

## Zpráva o plnění cílů projektu VISK 8/B

### 1. Název projektu:

*Zajištění provozu Jednotné informační brány v celonárodním měřítku a zpřístupnění informačních zdrojů Národní knihovny České republiky, zpřístupnění informačních zdrojů Národní knihovny České republiky a českých webových zdrojů*

### 2. Název a adresa žadatele/předkladatele zprávy

Národní knihovna České republiky, Klementinum 190, 110 00 Praha 1

### 3. Hodnocení plnění cílů projektu – tabulka + text

#### Tabulka

Cíl	realizováno (termín)	dostupnost výsledku (místo)	nerealizováno (důvod)	poznámka
Zajištění provozu Jednotné informační brány (JIB) v celonárodním měřítku	2004	<a href="http://www.jib.cz">http://www.jib.cz</a>		
Zpřístupnění informačních zdrojů Národní knihovny ČR	2004	<a href="http://www.jib.cz">http://www.jib.cz</a>		
Zpřístupnění významných českých webových zdrojů	2004	Báze WEB – katalog NK ČR; info viz server webarchiv <a href="http://sigma.nkp.cz/cze/web">http://sigma.nkp.cz/cze/web</a> <a href="http://webarchiv.nkp.cz">http://webarchiv.nkp.cz</a>		
Zahájení budování tematických bran v ČR	2004	<a href="http://jib-info.cuni.cz/dokumenty/dokumenty.php.iso-8859-2?sekce=tematbrany">http://jib-info.cuni.cz/dokumenty/dokumenty.php.iso-8859-2?sekce=tematbrany</a>	*	*
Optimalizace věcného zpřístupnění s ohledem na integraci českých zdrojů do mezinárodního kontextu	2004	**		

\*

V návaznosti na schválení prostředků na testování a rozvoj standardů pro budování oborových informačních bran v ČR částečně v rámci výzkumného záměru NK a částečně v rámci projektu 1N a naopak neschválení technických prostředků nezbytných pro budování oborových informačních bran, byl z prostředků VISK 8/B financován pouze nákup dvou samostatných SFX instancí nezbytných pro poskytnutí diferencovaných přidaných služeb, nikoli testování a rozvoj knihovnických standardů pro budování oborových informačních bran v ČR.

\*\*

Ve spolupráci byly vytvořeny aplikace schématu Konspektu pro

staré tisky [http://www.nkp.cz/pages/page.php3?page=fond\\_startisk\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_startisk_kospekt.htm)

hudebniny [http://www.nkp.cz/pages/page.php3?page=fond\\_hudeb\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_hudeb_kospekt.htm)

lékařství [http://www.nkp.cz/pages/page.php3?page=fond\\_lekarstvi\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_lekarstvi_kospekt.htm)

Bylo dokončeno ladění konkordance MDT a DDT

[http://www.nkp.cz/pages/page.php3?page=fond\\_Mdt\\_tabulky1.htm](http://www.nkp.cz/pages/page.php3?page=fond_Mdt_tabulky1.htm)

Byly zapracovány potřebné úpravy do schématu Konspektu

[http://www.nkp.cz/pages/page.php3?page=fond\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_kospekt.htm)

Byly připojeny anglické ekvivalenty v autoritních záznamech tematických autorit

[http://sigma.nkp.cz/F?func=file&file\\_name=find-a&local\\_base=AUT10](http://sigma.nkp.cz/F?func=file&file_name=find-a&local_base=AUT10)

## Text

### **Stručné zhodnocení použití grantu:**

**Cíle projektu Zajištění provozu Jednotné informační brány v celonárodním měřítku a zpřístupnění informačních zdrojů Národní knihovny České republiky byly v roce 2004 rozděleny do několika oblastí**

1. Zajištění **provozu a rozvoje** Jednotné informační brány (JIB) v celonárodním měřítku
2. **Zpřístupnění informačních zdrojů** Národní knihovny České republiky
3. **Zpřístupnění významných českých webových zdrojů**
4. **Příprava, testování a rozvoj standardů pro budování oborových informačních bran** v ČR
5. **Optimalizace věcného zpřístupnění** s ohledem na integraci českých zdrojů do mezinárodního kontextu

Všechny dílčí cíle byly splněny.

