

VEŘEJNÉ INFORMAČNÍ SLUŽBY KNIHOVEN

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

Obálky knih.cz - rozvoj projektu v roce 2020

*Jihočeská vědecká knihovna v Českých Budějovicích
leden 2021*

Zhodnocení projektu

Projekt Obálkyknih.cz sdružuje různé zdroje informací o dokumentech do jedné, snadno použitelné webové služby. Databáze aktuálně obsahuje přes **2,26 miliónu obálek** (nárůst za rok 2020 o cca 170 000 obálek), **568 tisíc obsahů** českých a zahraničních dokumentů (nárůst za rok 2020 o cca 60 000 obsahů) a **10 100** seznamů doporučené literatury. Dále poskytuje přes **500 tisíc anotací, 3,57 miliónu hodnocení u 174 tisíc titulů, 8,5 tisíc komentářů, 57 tisíc fotografií autorit** a cca. **0,6 miliónu vygenerovaných citací** dle normy ISO 690. Rok 2020 byl velmi ovlivněn zdravotní situací v ČR, což ovlivnilo i nárůst počtu záznamů v projektu.

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.

Přehled vlastností projektu:

- hlavní servery jsou provozovány v **Jihočeské vědecké knihovně v Českých Budějovicích**, záložní server je v **Moravské zemské knihovně v Brně**
- 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 2020:

Počet odeslaných dokumentů	180 289
Počet uložených obálek (COVER)	170 970
Počet uložených stran obsahů (TOC)	120 833
Počet uložených fotografií autorit	870
Počet uložených stran seznamů lit.	13 912

Počty odeslaných stran a titulů přes skenovacího klienta dle jednotlivých knihoven (rok 2020):

STRAN	TITULŮ	SIGLA	NÁZEV
19055	8708	CBA001	Jihočeská vědecká knihovna v Českých Budějovicích
15859	4272	ABA013	Národní technická knihovna
15455	7326	OLD012	Knihovna Univerzity Palackého v Olomouci
12995	5493	ABA004	Národní knihovna ČR - Slovanská knihovna
12934	6683	BOA001	Moravská zemská knihovna v Brně
12002	3434	ABB022	Středisko vědeckých informací Fyziologického ústavu AV ČR, v.v.i.
11331	2196	BOD010	Masarykova univerzita - Právnická fakulta
10937	7465	ABA001	Národní knihovna ČR
9342	3417	ABA008	Národní lékařská knihovna
9294	5137	OLA001	Vědecká knihovna v Olomouci
7348	6737	ULG001	Severočeská vědecká knihovna

6315	2252	ABB019	Knihovna Sociologického ústavu AV ČR, v.v.i.
5474	5474	ABD001	Knihovna Ústavu Dálného východu Filozofické fakulty Univerzity Karlovy
4760	1192	ZLD002	Univerzita Tomáše Bati ve Zlíně
4545	3579	HKA001	Studijní a vědecká knihovna v Hradci Králové
4405	2356	KVG001	Krajská knihovna Karlovy Vary
4392	1083	ABA006	Vysoká škola ekonomická v Praze
4374	1836	BOE020	Knihovna Ústavního soudu
3812	1582	CBD007	Akademická knihovna Jihočeské univerzity
3351	1114	BOD031	Masarykova univerzita, Fakulta sociálních studií, Ústřední knihovna
2537	1734	PNA001	Studijní a vědecká knihovna Plzeňského kraje
2395	1461	BOD003	Ústřední knihovna Pedagogické fakulty MU
2387	1614	UHG001	Knihovna Bedřicha Beneše Buchlovana
2264	1272	LIA001	Krajská vědecká knihovna v Liberci
2236	480	BOD004	Ústřední knihovna Přírodovědecké fakulty MU
2115	740	ABD103	Fakulta sociálních věd
1849	1292	BOE451	Knihovna Biskupství brněnského
1403	793	SMG506	Městská knihovna Antonína Marka Turnov
1281	444	ABA007	Knihovna Akademie věd
1227	837	PAG001	Krajská knihovna v Pardubicích
1097	417	LID001	Knihovna Technické univerzity v Liberci
955	292	BOD022	Středisko vědeckých informací ESF MU
776	206	ABD100	ÚK ČVUT
754	396	PNG001	Knihovna města Plzně
727	364	KLG001	Středočeská vědecká knihovna v Kladně
645	435	OSA001	Moravskoslezská vědecká knihovna v Ostravě
553	334	OSG002	Knihovna města Ostravy
527	195	ULE301	Muzeum města Ústí nad Labem
484	359	ZLG001	Krajská knihovna Františka Bartoše ve Zlíně
407	415	TAG001	Městská knihovna Tábor
311	100	ABA011	Parlamentní knihovna
300	233	SOG504	Městská knihovna Chodov
274	151	HBG001	Krajská knihovna Vysočiny
214	65	BOD006	Informační centrum, ústřední knihovna Mendelovy univerzity v Brně
213	148	BOD001	Ústřední knihovna filozofické fakulty MU
180	93	ULD001	Ústřední knihovna UJEP
175	97	ABA100	Všenorská knihovna a informační centrum Berounka
139	43	BOD018	Knihovna FI MU
104	104	ABD027	Evangelická teologická knihovna UK
99	23	ABD010	Knihovna MFF UK - půjčovna skript a učebnic
96	40	JID501	Knihovna Univerzitního centra Telč Masarykovy univerzity
53	16	ABG312	Knihovna Jabok
44	24	CBA903	Archiv
29	14	UOE802	Knihovna Regionálního muzea ve Vysokém Mýtě
26	5	OSD002	Ústřední knihovna VŠB - TU Ostrava
22	11	ABB001	Knihovna Archeologického ústavu AV ČR, Praha, v.v.i.

