

BLOK — RETROBI

Retrospektivní bibliografie české literatury 1770–1945

Retrospektivní bibliografie české literatury byla a je zřejmě nejcennější materiálovou sbírkou uchovávanou a zpracovávanou v Ústavu pro českou literaturu AV ČR či jeho předchůdcích. Během bezmála šedesáti let pro ni bylo vyexcerpováno 525 časopiseckých či novinových titulů vycházejících na území českých zemí v češtině či němčině, přičemž velkorosý excerpční záběr sledoval nejen reflexi české literatury a studie o ní tak, jak je tomu u návazných bibliografických bází Česká literární věda pro období od druhé poloviny 20. století až do současnosti, ale též otisky beletrie včetně překladů a kromě literatury české byly excerpovány i texty k literaturám jiných národů, popř. i texty z blízkých oborů — zejména žurnalistiky, historie, divadla a divadelní vědy atp. Nejen pro svůj rozsah (bezmála 1 700 000 excerpčních lístků) a chronologický záběr (1770–1945, s návaznými bázemi dovedena až do současnosti), ale též pro promyšlenou metodiku svého zpracování jen obtížně hledá ekvivalent jak mezi oborovými bibliografiemi českými, tak literárněvědnými bibliografiemi v blízkém zahraničí. Dlouhá léta však zůstávala badatelské veřejnosti přístupná pouze prostřednictvím osobní návštěvy v ÚČL, což značně omezovalo možnosti jejího využití.

S nástupem moderních digitálních technologií začal být tento nedostatek pocítován stále zřetelněji a úvahy o její digitalizaci byly v rámci ústavu vedeny již od poloviny devadesátých let minulého století. Tehdy také vznikl pokus digitalizovat ji přepisem do strukturované databázové podoby. Postupně se však ukázala zásadní úskalí tohoto postupu: vedle jeho nepochybné časové a finanční náročnosti si stále větší čas vyžadovala redakce databáze, vznikající na základě v té době ještě prakticky nezredigované předmětové části. Kvůli zařazování odkazů na tentýž článek na různá místa v lístkové kartotéce se totiž do databáze v rostoucí míře dostávaly nežádoucí duplicity a tehdejší softwarové řešení prakticky neumožňovalo jejich smysluplnou eliminaci. Pokus o strukturovaný přepis Retrospektivní bibliografie pak na přelomu tisíciletí

pro nedostatečnou personální kapacitu ustrnul. Jeho výsledkem se stala databáze RET (<http://isis.ucl.cas.cz/index.jsp?form=ret>), jejíž jádro tvoří torzo přepisu předmětové části Retrospektivní bibliografie, obohacené o průběžně pořizovanou novou excerpici pro Retrospektivu. Celkem takto bylo pořízeno cca 95 000 záznamů, což sice v absolutních číslech představuje dosti vysokou hodnotu, avšak v relativním poměru šlo pouze o cca 5 % celkové počtu lístků a o cca 60 % záznamů z předmětové části. Bylo tak zřetelné, že tímto tempem by převod kartotéky přesáhl jakýkoli smysluplný časový horizont. Po neúspěchu těchto pokusů byly další úvahy o digitalizaci lístkové kartotéky Retrospektivy počátkem nového tisíciletí na čas odloženy a uvolněné pracovní kapacity byly napříměny k redigování a definitivnímu rozřazení stávající kartotéky.

Konkrétnější podobu následné plány na digitalizaci kartotéky získaly až na jaře 2009, kdy bylo rozhodnuto o podání grantového projektu digitalizace kartotéky do soutěže Ministerstva školství, mládeže a tělovýchovy v rámci programu INFOZ. V září téhož roku bylo ze strany ministerstva rozhodnuto o přidělení podpory a v závěru roku tak mohla započít realizace samotného projektu.

Formálně bylo jeho řešení ukončeno úspěšnou obhajobou v únoru 2012, nicméně dílčí upgrady softwaru byly uzavřeny až v březnu 2012. Tou dobou bylo prakticky ukončeno i skenování kartotéky¹ a výsledek projektu — webová aplikace RETROBI — začal být plnohodnotně k dispozici veřejnosti a byl následně též prezentován na odborných knihovnických seminářích a konferencích. Bohemistické veřejnosti byl pak asi nejsouvisleji prezentován v rámci „Dne s Retrospektivní bibliografií“ II. prosince 2012 v prostorách ÚČL.

Následující trojice textů by měla celý projekt přiblížit i čtenářům časopisu *Česká literatura*. Příspěvek Daniela Řeháka by měl představit Retrospektivní bibliografii jako lístkový katalog, a to jak z historického, tak metodologického hlediska, následný text Vojtěcha Malínka se pak soustředí na prezentaci její digitalizované podoby v systému RETROBI s důrazem na možnosti jejího využití pro odbornou veřejnost a závěrečná stať Elišky Davidové pak ve stručnosti přiblíží situaci literárněvědné článkové bibliografie v blízkém zahraničí.

(vm)

1 V původních předpokladech v přihlášce bylo kalkulováno s cca 300 000 oskenovanými lístky a s případným pozdějším kompletním dokončením digitalizace již mimo rámec projektu. Tento propočít byl nakonec již v průběhu řešení grantu čtyřnásobně překročen a celá kartotéka byla v momentě obhajoby projektu seskenována již prakticky celá.

K historii a struktuře Retrospektivní bibliografie

Daniel Řehák

Retrospektivní bibliografie české literatury 1770–1945 byla vytvářena po několik desetiletí za účelem základního výzkumu materiálu z novin a časopisů k ústavním pracím literárněhistorickým, literárněvědným a lexikografickým. Excerptce opatrně začala již roku 1953, ale několik let se doladovaly strukturální problémy a metodika práce. Intenzivní fáze excerptce trvala od roku 1957 do roku 1969 s tím, že se ročně odevzdávalo k zařazení mezi 60 000 až 100 000 záznamy a počet externích excerptorů-spolupracovníků dosahoval v letech nejvyšší výkonnosti až ke čtyřiceti lidem. Masivní excerptce založená především na externistech však postupně slábla, a roku 1973 byla externí excerptce ukončena. Roku 1975 pak byla konstatována praktická nemožnost pokračovat v excerptci za stávajícího finančního zajištění. Od počátku sedmdesátých let se tak excerptovalo především doplňkově a starší materiál se spíše redigoval, rozepisoval a řadil, přičemž přírůstek lístků do katalogu činil cca 20 000 až 30 000 záznamů ročně. Navíc se část bibliografického týmu plně přeorientovala na vytváření ústavních bibliografických ročenek mapujících aktuální literárněvědnou a kritickou produkci, a tak objem práce na Retrospektivní bibliografii soustavně klesal. Poslední záznamy o pracích na bibliografii pocházejí z roku 1985, kdy byla patrně správa bibliografického katalogu organizačně propojena s oddělením *Lexikonu české literatury*, k jehož vytváření sloužila především. Okleštěné personální zajištění spojené s upřednostňováním jiných úkolů vedlo v osmdesátých letech prakticky k ochromení prací na Retrospektivní bibliografii.

Nového impulzu se pracím na kartotéce dostalo až počátkem devadesátých let, kdy byl v rámci Střediska literárněvědných informací ÚČL pro její správu vyčleněn jeden úvazek, a tak se znovu rozeběhly práce směřující k jejímu dokončení a zároveň byly učiněny první kroky k jejímu převodu do elektronické podoby. V této době se také znovu začíná prostřednictvím velmi omezeného množství financí na pokrytí externích pracovníků s doplňkovou

excerpcí, která však byla s to sanovat jen její nekřiklavější výpadky z předelších let. Pracovat na Retrospektivě s takovou intenzitou jako v padesátých a šedesátých letech bylo v této době nemožné.

Dnes, po skončení prací na katalogu, lze konstatovat, že byla vytvořena jedna z nejrozsáhlejších českých článkových bibliografií, pro niž bylo zpracováno přes 500 excerpovaných časopisů, přičemž celkový počet záznamů se v době definitivního seřazení katalogu odhadoval na 1 750 000 záznamů. Dnes díky seskenování všech částí lístkového katalogu víme, že i tento odhad byl skoro o padesát tisíc vyšší než reálný stav. Pokusme se nejprve říci něco o struktuře katalogu ve smyslu jeho historie, metody a vrstvení obsaženého materiálu.

Struktura a organizace

Vybraný časopisecký a novinový materiál se excerpoval více méně ve třech fázích. **První fázi** budování bibliografie české literatury představovala excerpce periodik z první poloviny 19. století a ojedinele též časopisů starších, vycházejících už na konci 18. století. Tato vrstva je zpracována především rukopisně, tj. prostřednictvím rukou psaných lístků, a je zřejmé, že v této době se vytvářela metodika dalšího zpracovávání a že tady patrně docházelo k poučení z nedůsledností. V **druhé fázi** se excerpovala periodika z druhé poloviny 19. století a doplňkově k předchozímu období též revoluční roky 1848 a 1849. Zde se již plně uplatňovaly záznamy vytvořené na psacím stroji, ale tu a tam se stále ještě objevovaly lístky psané rukou. Je zde patrná snaha po širší excerpční základně. Ve shodě se sledovanou dobou jsou tak excerpovány i širší události národního obrozenského zápasu. **Třetí fázi** představovala excerpce novin a časopisů vycházejících po roce 1900, v jejímž průběhu se v rámci přirozené kontinuity sledovaly prvopočátky hlavních deníků prvorepublikových a poválečných (*Lidové noviny*, *Právo lidu*, *Venkov* nebo později *České slovo*). Zde se také při excerpci využívalo paměti někdejších redaktorů těchto časopisů a listů k autorizaci šifer a pseudonymů. V některých případech se někdejší redaktoři a příspěvatelé dokonce přímo podíleli na excerpci (např. A. C. Nor pro *České slovo* či později Vlasta Megová pro doplňkovou excerpce katolických kulturních časopisů). Ne všechny z těchto časopisů však byly doexcerpovány až ke konci svého vydávání nebo k historicky významným předělům. Většina záznamů z této vrstvy je zpracována na stroji. Jakkoli jsou však primární strojové záznamy dobře čitelné, jejich rozpis přes kopírovací papír dovádí některé záznamy k nekvalitě velmi blízké záznamům rukopisným.

Daleko důležitější než rozvržení prací a kvantitativní nárůst excerpce se zdá být především **metodický návrh** zpracování excerpce z pera Emanuela Macka (byl několikrát upravován, knižně vyšel až roku 1969), který je perfektní ukázkou přesných a inspirativních úvah, které stály u zrodu tohoto lístkového katalogu. Navržená pravidla jsou založena na dlouhodobé zkušenosti

s problémy, které obor přináší ve všech polohách své náročnosti a u nichž často později litujeme, že nebyly řešeny již na úrovni základního výzkumu. V tomto případě řešeny jsou, a to dokonce s důrazem na dvoustupňovou redakční kontrolu kvality excerpce. Retrospektivní bibliografie je zároveň soubor, jehož záznamy jsou často doprovázeny obsáhlými **anotacemi**, které svým rozsahem výrazně překračují obvyklé výpisy jmen k případnému rozepisování pro odkazovou část. Někdy to může vést dokonce k tomu, že informace, po které se pátrá, je nalezena již na lístku a článek se tak ani nemusí dohledávat. Důležitý je rovněž velký důraz na plný **zápis údajů** vyznačujících umístění článku či studie v novinách či časopise. V mnoha lístkových katalozích tohoto druhu se u novin uvádí prostě jen datum a stránka, někdy dokonce jen rok a číslo. V našem případě je v excerpčních pravidlech stejný důraz i na další součást bibliografické citace, tj. na ročník a číslo, což je vskutku důležitá pojistka proti chybám, které se při průběžné excerpci stanou velmi snadno a po vřazení do abecedního souboru již nelze zjistit údaje správné. **Záběr excerpce** rovněž překračuje mantinely ústavních úkolů a příslušného vědního oboru. Sledovala se totiž nejen literatura a divadlo (původní i překladové), ale okrajově se přiváděly do katalogu také články o filmu, hudbě a výtvarném umění — samozřejmě s podmínkou minimálního vztahu k literatuře nebo ke sledovaným autoritám literárním.

