

BÁZE DAT *CHEMISCHES ZENTRALBLATT* SE STÁVÁ SOUČÁSTÍ PROGRAMU *SCIFINDER*

JAROSLAV ŠILHÁNEK[†]

Ústav organické technologie, Vysoká škola chemicko-technologická v Praze, Technická 5, 166 28 Praha 6

Klíčová slova: chemické báze dat, SciFinder, Chemisches Zentralblatt, ChemZentTM, referátové časopisy, historie

O co jde

V tomto časopise¹ bylo v roce 2009 referováno o mimořádné aktivitě společnosti *Fachinformationszentrum Chemie GmbH* v Berlíně, která se podjala úkolu zachránit pro chemii jediný zdroj bibliografických informací převážně z 19. století, referátový časopis *Chemisches Zentralblatt*². Vydávání tohoto, po dlouhá léta jediného a i pak jednoho z nejdůležitějších zdrojů chemických informací, skončilo v r. 1969 a celkem logicky během následujících let jednotlivé svazky postupně mizely z regálů chemických knihoven, ať už z důvodů prostorových, nebo jednoduše z nepochopení stálé platnosti a užitečnosti tam uvedených informací, i když později vyvrácených nebo opravených, ale jednou nalezených a formou vědeckých článků zaznamenaných výsledků chemického bádání. Mimochodem, jedním z argumentů pro získání prostředků pro takovou akci byla možnost, že se v tomto souboru mohou skrývat velmi zajímavé zmínky, nebo i podrobnější údaje o biologických aktivitách hlavně malých molekul. Způsob, jakým byla tato akce realizována, spočíval v jednoduché (i když technicky pochopitelně značně náročné) digitalizaci celého souboru vydaných svazků od r. 1830 do roku 1969 od první do poslední stránky, včetně svazků rejstříků. A v této podobě byl výsledek nabídnut chemikům k využití, jako báze dat *Chemisches Zentralblatt*.

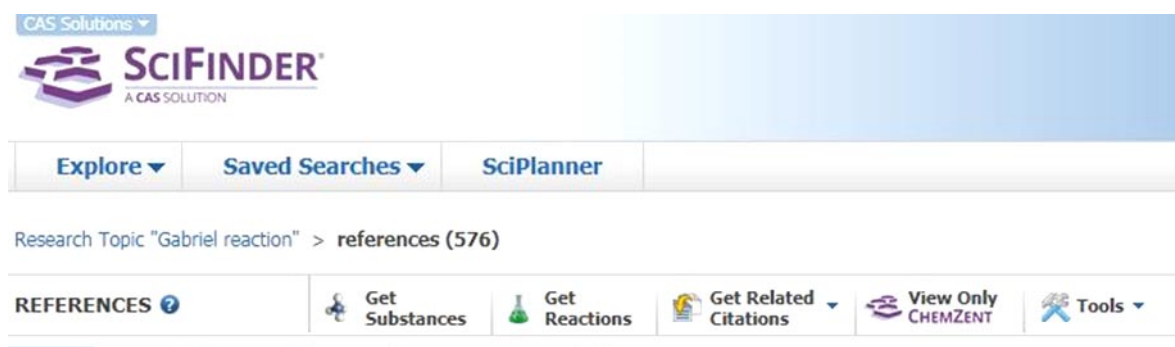
Kdo ale využil této skutečnosti (mimo jiné i díky tomu, že na základě vstřícnosti Spolku německých chemiků (*Gesellschaft Deutscher Chemiker*) byl členům České chemické společnosti umožněn přístup zdarma), zjistil, že možnosti jsou nepochybně lákavé, ale práce s prostou digitalizací všech stránek jednotlivých svazků není taková, na jakou už jsme zvyklí při operacích s většinou současných chemickýchází dat. V podstatě by bylo správné pracovat s digitalizovaným souborem stejně jako s tištěnou verzí, tedy začít s příslušnými rejstříky, najít zajímavé odkazy a nakonec otevřít příslušné stránky, tentokrát v podobě pdf formátu. Digitalizovanou bázi doplnila nástroji umožňujícími rozpoznávání chemických sloučenin na základě jejich systematických názvů německá společnost *InfoChem* a tato verze pod názvem *Chemisches Zentralblatt Structu-*