18	8	BOE303	Knihovna Moravské galerie v Brně
8	8	NAE951	Knihovna (Římskokatolická farnost – děkanství Nové Město nad Metují)
8	3	ABD009	Knihovna MFF - Matematické oddělení
4	2	ABC023	Česká geologická služba

Úkoly řešené v rámci projektu v roce 2020:

ZMĚNA DATOVÉ STRUKTURY

Projekt obálkyknih.cz poskytuje služby knihovnám a jiným institucím již více než 11 let. Od doby vzniku využívá původní strukturu databáze. Nové funkce projektu byly napsány a implementovány na stávající strukturu, která byla zachována s ohledem na zachování původní funkčnosti a SW implementace.

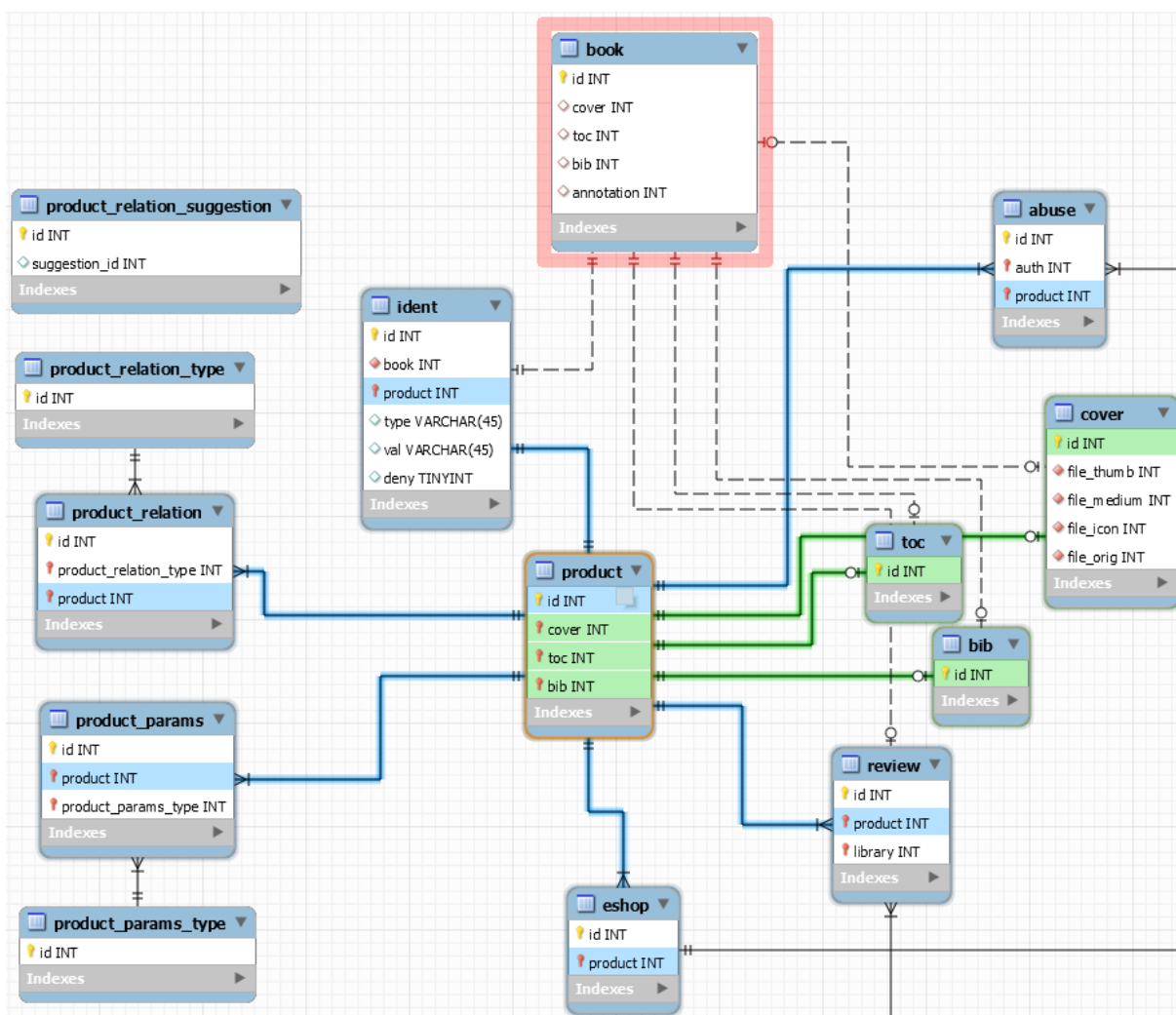
Stávající datová struktura nevyhovovala novým požadavkům na poskytované služby a zejména na snadnost správy projektu a údržby databáze. Hlavní cíle úprav byly:

