

VEŘEJNÉ INFORMAČNÍ SLUŽBY KNIHOVEN

**Koordinální centrum programu a implementace Koncepce rozvoje knihoven
v České republice**

Obálky knih.cz - rozvoj projektu v roce 2022

**Jihočeská vědecká knihovna v Českých Budějovicích
leden 2023**

ZHODNOCENÍ PROJEKTU

Projekt Obálkyknih.cz sdružuje různé zdroje informací o dokumentech do jedné, snadno použitelné webové služby. Databáze aktuálně obsahuje přes **2,47 miliónu obálek** (nárůst za rok 2022 o cca 160 000 obálek), **793 tisíc obsahů** českých a zahraničních dokumentů (nárůst za rok 2022 o cca 96 000 obsahů) a **15 000** seznamů doporučené literatury (nárůst za rok 2022 o cca 3 000 obsahů). Dále poskytuje přes **660 tisíc anotací, 13,35 miliónu hodnocení u 270 tisíc titulů, 8,5 tisíc komentářů, 60 tisíc fotografií autorit** a cca. **0,9 miliónu vygenerovaných citací** dle normy ISO 690.

API služby projektu využívá většina knihoven v České republice, muzea, archivy, oborové projekty, CPK, aj.

Správce projektu Obálkyknih.cz je Jihočeská vědecká knihovna v Českých Budějovicích a projekt provozuje ve spolupráci s Moravskou zemskou knihovnou v Brně. JVK i MZK do projektu z vlastních zdrojů vkládají nemalé lidské a finanční zdroje. Do spolupráce na projektu se zapojují taktéž desítky knihoven z celé republiky (přispívání obálek a obsahů, kontroly dat, připomínkování vývoje, ...).

Přehled vlastností projektu:

- hlavní servery jsou provozovány v **Jihočeské vědecké knihovně v Českých Budějovicích**, záložní servery jsou v **Moravské zemské knihovně v Brně** a v **Centrum vědecko-technických informací SR v Bratislavě**
- v případě výpadku jednoho ze serverů mají knihovní systémy možnost přejít během několika vteřin na záložní stroj bez ztráty dostupnosti služeb pro své čtenáře
- měsíčně hlavní server odbaví průměrně **50 miliónů** požadavků, cca. **2 milióny denně**, průměrně **20 dotazů za vteřinu**
- ve špičkách (9-15:00) odbavují servery **40-80 požadavků za vteřinu**
- denně do databáze je nově nahráno nebo upraveno průměrně **500 dokumentů**
- denní přírůstek dat činí **8 GB**, z nich se následně generují náhledy obálek v různých rozlišeních, PDF dokumenty s obsahy a rozpoznává se text pomocí OCR
- **20 Mbit za vteřinu** je datový tok ven ze serveru a na server což představuje cca. 85% všech dat, které projdou internetovým připojením JVK

Detailní statistiky exportu dokumentů přes skenovacího klienta za období leden - prosinec 2022:

Počet odeslaných dokumentů	138 105
Počet uložených obálek (COVER)	125 257
Počet uložených stran obsahů (TOC)	144 900
Počet uložených fotografií autorit	1 185
Počet uložených stran seznamů lit.	19 961

Počty odeslaných stran a titulů přes skenovacího klienta dle jednotlivých knihoven (rok 2022):

STRAN	TITULŮ	SIGLA	NÁZEV
42210	26006	CBA001	Jihočeská vědecká knihovna v Českých Budějovicích
23964	11610	OLD012	Knihovna UP
22161	6023	ABA013	Národní technická knihovna