Základní rozvržení lístkového katalogu na **autorskou část** (autoři, překladatelé, citování autoři) a **odkazovou část** (osoby odkazované přímo nebo nepřímo; osoby, jimž jsou věnovány básně), kvůli nimž obvykle katalogy tohoto druhu vznikají, doplňují další sledované okruhy, jako jsou **ikonografické části**, obsahující odkazy na portréty a karikatury spisovatelů a publicistů nebo odkazy na autory těchto kreseb, karikatur, fotografií či v textu zobrazovaných maleb, soch či pomníků. Excerptorům bylo zdůrazňováno sledování věcných témat, která se zvláště rozepisovala a jsou uspořádána v **předmětové části**, a to především za účelem sledování mediálních témat (ohlasy na časopisy, noviny, spolky, literární žánry a další obecnosti či jednotlivosti související s literární vědou, které nebylo možné personifikovat). Tato část byla již před lety částečně převedena do databáze s názvem RET, přednostně šlo o záznamy týkající se novin, časopisů a spolků. Způsob řazení lístků v těchto oddílech je samozřejmě abecední, není však důsledně chronologický až na úroveň měsíců a dnů. V rámci jednoho autora jsou totiž záznamy řazeny podle roků, ale uvnitř jednotlivých roků jsou excerpta řazena opět abecedně podle názvů periodik, tedy tak, aby záznamy z jednoho periodika zůstaly pohromadě, což se ukázalo jako velmi praktické při větších přesunech v rámci katalogu.

Největším doplňkovým souborem katalogu je odnož autorského oddílu nazývaná **dešifrační část**, která obsahovala záznamy podepsané šiframi a pseudonymy autorit, jejichž jména nebyla zjištěna. Tato část je řazena abecedně podle šifer, dále podle názvů časopisů a nakonec podle roků vydání.

V poslední fázi pořádání této části byly vyřazeny všechny jmenné a slovní pseudonymy a přesunuty do autorské části. Pokud byli původci šifer a krycích jmen zjištěni, byl podle excerpčních pravidel vytvořen lístek, který odkazuje na první výskyt šifry nebo pseudonymu v daném ročníku časopisu či čtvrtletí deníku s vyznačením jména autority. Tyto odkazy na provedené dešifrace, a tudíž na lístky vřazené do autorské části, jsou seřazeny v **dešifratové části**. Řazeny jsou abecedně podle formy šifer a pseudonymů, dále podle jmen zjištěných autorit, dále podle časopisů či novin a konečně podle roků vydání.

Dešifrace a autorizace

Právě zpracovávání části katalogu zabývající se šiframi a pseudonymy, kterou jsme pořádali posledních pět let před ukončením prací na souboru v lístkové podobě, ukázalo některé nedostatky a slabiny lístkového katalogu. Je to jediný případ, kdy jsme postupovali proti pravidlům navrženým Emanuelem Mackem v metodické příručce *Bibliografie české beletrie a literární vědy* z roku 1969. Oddělení dešifrátů odkazujících na provedené dešifrace od nedešifrovaných záznamů a jejich abecední seřazení do jednoho celku je inovací oproti původním záměrům, podle kterých mělo být obojí ponecháno v jednom souboru a řazeno abecedně pouze v rámci souborů pro jednotlivé noviny a časopisy. Rozměrový nepoměr navrhovaných titulárních oddílů (12 zásuvek pro významný deník, jako jsou *Lidové noviny*, a 5 lístků pro dobře dešifrovaný časopis) a zkušenost, že jedna šifra se může pohybovat napříč několika časopisy i novinami, takže dešifrace v jednom periodiku nás může dovést k obdobným záznamům v jiném, nás přivedly k tomu, že byl vytvořen jeden soubor, kde jsou stejné šifry řazeny nejprve podle zdroje, a teprve potom chronologicky. Zároveň byly z tohoto souboru vyjmuty všechny navrhované pseudonymy (i slovní obraty) a byly vřazené do autorské části, protože bylo zjištěno, že u nápodob jmen nebylo někdy možné zjistit, zda (ne)jde o pseudonym, takže se po letech pořádání těchto souborů stalo, že stejný pseudonym měl zastoupení jak v části autorské, tak dešifrační, tj. na dvou místech, z nichž alespoň jedno nebylo možné předpokládat. Jakkoli nejde o řešení ideální, byla to zkrátka jediná možnost nějakého smysluplného řešení problému.

V rámci prací na dešifračních záznamech se také ukázalo, jak mylná je představa — téměř psychologická past — většiny badatelů, kteří se domnívají, že pokud mají ve své rešerši nějaké dešifrace, jsou patrně u příslušného autora provedeny všechny. Vzhledem k přibývání materiálu do katalogu v plynoucím čase je něco takového prostě nemožné, takže každá autorská rešerše by se měla tázat po všech autorových identitách, variantách jména, pseudonymech i šifrách. Při dopracovávání této části byly dále odhaleny rozsáhlé nesoustavnosti, které spočívaly především v tom, že při excerpci jednoho ročníku novin se někdy i po čtvrtích roku střídali excerptoři. To způsobilo značné kolísání kvality v tom smyslu, že na to, co dešifroval jeden, druhý

nemusel přijít, a podobně. Šifry z prvního kvartálu tak skončily v autorské, šifry z druhého zůstaly nedešifrované v dešifrační části katalogu. Ba co víc: jeden excerptor určil u šifry v příslušném ročníku jednoho autora, jeho kolega zase jiného. Obojí skončilo bez vzájemné komunikace v autorské části na odlišných místech a navíc případně zanašelo chyby do literárněvědné literatury. V rámci dokončovacích prací tak bylo vyvinuto velké úsilí o sjednocení těchto výkyvů a případné odstranění explicitních omylů při dešifracích, a to i v případě, pokud se tyto chyby dostaly v rámci rozpisu pro odkazovou část do jiných částí katalogu, než je autorská.

Tyto nároky na kvalitu dešifrací však přinesly jedno velmi důležité poznání, které vyvrcholilo konstruktivní kritikou stávajících příruček pro šifry a pseudonymy a konečně především stávajících souborů autorit a jejich osobních údajů poznamenaných vynechávkami a zámlkami z období nesvobody tisku. Vše vyplynulo během několika málo let z toho, jakým způsobem se měnil náš přístup k dešifracím. Nejprve se používalo — stejně jako po celá léta — standardních postupů, jako jsou prosté kontroly v rejstřících základních příruček (Vopravilův *Slovník pseudonymů a šifer*). Následovalo využití soupisu šifer a pseudonymů zaznamenaných v *Lexikonu české literatury*, protože jeho hesla byla vytvořena na základě rozsáhlých rešerší či archivního výzkumu a obsahují mnohdy nové poznatky. Zde však bylo zároveň konstatováno, že *Lexikon* zaznamenává také některé omyly převzaté z našeho katalogu (např. šifra *drb.* byla chybně připisovaná Václavu Brtníkovi, přitom patřila Jaromíru Boreckému). Takové případy vedly k větší obezřetnosti při dešifracích a jisté nezávislosti na běžně užívaných příručkách, konkrétně pak k rešerším, jimiž se dohledávala jména autorů šifer prostřednictvím sledovaných témat v katalogu samém, konkrétně v předmětové části, kde se k sobě v zúženém spektru přibližovaly šifry a jména v jednom periodiku a čase natolik, že bylo možné nabýt jistoty potřebné k dešifraci. Pomalu se pronikalo do dalších, někdy extrémních, možností dešifrace jako třeba úplné obrácení problematiky v dohledávání iniciálových šifer na základě jmen evidentních publicistů. Vše vyvrcholilo studiem archivních pramenů, zvláště redakčních archivů vybraných periodik (např. *Lidové noviny*, *Národní listy*), nepoužívaných dotazníkových souborů (*Kulturní adresář ČR*, *Československá biografie*), jmenných spisoven (Ministerstvo lidové osvěty, tiskový odbor MZV), redakčních zápisníků (např. Eduard Bass) nebo studiem tematických výstřižkových souborů (ZTA nebo MZV VA v Národním archivu). Jakkoli byly na začátku těchto prací pochybnosti o míře jejich vlivu na celek katalogu, návratnost časové investice v podobě množství záznamů převedených z dešifrační do autorské části tyto otázky zodpověděla velmi přesně. K dnešnímu dni byly provedeny dešifrace tisíců záznamů v počtech, které odpovídají přírůstku katalogu z běžné roční excerptce. Navíc se zdá, že tyto práce budou mít zcela zásadní přínos metodický, protože se na základě zkušeností

s dešifracemi podařilo vytvořit jakýsi vzorec pro správnou dešifraci, který spočívá ve splnění několika kritérií: šifra musí odpovídat jménu v souvislosti se střídáním jména a šifry v jednom časopise a jednom roce při sledování stejného nebo podobného tématu. Splněním těchto kritérií, stejně jako revizí chyb ve stávajících příručkách by se mělo v budoucnu riziko špatné dešifrace při excerpci zcela minimalizovat.

Nedůslednosti a chyby

Tento obraz lístkového katalogu Retrospektivní bibliografie možná působí i pro zasvěcené bibliografy značně monumentálně. S monumentalitou a promyšleností jde však provždy ruku v ruce také monstróznost a nepřesnost, základní projevy lidského faktoru. Po zkušenostech se zpracováním katalogu nelze než konstatovat, že je jistě skvělé všechno dobře promyslet hned na počátku, ale záleží rovněž na dobré organizaci těchto prací v průběhu let a na průběžné kontrole odevzdávané práce. Zde se patrně nachází — jako u všech ostatních lístkových katalogů — pro uživatele utajená problematika práce s celým souborem, kterou je však důležité zpřehlednit, protože nám může umožnit neopírat se s přehnanou důvěrou o tento olbřímí soubor dat, který může být stejně velkolepou kupou informací jako podivnou hromadou vybraných údajů, které dokážou notně vychýlit badatelský náhled na sledovaný problém. Vždy záleží na úhlu pohledu a seznamu excerpaných periodik by se jako důležitá uživatelská příručka neměl podceňovat. Zkrátka to, že na daný dotaz nenalezneme žádný záznam v katalogu, neznamená, že se hledanému tématu v publicistice nikdo nevěnoval.

K základním problémům spojených s katalogem, na které jsme v průběhu let narazili, patří především následující:

Pokud je známo, rozsah zamýšleného celku ani taktika výběru časopisů nebyly přesně stanoveny, odvíjely se částečně od momentálních potřeb ústavních úkolů. V tomto ohledu je přítomnost nebo absence některých titulů v souboru téměř nahodilá. Tím také dochází k velkému výkyvu pokrytí určitých roků, což se samozřejmě odráží v četnosti zpracování recenzí na určité knihy nebo překlady vydané v tom nebo onom roce. Je vlastně velmi zvláštní, že metodika zpracování byla propracována tak pečlivě, když zároveň excerpcí základna utonula v naprosté bezbřehosti. K tomu přistoupil také problém materiálu, který byl odevzdán, prošel kontrolou, ale nebyl desítky let vřazen. A když, tak pouze na základě náhodně odevzdané práce nebo podle pořadí časopisů uložených v pracovním katalogu. Proč se o tom vůbec zmiňovat? Tento problém se totiž ve vlnách zpracování přeléval do oborových publikací prvního řádu, jako byl především *Lexikon české literatury*. Rozdíl mezi prvním, respektive druhým, a posledním dílem této literárněhistorické příručky je vzhledem k postupu dokončovacích prací na řazení kartotéky, které byly vykonány mezi roky jejich vydání, velmi znatelný.