- Zjednodušení a normalizace relačních vazeb. V původním řešení existovali dvojice vazeb mezi tabulkami, které odkazovaly vzájemně na sebe, čím bylo prakticky znemožněno mazání, nebo slučování takovýchto záznamů.
- Zrušení účelu tabulky book jako datového skladu. Nejednalo se o smazání tabulky, ale změnu účelu a práce s tabulkou. Úlohou tabulky book je v současnosti deduplikace a agregace dat nacházejících se v tabulce product. Účelem tabulky product je právě uschování originálních dat získaných při plnění jedním ze standardních způsobů (importem pomocí skenovacího klienta, nebo skriptem pro sběr dat).
- Ukládání identifikátorů (kterými jsou isbn, issn, ean, nbn, oclc, uuid, vlastní identifikátor, ...) na jediném místě (tabulka ident). Zároveň identifikace neplatných identifikátorů bibliografických záznamů.
- Změna vazby mezi tabulkami produkt, cover, toc a bib na vazbu s relacemi typu 1:1. Tím došlo k duplikaci dat, kdy v předchozí verzi datové struktury bylo možné přiřazovat jednu obálku vícero knihám, nebo autoritním záznamům. To se v praxi ukázalo jako komplikace při správě dat, kdy slučování, nebo mazání takto svázaných záznamů bylo komplikované a při potřebě spravovat objemnou databázi se ukázalo jako v praxi nereálné.
- Změna vazby mezi tabulkou book a tabulkami cover, toc a bib na neidentifikující typ vazby. Účelem bylo vytvoření reprezentativních náhledů obálek, TOC a seznamu literatury tak, aby vazba neznemožňovala správu databáze.

Hlavním cílem bylo postavení tabulky product do role datového skladu. Co byl zásadní rozdíl v porovnání s předchozím řešením, kdy obdobnou roli plnila i tabulka book a vzájemné vazby mezi sebou a navíc s tabulkami cover a toc tvořily vazbu s minimálními možnostmi na spravování ve větším měřítku. V novém řešení je roli tabulky book možné považovat za líně plněnou cache tabulku, která se stará o deduplikaci a agregaci dat získaných při importu a sběru. Tj. tabulka book je zdrojem pro následnou prezentaci buďto na webu obálek knih, nebo pomocí API a tabulka product (společně s tabulkou ident) je zdrojem raw dat, která byla sklizena s vícero zdrojů a může se jednat (ve většině případů se i jedná) o data toho samého titulu. Líným plněním se rozumí vlastnost, kdy je data obsažena v tabulce book možné promazat a k obnově původního stavu záznamu dojde při opakovaném požadavku o poskytnutí dat buďto ve formě webové stránky obálek knih, nebo API. Výhodou tohoto procesu je např. možnost opravy prezentačních dat titulu v případě nalezeného problému ve zdrojových datech, nebo v některém z interních procesů. V takovém případě je možné postupovat stylem – oprava dat, nebo procesu způsobujícího chybový stav – výmaz prezentačních dat záznamu – automatické naplnění procesem líného plnění. Získává se tím určitá volnost ve správě dat a interních procesů backend systému obálek knih v porovnání s dřívějším řešením. Byla taky

zavedena vlastnost zamykání prezentačních záznamů tj. obdoba revize. Záznamy u kterých dojde k revizi je možné zakonzervovat a brát je jako konečný stav.

ER-D model části databáze – obr. 1. Modrou a zelenou barvou jsou zvýrazněné relace a pro zjednodušení jsou zahrnuty pouze entity primárních a cizích klíčů.



Obr. 1: Návrh nové struktury databáze

Při importu dat pomocí vkládacího API, buďto pomocí skenovacího klienta, nebo crawlerem / harvestorem se data vkládají pouze do tabulky product a souvisejících tabulek (jako identifikátory bibliografického záznamy do tabulky ident, mediální data do tabulek cover, toc a bib - zvýrazněných na obr. 1 modrou a zelenou barvou). Neplní se do tabulky book. Tabulka book představuje:

- Deduplikovanou podobu tabulky product.
- Je jí možné promazávat a díky procesu líného plnění je regenerována. Tato vlastnost bude i v budoucnosti využita při změně logiky deduplikačního algoritmu, nebo zdrojových dat. Líné plnění daty se už dříve ukázalo jako výhodné v případě frontend vrstvy projektu.
- Tabulka book neobsahuje striktní vazby na tabulky cover, toc a bib. Jsou vyměněny za neidentifikující vazby a slouží pro přiřazení reprezentativní obálky, obsahu, seznamu literatury, anotace atd. zvolených právě deduplikačním algoritmem.

Po technické stránce došlo k nahrazení doteď používaného frameworku Perl Catalyst za novější PHP Phalcon. Tím došlo i k výměně scriptovacího jazyka z původního Perl za současný PHP. Důvodem mezi jiným bylo i opouštění komunity toto řešení, kdy za poslední dekádu nedošlo k žádnému posunu a

vývojáři PERL modulů, na kterých je projekt závislý komunitu opouštěli, nebo dokonce komentovali ve veřejném CPAN repozitáři problémy s kompatibilitou s novými verzemi scriptovacího jazyka PERL a označovali modul za neudržitelný. Jako náhrada byl zvolen framework Phalcon, který je kompilovaný v jazyku C (není to jen skupina dalších .php scriptů) a tím se v době návrhu stal nejrychlejším PHP frameworkem. Dalšími pozitivními vlastnostmi je jeho minimalističnost a používání návrhového vzoru MVC, co je ve světě webových aplikací ještě i v současnosti standardem.