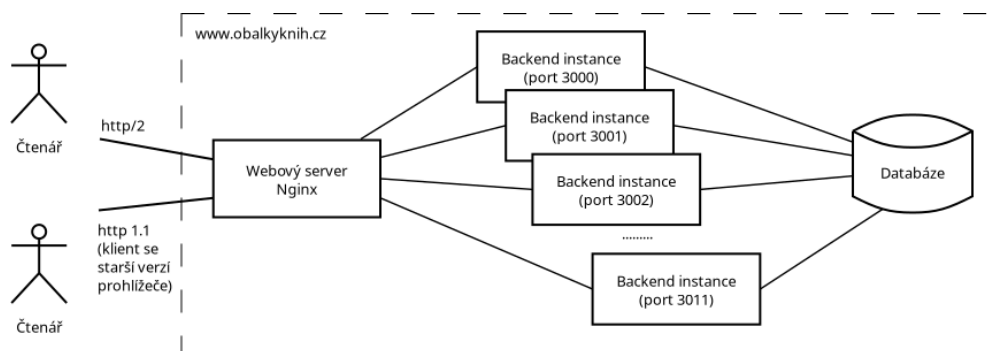
15894	2668	BOD010	Masarykova univerzita - Právnická fakulta
15750	7645	BOA001	Moravská zemská knihovna v Brně
15113	6782	ABA004	Národní knihovna
14150	9114	ABA001	Národní Knihovna ČR
9683	5078	OLA001	Vědecká knihovna v Olomouci
9249	2406	ABB022	Středisko vědeckých informací Fyziologického ústavu AV ČR, v.v.i.
8012	3667	ABD001	Knihovna Filozofické fakulty UK
6990	2923	ABA008	Národní lékařská knihovna
6919	4770	ULG001	Severočeská vědecká knihovna v Ústí nad Labem
5452	1582	ABA006	Vysoká škola ekonomická v Praze
4202	1871	BOE451	Knihovna Biskupství brněnského
3514	1395	BOE020	Knihovna Ústavní soud
3318	1073	ABB019	Knihovna Sociologického ústavu AV ČR, v.v.i.
3067	1472	CBD007	Akademická knihovna Jihočeské univerzity
2992	1993	HKA001	Studijní a vědecká knihovna v Hradci Králové
2971	1054	ABA007	Knihovna AV ČR
2635	604	ABD100	ÚK ČVUT
2579	601	ZLD002	Knihovna UTB ve Zlíně
2512	1199	PNA001	Studijní a vědecká knihovna Plzeňského kraje
2048	1187	LIA001	Krajská vědecká knihovna v Liberci
1928	625	BOD031	Masarykova univerzita, Fakulta sociálních studií, Ústřední knihovna
1870	1223	BOD003	Ústřední knihovna Pedagogické fakulty MU
1780	1442	UHG001	Knihovna Bedřicha Beneše Buchlovana
1666	917	OSG002	Knihovna města Ostravy
1454	334	OSD002	Knihovna VŠB-TUO
1173	353	BOD004	Ústřední knihovna Přírodovědecké fakulty MU
1012	605	HBG001	Krajská knihovna Vysočiny
955	484	PNG001	Knihovna města Plzně, p.o.
953	467	KVG001	Krajská knihovna Karlovy Vary
672	304	ULD001	Ústřední knihovna UJEP
610	614	PAG001	Krajská knihovna v Pardubicích
593	340	SMG506	Městská knihovna Antonína Marka Turnov
574	194	ABA003	Knihovnický institut
569	570	LID001	Knihovna Technické univerzity v Liberci
564	563	ABE313	NIPOS - Národní informační a poradenské středisko pro kulturu
494	149	BOD022	Středisko vědeckých informací, Ekonomicko-správní fakulta MU
426	429	OSA001	Moravskoslezská vědecká knihovna v Ostravě
397	113	ABD103	Centrum vědeckých informací Fakulty sociálních věd UK
375	128	ABA011	Parlamentní knihovna
374	321	ZLG001	Krajská knihovna Františka Bartoše ve Zlíně
281	141	ABG312	Knihovna Jabok

245	148	KMG001	Knihovna Kroměřížska - p.o.
244	82	BOD006	Informační centrum, ústřední knihovna Mendelovy univerzity v Brně
221	111	KLG001	Středočeská vědecká knihovna v Kladně
207	60	BOD018	Knihovna FI MU
191	169	SOG504	Městská knihovna Chodov
160	88	BOE303	Knihovna Moravské galerie v Brně
154	68	ULE301	Muzeum města Ústí nad Labem
103	45	ABE320	Oddělní knihovny GnŘ
80	36	JID501	Knihovna Univerzitního centra Telč Masarykovy univerzity
52	22	ABA100	Všenorská knihovna a informační centrum Berounka
46	16	ABB060	Knihovna Ústavu pro českou literaturu AV ČR
30	13	ABC023	Česká geologická služba
29	29	ABA010	Knihovna Národního muzea
20	10	ABB001	Knihovna Archeologického ústavu AV ČR, Praha, v.v.i.
17	7	CBA903	Archiv
9	2	ABD010	Knihovna MFF UK - půjčovna skript a učebnic
9	9	NAE951	Knihovna (Římskokatolická farnost – děkanství Nové Město nad Metují)

