PROJEKT MEZINÁRODNÍ BANKY SPEKTROSKOPICKÝCH DAT EUROSPEC*

LIDMILA VOLKOVÁ a KAREL VOLKA

Ústav analytické chemie, Vysoká škola chemicko-technologická v Praze, Technická 1905, 166 28 Praha 6
e-mail: Karel.Volka@vscht.cz

Došlo 2.4.03, přepracováno 20.4.03, přijato 6.5.03.

Klíčová slova:

Ke dni 26. března 2003 bylo v Chemical Abstracts registrováno 21 211 680 organických a anorganických látek, přičemž ke dni 21. září 2001 to bylo „jen“ 18 683 580 organických a anorganických látek (aktuální počet registrovaných chemických látek lze snadno zjistit na webové stránce Chemical Abstracts Service, http://www.cas.org/cgi-bin/regreport.pl). Během půldruhého roku tedy počet registrovaných chemických látek vzrostl o více než 2,5 milionu.

Porovnáme-li tyto počty s existujícími databázemi spekter, zjistíme několikařádový rozdíl mezi počtem známých sloučenin a počtem uložených spekter. Je přitom známo, že právě molekulová spektra (především spektra NMR, IČ, Ramanova a hmotnostní) jsou jedním ze základních zdrojů informací o nových látkách. Spektra jsou ve většině publikací zredukována na jednovětá konstatování, po případě na jednoduchou tabulku. Původní spektra v lepším případě uchovává autor, dříve či později jsou i tato spektra často zapomenuta či ztracena. Tabulka I shrnuje odhad počtu spekter, která byla zmíněna jen v prvních šesti číslech 67. ročníku časopisu Collection of Czechoslovak Chemical Communications.

V uplynulých letech jsme byli svědky řady pokusů vybudovat mezinárodní banku spektroskopických dat. Idea budování takové mezinárodní databáze není nová a byla úspěšně ověřena na bance krystalografických dat proteinů, do které jsou data ukládána téměř z každé recenzované proteinové krystalografické publikace. Tato banka dat je zpřístupněna krystalografické obci a je klíčovým zdrojem pro aktivity na tomto poli (http://www.rcsb.org/pdb/).

Tabulka I
Přibližný počet citovaných spekter v článcích časopisu Collection of Czechoslovak Chemical Communication *67* (2002) (spektra: IČ – infračervená, UV – ultrafialová, ESR – elektronové spinové resonance, NMR – nukleární magnetické resonance, MS – hmotnostní)

Číslo	IČ	UV	ESR	RAMAN	^{13}C-NMR	^{1}H-NMR	Jiná	MS
1	70	0	0	0	138	134	25	6
2	16	16	0	0	23	37	0	0
3	45	14	0	0	32	75	35	64
4	3	2	0	0	3	3	0	0
5	101	12	0	0	158	166	14	2
6	1	0	10	2	41	56	83	6
Celkem	*236*	*44*	*10*	*2*	*395*	*471*	*157*	*78*

V roce 1996 byl na konferenci „Spojení a interpretace spekter prostřednictvím molekulových struktur“ (Linking and Interpreting Spectra through Molecular Structures), konané ve Warwicku (Velká Britanie), podán návrh na založení obdobného zdroje spektroskopických dat. Hnací silou tohoto návrhu byl stále se zvětšující rozdíl mezi počtem známých sloučenin a počtem spekter uložených ve spektrálních databázích. V roce 1996 bylo evidováno 14 milionů chemických látek.

Po řadě diskusí na různých konferencích v letech 1996 až 1998 jak v Evropě, tak i v USA byly v roce 1999 a 2000 předloženy dva návrhy vyjádření zájmu o 5. rámcový program založení spektroskopické databanky. Duší celého projektu se stal Dr. Antony N. Davies, University of Glamorgan (Wales) a Creon Lab Control A.G., Frechen (SRN) (dříve Institut für Spektrochemie und Angewandte Spektroskopie, Dortmund (SRN)). Oba návrhy byly pozitivně posouzeny, prvnímu však bylo vytknuto, že se nezabýval strukturou projektu po skončení jeho financování EU. V druhém kole byl projekt s názvem „Access to Research Spectroscopic Data and Associated Chemical Knowledge“ (Přístup k výzkumným spektroskopickým datům a doprovodným chemickým znalostem) (EUROSPEC) vysoce oceněn a schválen k financování jako tématická síť v rámci programu Konkurenceschopný a trvale udržitelný růst. Řešiteli projektu jsou Institut für Spektrochemie und Angewandte Spektroskopie, Dortmund (SRN), Creon Lab Control A.G., Frechen (SRN), Institute National Agronomique Paris-Grinon, Paris (Francie), Universidade de Aveiro, Aveiro (Portugalsko), Vysoká škola chemicko-technologická v Praze, Praha, LGC Ltd., Runcorn (Velká Británie), IM Publications, Chichester (Velká Británie) a Specs and Biospecs B.V., Rijswijk ZH (Nizozemsko). Prosazení tohoto projektu je velkým úspěchem nejen pro zúčastněné konsorcium, ale pro spektroskopii jako celek, neboť potvrzuje její nepostradatelnou úlohu ve sféře priorit EU.