Změna v případě dotazovacího API – změna logiky vyhledávání podle identifikátorů bibliografických záznamů v souvislosti s popisovanou změnou datové struktury. Nově se identifikátory titulu nenachází přímo v tabulce book, ani product_param, jak tomu bylo v předchozím řešení. Identifikátory jsou vkládány do tabulky ident a ta je spojena vazbou s tabulkou product. Navíc byl přidán příznak identifikující neplatné identifikátory. Deduplikační algoritmus v případě dotazu na konkrétní titul vyhledává podle identifikátorů jeden, nebo více záznamů product. Na základě nich se rozhoduje, které informace z vyhledaných záznamů použije. Výsledkem deduplikačního algoritmu u jednoho záznamu titulu je právě jeden výsledný deduplikovaný záznam v tabulce book a použité záznamy product jsou s ním svazovány.

Z původního návrhu byla vypuštěna relační tabulka b2product (resp. b2ident) navržena z důvodu, že deduplikační algoritmus se může rozhodnout pro podskupinu identifikátorů a nemusí použít všechny. Bylo tomu tak zejména pro použití u seriálů a částí monografií, co ale je možné řešit i příznakem, že se jedná o záznam seriálu, nebo vícesvazkové monografie, tomu uzpůsobit logiku a není nutné zakládat vazební tabulku. Deduplikační algoritmus je tomuto chování u seriálů a vícesvazkových monografií možné přizpůsobit.

Pokud už dotazovaný identifikátor existuje mezi záznamy svázanými pomocí tabulky ident, nebude se deduplikace opakovat, ale použijí se už vygenerovaná data.

Vkládací API – plní data z různých zdrojů do tabulek product, ident a dalších souvisejících tabulek určených pro mediální data. Navíc současně dochází i k promazávání souvisejících už deduplikovaných záznamů v tabulce book kvůli aktualizaci jejich dat pomocí opakovaného spuštění deduplikačního procesu. S vkládacím API a promazáváním záznamů v tabulce book souvisí i proces promazávání dat uložených na cache serverech cache1 a cache2. Toto promazávání se řídí skupinou tabulek fe_sync popsanych na obr. v textu níže.

Vazba mezi tabulkami product a skupinou tabulek cover, toc a bib je typu 1:1. Na straně product 1 a právě 1, na straně cover / toc / bib 1, anebo 0. Jeden záznam product má jednu, nebo žádnou obálku. Pokud existuje záznam cover, musí mít přiřazený právě jeden záznam product.

Technický popis úprav

Součástí úprav provedených v roce 2020 byla i obnova použitého SW řešení a reinstalace virtuálních serverů infrastruktury obálek knih.

Záměrem bylo prodloužení doby podpory použitého operačního systému a aktualizace SW součástí. V době návrhu byla zvolena jako vhodná distribuce Centos verze 8, co bylo i provedeno. Po předčasném zkrácení podpory distribuce Centos 8 (do konce 2021) je následně nutná neplánovaná aktualizace na Centos Stream 8.

Celkem byly instalovány 3 instance serverů na kterých byla provedena aktualizace veškerého SW vybavení na nejaktuálnější stabilní verze. V případě frontend serverů se jedná o:

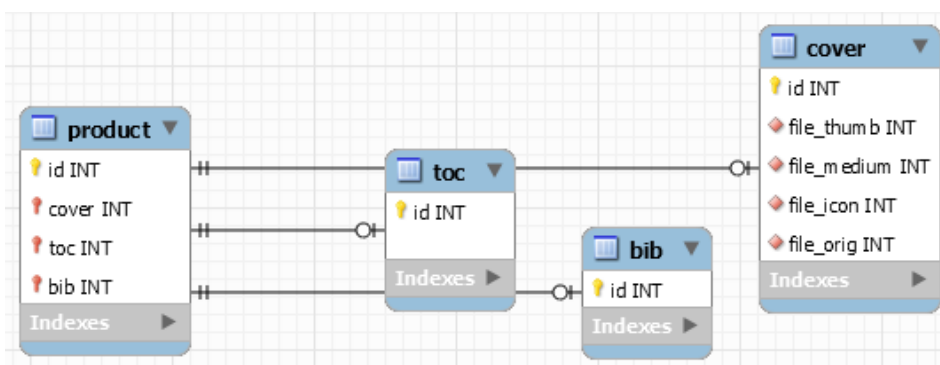
- CentOS Linux release 8.2.2004 (Core) – nutné neplánovaně aktualizovat na Centos Stream 8
- Node.js v14.15 (současně s podporovaným npm v6.14 a pm2 v4.5)
- MongoDB 4.4.2

V případě backend serveru:

- CentOS Stream release 8
- Perl v5.26 (kvůli současnému běhu scriptů původní i nové verze backend systému, včetně aktualizace všech modulů vyžadovaných frameworkem Catalyst)
- MySQL 5 vyměněné za výkonnější MariaDB v10.3
- PHP 7.4 s moduly pro framework Phalcon v4.0.6

U přechodu z původní na novou strukturu byla provedena i jednorázová transformace dat. Tomu předcházelo čištění dat, zjišťování souvislostí v původních zdrojových kódech backend systému, úprava struktury tabulek apod.:

- Do tabulky product bylo potřeba z protějšků záznamů z tabulky book doplnit identifikátory part_year, part_volume, part_no, part_name, part_type a taky protějšky s prefixem „orig_“ pro každý takovýto záznam z tabulky product. Také sloupce doc_type, edition, other_relation_text a review. Jedná se o identifikátory, které se nachází pouze v tabulce book, ale bylo je potřebné zachovat.
- Sloupce cached_rating_sum, cached_rating_count a sloupce s prefixem „citation_“ nebylo nutné přesouvat. Vytvoří se automaticky v rámci procesu deduplikace.
- Záznamy bez užitečné informace, jako vazby na obálku, TOC, seznam literatury, anotaci apod. byly vymazány. Zbylé záznamy byly ponechány.