Dále je zřejmé, že výběr titulů k excerpci byl v počáteční fázi ovlivněn ideologicky a dosud se nepodařilo tuto „závadu“ zcela odstranit, což je patrné zvláště u periodik vycházejících za první republiky. Příklad za všechny — excerpce časopisu *Přítomnost* ve shodě s obstavením tohoto periodika v československých knihovnách neproběhla. Nicméně, největší dluh zde představuje především nedokončená excerpce dvou významných deníků s přesahem do období protektorátu. Jde o *Národní listy* (pouze do roku 1922) a *České slovo* (pouze do roku 1934). Zdá se, že v některých z těchto případů došlo dokonce ke ztrátě již vypracované excerpce během stěhování ústavu ze Strahova na Náměstí republiky. Dále je znatelná kompletní nepřítomnost významných deníků moravských (*Moravsko-slezský deník*) nebo katolických (*Lidové listy* nebo *Našinec*). O mnoha literárních časopisech, které se nestihly vyexcerpovat, ani nemluvě. Zároveň je třeba zdůraznit, že jistým zázrakem byla ve své době vyexcerpována — alespoň z bohemistického hlediska — německy psaná periodika jako *Ost und West*, *Bohemia*, *Prager Zeitung*, *Prager Abendblatt*, *Prager Abendzeitung*, *Prager Presse* nebo *Prager Tagblatt*. Podle původního záměru však měly být meziválečné německé noviny uloženy zvláště, a to navzdory stejnému způsobu zpracování i tematickému záběru. Tento záměr jsme přehodnotili a celý soubor prvorepublikových německých deníků byl vřazen do katalogu k excerptům z periodik českých.

Dalším velkým problémem katalogu je nevyjasněný časový rozsah jeho celku a nevyváženost záběru excerpce v určitých obdobích. Katalog měl programově končit rokem 1945 (to není úplně vhodně zvolený předěl vzhledem k tomu, že definitivní osvobození Československa přišlo v květnu, tj. v polovině tohoto roku). Nicméně excerpce některých časopisů byla ve své kontinuitě dovedena až do roku jejich zániku, tj. roku 1948 (např. *Kritický měsíčník*, *Střední škola*), nebo tam naopak dovedena nebyla. To se týká například časopisů *Akorď* nebo *Archa*, které byly doexcerpovány až při celkovém dokončování katalogu a přípravě k jeho skenování. Je však předpoklad, že období 1945–1948 by alespoň v rozsahu sekundární oborové bibliografie mělo být pokryto do konce roku 2015 v rámci stávajícího grantového projektu „Bibliografie české literární vědy 1945–1960“. S tím souvisí také obrovský počet málo vysvětlitelných mezer v chronologii excerpce. Zdá se samozřejmé, že deník nebo časopis se prostě excerpuje tak, že se články vypisují od prvního do posledního ročníku. Tak se to však vůbec nedělo. Některé tituly byly excerpovány doslova ob rok, jeden ročník ano, druhý ne, další ano atd. Těchto mezer obsahoval katalog podle databáze excerpovaných časopisů ECR kolem stovky. Nejhorší z těchto zlořádů byly odstraněny během posledních pěti let před skenováním — některé doslova těsně před ním. Se špatným přehledem po skutečně a budoucí excerpci také souvisí naprosto neblahá skutečnost dvojí excerpce stejného titulu a roku, neblahý to znak rozsáhlosti prací a stupňující se dezorganizace. Tento politováníhodný jev se čas od

času přihodil prostě na základě absence excerptční karty, která by mu mohla zamezit. Na podobné duplicity se pak přišlo až při vřazování rozepsaných a předřazených záznamů do katalogu. Pak nezbývalo než vybrat kvalitnější lístek a horší odstranit, nebo ponechat oba, pokud se záznamy svým popisem výrazně lišily.

Také se bohužel nelze zcela spolehnout na výše uvedené programové proklamace o propracovaném rozpisu a systematickosti v katalogu. Podle praktických zjištění ne všichni překladatelé jsou rozepsáni, ne po všech provedených dešifracích jsou zachovány dešifrační lístky, ne všechny odkazované osoby jsou rozepsány, ne všechny odkazové lístky jsou v odkazové části, ne všechny záznamy obsahují v bibliografické citaci i čísla a ročníky zdrojových dokumentů, ne všechny záznamy obsahují čísla stran, pokud bylo nutné je dopočítat (*Hlas národa, Rudé právo*). Zkrátka a dobře: každý excerptor chť nechtě nakonec vyexcerptuje tak trochu něco jiného a nějak jinak než jeho kolega. Ne snad výrazně, ale přece jen se excerptce dvou lidí musí nutně lišit nejen množstvím vytvořených lístků, ale právě především kvalitou, množstvím dohledaných údajů, dešifracemi nebo připomenutím souvislostí v anotaci, zvláště pokud je excerptční základna tak široká jako v případě Retrospektivy, kde je odklon od čistě literární tematiky někdy opravdu veliký.

Jakákoli kontrola či redakce odevzdaných lístků — byť by byla součástí excerptčních pravidel jako v našem případě — nemůže těmto zlořádům zabránit. Například zamýšlená redakce všech záznamů poměrně brzy — navzdory dochovaným výkazům — zkolabovala, protože jejími nároky se překročily již tak napjaté časové možnosti tohoto velkorysého projektu, který by dnes nebo vůbec kdy jindy patrně vůbec nevznikl. Redakční práce se totiž ukázaly velmi náročným doplňkem základních rozepisovacích a rozřazovacích prací na souboru a po drastickém omezení finanční dotace na správu kartotéky na počátku sedmdesátých let pak bylo udržení smysluplného tempa všech těchto prací nad síly vyčleněných pracovníků.

Opravdu velký problém představuje předmětová část katalogu, u níž lze s ohledem na její rozsah jednoho katalogu sotva uvěřit, že jde skutečně o věcný popis 18 katalogů autorské a odkazové části. Zdá se, že mnohé věcné označení v pravém horním rohu lístku vybízející k rozpisu zůstalo oslyšeno, tedy pokud bylo vůbec vyznačeno. Tyto nedostatky by snad mohly být částečně sanovány právě skenovanou verzí katalogu a jejím OCR vyhledavačem. V tomto ohledu má nicméně katalog do budoucna velké zastání v digitálních knihovnách Kramerius, které jednotlivě provozuje většina významných českých knihoven (Národní knihovna jako koordinátor digitalizačního projektu VISK přitom většinu těchto dat přebírá do své digitální knihovny). V nich lze pomocí fulltextu vyhledávat až na úroveň jednotlivého slova. Věřit v samospasitelnost podobných souborů skenů a jejich vyhledavačů by však bylo neopatrné, protože tyto archivační systémy nejsou a patrně ještě dlouho ne-

budou schopny řešit problematiku šifer a pseudonymů, stínových to identit mnohých publicistů, natož pak homonymii jednotlivých věcných hesel či osobních jmen atd.

Závěr

Máme-li shrnout práci posledních dvaceti let, můžeme — co se týče kvantity — konstatovat, že bylo do katalogu přivedeno několik stovek tisíc dosud nevyužívaných záznamů k aktivní činnosti, a tím byly posíleny autorská a odkazová část celkem o další čtyři katalogové bedny a k dispozici je nyní také abecedně seřazený soubor nedešifrovaných záznamů a menší soubor zaznamenávající provedené dešifrace. Spolu s výraznými posuny uvnitř katalogu došlo k celkovému navýšení počtu záznamů o několik set tisíc záznamů. Je třeba poděkovat všem, kteří se podíleli na dokončování katalogu Retrospektivní bibliografie za jejich osobní podíl na věci, ať už byl jakkoli velký. Mohou si považovat za čest, že jsme se mohli zúčastnit něčeho tak významného, jako je dovést do více méně optimálního stavu největší oborovou článkovou bibliografii v této zemi, a zpřístupnit ji tak badatelům v snadno dosažitelné podobě, jako je internetové prostředí, a konečně nezůstat v posledku pouze u rozepisování a řazení lístků, ale v rámci poučení z chyb předchůdců a správného kladení otázek přinést něco nového k minulosti a doufejme i budoucnosti článkové bibliografie.

RETROBI — Digitalizovaná Retrospektivní bibliografie české literatury 1770–1945 a možnosti jejího využití pro literárněvědný výzkum

Vojtěch Malínek

Projekt s názvem „Digitalizace lístkového katalogu Retrospektivní bibliografie české literatury 1770–1945“ (VZ09004) byl v ÚČL řešen od října 2009 do konce roku 2011 a jeho realizace se záhy rozdělila do tří hlavních částí: finální redakce kartotéky spojené s poslední dávkou excerptce, samotné digitalizace lístků a do vývoje prezentačního software.

Již před podáním grantové přihlášky bylo rozhodnuto o digitalizaci kartotéky seskenováním, která se jevila jako rychlejší a efektivnější řešení než ruční přepis. Byť bylo v tomto ohledu možné částečně se opřít o předchozí zkušenosti jiných institucí, zejména velkých knihoven, přistupoval ÚČL ke skenování své kartotéky cca 8–10 let po hlavní vlně digitalizace lístkových kartoték v českých knihovnách přece jen v odlišné situaci — zaprvé se nabízely nové hardwarové nástroje, tj. zejména skenery, a zadruhé před námi stála možnost využít výtvarných, které relativně nově nabízejí moderní technologie — zejména šlo o OCR rekognoskaci textových dat. Využití OCR sice nebylo závaznou součástí původního návrhu projektu, s postupem analytických prací se však řešitelský tým rozhodl OCR přepisů jednotlivých lístků využít už jako přímé součásti softwarové aplikace, která měla v rámci projektu vzniknout dle původního návrhu jen s tím, že bude na implementaci OCR struktury v budoucnu předpřipravena.

Využití OCR technologie však neznamenalo jediný rozdíl oproti běžně rozšířeným systémům pro prezentaci lístkových katalogů. Oproti lístkovým katalogům knihoven byl totiž specifický i sám lístkový materiál v kartotéce Retrospektivy. Vyskytují se v ní totiž lístky jednoobrázkové, excerpta psaná po obou stranách lístku i záznamy vícelístkové, lístky různé kvality i barvy písma (rukopis, strojopis, tisk), excerptní záznamy byly v průběhu času porizovány na různé druhy papíru (od průklepového až po těžší škartkový) atp. a zejména — zařazení lístku v kartotéce není jednou provždy dané, ale oproti jiným lístkovým katalogům v ní dochází k značným pohybům, zejména

kvůli novým dešifracím. Navíc byl unikátní i sám materiál: zatímco knihovny mohly počítat s tím, že v rámci sdílené katalogizace budou moci do budoucna využít digitalizované záznamy jiných institucí, oborové článkové bibliografie — obzvláště pro období před rokem 1945, odkdy začínají nejstarší excerpta z výběrové článkové bibliografie národní — srovnatelný korelativ v českém prostředí nemají.