ÚKOLY ŘEŠENÉ V RÁMCI PROJEKTU V ROCE 2022

Úpravy backend vrstvy obálek knih

Z důvodu výskytu problémů s odezvou webového serveru ve chvílích se zvýšenou zátěží, zejména při delší obsazenosti importním API komunikujícím se skenovacím klientem a ve chvílích intenzivního skenování webovými crawlery indexující web bylo přistoupeno k výměně webového servera http Apache za řešení Nginx. Webový server Nginx byl nasazen jako reverzní proxy balancující mezi vícerymi instancemi backend serveru obálek knih se zachováním funkcionality při potřebě klienta přihlášení se k webovému administračnímu rozhraní obálek knih. V průběhu testování bylo voleno více konfigurací hlavně co se týče počtu instancí backend serveru a mechanismu komunikace. Výsledným řešením je 12 full-stack instancí backend serveru komunikujících protokolem HTTP/1.1, které jsou dále poskytovány veřejnosti v podobě šifrované komunikace protokolem HTTP/2.0 a s podporou TLS 1.3. Jedná se tedy o standardní reverzní proxy s použitím současných podporovaných standardů pro komunikaci mezi klienty v prostředí internetu.



Dotazy na více identifikátorů

Byla přidána možnost dotazování se na více identifikátorů stejného typu. Identifikátorem může být jedna hodnota, nebo pole hodnot. Použitelné pro případy, kdy dokument obsahuje více identifikátorů ISBN. Příklad použití v případě, že katalogový záznam obsahuje 2 identifikátory:

```
https://cache.obalkyknih.cz/api/books/?multi=[{"isbn":["9788073060176","9788024206394"]}]]
```

Obecnější příklad použití:

```
https://cache.obalkyknih.cz/api/books/?multi=[  
  {"nbn":"cnb000154538", "isbn":["9788073060176","9788024206394"]},  
  {"isbn":"978-80-86964-09-6", "nbn":"cnb000154538", "oclc":"(OCoLC)311597120"}  
]
```

Dotazování na pole identifikátorů stejného typu je možné pomocí backend API (pro starší systémy používající tento typ API) i pomocí cache frontend servery poskytující metadatové API (doporučený způsob dotazování se na metadata, obálky, toc, ...).

Úpravy frontend vrstvy obálek knih

Kromě výše uvedené úpravy podporující dotazy na více identifikátorů stejného typu byly na úrovni frontend vrstvy projektu obálek knih průběžně prováděny aktualizace knihoven a SW součástí na aktuální verze. Jednalo se zejména o aktualizace běhového prostředí node.js a knihoven používaných frontend scripty.

Z důvodu podpory protokolu HTTP/2 při komunikaci s webovými prohlížeči bylo upuštěno od komunikace s frontend skriptem používajícím http2 knihovnu na přímo, kdy veškerá komunikace byla předávána beze změny a bylo nasazeno obdobné řešení jako v případě backend vrstvy za pomocí webového serveru Nginx v roli reverzní proxy. Výjimkou je, že na běh frontend API je využita jedna instance, která kapacitně postačuje jelikož běh node.js není blokující. Výhodou tohoto řešení je zajištění zpětné kompatibility s klienty komunikujících starší verzí protokolu HTTP/1.1 a také komunikace protokolem TLS1.3. V případě frontend scriptu se obdobná podpora kompatibility jakou nabízí webový server Nginx zajišťuje hůře a docházelo také k pádům a následným restartům webové aplikace.

Samostatným úkolem bylo nasazení nové instance frontend serveru cache3.obalkyknih.cz a úprava scriptu v prostředí, kde autentizace na základě volající IP strany je nutné ověřovat hlavičkou x-forwarder-for HTTP protokolu.