ral Database už byla uživatelsky podstatně přijatelnější. Ovšem stále to nebyla standardní chemická báze dat, tedy báze s rozložením dat do jednotlivých polí s možnostmi přesné konstrukce dotazů, jejich kombinace a následného upřesňování a ladění. Tohoto úkolu se nakonec na základě dohody ujala služba *Chemical Abstracts Service* a výsledek zařadila jako volitelnou součást programu *SciFinder*, v podobě samostatného modulu s možností oddělit záznamy z bázi *Chemical Abstracts* od převzatých záznamů z báze *Chemisches Zentralblatt*. Modul převzatých záznamů z digitalizované verze *Chemisches Zentralblatt* je označen jako *ChemZentTM* a vyžaduje samostatnou licenci v rámci licence pro přístup k programu *SciFinder*.

Jak to funguje?

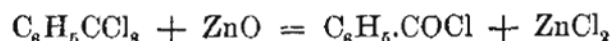
Chemisches Zentralblatt je samozřejmě bibliografická báze dat, se kterou se v tištěné verzi pracovalo zcela obdobně jako s tištěnými sešity *Chemical Abstracts*, tedy prostřednictvím rejstříků. Rejstříky k referátovému časopisu *Chemisches Zentralblatt* byly prakticky stejné jako k tištěné verzi *Chemical Abstracts*, tedy autorský, předmětový včetně chemických názvů a vzorcový, typicky rejstřík molekulových vzorců. Digitalizovaná verze samozřejmě nemůže obsahovat žádné jiné rejstříky, takže odkazy za výše uvedené časové období můžeme očekávat jen v analogických případech. Tedy budeme-li hledat jména autorů, předmětová hesla z textů abstraktů a názvy chemických sloučenin nebo jejich molekulové vzorce. Nemůžeme ale očekávat výsledky dotazů využívajících možností elektronickýchází dat, tedy např. odkazy na články, nebo jejich identifikátory, názvy komerčních společností a některé další. Principiálně lze těžit z obecně obsáhlejších textů abstraktů, problém je v jejich překladech do angličtiny. Je logické, že odpovědi z tohoto modulu můžeme očekávat jednak za období nepokryté bázi *SciFinder*, tedy za období 1830 až 1907 a dále za období 1907–1969, kdy se určitý objem záznamů bude překrývat. Máme-li platnou licenci pro práci s modulem *ChemZentTM*, jsou jakékoliv rešeršní dotazy realizovány automaticky i s tímto modulem. Pokud výsledek takového dotazu bude obsahovat i bibliografické záznamy z modulu *ChemZentTM*, jsme na tuto skutečnost upozorněni novým tlačítkem na horní liště *View Only ChemZentTM* (obr. 1), které, máme-li zájem, umožní oddělit a zobrazit jen bibliografické odkazy z báze *ChemZentTM*. Kromě toho můžeme omezit výsledky rešeršní standardní nabídkou na upřesnění souboru nalezených hitů (*Refine*) stejně, jako např. modul *Medline*.

Ilustrujme si převedení bibliografického záznamu z časopisu *Chemisches Zentralblatt* (obr. 2) do báze *SciFinder*.



Obr. 1. Vstupní obrazovka programu *SciFinder* s možností zobrazit výsledky rešerše v bázi *Chemisches Zentralblatt*. Published with permission of the American Chemical Society, Copyright © 2013. American Chemical Society (ACS). All rights reserved.