Plánovaná webová aplikace proto dle vstupních rozvah řešitelského týmu — Jana Anděla, vedoucího Střediska výpočetní techniky ÚČL, který se podílel na analytických pracích a návrzích jednotlivých komponent systému, Daniela Řeháka, dlouholetého správce katalogu a jeho největšího znalce, Petry Černohorské a Pavla Miky, kteří měli na starost samotné skenování kartotéky, Ondřeje Komendy a Vojtěcha Hordějčuka z firmy inSophy, autory softwarového řešení programu, manželů Elišky a Stanislava Kosových, de facto posledních excerptorů lístků pro kartotéku, a autora tohoto článku jako řešitele projektu — neměla sloužit pouze jako prostá digitální obdoba původní lístkové kartotéky, ale byla koncipována jako svébytné badatelské prostředí, které by za pomoci moderních softwarových nástrojů umožnilo badatelům co nejvyšší kvantitativní i kvalitativní využití prezentovaných dat. Její základní jednotkou se tak neměl stát jen grafický soubor reprezentující původní lístek, ale jakýsi virtuální lístek, který by v sobě vedle informací grafických (původní předloha) a systémových, identifikujících daný virtuální lístek v rámci systému, obsahoval i informace textové, které by umožnily prohledávání kartotéky i jinak než pouhým mechanickým listováním. Řešitelský tým se rozhodl připravit takové řešení, které by jednak umožnilo aktivní zapojení uživatelů při tvorbě a zejména korekcích obsahu webové aplikace a které by zároveň bylo co nejvíce otevřené pozdějším úpravám a nástavbám, jež by mohly v budoucnu napomoci převodu dat Retrospektivy do podoby strukturované databáze.

Webová aplikace RETROBI

Webová aplikace RETROBI je primárně určena k prohlížení oskenované lístkové kartotéky Retrospektivní bibliografie v elektronické podobě, zároveň však obsahuje řadu pokročilých softwarových funkcionalit pro její další komfortní badatelské využití. Systém je volně a bez jakýchkoli omezení dostupný online na <http://retrobi.ucl.cas.cz/>. Zároveň existuje volitelná možnost vytvořit si v rámci aplikace vlastní uživatelský účet (nabídka „Registrace“). Registrace však není koncipována jako povinná a slouží především k možnosti ukládat si v rámci systému individuální data. Jednotlivé funkce celé aplikace, především s důrazem na jejich praktické využití pro literárněvědného badatele, si nyní můžeme představit podrobněji.¹

1 Při představování jednotlivých funkcionalit zůstaneme na obecné rovině, pro popis konkrétních technických postupů odkazujeme na vlastní nápovědu webové aplikace.

Každý, tj. i neregistrovaný, uživatel může v základní verzi procházet celou bibliografii tak, jako by měl před sebou původní lístkovou kartotéku (nabídka „Katalog“). Zároveň je mu však umožněno i fulltextové vyhledávání nad textovými podobami jednotlivých lístků (nabídka „Vyhledávání“), jemuž se bude věnovat podrobněji v dalším. V rámci aplikace si uživatel může též zpracovat, seřadit a vyexportovat libovolnou rešerši, přičemž je mu k dispozici konfigurovatelné nastavení typu, formátu a kvality zvolených dat (nabídka „Schránka“).

Řada funkcí je pak navázána přímo na jednotlivý lístek. Z nich je asi nejzajímavější možnost editovat a upravovat původní OCR přepisy jednotlivých lístků, a to buď opravou stávajících chyb, nebo jejich zcela novým přepisem, a tím zkvalitňovat a dále strukturovat množinu textových dat v systému (viz dále). Tyto úpravy jsou otevřeny každému registrovanému uživateli. Vedle toho může uživatel zanechat u jednotlivého lístku zprávu administrátorům systému pomocí speciálního modulu pro hlášení chyb (např. v případě nekvalitní grafické předlohy, chybného řazení či věcného omylu na lístku) či uložit si k němu v rámci aplikace svůj vlastní komentář.

K dispozici má uživatel systému i základní statistiky. Ve stejnojmenné nabídce mu jsou vedle základních číselných přehledů o počtu lístků dle jejich stavu či dle jednotlivých částí databáze připravovány i pravidelně každodenně aktualizované rejstříky pro jednotlivá pole strukturovaného položkového rozpisu. Do budoucna se počítá s automatickou indexací polí „Název zdroje“, „Rok vydání“ a „Charakteristika“, která jsou jednak nejspíše indexovatelná, a zároveň lze očekávat, že právě tato pole budou nejspíše nejvíce využívána při položkovém vyhledávání.

Pokročilejší úpravy (editace databázových polí, úprava obrazových dat aj.) jsou pak dostupny administrátorům systému, kteří mají k dispozici poměrně robustní rozhraní pro správu celé aplikace. Jeho nejzajímavější součástí představují zřejmě moduly pro hromadnou editaci dat vybrané skupině lístků, díky němuž by do budoucna mělo být možné např. alespoň nad určitou množinou lístků zprovoznit strukturované položkové vyhledávání. V administrativním modulu je dále možné zpracovávat jednotlivá uživatelská hlášení a uživatelské účty, nastavovat systémové parametry báze či získat statistické přehledy o jejím využití.

Fulltextové vyhledávání a jeho možnosti

Jednou z hlavních inovací, kterou aplikace RETROBI nabízí, je bezesporu možnost fulltextového vyhledávání, díky němuž je značně rozšířena škála dotazů, které je možno nad záznamy Retrospektivní bibliografie provádět. Avšak dříve, než se pokusíme nastínit možnosti, které fulltextové vyhledávání v RETROBI badatelům nabízí, je třeba upozornit na několik základních, avšak pro nepoučeného uživatele možná ne zcela samozřejmých omezení, která je třeba mít před jeho využitím na zřeteli.

Především je třeba zdůraznit, že **vyhledávání v aplikaci RETROBI je primárně založeno na fulltextovém prohledávání v OCR prepisech jednotlivých lístků a tyto OCR prepisy mají samozřejmě dosti rozdílnou kvalitu** podmíněnou vlastnostmi původní grafické předlohy. Zatímco lístky tištěné na tiskárně, popř. originály strojopisu, byl OCR program s to přečíst s chybou v řádu jednotek procent, u horší kvality předlohy (např. průklepové kopie strojopisu) chybovost samozřejmě stoupá a u lístků psaných ručně nebylo možné snad s výjimkou tiskacích písmen v záhlaví lístku přečíst vůbec nic. V takovýchto případech obsahuje prepis lístku hodnotu „\$bez_ocr\$“.² Při stanovení množiny ručně psaných lístků lze jen obtížně generalizovat, avšak obecně lze tvrdit, že rukou byla zapisována především nejstarší excerptce, tj. zejména periodika první poloviny 19. století, ale v závislosti na osobě excerptora můžeme narazit i na ručně psané lístky s excerptcí některých deníků z 20. století (např. některé ročníky *Lidových novin*) či určitých okruhů časopisů (např. většina periodik avantgardy z 20. let či prvorepublikové časopisy katolické). Hrubým odhadem jsou však pro fulltextové vyhledávání v rozdílné kvalitě přístupny cca 2/3 až 3/4 všech lístků a lze odůvodněně předpokládat, že se — pomineme-li průklepové kopie — v kartotéce nachází ve velké většině případů alespoň jeden exemplář lístku ve slušné grafické kvalitě, která umožnila kvalitní OCR rekognoskaci. I pokud zůstaneme u dolních hranic odhadovaného množství takto prohledatelných lístků, dostáváme se k alespoň 500 000 jedinečných článků, s nimiž lze v Retrospektivní bibliografii pomoci fulltextového vyhledávání pracovat. Další limitující faktor pro fulltextové vyhledávání pak představují různé doplňky a vpisky vzniklé při redakci záznamu či redakci kartotéky — z pohledu OCR rekognoskace naneštěstí rukopisné.

Při zohlednění výše popsanych skutečností je třeba znovu podtrhnout, že **fulltextové vyhledávání není** — alespoň prozatím — **primární metodou prohlížení stávajících dat, obzvláště pakliže typu daného dotazu lépe vyhovuje výchozí řazení kartotéky**. Hledá-li tedy uživatel texty od nějakého konkrétního autora či o nějaké konkrétní osobě, poslouží mu mnohem lépe listování v kartotéce, která byla dle těchto kritérií řazena, a dotaz přes fulltext mu může maximálně napomoci odhalit chybně zařazené lístky, případně multiplikáty ztracených lístků — což ovšem nemusí být vždy zcela zbytečné kontrolovat.

Svůj význam však už fulltextové vyhledávání má při vyhledávání určitého specifického věcného předmětu či tématu nebo při vyhledávání časových údajů či některých selekčních termínů typu název časopisu, charakteristika článku atp., protože v těchto případech je fulltextové vyhledávání jedinou možností, jak Retrospektivní bibliografii vytežit, a samozřejmě pak při omezení dotazu na uživatelem specifikovanou množinu lístků.

² Takovýchto lístků je v systému 74 951, u zbylých rukopisných pak byla prostřednictvím OCR rozpoznána alespoň náhodná nesouvislá sada znaků.

Obecně platí, že relevantní informace systém poskytne především v případech, kdy hledáme nějaký úzce specifický termín (např. cenzura, daktyl, manýrismus atp.), nebo naopak hodnotu, která se v zápise na excerpčním lístku dle excerpčních norem musí vyskytovat vždy (např. název zdrojového periodika či chronologický údaj) nebo jejíž zápis bývá pevně normován, tj. spojen s nějakou pevně danou sekvencí znaků (údaje o jazyce originálu, charakteristice dokumentu, ale opět i název periodika aj.). Tyto případy se nyní pokusíme rozebrat detailněji.

Vyhledávací algoritmus a syntax dotazovacího jazyka

Pro fulltextové vyhledávání v aplikaci RETROBI je používán vyhledávací engine Lucene.³ I před položením dotazu ve fulltextovém vyhledávání je třeba uvědomit si, že při prohledávání systém pracuje — alespoň v drtivé většině — nejen s neverifikovanými (viz výše), ale též s nestrukturovanými daty a je tak třeba počítat s odlišným chováním, než na jaké je běžný uživatel zvyklý při prohledávání strukturovanýchází (např. knihovních katalogů atp.), a především i s odlišnými výsledky. Ne vše, co na papíře vypadá stejně, musel totiž zcela identicky rekognoskovat i OCR program a jeden drobný rozdíl v rekognoskaci (přečtena tečka namísto čárky; přečtena mezera navíc, záměna jednoho písmene za druhé atp.) může výsledky vyhledávání značně ovlivnit. Vyhledávací engine navíc při zpracování řešerše nepracuje se žádnými sémantickými úpravami, ani neindividualizuje výslednou odpověď dle koncového uživatele jako např. Google, ale nabídne vždy všechny záznamy, které plní danou podmínku. Při vyhledávání tedy fulltextový engine prochází jednotlivé nestrukturované a nehierarchizované textové řetězce (resp. v drtivé většině nestrukturované, protože strukturace je umožněna jednotlivým registrovaným uživatelům systému a vzniká tedy až následně) a není proto s to odlišit např. homonymní termíny či výsledky řešerše dále hierarchizovat.

Vyhledávací engine však obsahuje řadu pokročilých nástrojů, které umožňují dotaz zpřesňovat a specifikovat a díky tomu představuje, domnívám se, velmi dobrý nástroj pro prohledávání textových dat. Kladení „náročnějších“ dotazů však vyžaduje určitou obeznamenost s interpunkcí vyhledávacího dotazu, která sice může na první pohled pro nového uživatele představovat problém, ale to spíše proto, že se dosud s takovýmto nástrojem nesetkal, než kvůli tomu, že by byl nástroj složitý.