Úprava aplikace skenovacího klienta obálek knih

Příprava v zavádění podpory aplikace skenovacího klienta pro skenování pomocí protokolu TWAIN. Příprava spočívala ve výběru knihovny, přípravy testovacího klienta aplikace skenovacího klienta a testování. K této úpravě vedla potřeba podpory většího počtu zařízení skenerů, kdy se v praxi vyskytují případy pořízení skenovacího zařízení knihovnou s cílem zapojení se do skenování obrazového materiálu pro projekt obálky knih. Kdy se množí výskyty modelů skenovacích zařízení se slabší, nebo žádnou podporou API Windows Image Acquisition. WIA je jediné v současnosti podporované API pro komunikaci mezi skenovacím

klientem a zařízením skeneru. Při skenování vzniká vadné barevné podání, nebo v některých případech není skenovací zařízení rozeznáno. Zavedením komunikace skenovacího klienta se skenerem pomocí API TWAIN bude sníženo riziko problémů se skenovacím zařízením. Bude také rozšířena podpora skenovacího klienta o výběr parametrů rozhraní jakými jsou obrazová kvalita, jas, kontrast, gamma korekce, případně dalších parametrů.

Bylo řešeno:

- Volba knihovny pro komunikaci s rozhraním TWAIN pro prostředí frameworku .NET. Byla vybrána knihovna NTwain, která vyhovuje pro použití s aplikací skenovacího klienta a je volně šiřitelná licencována pod licencí MIT.
- Příprava testovací aplikace skenovacího klienta a testování.

Další úpravou co se týče aplikace skenovacího klienta byla migrace zdrojových kódů z doposud používaného prostředí Visual Studio Express 2010 do aktuálního prostředí Visual Studio Community 2022 a použití platformy .NET Framework 4.8, která je podporována v posledních dvou verzích OS Windows.

Hromadný výstup dokumentů volně přístupných a online dostupných děl

Přidán nový výstup dokumentů volně přístupných a online dostupných děl z digitálních knihoven (DNNTO) v rámci služby [DNNT](#). Export popisných dat, anotací, fulltext obsahů dokumentů volně přístupných a online dostupných děl z digitálních knihoven (DNNTO) - po přihlášení čtenáře pomocí systému eduID. Data dostupná v rámci služby [DNNT](#). Popis viz <https://servis.obalkyknih.cz/vystupy.php>

Fotografie knihoven

V projektu je možné ukládat a zobrazovat fotografie sídel knihoven. Jako identifikátor se používá sigla knihovny dle vzoru sigla-sigla (příklad CBA001-CBA001) <https://www.obalkyknih.cz/view?nbn=cba001-cba001> .

Bylo využito v Centrálním portálu knihoven (viz náhled) a v regionálních systémech knihoven.

Národní zdroje Zahraniční zdroje Adresář knihoven

Zadejte hledaný dotaz, např. název knihovny, město, ... Ve všech polích **Hledat** Pokročilé vyhledávání

Upřesnit hledání Výsledků na stránku 20 1 - 20 z 6 521 Řadit dle relevance


Zapojení do projektu

- DNNT: Díla nedostupná na trhu 148
- Knihovny.cz 93
- Služba Získej 44

Umístění

- Středočeský kraj 878
- Jihomoravský kraj 774
- Kraj Vysočina 585
- Jihočeský kraj 583
- Plzeňský kraj 502
- Olomoucký kraj 479
- více...

Vybrat vše | Vybrané: [E-mail](#) [Export](#) [Tisk](#) [Přidat do oblíbených](#) [Vložit do schránky](#)


 **Národní knihovna České republiky - Knihovní fondy a služby - Odbor služeb**

Klementinum 190, 110 00 Praha 1

[Vyhledávat v knihovně](#)

[Vložit do schránky](#)

[Přidat do oblíbených](#)

 **Národní knihovna České republiky - Slovanská knihovna - oddělení služeb**

Klementinum 190, 110 00 Praha 1

[Vyhledávat v knihovně](#)

[Vložit do schránky](#)

[Přidat do oblíbených](#)

Vedlejší vývojové změny:

- Zprovoznění cache3 serveru projektu v Centrum vědecko-technických informací SR v Bratislavě
- Přidání možnosti uploadu obálek s identifikátorem oclc pomocí webového rozhraní na webu obalkyknih.cz.
- Zlepšení kvality generovaných náhledů v rozlišení s výškou 510px.
- Vývoj scriptu pro komunikaci s API Pero za účelem získání TXT ze skenované předlohy obsahu titulu. Cílem byla tvorba PDF s textovou vrstvou pro zpřístupnění jako TOC. V průběhu vývoje nebylo cíle dosaženo z důvodu nevyřešení problému s tvorbou textové vstvy z předlohy popisného souboru ALTO. Návrhem je postupovat podobně jako v případě řešení obdobného problému v rámci projektu Kramerius.
- Přidání odkazu na „Weby o knihách“ (cbdb.cz, databazeknih.cz) do datového kontajneru