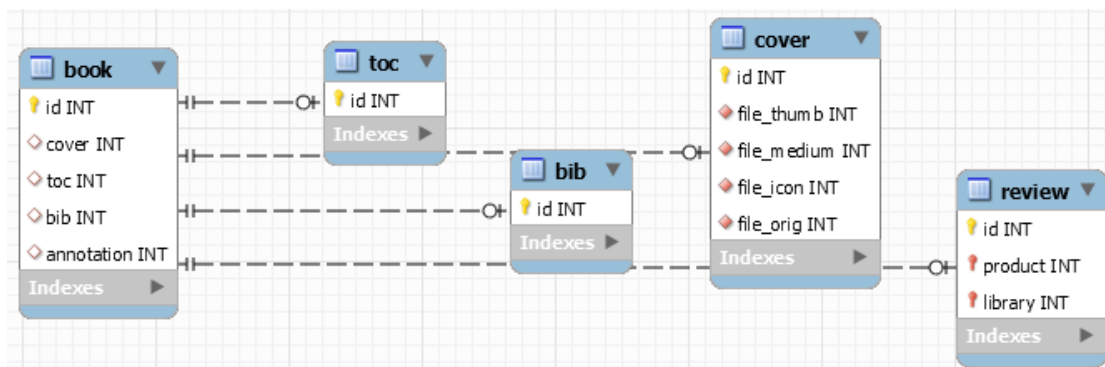
Obr. 2: Tabulky product, toc, bib, cover vyjmuté z celku datového modelu.

U tabulky cover byla odstraněna vazba 1:N a změněna na 1:1 a taky zrušen proces deduplikace obálek v tabulce cover, kdy se podle hash řetězce uploadovaného obrázku hledal existující záznam a pokud se našel, svázal se s novým záznamem product s existujícím záznamem cover.

V souvislosti s přechodem na novou strukturu byla provedena i jednorázovou transformace dat – přesun dat. V případě tabulky cover:

- Vymazání záznamů v tabulce fileblob, které nesměrovaly na žádný záznam v tabulce cover.
- Vymazání záznamů v tabulce cover, které nesměrovaly na žádný záznam v tabulkách product, nebo auth.
- Pokud existoval záznam cover, který směřoval na více záznamů product, nebo auth, bylo potřebné zduplikovat výskyty z důvodu vazby 1:1 – pro každý záznam v tabulce product, nebo auth vytvořen vlastní záznam cover + čtveřice záznamů fileblob.
- Tabulka fileblob nabývala z důvodu ukládání binárních dat velkých rozměrů. Data přesunuta na souborový systém.

Tabulka book s neidentifikujícími vazbami. Tabulka book odkazuje na reprezentativní obálku, obsah TOC a seznamu literatury, případně anotaci, nebo hodnocení titulu čtenáři. Záznamy v tabulkách toc, bib, cover, review budou existovat i v případě, že neodkazují na žádný záznam v tabulce book. Identifikující je ale vazba mezi tabulkami cover, toc, bib, review a mezi tabulkou product.



Obr. 3: Tabulky book s neidentifikujícími vazbami na cover, toc, ... vyňaté z celku datového modelu.