Z hlediska uživatele je pak třeba ještě zdůraznit, že vyhledávací engine prohledává primárně vždy určitou ucelenou sekvenci znaků od mezery do mezery. Hledání víceslovných výrazů či slovních frází je též možné, ale pro takovýto dotaz je potřeba použít speciální interpunkci (zde konkrétně

3 Oficiální dokumentace je přístupná na adrese <http://lucene.apache.org/core/>.

uzavřít celou víceslovnou frází do uvozovek). V rámci jednoho souvislého řetězce znaků od mezery k mezeře však systém umí pracovat i s tzv. divokými kartami, tj. znaky, které z hlediska interpunkce vyhledávacího dotazu zastupují znak či znaky jiné (konkrétně ? zastupuje právě jeden znak, * pak libovolné množství znaků od nuly do nekonečna; oba znaky je přitom možno v rámci jednoho dotazu vzájemně kombinovat). Tato funkce je pro prohledávání neverifikovaných textových dat typu OCR rekognoskací velmi šikovná, neboť umožňuje odfiltrovat řadu chyb vzniklou chybnou OCR detekcí). Vyhledávací engine dále nabízí možnost pokládat vyhledávací dotaz s využitím základních logických operátorů ANO (znak +), NE (znak -), NEBO (znak ||) a tyto logické operátory vzájemně kombinovat. Další velmi vítanou možností je tzv. intervalové vyhledávání, kdy na základě znalosti abecedního rozsahu je možno položit dotaz typu „najdi všechny lístky, na nichž je libovolné slovo, začínající písmenem, které je v abecedě mezi E a H“. Pro práci s RETROBI je pak intervalové vyhledávání použitelné zejména pro vyhledávání číselných rozsahů a umožňuje tak nad materiálem Retrospektivní bibliografie klást dotazy typu „najdi všechny záznamy obsahující číslici mezi 1887 a 1893“.⁴ Vítaná je i možnost aktivovat či deaktivovat rozlišování velkých a malých písmen. Okrajovější využití pro data Retrospektivní bibliografie pak budou zřejmě mít funkce vyhledávání slov dle jejich podobnosti (při „rozšíření“ pásma podobnosti se objem vzniknuvší rešerše násobně a vzhledem k objemu dat v Retrospektivě často až neúměrně zvětšuje) či vyhledávání slov (řetězců) vzdálených od sebe n jiných slov (řetězců), ale i pro tyto možnosti vyhledávacího engine lze jistě nalézt v určitých případech využití.

Všechny tyto jednotlivé funkce jsou umožněny vložením speciálních interpunkčních znaků do vyhledávacího řetězce⁵ a jejich stručný přehled podávají tabulky na následující straně.

Vedle základního prohledávání OCR přepisů je systém připraven i pro prohledávání konkrétněji vymezených množin dat, ať již jde o vyhledávání dle stavu lístku, jímž např. lze vyhledávání omezit jen na lístky přepsané, nebo omezení vyhledávání jen na určitou vybranou množinu záznamů (předdefinovaná část či části katalogu; uživatelská schránka atp.), či dokonce omezení na určitou konkrétní hodnotu položkového rozpisu, pro jehož vyplňování je

4 Vyhledávací engine ovšem nevyhledává jen všechny číslice ze souvislé řady, ale z jakéhosi „abecedního“ řazení, v němž se nezohledňuje různý počet řádů jednotlivých čísel. Systém tedy řadí 18, 180, 1800, ..., 181, 182, ..., 189, 19 atd. V našem konkrétním případě by tak kromě čísel s pěti a více řády našel hodnoty 1887, 1888, 1889, 189, 1890, 1891, 1892 a 1893.

5 Před znaky, které mají význam z hlediska strukturace rešeršního dotazu, je v případě, že je chceme vyhledat v textovém řetězci, potřeba napsat zpětné lomítko.

Interpunkce vyhledávacího dotazu

DIVOKÉ KARTY VS. OHRANIČENÍ FRÁZE

Znak	Význam	Příklad dotazu	Příklad odpovědi
*	Libovolně dlouhá souvislá řada znaků (o–nekonečno)	st*l	stal, strojil, stůl, stodol
		*st*l*	přistáli, ustrojili, stoly, stodola
?	Právě jeden libovolný znak	st?l	stůl, stal, styl
""	Ohraničení fráze	"písně kosmické"	Najde všechny výskyty této slovní sekvence

VYHLEDÁVACÍ ZNAKY

Znak	Význam	Příklad dotazu	Příklad odpovědi
""	Ohraničení fráze	"písně kosmické"	Najde všechny výskyty této slovní sekvence (nelze využít divoké karty)
~	Podobnost	stůl~0.8	Najde všechna slova podobná výchozímu z 80 %
""~	Vzdálenost	"písně kosmické"~10	Najde všechny případy, kdy jsou obě slova vzdálena maximálně n slov od sebe
{ }	Exkluzivní interval	{1901 TO 1905}	Najde všechny případy, kdy se v definované množině textů vyskytuje řetězec spadající do uvedeného intervalu s vyloučením krajních hodnot
[]	Inkluzivní interval	[1901 TO 1905]	Najde všechny případy, kdy se v definované množině textů vyskytuje řetězec spadající do uvedeného intervalu včetně krajních hodnot

systém připraven. V tomto momentě je pak již možné pokládat dotaz tak, jak je uživatel zvyklý z jiných databázových systémů, tj. bez znalosti syntaxe dotazu vyhledávacího engine, neb v tomto momentě jsou již prohledávána data, která do systému vložil — snad poučený — uživatel a která prošla kontrolou administrátora.⁶ Jejich množství je však v současnosti nevelké a kromě některých polí,

6 Až do definitivního schválení přepisu administrátorem jej jakýkoli registrovaný uživatel může dále upravovat a měnit a jakákoli z těchto změn se okamžitě promítne do fulltextového vyhledávání.

ZNAKY PRO KOMBINOVANÝ DOTAZ

Znak	Význam	Příklad dotazu	Příklad odpovědi
+ / AND / &&	slučování	+Němcová +Babička	Najde všechny záznamy, kde se objevují oba řetězce současně
- / NOT	negace	+Němcová – Babička	Najde všechny záznamy, kde se objevuje první, ale nikoli druhý řetězec
OR /	alternace	+Němcová OR Babička	Najde všechny záznamy, kde se objevuje první nebo druhý řetězec
()	výběr části	+(Němcová OR Babička) +I855	Najde všechny záznamy, které obsahují třetí a buď první, nebo druhý řetězec

kteřá budou postupně naindexována automaticky, nelze v dohledné době bez větší finanční dotace v tomto směru předpokládat větší rozvoj systému.

Doporučené typy dotazů vhodné pro fulltextové vyhledávání

V následujícím přehledu se zaměříme již na konkrétní ukázky, jak v systému RETROBI najít vybrané typy záznamů. Nejprve upozorníme na způsob zápisu daného typu údaje na původní excerpční lístky, od něž se podoba vyhledávacího dotazu odvíjí, dále pak ukážeme konkrétní podoby toho, jak dotaz pokládat, a upozorníme i na základní rizika či nepřesnosti, na něž můžeme ve výsledné odpovědi narazit. Některé z těchto zkušeností, které se systémem RETROBI získali jeho administrátoři, platí obecně pro jakýkoli typ dotazu (vhodnost vyzkoušet různé varianty hledaného slova a různé rozmístění „divokých karet“), jiné platí singularně pouze právě pro daný typ dotazu.

1) Vyhledávání chronologických údajů

Chronologické údaje lze vyhledávat buď jako jednotlivé řetězce, nebo jako jejich ucelenou sekvenci. Znaký jednotlivých číslic se pro OCR program ukazují jako poměrně spolehlivě rozpoznatelné, snad s výjimkou občasných záměn párů 3 a 8, o a O, l a I

Je nutno počítat s následujícími nepředpokládanými odpověďmi programu: — u periodik, u kterých ročník neodpovídá kalendářnímu roku, se můžeme setkat mj. i s následujícími typy zápisů: 1936/1937; 1936–37 atp. — na excerpčním lístku může být i jiný letopočet než rok vydání zdrojového periodika, např. biografická data odkazované osoby, rok vydání recenzované knihy, datum historické události atp. — čtyřmístná číslovka může mít ojedinelé jiný význam než chronologický údaj (např. stránku, údaj o množství atp.)

Typy vzorových dotazů

- *1933* či *1*9*3*3* — najde všechny záznamy s hodnotou 1933
 (*1933* OR *1937* OR *1942*) — najde záznamy právě ze zvolených let
 [1901 TO 1912] — najde záznamy ze zvoleného časového rozmezí
- *186* — najde záznamy z let 1860–1869
 (ale i záznamy ze strany 186...)

Nabízí se i možnost položit dotaz na přesnější datum. Protože konkrétní data byla zapisována obvykle buď slovním vyjádřením měsíce u měsíčníků, nebo kombinací „den/měsíc“, tj. např. 16/2 pro 16. února, pak je třeba k dotazů na konkrétní rok/léta doplnit např.:

- Pro měsíce: *srpen* nebo */8* — najde záznamy článků vydaných v srpnu
- Pro konkrétní den: *16/2* — najde záznamy vydané 16. února

2) Vyhledávání dokumentů dle jejich charakteristiky

Každý jednotlivý podchycený článek měl být dle excerptních zásad na katalogizačním lístku jedním, popř. několika slovy charakterizován. Tato charakteristika byla pravidelně zapisována do hranatých závorek (hrozí přitom možné chybné přečtení prvního a posledního znaku jejího slovního vyjádření kvůli ručně doplňovaným patkám k lomítkům) a můžeme tak narazit na hodnoty typu básně, fejeton, zpráva, povídka atp.

Prostřednictvím fulltextového vyhledávání je tedy možné položit dotaz na množinu excerpt dle jejich charakteristiky, přitom je však vhodné zkusit různé alternativy dotazu (celé slovo, slovo bez prvního a posledního znaku atp.). Chceme-li najít např. všechny studie, ptáme se kombinacemi: *studie*, *s*t*u*d*i*e*, *tudi*, *[*s*t*u*d*i*e* atp.⁷

Protože se slovní vyjádření charakteristiky může objevit i v jiném pádě než nominativu singuláru, je třeba při formulaci dotazu zohlednit i možné tvary v jiných pádech, zejména pak v genitivu singuláru a nominativu plurálu. Např. na básně se tedy kvůli výskytu hodnot jako básně, dvě básně, tři básně, překlad básně, úryvek básně atp. ptáme kombinacemi:

(*básně* OR *básně*), (*b*á*s*e*ň* OR *b*á*s*n*ě*), *bás*, (*á*s*e* OR *á*s*n*) atp.

7 Při příliš obecné formulaci dotazu a v případech, kdy je v relativně krátkém textovém řetězci rozmístěno několik divokých karet, objem nalezených výsledků samozřejmě řádově stoupá.

Zohlednit je třeba též starší pravopisné normy. Fejton tak můžeme v kartotéce najít zapsaný též jako feuilleton, podobný případ je např. esej/essay atp. Velmi detailně byly rozlišovány jednotlivé žánry žurnalistické: kromě obecných označení typu zpráva a článek najdeme mj. charakteristiky jako rozhlásek, causerie, soudnička, entrefillet, kurziva a kurzivka atp.

Specifickou množinu lístků tvoří záznamy recenzí a referátů, které měly předepsaný zčásti odlišný způsob zápisu. Pro účely vyhledávání však stačí uvést, že takové lístky obsahovaly pravidelně zkratku *Rf*: či *Ref.*, která nemusela být ve všech případech uzavřena hranatou závorkou.

Nepředpokládané odpovědi na vyhledávací dotaz

Slovní vyjádření charakteristik bývají bohužel většinou dosti obecná označení, popř. oborově velmi frekventované pojmy, proto v odpovědích na dotazy nalezneme často případy, kdy je hledané slovo na lístku zapsáno nejen jako charakteristika, ale např. též v názvu excerpovaného článku, v jeho anotaci atp.

3) Vyhledávání dokumentu dle zdrojového časopisu

Další možností, kterou fulltextové vyhledávání nabízí, je vyhledávání dle názvu zdrojového periodika. Abecední soupis excerpovaných časopisů je k dispozici na adrese <http://retrobi.ucl.cas.cz/retrobi/projekt.4#toc-casopisy>, v databázové podobě s podrobnějšími údaji o stavu excerpce pak na adrese <http://isis.ucl.cas.cz/index.jsp?form=ec>.