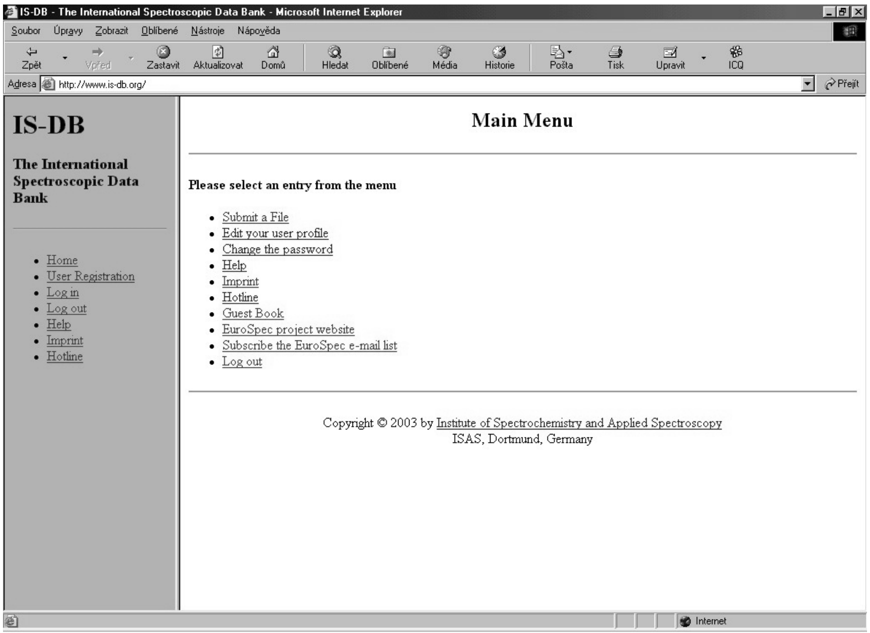
Jedním z klíčových momentů přípravy celého projektu byla všeobecná podpora a příslib účasti řady vědeckých i komerčních vydavatelů časopisů i vědeckých společností. I přes silný atak ze strany komerčních organizací, zabývajících se tvorbou a prodejem spektrálních databází, projekt získal silnou podporu celé řady velkých i malých vydavatelství, a to České chemické společnosti, EDP Sciences, Elsevier Science, IM Publications, International Society for Ion Mobility Spectrometry, Spektroskopické společnosti Jana Marka Marci (Ioannes Marcus Marci Spectroscopic Society), John Wiley & Sons, Royal Society of Chemistry, Society for Applied Spectroscopy a Springer-Verlag Heidelberg. Projekt počítá s připojením dalších vydavatelů časopisů, včetně časopisů východní Evropy (tab. II).

Vlastní tříletý projekt byl zahájen 1. ledna 2002 a v prvním roce měl dva klíčové úkoly: připravit infrastrukturu projektu a přesná pravidla ukládání a manipulace s daty. Vydavatelé časopisů by měli vyzvat svoje autory k zasílání dat do této mezinárodní databáze a archivu (International Spectroscopic Databank and Archive, IS-DB) již v roce 2003.

* Volně zpracováno podle článku A. N. Daviese: An update on the International Spectroscopic Data Bank Project. Spectroscopy Europe *13* (5), 24–26 (2001).