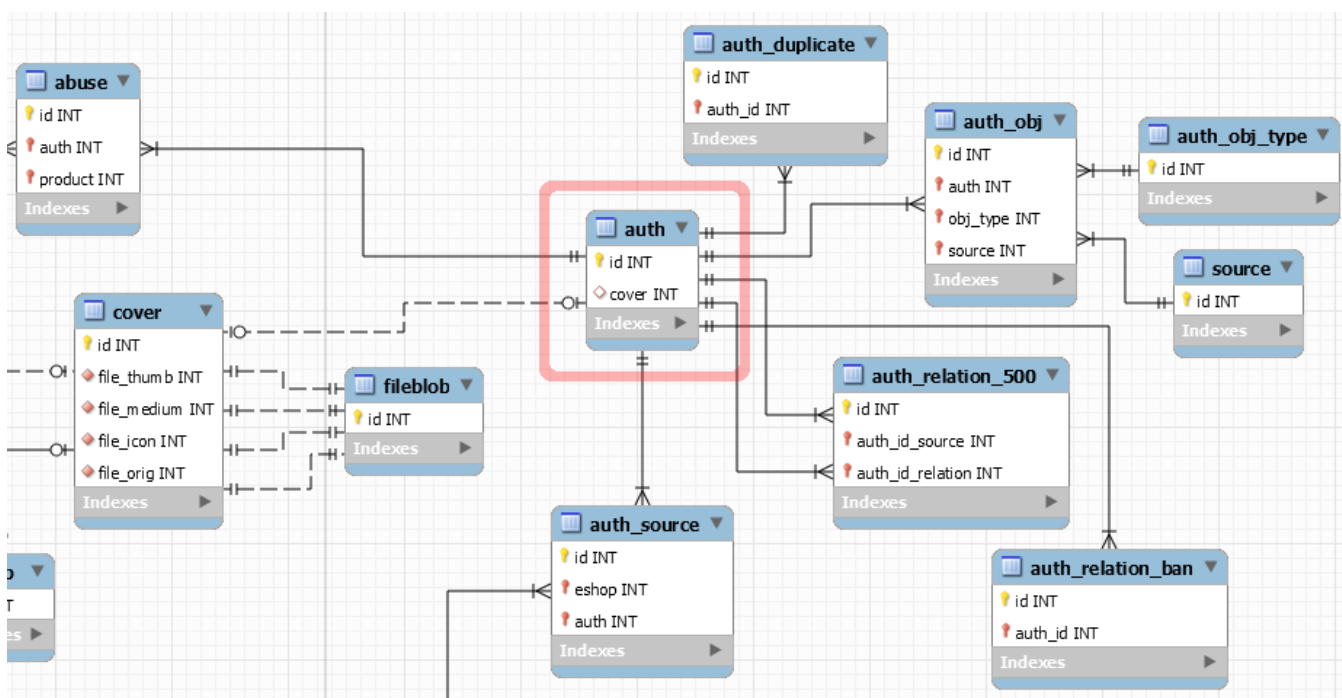
Příklad: Existuje více záznamů v tabulce product reprezentujících stejný titul. Jeden vznikl skenovacím klientem, druhý stažením z Krameria, třetí stažením z některého e-shopu. Každý z nich bude obsahovat právě jednu obálku (vazba 1:1 mezi product a cover), a jeden, nebo více identifikátorů (vazba 1:N mezi product a product_params). Úlohou deduplikačního algoritmu je vytvořit z těchto záznamů jednu deduplikovanou verzi a uložit ji do tabulky book. Deduplikační algoritmus se rozhoduje, kterou z obálek upřednostnit podle kvality a zdroje původu. Odkaz na zvolený obrázek obálky se ukládá jako vazba mezi tabulkami book a cover.

Další tabulky nacházející se na obr.1:

- product_params_type – jedná se o číselník jiných parametrů importovaných dat, než bibliografické identifikátory titulu např. url adresa e-knihy v různých dostupných formátech, ... v takové podobě už existuje dnes a používá se právě na e-knihy sklizené z Městské knihovny v Praze
- product_params – konkrétní hodnota parametru. Jeden produkt může mít více parametrů a toto bude jediným úložištěm parametrů (kromě bibliografických identifikátorů titulů), což je rozdíl v porovnání s původním řešením, kdy identifikátory obsahuje i tabulka book a product.
- product_relation, product_relation_type, product_relation_suggestion – používá se na svazování podobných dokumentů např. edicí, nebo svazování záznamů e-knih s papírovými tituly. Tabulka product_relation_suggestion obsahuje automaticky generované návrhy v současnosti na vazby typu e-kniha – papírová kniha, které se dále pomocí admin rozhraní schvalují. V případě těchto tabulek byla potřebná transformace (změna relací se záznamem knih na záznamy product), neboť v původní struktuře byly svázané se záznamy book a ne product. (původně pojmenované jako book_relation, book_relation_type a book_relation_suggestion).
- abuse – tabulka s hlášenými špatnými obálkami knih, nebo foto autorů. Při změně datové struktury byla taky nutná změna z vazeb namísto tabulky book na nově tabulku product.
- review – komentáře, hodnocení a anotace, po zrušení vazby na book zůstala prakticky ve shodné podobě jako v původní verzi.
- eshop – identifikace zdroje záznamů v tabulkách product i auth (u auth resp. tabulak source).

Datová struktura autoritních záznamů.

Od původní podoby se liší pouze vazbou mezi tabulkou cover, kde se jedná taktéž o vazbu 1:1. Každý autoritní záznam bude mít nejvýše jeden záznam obrázku.



Obr. 4: Tabulky a vazby u autoritních záznamů.

Vkládací API v případě autoritních záznamů bylo doplněné o funkcionalitu, kdy se foto autora (např. při přeskenování) kopíruje i na zbylé záznamy cover autoritních záznamů, které jsou navzájem svázané tabulkou auth_relation_500. Tato tabulka obsahuje relace autoritních záznamů, stejných autorů. Autoři vystupují pod různými jmény – pseudonymy. Jedná se přitom o stejnou osobu tj. i stejné foto. V původní datové struktuře toto nebylo nutné, protože všechny takovéto autority odkazovaly na stejnou obálku. Nový model s vazbou 1:1 i když na první pohled vypadá být nevýhodný, není tomu tak docela. Výhodné je to naopak u autoritních záznamů, které jsou svázané se stejnou foto, ale i přesto se nejedná se o stejné osoby (toto se běžně stává a není tomu jednoduchým způsobem zabránit). Příkladem jsou např. autoři publikující za jednu organizaci, kdy namísto jejich foto je logo organizace ve které působí apod. Tady je vazba 1:1 nové datové struktury výhodná, protože s uploadnutím fotografie jednoho z takovéto skupiny autorů se obálka nepravdivě nezobrazí i u dalších autorů.

Ostatní tabulky zůstaly beze změny.

Pravidelné čištění tabulek datové struktury:

- Týdně se promazává obsah synchronizačních tabulek fe_sync, fe_sync_param a fe_sync2param. Pouze záznamů s příznakem flag_synced = 1. Obsah tabulek není nutné archivovat.
- Tabulky visitor a request byly nahrazeny počítadly, protože účelem těchto tabulek bylo pouze zjišťování počtu requestů na API a webové stránky projektu.
- Týdně se promazává obsahy tabulky upload a protějšky souborů na souborovém systému.