Vyhledávání dle názvu zdrojového dokumentu dále usnadňuje fakt, že se na excerpčních lístcích jejich názvy pravidelně objevují po kombinaci lomítka (/), popř. okrajově jiná závorka, a znak =, což umožňuje odfiltrovat případné výskyty daného řetězce jinde na lístku. Problém však může způsobovat mezera, která se může, ale nemusí po znaku = vyskytovat. Pokud totiž po znaku = mezera následuje, není systémem možno položit dotaz obsahující divoké karty.

Pro vyhledání časopisu se tedy tážeme např.:

Akord, *A*k*o*r*d*, */*=*Akord*, *=*Akord*, popř. s využitím vyhledávání slovni fráze „/= Akord“

Vyhledání časopisu může zkomplikovat případné využití zkratek pro název zdrojového periodika. Tato situace se týká většinou obecně známých zkratkou názvů hlavních deníků. Je však třeba myslet na to, že na různých lístcích může být název daného periodika uveden jak zkratkou, tak ve formě plného názvu.

Seznam nejpoužívanějších zkratkou názvů periodik

České slovo — ČS

Právo lidu — PL

Lidové noviny — LN

Rudé právo — RP

Národní listy — NL

Rudý večerník — RV

Pakliže měl daný deník svůj večerník, může se stát, že se jeho název dostane do jiného pádu — např. *Večerník Práva lidu* atp. Mohou se však vyskytnout též podoby *Večerník PL* či dokonce *VPL*.

Pakliže systém u dotazu na zdrojové periodikum nenabídne žádnou odpověď, je možné pro kontrolu nalézt některý z článků dle autora a tak ověřit, zda excerpta z daného časopisu či konkrétní excerptce z daného roku nejsou pouze rukopisná, tj. pro fulltextové vyhledávání nepřístupná.

4) Vyhledávání dle jazyka překladu

Poměrně dobře jsou prostřednictvím fulltextového vyhledávání identifikovatelné i záznamy excerpt překladů z jednotlivých zahraničních literatur. Dle excerptních norem totiž tyto lístky musejí v názvové informaci obsahovat zkratku přel. doplněnou o zkratku, okrajově i nezkrácenou podobu příslušného zdrojového jazyka. Na překlady z jednotlivých národních literatur je pak možné zeptat se kombinací na zkratku přel. a zkratku požadovaného jazyka, tj. pro angličtinu např.:

+*přel.* / přel. / *přel* / *p*ř*e*l* / ... +*angl.* / *a*n*g*l* / ...

Zkratky nejčastějších jazyků

angl. — anglicky

bul. — bulharsky

čín. — čínsky

dán. — dánsky

fin. — finsky

fra. — francouzsky

ital. — italsky

jap. — japonsky

něm. — německy

nor. — norský

ruš. — rusky

pol. — polsky

port. — portugalsky

rum. — rumunsky

sloven. — slovensky

slovin. — slovinsky

srbch. — srbochorvatsky

šp. — španělsky

švéd. — švédsky

ukr. — ukrajinsky

5) Vyhledávání dle věcného tématu

Fulltextové vyhledávání je samozřejmě ideálním nástrojem pro přípravu rešerší na jednotlivá věcná témata. Zde je třeba znovu zmínit existenci předmětové části Retrospektivní bibliografie, která měla záznamy k jednotlivým věcným tématům cíleně podchycovat. Při tvorbě lístkové kartotéky se však takové řešení ukázalo jako velmi problematické, protože zde více než jinde hrál důležitou roli individuální přístup jednotlivých excerptorů. Některá předmětová hesla (např. *RKZ*) byla samozřejmě zaznamenávána cíleně, ta okrajovější pak však jen dle individuálního rozhodnutí jednotlivých excerptorů a výsledný stav ovlivnila i kvalita redakčních a rozpisových prací. K tomu se přidávaly problémy se sdružováním hesel do větších tematických

celků, s jejich vnitřním řazením atp. Vznikuvší předmětová část Retrospektivní bibliografie je tedy s přihlédnutím k těmto okolnostem nejméně spolehlivou či vypovídající částí původní lístkové kartotéky. Nicméně i tak vznikla předmětová kartotéka o cca 150 000 lístcích, která může sloužit jako zajímavý, byť vzhledem k vyřčenému ne vždy stoprocentně spolehlivý zdroj informací.

Pro prohledávání v předmětové části existují vedle fulltextového vyhledávání dvě možnosti. Nejprve je tu varianta využití databázi RET, do níž byla kdysi tato část kartotéky, konkrétně především lístky s excerpty k literárním časopisům a spolkům, převáděna v podobě strukturované databáze přístupné na: <http://isis.ucl.cas.cz/index.jsp?form=ret>. Tyto lístky nebyly pro účely báze RETROBI seskenovány, jejich případné multiplikáty jsou v bázi označeny razítkem s textem „Zadáno do databáze RET“. Druhou možností je pak využít témat vydělených ve zbytku předmětové části původní lístkové kartotéky a předmětovou část kartotéky procházet dle jejího původního řazení. Kompletní soupis témat z předmětové části je k dispozici na adrese <http://isis.ucl.cas.cz/ret-cat.jsp>.

Jednotlivá věcná témata je samozřejmě možno vyhledávat i fulltextově. V tomto případě obecně platí, že čím užší/specifičtější téma, tím relevantnější výsledek je možno získat. K jednotlivým skupinám věcných témat jen několik dílčích poznámek:

- při vyhledávání geografických pojmů mohou výsledek značně zkreslit místa vydání knih zapisovaná na odkazových referátových lístcích, což platí zejména pro rešerše typu Praha, Brno, Olomouc aj.
- u vyhledávání informací k různým kolektivním dílům či událostem je vhodnější začít rešerši od osob, které se jich účastnily, s velkou pravděpodobností tam bude nějaký odkazový lístek k dispozici.

Zpracování a export rešerší

Pro zpracování a export rešerší byla v systému RETROBI vytvořena tzv. schránka. Do ní si registrovaný uživatel může uložit libovolné množství lístků a ty si jejím prostřednictvím dle své potřeby seřadit a vyexportovat. Export dat je koncipován jako dosti variabilní a uživatel si může zvolit, zda chce exportovat lístky v podobě jednotlivých souborů, nebo jako jeden editovatelný textový dokument, může si vybrat kvalitu výstupních obrázků či specifikovat typ exportovaných textových dat (původní OCR, jen uživatelské přepisy, položkový rozpis, osobní komentáře). Asi nejzajímavější je pak možnost vzniklou rešerši exportovat na vlastní počítač ve formátu *.rtf v podobě textového souboru, který může daný uživatel dále upravovat i mimo aplikaci RETROBI. Kromě toho si může uživatel danou schránku uložit i v rámci systému a kdykoli později se k ní v případě potřeby vrátit.

Přepisy lístků

Významnou přednost systému RETROBI představuje možnost uživatelských přepisů jednotlivých lístků. Díky němu je možno korigovat textová data vzniklá OCR rekognoskací původních lístků a opravit v nich případné omyly a nepřesnosti. Takto korigované lístky jsou v systému okamžitě k dispozici pro případný export či pro fulltextové vyhledávání.

Možnost editace jednotlivého lístku je koncipována jako několikastupňová. V základní verzi může každý registrovaný uživatel lístek jako celek bez dalšího dělení přepsat jako jeden souvislý text. Vedle toho je však možné lístek tzv. segmentovat, tj. rozdělit jej na několik dílčích částí (názvové údaje, bibliografická citace, anotace) a tím lístek zpřístupnit pro strukturovanější vyhledávání a umožnit přehlednější grafické členění jeho přepisu. Segmentace funguje poměrně jednoduše, a to na základě vložení speciálního znaku dle definovaných pravidel na max. 4 místa v textu (blíže je popsána přímo v nápovědě aplikace). Kontrolní mechanismus pak automaticky zkontroluje, byly-li segmentační znaky zapsány na náležité pozice. Pokud přepis z hlediska segmentace vyhovovat nebude, na dílčí části se sice nerozdělí, ale uloží se alespoň jako prostý přepis OCR.

Přepisy lístků včetně jejich segmentací mohou registrovaní uživatelé i nadále měnit, a to až do úplného uzamčení lístku administrátory systému. Pak-liže se na některém lístku uzamčeném pro zpracování i přesto objeví chyba, je možno ji administrátorům systému nahlásit prostřednictvím systému tzv. uživatelských hlášení.

Vyšší úrovně editace lístku jsou přístupny již jen uživatelům s vyšším administrátorským oprávněním — ti mohou pořídit detailní přepis lístku do strukturované databázové podoby, popř. provádět tzv. hromadnou editaci, tj. přidat určitou hodnotu konkrétního atributu libovolně zvolené množině lístků.

Veškeré tyto úpravy textových dat využije běžný uživatel především při vyhledávání v systému, nebo pak i při exportu dat, kdy takto může ze systému získat přehlednatou a pravidelně členěnou rešerši. Všechny parametry prohledávaných dat lze nadefinovat přímo ve vyhledávací masce. Pomocí roletky „Typ textu“ lze omezit množinu prohledávaných dat jen na určitou segmentaci vydělenou část (tj. např. na názvové údaje nebo na bibliografickou citaci) či konkrétní pole. V těchto případech je k dispozici samozřejmě neporovnatelně menší množství lístků, avšak prakticky odpadá problém s (ne)spolehlivostí OCR rekognoskace. Segmentovaných lístků je momentálně k dispozici cca 10 000, výhledově se počítá se zahájením indexace některých vybraných polí (nejdříve zřejmě Rok vydání, Zdrojové periodikum a Charakteristika). Případný další vývoj lze průběžně sledovat v menu Statistiky, kde lze nalézt jednak průběžně aktualizované počty lístků dle stavu a jednak funkci rejstřík, kde si lze vytvořit rejstřík hodnot pro každé jednotlivé databázové pole včetně počtu lístků, které danou hodnotu v daném poli obsahují.

Bibliografie české literatury — výhledy do budoucna

Digitalizace Retrospektivní bibliografie znamená výrazný posun nejen pro bibliografickou práci ÚČL, ale též pro literárněvědnou bohemistiku vůbec. Badatelské veřejnosti se její digitalizací a veřejným online zpřístupněním nabízí pramen unikátního rozsahu a významu, který by do budoucna mohl a měl znamenat základní informační zdroj pro literárněvědný výzkum období před rokem 1945.

Pro další vývoj bibliografického oddělení ÚČL digitalizace Retrospektivní bibliografie znamená výrazný přelom a definitivní tečku za excerpcí dat do lístkové kartotéky. To by ovšem nemělo v žádném případě znamenat ústup od plánů na další bibliografické pokrytí období před rokem 1945: je však evidentní, že toto nelze, obzvláště s přihlédnutím k původnímu velkorysému založení celého projektu, ve smysluplném rozsahu provádět ve stávající personální kapacitě oddělení, ale bude třeba uvažovat o zisku účelového financování. Jakákoli budoucí excerptce se však bude muset provádět do strukturované bibliografické databáze respektující základní knihovnické normy a předpisy a zohledňující oborové potřeby literárněvědné bohemistiky. Pro nejbližší období bude třeba provoz aplikace RETROBI alespoň ustálit a získat co nejširší zpětnou vazbu od laické i odborné veřejnosti.