KONTROLY A DALŠÍ PRÁCE

V roce 2022 pokračovaly kontroly záznamů projektu v našem kontrolním rozhraní (<https://servis.obalkyknih.cz/abuse/>). Do kontrol se opětovně zapojovali knihovníci, kteří pracovali z domova v rámci homeoffice a dobrovolníci . Aktuálně je v systému zaregistrováno 178 uživatelů (aktivně kontrolovala v roce 2022 cca 35 z nich ... bohužel většina „homeoffice“ uživatelů už z domova nepracuje)

Při kontrolách záznamů dochází k doplňování chybějících identifikátorů (číslo cnb, oclc, isbn), doplňování popisných metadat pro zobrazování na stránkách projektu (název, autor, rok vydání), doplňování či úpravě anotace, výběru nejkvalitnější obálky (rozlišení, bez nálepek,

...). Zároveň se opravovaly chybné záznamy a propojovaly se související záznamy (dílové tituly).

V roce 2022 bylo již zkontrolováno celkem přes 486 000 záznamů (viz tabulka).

Tabulka počtu opravených záznamů dle zapojených institucí (stav k 9.1. 2023):

Přehled počtu zkontrolovaných záznamů dle zapojených knihoven	
Jihočeská vědecká knihovna v Českých Budějovicích	275 944
Městská knihovna v Praze	179 894
Severočeská vědecká knihovna v Ústí nad Labem	10 222
Moravskoslezská vědecká knihovna v Ostravě	4 592
Krajská knihovna Karlovy Vary	4 262
Krajská knihovna Františka Bartoše ve Zlíně	3 202
Moravská zemská knihovna v Brně	3 139
Krajská vědecká knihovna v Liberci	2 566
Krajská knihovna Vysočiny	932
Národní Knihovna ČR	620
UISK	508
Knihovna Třinec	500
Městská knihovna Tábor	365
Knihovna města Plzně, p.o.	161

Statisticky vychází, že u cca třetiny záznamů dochází k nahrání kvalitnější obálky a u cca. 25% záznamů je doplněna či upravena anotace.

Knihovny takto získaly kvalitnější záznamy s většími obálkami.

Dále byly v roce 2022 prováděny úkoly související s kontrolou anotací a hodnocení, propojení na další vydání titulu a propojení vazeb e-knih.

Jednalo se o práce, které nelze provádět plně automatizovat a je nutné je provádět pomocí zaškolené obsluhy.

Výsledkem práce bylo:

- 1) napojení e-knih na klasické papírové tituly – aktuálně je dostupných více jak 2800 e-knih z produkce Městské knihovny v Praze
- 2) propojení na další vydání titulu - propojení jednotlivých vydání shodného titulu nebo navazující série titulů – řešeno ve spolupráci s CPK
- 3) opravy a schvalování anotací (stahovaných automaticky či ručně) k dokumentům – za rok 2022 přibylo v projektu 50 tisíc anotací, které byly formálně zkontrolovány a doplněny k titulům,

Další úkoly řešené v roce 2022 mimo projekt:

- stahování obálek a obsahů z Kramériů knihoven ČR – do dohodě s jejich administrátory
- údržba a podpora skenovacího klienta pro nahrávání dat knihovnamí do projektu
- nastavení individualních exportů knihovnam pro doplnění obálek a obsahů vlastního fondu (Knihovny Masarykovy univerzity, Národní památkový ústav, Masarykův ústav a Archiv AV ČR, NIPOS – Národní informační a poradenské středisko pro kulturu)
- aktualizace webových stránek projektu
- servis a rozšíření servisních stránek projektu <https://servis.obalkyknih.cz/>
- aktualizace metodických pokynů a manuálu pro knihovny a knihovní systémy
- prezentace projektu na knihovnických akcích (Seminář uživatelů ARL, FF UK, Výroční seminář SK ČR)
- emailová a telefonická podpora projektu, spolupráce s tvůrci KIS, CPK

Popis řešení a veškeré kódy aplikace jsou volně dostupné jako opensource na adrese <https://github.com/cbvk/obalkyknih/wiki> a <https://www.obalkyknih.cz>.

V Českých Budějovicích 9. 1. 2023

Ing. Jiří Nechvátal
Jihočeská vědecká knihovna v Českých Budějovicích