Eduard Hoffmann u. Victor Meyer, *Zur Kenntnis der Benzoylverbindungen.*
 Im käuflichen Benzoylchlorid fanden Vf. Benzaldehyd, der bei der Einw. von NH_3 auf dies Benzoylchlorid die Bildung einer Verb. $\text{C}_{21}\text{H}_{18}\text{N}_2\text{O}_2$, Nadeln aus h. A., F. 225° , veranlaßt, die Benzylidendibenzamid, $(\text{C}_6\text{H}_5\cdot\text{CO}\cdot\text{NH})_2\text{CH}\cdot\text{C}_6\text{H}_5$, ist. Der für diese Verb. bisher angenommene F. 197° ist unrichtig. Da Benzoylchlorid technisch nach:



herstellbar ist, würde ein Gehalt des Benzotrighlorids an Benzalchlorid das Auftreten des Benzaldehydes erklären. (B. 25. 209–13. 8/2. [21/1.] Heidelberg. Univ.-Lab.)

Obr. 2. Ilustrace abstraktu primárního článku v časopise *Chemisches Zentralblatt*. Je zřejmé, že praxe referátového časopisu *Chemisches Zentralblatt* respektovala snadnost získání originálu abstrahovaných článků, časopis *Berichte der deutschen chemischen Gesellschaft* byl natolik známý, že stačila jednopísmenná citace a krátký abstrakt. Původní článek má totiž 5 stran a obsahuje řadu dalších informací včetně experimentálních detailů a rovnic, kromě v abstraktu uváděné redukce benzotrighloridu. Published with permission of the American Chemical Society, Copyright © 2013. American Chemical Society (ACS). All rights reserved.

Tento záznam je nyní v bázi *ChemZent*TM k dispozici v následující podobě² (obr. 3).

Překlad německého abstraktu není prováděn „ručně“, ale za pomoci počítačových překladačů a nelze být proto na přesnost překladu příliš přísný. Důležitější je další klíčový aspekt, a to bibliografický odkaz na původní časopis. Ten byl samozřejmě rovněž poplatný své době, v uváděném příkladu se jedná o časopis *Berichte der deutschen chemischen Gesellschaft*, který byl tak všeobecně znám, že stačila jednopísmenná zkratka B. Je opět nutné ocenit programátorskou práci, která dokázala ve většině případů najít bibliografickou citaci v *Chemisches Zentralblatt* a učinit z ní součást záznamu v *ChemZent*TM. V daném případě v podobě na obr. 4.

Všimněme si ale, že v modulu *ChemZent*TM je bibliografický záznam identifikován jako záznam v tištěné verzi referátového časopisu a nikoliv odkazem na *Chemische Berichte*. Je to opět důsledek skutečnosti, že v bibliografických záznamech v časopise *Chemisches Zentralblatt* neby-

ly odkazy na primární zdroje ukládány separátně v té době na kartotéční lístky, zatímco při excerpcích pro báze *Chemical Abstracts* už tomu tak bylo. Proto nemůžeme očekávat nalezení relevantních odkazů v modulu *ChemZent*TM při požadavku vyhledání konkrétní bibliografické citace (pole *Journal* v bázi *SciFinder*), ani nemůžeme nalezený soubor bibliografických odkazů v modulu *ChemZent*TM analyzovat podle periodik. Všechny citace v *ChemZent*TM mají odkaz na příslušný ročník, svazek a stránky časopisu *Chemisches Zentralblatt*³.