Digitalizace Retrospektivní bibliografie však byla jen prvním ze čtyř velkých úkolů, jimiž zpracování článkové bibliografie v ÚČL aktuálně prochází. Druhý z nich — přechod na modernější databázový systém Aleph — je už též v zásadě splněn: s platností od 1. srpna 2012 do něj kompletně přešla nová excerptce bibliografického oddělení ÚČL. Jasnou prioritu tak pro období do roku 2015 mají úkoly jiné: grantový projekt Bibliografie české literatury 1945–1960, v jehož rámci by mělo být do konce roku 2015 pokryto poslední dosud nezpracované chronologické období ve vývoji novodobé české literatury, a konverze komplexu stávajících databází Česká literární věda (1961–2012) do nového systému Aleph (dostupný na: http://aleph20.lib.cas.cz/F/?func=file&file_name=find-b&local_base=UCLA). Po jejich dokončení by tak ÚČL mohl nabídnout chronologicky ucelené bibliografické zdroje pokrývající kompletní období vývoje novodobé české literatury, tj. cca od roku 1770 do současnosti, ve dvou návazných systémech: v RETROBI pro období do roku 1945 a v Aleph pro období od roku 1945. Je však nepochybné, že i poté se bude třeba snažit o další trvalý rozvoj stávajících zdrojů, a to jak po kvalitativní (průběžná redakce), tak kvantitativní stránce (zpětná i nová excerptce) tak, aby stav bibliografických databází co nejvíce odpovídal aktuálním potřebám oboru a doby.

Retrospektivní bibliografie v Německu, Rakousku, Polsku a na Slovensku

Eliška Davidová (ed.)

Během dokončování digitalizace Retrospektivní bibliografie jsme si kladli otázku, jaké podmínky pro literárněvědné bádání mají kolegové věnující se literaturám jiných národů a jazyků a jak vypadají článkové bibliografie v jiných zemích. Obrátili jsme se přirozeně v první řadě do oblastí, které nám jsou kulturně i geograficky nejbližší, tedy k našim bezprostředním sousedům. Rešerše, které se na tuto otázku pokoušejí odpovědět, vytvořili pracovníci z ÚČL AV ČR: Marie Havránková pro Polsko, Michal Jareš pro Slovensko, Petr Píša pro Německo a Michael Wögerbauer pro Rakousko.

Německo

Soustavné bibliografie literárněvědné germanistiky vznikaly v Německu pod různými názvy už od osmdesátých let 19. století a dohromady souvisle pokrývají celé toto období až do r. 1950. Zaměřovaly se především na texty k novější německé literatuře (tedy po r. 1624), výzkum středověké literatury byl mapován v paralelních bibliografiích jazykovědných. Struktura, excerpční základna a kvalita těchto pomůcek se postupem času proměňovala. Nejpodrobněji jsou zpracovány ročenky pro období 1921–1939 (*Jahresbericht über die wissenschaftlichen Erscheinungen auf dem Gebiete der neueren deutschen Literatur. Neue Folge*), jejichž zásadním problémem je však absence údajů o soudobých autorech.

Na tuto řadu navázala v padesátých letech Bibliografie německé literární vědy (*Bibliographie der deutschen Sprach- und Literaturwissenschaft*, zkráceně BDSL; do r. 1968 společná s jazykovědnou bibliografií), známá též jako *Eppelsheimer — Köttelwelsch* podle někdejších dvou hlavních autorů. Ti nejprve popsali materiál za léta 1945–1953, další ročníky byly zpracovávány a vydávány průběžně. Od počátku devadesátých let vzniká paralelně elektronická databáze a tištěné ročenky, zpětně byly zdigitalizovány údaje od roku 1985. Volně přístupné online je však pouze období 1985–1995 (<http://www.bdsl-online.de>), novější roční-

ky jsou placené, přičemž licenci pro přístup má cca 200 pracovišť (především veřejné a univerzitní knihovny v německojazyčném prostoru). Tato bibliografie zaznamenává sekundární texty k německé literatuře bez jazykového či teritoriálního omezení. Je ovšem založena na autopsii — veškeré excerpané texty (monografie, sborníky a články) musí být fyzicky přítomny v Univerzitní knihovně ve Frankfurtu nad Mohanem — a je výběrová (roční přírůstek v současnosti obnáší cca 15 000 záznamů). Její nespornou výhodou je přívětivé uživatelské prostředí online verze, nabízející propojení s dalšími elektronickými databázemi, s virtuální knihovnou germanistiky, s *Wikipedií* apod.; nevýhodou je pochopitelně především placený přístup.

Kromě této v současnosti základní pomůcky je poměrně široce využívána také Mezinárodní bibliografie humanitních a společenských věd (*Internationale Bibliographie der geistes- und sozialwissenschaftlichen Zeitschriftenliteratur*, zkratka *IBZ*). Její protoverze začala vycházet v Německu už v roce 1897 a dnešní excerpční základna je skutečně úctyhodná — zahrnuje okolo 11 500 titulů ze 40 zemí. K německé literatuře se však vztahuje jen 1 000 až 1 500 záznamů ročně, což je oproti *BDSL* asi desetina. Od roku 1983 funguje také digitálně (<http://www.degruyter.com/view/db/ibz>) a rovněž její používání vyžaduje licenci, kterou vlastní cca 120 pracovišť.

Další mezinárodní bibliografii vznikající v německé oblasti je *Germanistik*, jejímž tematickým profilem je veškerá germanistika v širším slova smyslu (literární- i jazykovědná, včetně nederlandistiky a skandinavistiky) — v oblasti německé literární vědy je zato méně podrobná než *BDSL*. Jejím nepochybným kladem jsou však obsáhlé abstrakty. Ve formě čtvrtletníků vychází od roku 1960, v roce 2009 byla online zpřístupněna elektronická databáze údajů ve všech 50 ročnících (s průběžnou aktualizací), přístupná ovšem pouze v sedmi německých knihovnách.

Kromě toho mají germanisté k dispozici dva rozsáhlé knižní bibliografické zdroje poněkud odlišného typu: starší z nich je Goedekeho *Nástin dějin německé literatury na základě pramenů* (*Grundriß zur Geschichte der deutschen Dichtung aus den Quellen*), příručka na pomezí bibliografie, autorského lexikonu a literárních dějin, mapující německou literaturu od nejstarších dob do r. 1830, vydávaná opakovaně v aktualizovaných verzích v letech 1859–1991. Svazky týkající se středověku jsou považovány za překonané, ostatní jsou však jako bibliografický zdroj nadále běžně používány. V současnosti na něj formátem i strukturou navazuje Lexikon německých spisovatelů 1830–1880 (*Deutsches Schriftsteller-Lexikon 1830–1880*), z jehož plánovaných dvanácti svazků vyšlo od devadesátých let zatím osm; po kompletním knižním vydání má být zveřejněn také online.

Druhou podobnou publikací je Lexikon pramenů k dějinám německé literatury (*Quellenlexikon zur deutschen Literaturgeschichte*), bibliografie sekundárních textů k titulům vydaným v letech 1945–1990 (a vedle toho rovněž

k osobnostem z oblasti společenských věd), řazená podle autorů, případně názvů děl, zhruba na způsob *Lexikonu české literatury*. Je dostupná také na CD-ROMu, k jehož obsahu poskytuje přístup asi 40 německých vědeckých a univerzitních knihoven.

Alternativou k neexistující celoněmecké obecné článkové bibliografii je Virtuální německá zemská bibliografie (Virtuelle Deutsche Landesbibliographie), metavyhledávač v 15 zemských bibliografiích, zaměřených na vlastivědu, regionální dějiny a regionální literaturu a zahrnujících primární i sekundární texty. Aplikace je volně přístupná online (<http://www.ubka.uni-karlsruhe.de/landesbibliographie/index.html>) a obsahuje vesměs záznamy od osmdesátých nebo začátku devadesátých let (většina bibliografií navazuje na starší tištěná díla, která v jednotlivých spolkových zemích započala od dvacátých do osmdesátých let 20. století), v různé formě a s různou možností prohledávání v závislosti na konkrétních podmínkách a stavu jednotlivých databází.

Vedle germanistických bibliografií vzniká na německém území od konce padesátých let také Serbska bibliografija se záznamy monografií a článků v lužickosrbském jazyce a jinojazyčných textů vztahujících se k dějinám a kultuře Lužických Srbů, která je od r. 1996 průběžně zveřejňována online (<http://80.153.223.118/si-bib/>), zpětně byly na internetu zpřístupněny údaje od roku 1986.

Rakousko

Rakouská všeobecná národní ani specializovaná literárněvědná bibliografie neexistuje. Jsou zde ve velké míře využívány německé, obecně germanistické bibliografické zdroje (viz část věnovanou Německu), které však evidují většinou jen odbornou literaturu, nikoli deníky a běžné časopisy vydávané na území Rakouska.

V devadesátých letech 20. století vznikl projekt Rakouské retrospektivní bibliografie (Österreichische retrospektive Bibliographie, zkratka ORBI), jehož cílem bylo vytvořit všeobecnou bibliografii rakouské knižní a časopišské produkce od 16. století. Z plánovaných čtyř řad se však podařilo vytvořit pouze soupis denního tisku (vydán v letech 2001–2003), poté byl projekt v roce 2005 zastaven.

Pro starší literaturu jsou využívány především všeobecné germanistické pomůcky vzniklé v Německu, jak už bylo řečeno výše. Výjimku představuje kartotéka pražského rodáka Maxe von Portheim, jenž podrobně zmapoval šířeji vymezené josefínské období, tj. poslední třetinu 18. století. Kartotéka, uložená v Městské knihovně ve Vídni, odkazuje na primární i sekundární literaturu a lze v ní vyhledávat podle personálních, bibliografických a věcných hesel, přičemž poslední část je k dispozici také na internetu (<http://www.wienbibliothek.at/bestaende-und-sammlungen/portheim-portal/katalog.html>).

Badatelé v oblasti novější literatury mají k dispozici výstřižkový archiv v Literárním domě ve Vídni (Literaturhaus Wien — Dokumentationsstelle für neuere Österreichische Literatur), který byl založen r. 1965 a soustřeďuje se především na poválečné období. Tento archiv odnedávna funguje také digitálně (<http://www.literaturhaus.at/index.php?id=6542>), nahlédnout do něj je však možné pouze na místě. Podobné služby ovšem poskytuje pro veškerou německy psanou literaturu literární archiv v Marbachu am Neckar, kde se zpracovávají literární rubriky nejdůležitějších německojazyčných novin.

Pro nejnovější dobu je třeba brát v potaz také online přístup k velkému množství digitalizovaných periodik na portálu ANNO (rakouská obdoba Krameria v NK; <http://anno.onb.ac.at>), kde je dostupná mj. i řada německojazyčných periodik vycházejících v českých zemích.

Polsko

Základní polonistickou pomůckou pro poválečné období je Polská literární bibliografie (*Polska Bibliografia Literacka*), zpracovávaná na poznaňském pracovišti Instytutu Badań Literackich Polské akademie věd. Začala vznikat v r. 1948 a do r. 2000 vycházela tiskem v jednotlivých ročenkách, které pokrývají období let 1944/1945–1988. Bibliografie za rok 1989 byla vydána na CD-ROMu, záznamy z dalších let jsou volně dostupné online (<http://pbl.ibl.poznan.pl/dostep>). V současnosti jsou zde přístupná data za období 1988–2000.

Tato báze je koncipována velice široce. Kromě literárněvědných textů mapuje také oblast teatrologie a recepce cizích literatur v Polsku, zachycuje i beletrii (včetně překladů do cizích jazyků) vydanou časopisecky i knižně, ale také například namluvenou na gramofonových deskách, scénáře hraných filmů, zápisy rozhlasových a televizních literárních pořadů apod.

Elektronická databáze je rozčleněna na čtyři části: teoretickou, literárně-historickou, část věnovanou cizím literaturám a část se záznamy o divadelní, filmové, rozhlasové a televizní tvorbě. Po vzoru tištěných ročenek zde lze prohlížet materiál sebraný v jednotlivých oddílech nebo podle běžných kritérií prohledávat napříč databázi.

