupem času se to částečně podařilo, částečně uživatelé sami zjistili stylem „pokus – omyl“, jak mají dotazy formulovat. Po půl roce používání se tak mezi dotazy s nulovými výsledky vyskytují v převážné většině už jen překlepy.

Příklady dotazů s nulovým výsledkem (některé úsměvné☺): *st12345@student.upce.cz* (studentská e-mailová adresa), *st12345* (identifikátor studenta), *obraz doryana graye ☺*, *modelování a symulace ☺*, *web of science*, *prodejna skript*, *problémy s VPN*, *bakalářské práce*,

prdí hadi ☺, Horáková Jsem si vědoma toho, že má bakalářská práce nemohla vysvětlit a uchopit problematiku ☺, bakalářska práca-Bártová, vydavatelství pingvin ☺, mary hyggins clark ☺, Baskitsko ☺ a trochu vybočující pujčení a předdnašky.

Uvedené dotazy jsme pravidelně sledovali a v případě některých jsme dokázali navést uživatele na správné místo použitím funkce „Did you mean“. Do vnitřních vyhledávacích algoritmů Prima se přidala určitá slova (tzv. Hotwords) a definovala se reakce Prima na jejich zadání. V případě zadání názvu databáze se nabídne: *Chtěli byste přejít rovnou do Web of Science?*, přičemž název je prolinkován rovnou do rozhraní Web of Knowledge. Pokud je zadán dotaz na prodejnu skript, zobrazí se: *Hledáte prodejnu skript?*, opět s linkem na stránky prodejny. Podobně jsme přistoupili i k dotazům typu bakalářská či diplomová práce. Byly nadefinovány spisovné české varianty těchto slovních spojení v množném i jednotném čísle a po jejich zadání se Primo ptá: *Chtěli byste přejít rovnou do Digitální knihovny UPa?*, název samozřejmě linkuje na web digitální knihovny. A proč odkazujeme uživatele do nativního rozhraní, když jsou kvalifikační práce indexovány v lokálním indexu? Metadatový standard Dublin Core, resp. jeho nekvalifikovaná verze, ve které jsou do Prima stahovány záznamy o těchto pracích, v sobě totiž neobsahuje prvek „typ práce“. Na první pohled tak v Primu nelze rozeznat bakalářskou práci od diplomové. Výjimkou je, pokud se zmínka o typu nachází v abstraktu. Takové práce Primo samozřejmě nabídne v přehledu výsledků.

Dotazy typu *pujčení a předdnašky* jsem uvedla, protože nejsme bohužel schopni určit, co vůbec uživatelé hledali, co měli na mysli. Můžeme se jen dohadovat. Jsou naštěstí velmi ojedinělé, a proto je nijak neřešíme.

Co na to uživatelé?

Primo jsme se snažili propagovat na všech frontách. Na stránkách knihovny, na studentském i zaměstnaneckém intranetu, v univerzitním zpravodaji, na [facebookovém profilu knihovny](#), prostřednictvím komiksových stripů (viz obr. 8 a 9). Průběžně jsme vypisovali termíny hodinového školení, kde bylo Primo stručně představeno a kde se uživatelé dozvěděli užitečné tipy pro vyhledání relevantních dokumentů.

Potřebovali byste se naklonovat, když hledáte v informačních zdrojích?

Jde vám hlava kolem ze všech těch databází, katalogů a digitálních knihoven?

Přesně pro vás máme PRIMO!

Jednotné vyhledávací rozhraní pro všechny informační zdroje Univerzity Pardubice

Primo
vyhledávání ve zdrojích Univerzity Pardubice

články, knihy, časopisy, závěrečné práce, příspěvky ze sborníků

primo.upce.cz

Univerzita Pardubice

Primo
vyhledávání ve zdrojích Univerzity Pardubice

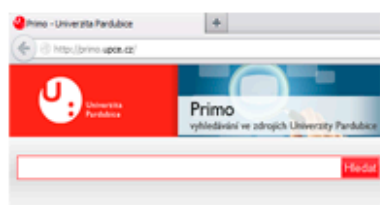
Obrázek 8 První komiksový strip

Akademický rok 2012/2013 a nové přírůstky v Univerzitní knihovně Pardubice

Univerzitní "Google" NA WEBU

Primo

od 1. 10. 2012



Vítejte v Primo pro Univerzitu Pardubice, jednotném rozhraní pro vyhledávání ve všech zdrojích.

Například zde:

- zejména elektronické informační zdroje přetřepávané univerzitou
- články, přílohy ze časopisů, e-knihy
- knihy a časopisy z regionální univerzitní knihovny
- závěrečné práce a publikace z Digitální knihovny LIFE

Primo se zatím nachází v testovací provozu, a proto uvítáme vaše připomínky a návrhy na zlepšení. Pište na banka.primo-usa@ucpa.cz

"Karkulka" PŘED KNIHOVNOU

Bibliobox

od 6. 3. 2013



Miminko NA DUKLE

Alice

od 3. 3. 2013



Obrázek 9 Druhý komiksový strip

Místo závěru...

Práce s Primem je velmi jednoduchá a intuitivní, prostředí je uživatelsky přívětivé, takže na sebe pozitivní reakce nenechaly dlouho čekat. Zde jsou dva názory za všechny.

Cit. „...s velkou radostí jsem vyzkoušel nový portál Primo. Chci Vám vyjádřit velkou spokojenost s tímto vyhledávačem, který je krásně propojený s osobním účtem v knihovně...“ (student 2. ročníku Fakulty chemicko - technologické).

Cit. „Gratuluji k discovery systému! :-) Bude slyšet z knihovny bouchání šampusu? ;-) Vypadá to dost dobře - všechny zdroje pohromadě, rychlá odezva, budu ho často používat. Stejně, běžný uživatel vůbec nemá ponětí, kolik času a energie je v tom utopené. Jen přijde k hotovému...“ (akademický pracovník Fakulty chemicko - technologické).

I knihovníci si Primo velmi pochvalují, protože se informace o tom, zda je kniha volná k vypůjčení a kde se v regálech nachází, zobrazí okamžitě již v seznamu výsledků, kdežto při vyhledávání v online katalogu je nutno se k těmto informacím proklikávat.

A samotná implementace? Až na drobnosti proběhla hladce, bez problémů a včas. Jsme tedy velmi spokojenou první univerzitní knihovnou v České republice užívající a svým „klientům“ nabízející discovery systém.