### **Zajištění provozu a rozvoje JIB**

V roce 2004 došlo díky finanční podpoře v rámci VISK 8/B k rozšíření možnosti využití programu MetaLib pro vyšší počet současně pracujících uživatelů a zároveň legalizaci možnosti jeho využití v národním měřítku.

Hlavní činnosti související s provozem a rozvojem JIB se dělí na činnosti technického rázu, které zajišťují externí pracovníci, a na činnosti knihovnické, které zajišťují převážně pracovníci referenčního centra NK.

### Technické činnosti

Hlavní činnosti správy a rozvoje MetaLibu a SFX v rámci JIB můžeme rozdělit do 5ti základních oblastí:

- Administrace a údržba systémů MetaLib a SFX
- Správa stávajících informačních zdrojů, připojování nových zdrojů a knihoven

- Správa a rozvoj Z39.50 rozhraní MetaLib JIB pro stahování záznamů
- Tvorba informačních stránek, Informačního zpravodaje JIB, školení a prezentace JIB
- Příprava rozšíření JIB o oborové informační brány a konzultace

#### *Administrace a údržba systému MetaLib a SFX*

V roce 2004 nedošlo ke změně verze MetaLibu (po celý rok jsme používali verzi 2), aplikovali jsme ale množství revizí, které opravovaly některé nedostatky nebo přinesly nové funkce a rozšíření. Hlavní rozvoj funkcionality se dočkal modul pro stahování záznamů (Z39.50 server MetaLibu a navazující aplikační část). SFX po celý rok rovněž pracoval ve verzi 2, podobně stejně jako v případě MetaLib byla aplikována řada revizí jak programové částí i znalostní báze. MetaLib a SFX po celý rok pracovaly bez vážných výpadků a jejich provoz lze označit za velmi stabilní.

#### *Správa stávajících informačních zdrojů, připojování nových zdrojů a zapojování knihoven*

V průběhu roku 2004 jsme udržovali v chodu všechny stávající informační zdroje, domácí i zahraniční. Vzhledem k velkému počtu připojených zdrojů docházelo na straně vzdálených zdrojů k výpadkům (nedostupnost Z39.50 serverů atp.), které jsme operativně řešili se správci těchto zdrojů. Tato činnost průběžně představuje poměrně velký objem prací a tvoří hlavní know-how projektu: na jejím výstupu vzniká cenná znalostní báze o zdrojích a jejich nastavení. V současné době nabízíme 157 aktivních zdrojů, z nichž některé kumulují stovky dílčích informačních zdrojů (např. Ebsco a další integrované zdroje). V roce 2004 se podařilo úspěšně připojit několik významných zdrojů, z nichž nejdůležitější jsou National Library of Medicine, California Digital Library, National Library of Canada, HighWire, ÚZPI a katalog STK. Rozpracováno zůstává připojení plnotextového archivu českých článků Anopress, digitální sbírky Manuscriptorium a generátoru citací provozovaném Masarykovou univerzitou v Brně. Byla provedena kompletní revize licencovaných zdrojů podle nových konsorciálních licencí a byly ověřeny a opraveny všechny evidované rozsahy IP adres zapojených institucí. K projektu JIB bylo připojeno osm nových knihoven, které prostřednictvím JIB mohou využívat předplacené informační zdroje.

#### *Správa a rozvoj Z39.50 rozhraní JIB pro stahování záznamů*

V oblasti stahování záznamů přes JIB se podařilo v roce 2004 vyřešit řadu problémů a službu začaly využívat desítky českých knihoven (44 knihoven k 13.12.2004). Byl nasazen konvertor bibliografických záznamů mezi formáty UNIMARC a MARC 21, což umožnilo začít využívat české i zahraniční záznamy bez ohledu na zdrojový a cílový formát. S postupným přechodem knihoven v ČR na formát MARC 21 se stalo přebírání záznamů v JIB pro mnohé knihovny jedinečným nástrojem pro získání záznamů z českých knihoven s odlišným formátem. K dispozici je také konverze znakových sad pro knihovny s AKS nepodporujícími UTF-8. Přebírání záznamů JIB tak může používat knihovna s libovolným AKS. Proces tvorby profilů pro stahování jsme částečně automatizovali pomocí webového generátoru. Modul pro stahování záznamů obsahoval několik vážných chyb, z nichž se nám podařilo ty nejvážnější odstranit aktivní spoluprací s firmou Ex Libris. S nárůstem uživatelů přebírání záznamu vzrostl značně i počet různých dotazů a žádostí o úpravu existujících profilů či přidání nových profilů. Pro včasné informování o novinkách ve službě přebírání záznamů byla založena samostatná elektronická konference [prebirani@cuni.cz](mailto:prebirani@cuni.cz). V té je v současné době registrováno 62 uživatelů služby. V první polovině prosince 2004 byl průměrný počet dotazů na Z39.50 serveru JIB (počítáno na pracovní dny) téměř 800 denně. Dohromady bylo pro české knihovny během roku 2004 vytvořeno 94 profilů.