KONTROLY A DALŠÍ PRÁCE

V souvislosti s nouzovým stavem v ČR došlo v roce 2020 k zásadnímu rozšíření a úpravám rozhraní pro opravu nahlášených chyb v projektu (<https://servis.obalkyknih.cz/abuse/>). Vzniklo nové kontrolní rozhraní. Do kontrol se zapojilo více jak 100 knihovníků z krajských knihoven v rámci „homeoffice“. Vzniklo editační a administrativní rozhraní, systém přidělování záznamů, statistický modul, kontrolní výpis zkontrolovaných titulů, ...

Při kontrolách záznamů došlo 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 2020 bylo takto zkontrolováno přes 197 000 záznamů (viz tabulka).

U cca poloviny záznamů (86 167) byla nahrána kvalitnější obálka (viz ukázka). U cca. 30% záznamů byla doplněna či upravena anotace.

Knihovny takto získaly kvalitnější záznamy s většími obálkami včetně nové funkce – souvisejících (dílových) propojení titulů.

Tabulka počtu opravených záznamů dle zapojených institucí (stav k 31.12. 2020):

Přehled počtu zkontrolovaných záznamů dle zapojených knihoven	
Jihočeská vědecká knihovna	88 896
Městská knihovna v Praze	88 881
Moravskoslezská vědecká knihovna v Ostravě	4 584
Severočeská vědecká knihovna	3 228
Krajská knihovna Františka Bartoše ve Zlíně	3 202
Moravská zemská knihovna v Brně	3 139
Krajská vědecká knihovna v Liberci	2 432
Krajská knihovna Karlovy Vary	1 560
Krajská knihovna Vysočiny	566
Knihovna Třinec	500
Městská knihovna Tábor	365
Národní knihovna ČR	224

Rozhraní pro kontrolu dat:

Kontrola hlášení chybných obálek - abuse



[zobrazit/obnovit seznam uživatelů](#)

Vyhledej titul

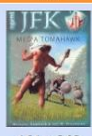
**A
B
U
S
E**

Problém s cover:
IP: 65499 referer STATUS: 2
Datum: 2020-03-25 16:42:28
Problém: **není**
ID: 88615
AUTH ID:

**C
O
V
E
R**


161 x 240
135 x 200

ID	ean13	isbn	oclc	autor	název	rok	hodnoc.	cover	top	bib	pomozna
110027101	9788072547470	cnb001641542		Meš a tomahawk - Miroslav Žamboch	2006	479259					
<p>S E R V I S</p> <p>HLEDAT OBÁLKU: Google (N+A) Google obr. (N+A) !!! Výmaz záznamu z DB !!! Vrátit TOC OPRAVENO</p>											

NKP ID	název	autor	rok	vydání	verze	nakladatel	SERVIS	DÍLY	identifikatory
88620061641542	Meš a tomahawk	Žamboch, Miroslav	2006	Vyd. 1.		[Praha] : ve spolupráci s EF vydalo nakl. Triton, 2006	DOPLNIT_POPIŠNÁ_DATA		cnb001641542 9788072547470 (OCLC)83979114

Anotace: [OPRAVIT ANOTACI](#)

!!! na obalkyknih.cz se vyskytuje DUPLIKÁT tohoto záznamu !!! - neřešte záznam sami a předejte jej adminovi - [PŘEDAT ADMINU](#)

DUPL	ID	ean13	isbn	oclc	autor	název	rok	cover	top	bib	pomozna	review	hodnoc.	servis
	127207603	9788072547470			Jiří W. Procházka	Meš a tomahawk						3115807	626376	

Záznam před opravou (zobrazuje se pouze obálka a obsah jednoho ze tří dílů titulu):

NKC - Úplné zobrazení záznamu



Vzvolte formát: [Standardní](#) -- [Katalogizační záznam](#) -- [Stručný záznam](#) -- [MARC](#) -- [Citace](#)

Záznam 1 z 1

Hlavní záhlaví [Fikejz, Miloš, 1959-2019](#)

Název [Český film : herci a herečky / Miloš Fikejz](#)

Vydání 1. vyd.

Nakl. údaje Praha : Libri, 2006-2008

Popis (rozsah) 3 sv. : portréty ; 21 cm

[Navrhnout na digitalizaci](#)

EXEMPLÁŘE 

Číslo nár.bibl. cnb001695000

ISBN
80-7277-331-3 (soubor ; váz.)
80-7277-332-1 (1. díl ; váz.) : Kč 690,00
978-80-7277-334-3 (2. díl ; váz.) : Kč 495,00
978-80-7277-353-4 (3. díl ; váz.) * 978-80-7277-354-4 (chybné) : Kč 760,00

Poznámka
Obsahuje rejstříky
3. díl obsahuje aktualizace k 1. a 2. dílu
2000 výt.

Obsahuje I. díl. A-K. 2006. 750 s. -- II. díl. L-Ř. 2007. 656 s. -- III. díl. S-Ž. 2008. 907 s.