Duplicitní záznamy

Na základě časového pokrytí zpracovávaného období je zřejmé, že bibliografické citace za léta 1907–1969 se alespoň pro významnější chemické časopisy budou překrývat, tj. budou zaznamenány jak v bázi *SciFinder* bez modulu *ChemZent*TM, tak v tomto modulu. Je to problém? Na

792. For the recognition of Benzoylverbindungen (Zur Kenntnis der Benzoylverbindungen)

By: Hoffmann, Eduard.; Meyer, Victor

In the käuflichen benzoyl chloride been authors Benzaldehyd at the effect of NH, on this benzoyl chloride forming one compound C₂H₁₈N, Os, needles from hot A., melting point 225 °, veranlafst, the benzylidene dibenz amide, (C₆H₅.CO.NH), CH.C, H₆. In this connection, hitherto assumed melting point 197 ° is incorrect. Since benzoyl chloride technical according to: CeH₆CC]+ Zn olanzapine*= C, H₅.COCl+ ZnCl₂, can be produced if the content of the Benzotrchlorids to benzal chloride the occurrence of benzaldehyde explain.

Indexing**Concepts**

melting point

Substances

98-07-7 Benzotrchloride 🔍
 98-87-3 Benzal chloride 🔍
 98-88-4 Benzoyl chloride 🔍
 100-52-7 Benzaldehyde 🔍

Obr. 3. **Záznam citace na obr. 2 v bázi ChemZent™**. Published with permission of the American Chemical Society, Copyright © 2013. American Chemical Society (ACS). All rights reserved.

SOURCE

Chemisches Zentralblatt
 Volume63 Book 1
 Issue10
 Pages435
 Journal
 1892
 CODEN:CHZEA6

Original Source: B. 25. 209—
 13. 8/2. [21/1.] Heidelberg.
 Univ.-Lab.

Obr. 4. **Ilustrace identifikace primárního časopisu z abstraktu v ChemZent™**. Published with permission of the American Chemical Society, Copyright © 2013. American Chemical Society (ACS). All rights reserved.

konkrétních příkladech je možné demonstrovat, že se jedná naopak o další zesílení informačního významu báze *SciFinder* pro chemické disciplíny. Excerptní praxe, i když v principu nepochybně velmi podobná, se přece jenom lišila v obou případech. Časopis *Chemisches Zentralblatt* měl za sebou už více než 70 let nepřetržitě činnosti a excerptní praxe byla svým způsobem „usazená“. *Chemical Abstracts* naproti tomu efektivní způsob a zásady vytváření bibliografických citací hlavně v období před 2. světovou válkou teprve hledala. Máme-li při možnosti vyhledávání informací porovnat bibliografické záznamy jak z časopisu *Chemisches Zentralblatt*, tak z *Chemical Abstracts*, narazíme často na velmi překvapivé rozdíly.

Tak např. v roce 1968 autoři L. R. C. Barclay a E. C. Sanford publikovali v časopise *Canadian Journal of Chemistry* článek³ pojednávající o mechanismu Friedelovy-Craftsovy reakce, který byl zpracován do časopisu *Chemical Abstracts* jako bibliografický záznam s abstraktem o délce 148 slov a excerptováno z něj bylo 6 sloučenin. Ale tentýž článek byl zpracován i v časopise *Chemisches Zentralblatt* a dnes je tedy součástí báze *SciFinder* a modulu *ChemZent™* a zde jeho bibliografický záznam obsahuje abstrakt v délce 448 slov a hlavně zaznamenává 24 sloučenin, ze kterých jen 3 jsou totožné se záznamem v *Chemical Abstracts*.

Kromě rozsahu je patrné, že abstrakt z časopisu *Chemisches Zentralblatt* obsahuje podstatně více experimentálních detailů a do jisté míry tak nahrazuje neexistenci fyzikálně-chemických a reakčních dat v bázi *SciFinder*. Užitečnost zařazení digitalizované verze *Chemisches Zentralblatt* jako modulu *ChemZent™* v bázi *SciFinder* je tak mimo jakékoliv pochybnosti.