Důležitým předpokladem správné komunikace Z39.50 v českém prostředí bylo schválení Profilu Z39.50 JIB, který definuje nastavení atributů Z39.50, mapování na UNIMARC/MARC 21 a přináší některá doporučení k indexování. Tento profil jsme navrhli a přednesli účastníkům ZIG na schůzce v únoru 2004. Profil přislíbili dodržovat všichni producenti a distributoři AKS na českém trhu.

#### *Tvorba informačních stránek, školení a prezentace JIB*

Průběžně jsme spravovali informační stránky na <http://jib-info.cuni.cz/>, připravili jsme několik technických dokumentů. Také jsme vydali 2 čísla Informačního zpravodaje JIB (<http://jib-info.cuni.cz/zpravodaj>), který je rozeslán elektronickou poštou uživatelům JIB. Pro rozesílání zpravodaje a pro včasné informování uživatelů o novinkách JIB slouží elektronická konference [jib-info@cuni.cz](mailto:jib-info@cuni.cz), ve které je k 9.12.2004 registrováno 843 uživatelů JIB (počet průběžně kolísá oběma směry – kromě růstu díky registraci nových uživatelů dochází i k úbytku těch, jejichž e-mailová schránka zanikla nebo byla delší dobu nedostupná). Proběhla celá řada školení a prezentací, jejichž soupis i příslušné dokumenty jsou k dispozici na [http://jib-info.cuni.cz/akce/akce.php?akce=ostatni\\_minule](http://jib-info.cuni.cz/akce/akce.php?akce=ostatni_minule).

#### *Příprava rozšíření JIB o oborové informační brány*

Zakoupení dvou dalších SFX instancí v rámci VISK 8/B došlo k zakoupení dvou dalších SFX instancí, které umožní diferenciaci přidanych služeb a zahájení budování dvou oborových informačních bran. Ve druhém kole programu 1N vyhlášeného MŠMT je zajištěno částečné financování odborných knihovnických prací souvisejících především s tvorbou metodiky oborových informačních bran a tvorbou záznamů, nikoli však k zajištění technických předpokladů jejich tvorby. Spojením prostředků získaných v rámci VISK 8/B a 1N tak dochází k účelnému propojení vícezdrojového financování oborových informačních bran. V závěru roku proběhly přípravné práce a analýzy pro rozšíření JIB o oborové informační brány. Byla navržena architektura, byli připraveny instance SFX a proběhly konzultace s odbornými garanty a řešiteli projektu.

#### Knihovnické činnosti

Oddělení referenčních a meziknihovnických služeb zajišťuje v rámci JIB nastavení, rozvoj a propagaci jak zdrojů, které jsou zapojeny do projektu, tak i služeb. Naším cílem je vytvořit v jednotném prostředí JIB takovou nabídku a klasifikaci (rozdělení) zdrojů, která by zabezpečovala uživatelům snadný přístup ke zdrojům ČR, výběrově i k nejvýznamnějším zdrojům světovým. Zároveň se zpřístupněním zkvalitňujeme nabídku přidanych služeb přes SFX prostředky, především v souvislosti s přístupem a získáváním plných textů. Během roku 2004 byla z pohledu služeb vytvořena nová kategorie „České články“, byly zapojeny nové zdroje, zkvalitněna nabídka SFX přidanych služeb a vytvořeny některé dokumenty, které uživatelům JIB slouží jako nápovědy či pomůcky při vyhledávání.

#### *Vytvoření kategorie „České články“*

V této kategorii jsou shromážděny článkové databáze (např. Baze ANL Národní knihovny ČR), regionální článkové bibliografie (např. KVK Liberec, SVK Hradec Králové) a oborové článkové bibliografie (např. Divadelní ústav, Ústav pro českou literaturu AV ČR). Některé zází článků je možné prohledávat přímo v prostředí JIB, jiné jsou pouze odkazové. Cílem vytvoření kategorie „České články“ bylo soustředění článkových bibliografií na jednom místě tak, aby uživatelé co nejnadhěji a nejrychleji získali přehled o existujících databázích.