Předmět. heslo
[filmoví herci -- Česko -- 20.-21. století](#)
[filmové herečky -- Česko -- 20.-21. století](#)

Chronolog. údaj * 20.-21. století

Forma, žánr * biografické slovníky

Skupina Konspektu
[791 - Film, Cirkus, Lidová zábava](#)
[929 - Biografie](#)

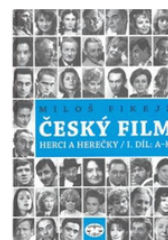
MDT
[929](#)
[791.635-051](#)
[791.635-055.2](#)
[\(437.3\)](#)
[\(031\)929](#)

Předm. hesla-angl. * film actors -- Czechia -- 20th-21st centuries * film actresses -- Czechia -- 20th-21st centuries

Forma, žánr-angl. * biographical dictionaries

[◀ Předchozí záznam](#) [▶ Další záznam](#)

Obalkyknih.cz:



Obsah

Předmluva	7
Úvod k 1. dílu	8
Seznam literárních organizací, institucí a zemí	9
Zkratky obecné	10
Herci a herečky / 1. díl A-K	13
Publikování a grafika	247
Obecní seznam herců na ukázkách z filmů	249

Záznam po opravě (doplněny a opraveny identifikátory – zobrazují se všechny díly titulu):

NKC - Úplné zobrazení záznamu



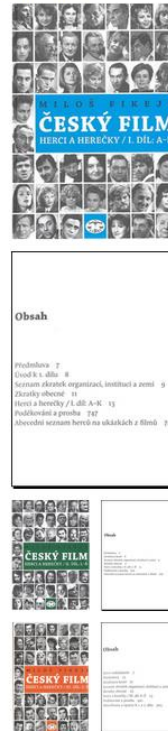
Zvolte formát: [Standardní](#) -- [Katalogizační záznam](#) -- [Stručný záznam](#) -- [MARC](#) -- [Citace](#)

Záznam 1 z 1

Hlavní záhlaví	● Fikejz, Miloš, 1959-2019
Název	● Český film : herci a herečky / Miloš Fikejz
Vydání	1. vyd.
Nakl. údaje	Praha : Libri, 2006-2008
Popis (rozsah)	3 sv. : portréty ; 21 cm
	Navrhnout na digitalizaci
EXEMPLÁŘE	
Číslo nář.bibl.	cnb001695000
ISBN	80-7277-331-3 (soubor ; váz.) 80-7277-332-1 (1. díl ; váz.) : Kč 690,00 978-80-7277-334-3 (2. díl ; váz.) : Kč 495,00 978-80-7277-353-4 (3. díl ; váz.) * 978-80-7277-354-4 (chybné) : Kč 760,00
Poznámka	Obsahuje rejstříky 3. díl obsahuje aktualizace k 1. a 2. dílu 2000 výt.
Obsahuje	I. díl. A-K. 2006. 750 s. -- II. díl. L-Ř. 2007. 656 s. -- III. díl. S-Ž. 2008. 907 s.
Předmět. heslo	● filmoví herci -- Česko -- 20.-21. století ● filmové herečky -- Česko -- 20.-21. století
Chronolog. údaj	* 20.-21. století
Forma, žánr	* biografické slovníky
Skupina Konspektu	● 791 - Film, Cirkus, Lidová zábava
MDT	● 929 - Biografie ● 929 ● 791.635-051 ● 791.635-055.2 ● (437.3) ● (031)929
Předm. hesla-angl.	* film actors -- Czechia -- 20th-21st centuries * film actresses -- Czechia -- 20th-21st centuries
Forma, žánr-angl.	* biographical dictionaries

◀ Předchozí záznam | Další záznam ▶

Obalky knih.cz:



Záznam před opravou (zobrazuje se pouze obálka o rozlišení 240 bodů, záznam neobsahuje anotaci):

Záznam po opravě (zobrazuje se pouze obálka o rozlišení 1024 bodů, záznam obsahuje nově i anotaci):



The screenshot displays the 'OBALKY KNIH' website. The main focus is the book cover for 'AEROFOTO ČESKÁ REPUBLIKA' by MIROSLAV KROB & JR. The cover art is a night-time aerial photograph of a large, illuminated castle with multiple towers and a complex roofline, surrounded by trees. The text on the cover is in white against a dark blue background. Below the cover, there is a section for 'Anotace' (Annotation) with a button to 'Upravit anotaci' (Edit annotation). The website header includes navigation links like 'Úvod', 'Přidat obálku', 'API a dokumentace', and 'O projektu'. A search bar is visible on the left, and a login section with 'E-mailová adresa', 'Heslo', and 'Přihlásit' is on the right.

Dále byly v roce 2020 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 2200 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) k dokumentům – za rok 2020 přibylo v projektu 20 tisíc anotací, které byly formálně zkontrolovány a doplněny k titulům,

Další úkoly řešené v roce 2020 mimo projekt:

- údržba a podpora skenovacího klienta pro nahrávání dat knihovny do projektu
- kontrola skenovaných periodik a opravy nalezených problémů, metodické vedení příspěvů
- upgrade použitých SW na serverech za účelem větší funkcionality a vyšší stability běhu služby
- obnova a upgrade SSL certifikátů serverů - z projektu Let's Encrypt
- aktualizace webových stránek projektu
- optimalizace archivních souborů exportních souborů na uložišti Cesnet

- úprava a rozšíření servisních stránek projektu <https://servis.obalkyknih.cz/>
- aktualizace metodických pokynů a manuálu pro knihovny a knihovní systémy
- 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>.

V Českých Budějovicích 9. 1. 2021

Ing. Jiří Nechvátal
Jihočeská vědecká knihovna v Českých Budějovicích