Dalším významným polským zdrojem pro literárněvědné bádání je Barova bibliografie polských časopisů 19. a 20. století (*Bibliografia Bara — Kartoteka Bibliografii Literackiej Zawartości Czasopism Polskich XIX i XX wieku*). Zpracovávána začala být retrospektivně za 2. světové války pod vedením Adama Bara a obsahuje téměř milion kartotéčních lístků se záznamy z periodik vydaných do r. 1939. Stejně jako výše zmíněná novější bibliografie zahrnuje také beletrii, recepci cizích literatur apod.

Kartotéka je uspořádána abecedně podle literárních osobností. Jednotlivá hesla jsou od r. 2003 elektronicky zpracovávána a publikována online (<http://www.bar.ibl.waw.pl/contents/main.php?page=index>), patrně v pořadí dle významnosti autorů či na přání uživatelů. Zpřístupněné části lze prohlížet jako

jednu bázi, jež umožňuje mj. i fulltextové prohledávání; prakticky zde však chybí věcný popis a bibliografické údaje jsou často neúplné.

Článeková bibliografie polských časopisů (*Bibliografia Zawartości Czasopism*) je součástí denně aktualizované elektronické národní bibliografie v polské Národní knihovně, zachycuje tedy výběrově články ze všech oborů, publikované v časopisech vydávaných v Polsku (bez ohledu na jazyk publikace), jejichž povinné výtisky jsou v knihovně ukládány. Excerpce probíhá od r. 1947, forma výstupů se v průběhu let měnila: data za některé ročníky existují pouze ve formě kartotéky (tyto záznamy jsou dnes digitalizovány), po většinu času vycházela bibliografie tiskem s měsíční periodicitou, od roku 1996 byla souběžně publikována i elektronicky, v roce 2005 bylo vydávání v tištěné podobě zastaveno a v současné době jsou tištěné měsíčníky přístupné online ve formátu PDF (vše na <http://www.bn.org.pl/katalogi-i-bibliografie/bibliografia-narodowa/bibliografia-zawartosci-czasopism>).

Slovensko

Základním zdrojem pro retrospektivní bibliografii slovenské literatury je obsáhlé dílo Ludovíta V. Riznera *Bibliografia písomnictva slovenského na spôsob slovníka od najstarších čias do konca roku 1900*, které zahrnuje především texty vznikající na Slovensku v daném období (dále také např. maďarské texty slovenských autorů vydané mimo slovenské území). První část prvního dílu této bibliografie vyšla v češtině (1917), jako celek ve slovenštině byla vydána v šesti svazcích v letech 1929–1934 a po válce doplněna ještě o osm svazků dodatků (1946–1972). Původních šest svazků bylo nedávno poskytnuto zdarma ke stažení ve formátu PDF na stránkách Digitálnej knižnice (<http://dikda.eu/l-vl-rizner-bibliografia-pisomnictva-slovenskeho-na-sposob-slovnika-od-najstarsich-cias-do-konca-r-1900>).

S prací na národní článkové bibliografii bylo na Slovensku započato ve čtyřicátých letech. Již v roce 1941 byla v rámci projektu vznikající Slovenskej národnej knižnice založena Slovenská bibliografická spoločnosť, která roku 1944 vydala první sešit *Slovenskej bibliografie*, obsahující článkovou bibliografii za leden téhož roku. Vzhledem k politické situaci se však ve vydávání pokračovalo až po tříleté přestávce a v pozměněné formě: od roku 1947 začíná bibliografie vycházet ve čtvrtletních sešitech s podtitulem *Sústaavný rozpis vybraných príspevkov z periodickej literatúry*; počínaje rokem 1954 odpovídají tyto soupisy způsobem členění analogickým českým bibliografickým katalogům. Jako doplňkový svazek byla současně vydána i *Bibliografia tlače na území Slovenského národného povstania* (29. 8. – 28. 10. 1944), chybějící roky 1945 a 1946 byly zpracovány zpětně v roce 1970.

Tyto článkové bibliografie vycházely pod měnicími se názvy tiskem do roku 1998, v letech následujících byly zpřístupňovány na internetu v databázích Slovenskej národnej knižnice a později sloučeny s bázi Slovenskej digi-

tálnej knižnice (<http://dikda.eu>), která by se měla soustředit jak na seriály, tak na elektronické zdroje. Na Slovensku tedy existuje souvislá bibliografická řada pokrývající celé poválečné období. Kromě těchto zdrojů zpracovávají státní vědecké knihovny a regionální knihovny své bibliografické soupisy, které spadají do existujícího konsorcia KIS3G. S konsorciem chce Slovenská digitálna knižnica spolupracovat a využívat jejich data.

Vedle aktuálních bibliografií začala r. 1956 vznikat také retrospektivní bibliografie článků ze slovenských novin a časopisů z let 1918–1945, navazující na výše zmíněnou Riznerovu bibliografii do r. 1900 (chybějící období 1900–1918 měla na starosti Štátna vedecká knižnica v Košicích), která se však zpočátku zaměřila především články z komunistického a odbojového tisku. Z tohoto projektu vznikla nejprve kartotéka článků z let 1918–1928, řazená podle autorových a předmětových hesel, na jejímž základě začaly koncem sedmdesátých let vycházet bibliografické soupisy, bylo však vydáno pouze několik svazků, týkajících se jiných oborů než literatury (tematicky nejbližší byl svazek věnovaný divadlu, filmu a rozhlasu). Pro literární vědce se tak zdá být nejpřírodnější *Bibliografia almanachov, ročeniek a zborníkov na Slovensku 1919–1944* (1991; díl popisující období 1945–1965 vyšel r. 1995).

Kromě toho je na internetu volně dostupná zmiňovaná Slovenská digitálna knižnica, do níž jsou nyní postupně zpětně začleňována data z poválečné *Slovenské bibliografie* (<https://www.kis3g.sk/cgi-bin/gw/chameleon?skin=clanky>; v současné době jsou zde k dohledání články zhruba od poloviny sedmdesátých let), není však jasné, zda by měla v budoucnu pojmout i materiál z let 1918–1938.

Závěr

Z výše uvedeného vyplývá, že stav článkových bibliografií využitelných jako zdroj pro literárněvědné bádání se v jednotlivých zemích výrazně liší. V Německu existuje celá řada různě široce pojatých databází, jež jsou však historicky roztržštěné a přístup do nich je často podmíněn licencí, kterou vlastní pouze některé knihovny; specializované pracoviště pro bibliografii literárněvědné germanistiky pokrývající celou její historii zde neexistuje. Rakousko je prakticky závislé na německých zdrojích, vlastní obecnou článkovou ani specializovanou bibliografii nemá. Velmi dobrá situace je naopak v Polsku — velkoryse založenou oborovou bibliografii pro období po r. 1945 doplňuje rozsáhlá Barova databáze, oproti české retrospektivě však stručnější. Mimoto je zde k dispozici i obecná bibliografie polských novin a časopisů. Národní článková bibliografie funguje také na Slovensku, specializovaná literárněvědná zde však chybí.

V České republice byla všeobecná bibliografie periodického tisku v Národní knihovně zpracovávána do r. 2011, její další osud je momentálně nejasný. Oborovou článkovou bibliografii pro českou literárněvědnou bohemistiku dlouhodobě zpracovává Ústav pro českou literaturu. V průběhu let se poda-

řilo vypracovat rozsáhlý, byť poměrně heterogenní korpus materiálů. Právě dokončená digitalizace Retrospektivní bibliografie české literatury 1770–1945, díky níž je tento zdroj alespoň v naskenované verzi přístupný online, je prvním krokem k zásadnímu přebudování ústavních databázových zdrojů. V současnosti jsou pak činěny kroky k sjednocení a zpřehlednění jednotlivých databází pro období po roce 1945: jednak byly zahájeny práce na komplexní excerpci období 1945–1960, zpracované zatím jen prostřednictvím tištěných bibliografických příruček, a vedle toho je od roku 2012 zaznamenávána aktuální článková produkce v modernějším databázovém systému Aleph, do něž by měly být v průběhu následujících let konvertovány všechny báze pro období let 1960–2012.

Na základě rešerší Marie Havránkové, Michala Jareše, Petra Piši a Michaela Wögerbauera zpracovala Eliška Davidová

Odkazy na online bibliografické zdroje

Česká republika

Retrospektivní bibliografie české literatury 1770–1945: <http://retrobi.ucl.cas.cz/retrobi/>
Bibliografie české literární vědy: <http://isis.ucl.cas.cz/index.jsp?form=biblio;> (1960–2012)
http://aleph20.lib.cas.cz/F/?func=file&file_name=find-b&local_base=UCLA
 (1945–1960 a od srpna 2012)

Německo

Bibliographie der deutschen Sprach- und Literaturwissenschaft (BDSL): <http://www.bdsl-online.de>
Internationale Bibliographie der geistes- und sozialwissenschaftlichen Zeitschriftenliteratur (IBZ): <http://www.degruyter.com/view/db/ibz>
Virtuelle Deutsche Landesbibliographie: <http://www.ubka.uni-karlsruhe.de/landesbibliographie/index.html>
 Rozcestník k dalším bibliografiím: <http://www.germanistik-im-netz.de/dbis/>

Rakousko

Porthem-Portal: <http://www.wienbibliothek.at/bestaende-und-sammlungen/porthem-portal/katalog.html>
Literaturhaus Wien — Dokumentationsstelle für neuere Österreichische Literatur: <http://www.literaturhaus.at/index.php?id=6542>

Polsko

Polska Bibliografia Literacka: <http://pbl.ibl.poznan.pl/dostep>

Bibliografia Bara — Kartoteka Bibliografii Literackiej Zawartości Czasopism Polskich XIX i XX wieku: <http://www.bar.ibl.waw.pl/contents/main.php?page=index>

Bibliografia Zawartości Czasopism: <http://www.bn.org.pl/katalogi-i-bibliografie/bibliografia-narodowa/bibliografia-zawartosci-czasopism>

Slovensko

Bibliografia písomnictva slovenského na spôsob slovníka od najstarších čias do konca roku 1900: <http://dikda.eu/l-vl-rizner-bibliografia-pisomnictva-slovenskeho-na-sposob-slovnika-od-najstarsich-cias-do-konca-r-1900>

Slovenská bibliografia: <https://www.kis3g.sk/cgi-bin/gw/chameleon?skin=clanky>

Résumé

This series of articles spotlights the Retrospective Bibliography of Czech Literature 1770–1945, the most valuable collection of material being kept and worked on at the ASCR Institute of Czech Literature. Over almost sixty years, 525 journal and newspaper titles published in Czech or German on the territory of the Czech lands were excerpted for it, this large-scale excerption range covering not only reflections and studies of Czech literature, as in the case of subsequent bibliographical bases for Czech literary studies for the period from the latter half of the 20th century to the present, but also fiction reprints including translations, while in addition to Czech literature, texts from the literatures of other nations were also excerpted, as well as texts from associated subjects, particularly journalism, history, theatre, drama studies and the like. Thanks to its size (almost 1.7 million excerpt cards) and chronological range (1770–1945; with follow-on bases up to the present), as well as its elaborate processing methodology, it is hard to find an equivalent among Czech specialist bibliographies or literary studies bibliographies in neighbouring countries. The three articles published in *Česká literatura* (3/2013), are meant to present the Retrospective Bibliography and its recently completed digitization project in greater detail to the specialist public. The initial article by Daniel Řehák presents the Retrospective Bibliography as a card catalogue, both from historical and methodological standpoints, while the subsequent text by Vojtěch Malínek focuses on presenting its digitized form in the RETROBI system, stressing its capacity for use by the professional public, and the final article by Eliška Davidová briefly presents the literary studies article bibliography situation in neighbouring countries.

Klíčová slova / Keywords

bibliografie — česká literatura — literární věda — Polsko — Rakousko — Německo — Slovensko

bibliography — Czech literature — literary studies — Poland — Austria — Germany — Slovakia