#### **Zpřístupnění informačních zdrojů Národní knihovny České republiky**

Bohaté informační zdroje NK byly zpřístupňovány uživatelům JIB, elektronické dodávání dokumentů z fondů NK fungovalo v rutinním provozu

### **Zpřístupnění významných českých webových zdrojů**

Z důvodu neefektivního přístupu k datům bylo ukončeno využívání páskového robota NK umístěného na digitalizačním pracovišti v CDH NK pro uložení archivovaných dat a řešitelé přešli na využití diskového pole, které umožní v příštím roce práci s větším objemem dat a bude dále rozšířeno. V polovině března 2004 byla spuštěna celoplošná sklizeň domény cz pomocí programu NEDLIB Harvester. V průběhu prvního pololetí 2004 se velikost celého komprimovaného webového archivu uloženého na diskovém poli v MZK přiblížila hranici 1TB a volné místo na tomto poli se snížilo pod 250 MB a začala tak akutně hrozit kolize s vlastními potřebami MZK. Zároveň začaly od počátku roku vycházet veřejné verze nového volně dostupného archivního crawleru Heritrix (vývoj Internet Archive), který od verze 0.2.0 uveřejněné v lednu dospěl až do listopadové verze 1.2.0, která už se vykazuje značnou stabilitou a mnoha vlastnostmi, které nemá ani starší NEDLIB Harvester. Všechny nové sklizně budou již prováděny výhradně s pomocí systému Heritrix. Při stěhování dat na diskové pole dojde i k jejich konverzi ze staršího formátu tar.gz NEDLIB Harvesteru do novějšího formátu arc.gz, používaného Heritrixem.

V oblasti zpřístupnění archivu se pozornost soustředila především na produkt NWA Toolset, který přímo podporuje formát arc generovaný Heritrixem. Také NWA Toolset letos již vyšel v několika verzích a přes určité problémy s lokalizací se nyní jeví jako vhodný nástroj pro fulltextové prohledávání webového archivu. K umožnění zpřístupňování archivovaných zdrojů z právního hlediska jsou uzavírány smlouvy s vydavateli vybraných online publikovaných zdrojů, poskytujících Národní knihovně ČR oprávnění ke zpřístupňování těchto publikací z digitálního archivu – k datu 30. 11. bylo uzavřeno 20 smluv. Pro správu údajů souvisejících se zpřístupňováním archivovaných dokumentů byl vytvořen komplexní administrativní systém, přístupný řešitelům ke sdílení na serveru webarchiv.

Významné domácí online zdroje jsou popisovány a zpřístupňovány prostřednictvím báze WEB v katalogu NK, kde je v případě zdrojů s povoleným přístupem do archivu zapisována rovněž URL adresa do archivu. Báze WEB by měla v příštím roce být konfigurována tak, aby umožňovala také popis a zpřístupňování zdrojů pro oborové informační brány. Záznamy jsou zde ukládány ve standardním formátu – do poloviny roku 2004 ve formátu UNIMARC, v polovině roku byla provedena konverze a záznamy jsou nyní ukládány ve formátu MARC21. Pro účely zpřístupňování zdrojů v oborových informačních branách budou záznamy konvertovány do formátu Dublin Core.

Všechny popsané výsledky jsou dostupné na <http://www.webarchiv.cz>

**Příprava, testování a rozvoj standardů pro budování oborových informačních bran v ČR**  
V návaznosti na schválení prostředků na testování a rozvoj standardů pro budování oborových informačních bran v ČR částečně v rámci výzkumného záměru NK a částečně v rámci projektu 1N a naopak neschválení technických prostředků nezbytných pro budování oborových informačních bran, byl z prostředků VISK 8/B financován pouze nákup dvou samostatných SFX instancí nezbytných pro poskytnutí diferencovaných přidaných služeb, nikoli testování a rozvoj knihovnických standardů pro budování oborových informačních bran v ČR.

## Optimalizace věcného zpřístupnění

K základním předpokladům **optimalizace věcného zpřístupnění** s ohledem na integraci českých zdrojů do mezinárodního kontextu patří především aplikace metody Konspektu pro jednotlivé typy dokumentů a obohacení tematických termínů o jazykové varianty, především v angličtině.