Strukturní rešerše

Jak jsme se již zmínili výše, už digitalizovaný text byl doplněn možností vyhledávání graficky zobrazených struktur, takže i při práci s modulem *ChemZent™* můžeme využívat grafický editor programu *SciFinder*, a to včetně substrukturních rešerší se všemi zabudovanými nástroji upřesňování zadání. Výsledek je prakticky stejný, samozřejmě příslušné bibliografické odkazy, které dostaneme naprosto stejným způsobem jako v programu *SciFinder*, jdou až do r. 1830. Záleží pak jen na nás, co s nimi dále uděláme. Možnost omezit soubor odkazů jen na publikace



Obr. 5. Ilustrace rozdílnosti abstraktů v programu SciFinder a modulu ChemZentTM. Published with permission of the American Chemical Society, Copyright © 2013. American Chemical Society (ACS). All rights reserved.

z ChemZentTM ovšem můžeme až po zobrazení odkazů pro nalezené sloučeniny, kde se zobrazí tlačítko View ChemZentTM only. Můžeme se ale setkat i s překvapujícími výsledky, kdy např. strukturální zadání s dvojnou vazbou nedefinované sterické konfigurace neposkytne žádný výsledek. Ten dostaneme až po určení ve struktuře, zda se jedná o *cis*-, nebo *trans*-konfiguraci.

Chemické reakce

V tomto ohledu modul ChemZentTM představuje rovněž zvýšení hodnoty programu SciFinder. Zahrnuje reakce přípravy a přeměn pro všechny sloučeniny, které byly ukládány na základě stejné zásady zaznamenávat existenci a všechny známé vlastnosti publikované v primárních zdrojích do kompendií Beilstein a Gmelin, tedy jdoucí daleko do minulosti. Výrazný přínos je především v tom, že obsahuje i reakce přípravy a přeměn anorganických sloučenin. Tedy něco, co by bylo jako nezávislá akce doplnění do programu SciFinder v současné době prakticky nemyslitelné. Jak je známo, reakce anorganických sloučenin, ale i řady jiných, nejsou v současné době excerpovány průběžně ani do báze CHEMREACT a zda dojde v tomto směru ke změně politiky na straně CAS, není zatím známo. Je nutné zdůraznit logický důsledek propojení, že totiž všude, kde je to možné, jsou informace o reakcích nalezených v modulu ChemZentTM doplněny informacemi z programu SciFinder a současně jsou možnosti vyhledávání rozšířeny i na funkci substrukturálního vyhledávání. Stručně řečeno, záznamy existence chemických reakcí pro velký počet chemických sloučenin od počátku existence vědecké chemie na konci 18. století už samy o sobě odpodat-

ňují doplnění licenčních přístupů i pro modul ChemZent.

Závěr

Je nutné vysoce ocenit, že po kompletní digitalizaci referátového časopisu *Chemisches Zentralblatt*, *Chemical Abstracts Service* zdvihla tuto hozenou rukavici a převedla celý materiál do podoby standardní chemické strukturální báze dat. Jednalo se o velmi náročnou akci sestávající nejenom z technických problémů, ale i otázek převodu tištěných rejstříků, překladů a spolehlivé identifikace abstraktů. Relativně úspěšně byla vyřešena identifikace primárních zdrojů. Svým způsobem se tak uzavírá kruh kompletní převedení chemických informací do podoby chemických bází dat, a to jak nejstarších, jen tištěných dat, tak i jejich doplnění a stále aktualizování v časovém horizontu téměř hodin. Ještě výrazněji se tak posiluje zcela mimořádné postavení chemie jako vědní disciplíny, která má absolutně nejúplněji zpracovaný systém svých poznatků, na kterých je pak možné spolehlivě vytvářet poznatky nové.

LITERATURA

- Šilhánek J.: Chem. Listy 103, 847 (2009).
- Hoffmann E., Meyer V.: Ber. Dtsch. Chem. Ges. 25, 209 (1892).
- Barclay L. R. C., Sanford E. C.: Can. J. Chem. 46, 3325 (1968).

Keywords: chemical data bases, SciFinder, *Chemisches Zentralblatt*, ChemZentTM, review journals, history