Ve spolupráci byly vytvořeny aplikace schématu Konspektu pro

staré tisky [http://www.nkp.cz/pages/page.php3?page=fond\\_startisk\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_startisk_kospekt.htm)

hudebniny [http://www.nkp.cz/pages/page.php3?page=fond\\_hudeb\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_hudeb_kospekt.htm)

lékařství [http://www.nkp.cz/pages/page.php3?page=fond\\_lekarstvi\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_lekarstvi_kospekt.htm)

Bylo dokončeno ladění konkordance MDT a DDT

[http://www.nkp.cz/pages/page.php3?page=fond\\_Mdt\\_tabulky1.htm](http://www.nkp.cz/pages/page.php3?page=fond_Mdt_tabulky1.htm)

Byly zapracovány potřebné úpravy do schématu Konspektu

[http://www.nkp.cz/pages/page.php3?page=fond\\_kospekt.htm](http://www.nkp.cz/pages/page.php3?page=fond_kospekt.htm)

Byly připojeny anglické ekvivalenty v autoritních záznamech tematických autorit

[http://sigma.nkp.cz/F?func=file&file\\_name=find-a&local\\_base=AUT10](http://sigma.nkp.cz/F?func=file&file_name=find-a&local_base=AUT10)

Důslednou aplikací těchto předpokladů vzniká integrovaný nástroj pro indexaci a vyhledávání heterogenních informačních zdrojů sloužící k řízenému zpřístupňování informačních zdrojů, který usnadňuje propojení jednotlivých informačních zdrojů a umožňuje univerzální přístup k informačním zdrojům a službám; hierarchická struktura řízeného souboru na úrovni skupin Konspektu slouží jako základ pro tvorbu předmětově orientovaných univerzálních, polytematických i monotematických bran.

## 4. Využití finančních prostředků z dotace

Žádáno: 2 850 000 Získáno: 2 850 000 (200 000 investice, 2 650 000 neinvestice)

Převody mezi položkami:

K převodům mezi různými kategoriemi rozpočtových položek sice nedošlo, oproti plánu se však změnila výše čerpání uvnitř jednotlivých položek, jak je zřejmé z následující tabulky.

Poznámka: Tabulka stručně shrnuje čerpání přidělených prostředků. Výsledky jednotlivých činností jsou popsány výše, detailní pohled na čerpání finančních prostředků přinášejí přiložené účetní doklady.

Cíl	částka v Kč	poznámka
<b>Investice</b>		
Diskové pole pro server JIB	200 000	Řádně zakoupeno o zprovozněno. Přepatek uhrazen z rozpočtu NK
<b>Drobný hmotný majetek</b>		
WebArchiv – rozšíření paměti serverů, periferní zařízení k PC sloužícím projektu	120 000	Pro WebArchiv bylo zakoupeno diskové pole z jiných zdrojů, prostředky na drobný hmotný

		majetek byly využity na zakupení nebo posílení výpočetní techniky a materiálu pro řešení projektu
<b>Služby</b>		
Ex Libris – MetaLib a SFX maintenance, údržba centrální znalostní báze	653 138	Vyšší poplatek byl placen až od srpna a činnosti související s údržbou centrální znalostní báze zajišťovali pracovníci ÚVT a další externí pracovníci, proto došlo k úspoře oproti plánu
Ex Libris – poskytnutí dvou dalších SFX instancí + maintenance	430 735	Navýšení ceny v důsledku nákupu celkového menšího množství SFX instancí v ČR oproti plánu
JIB – provoz a rozvoj – technické a programátorské činnosti	815000	ÚVT + externí pracovníci
WebArchiv – provoz a rozvoj, technické a programátorské činnosti	437 920	Větší rozsah prací – specifikace viz Text
Drobné technické a údržbové práce související s projektem	13 000	
<b>Ostatní osobní náklady</b>		
JIB – provoz a rozvoj – knihovnické činnosti	70 000	Vše bylo řádně provedeno
WebArchiv – provoz a rozvoj – knihovnické činnosti	40 000	Vše bylo řádně provedeno
Optimalizace věcného zpřístupnění	70 000	Vše bylo řádně provedeno

Zprávu připravil/a:

Dr. Bohdana Stoklasová, řešitelka

Zprávu předkládá

Mgr. Vlastimil Ježek, ředitel Národní knihovny ČR