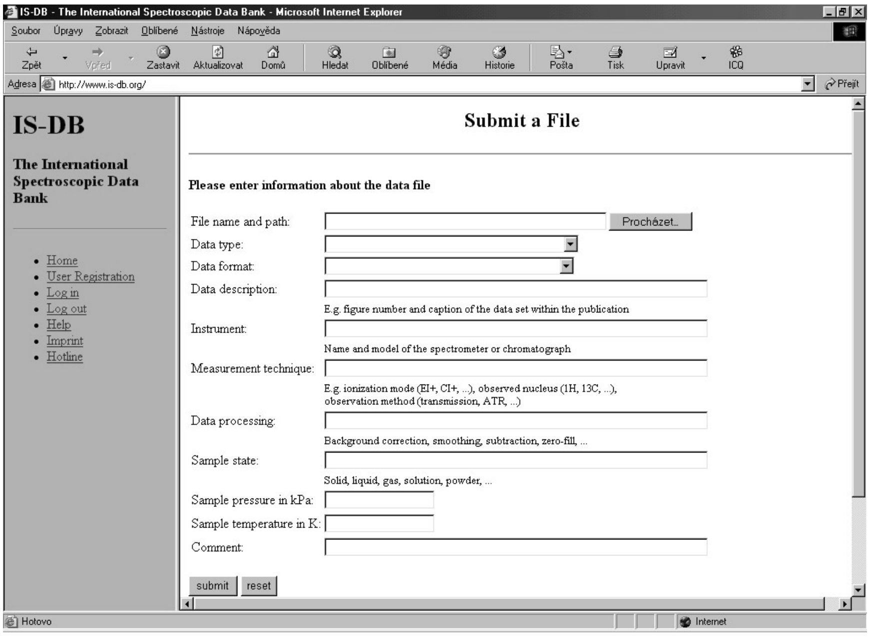
Tabulka II
Seznam participujících časopisů

Název časopisu	Název časopisu	Název časopisu
Actualité Chimique	Green Chemistry	Lab on a Chip
Analusis Magazine	International Journal for Ion Mobility Spectrometry	Magnetic Resonance in Chemistry
Analyst	International Journal of Mass Spectrometry	New Journal of Chemistry
Analytical and Bioanalytical Chemistry	Journal of Analytical Atomic Spectrometry	NIR News
Analytica Chimica Acta	Journal of Environmental Monitoring	PhysChemComm
Applied Spectroscopy	Journal of Mass Spectrometry	Physical Chemistry Chemical Physics
Bulletin České chemické společnosti	Journal of Materials Chemistry	Rapid Communications in Mass Spectrometry
Bulletin of the Ioannes Marcus Spectroscopic Society	Journal of Molecular Structure	Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy
Chemical Communications	Journal of the Chemical Society, Dalton Transactions	Spectrochimica Acta, Part B: Atomic Spectroscopy
Chemické listy	Journal of the Chemical Society, Perkin Transactions 1	Spectroscopy Europe
Collection of Czechoslovak Chemical Communications	Journal of the Chemical Society, Perkin Transactions 2	Surface and Interface Analysis
CrystEngComm	Journal of Near Infrared Spectroscopy	Talanta
European Journal of Mass Spectrometry	Journal of Pharmaceutical and Biomedical Analysis	Vibrational Spectroscopy
Faraday Discussions	Journal of Raman Spectroscopy	X-Ray Spectrometry
Geochemical Transactions		



Obr. 1. **Vstupní webová stránka Mezinárodní spektroskopické databáze**

Snahou řešitelů projektu je, aby vstup dat do databáze byl co nejjednodušší. Bude realizován pomocí webové stránky, která by časem měla být dostupná i v různých jazykových mutacích, včetně češtiny (obr. 1). Preferováno bude uložení originálních dat s dostatečnými informacemi pro identifikaci měřicího systému ve standardním formátu pro výměnu



Obr. 2. **Jedna ze stránek pro zaslání spektra do Mezinárodní spektroskopické databáze**

dat, jako je JCAM-DX pro spektra, MOLFILES pro struktury atp.

Uložená data budou v první fázi přístupná pouze vydavatelům a po zajištěné lince recenzentům určitého článku. Přístup k úplným spektrálním datům by měl recenzentům usnadnit jejich práci a odhalit chyby v interpretaci spekter. Po přijetí článku k uveřejnění budou spektroskopická data a doprovodné chemické informace (název, chemická struktura, reaktivita, zdraví a bezpečnost atp.) zpřístupněny vhodnými odkazy v elektronické verzi článku nebo adresou webového prohlížeče v tištěné verzi. Spektroskopická data (obr. 2) spolu s doprovodnými chemickými informacemi budou v databance uložena jako jednotlivé soubory, projekt však nepočítá s použitím prostředků na vyhledávání spekter. Tento úkol bude přenechán organizacím, které jsou na jeho plnění lépe připraveny a bude řešen spolu se zajištěním provozu databáze po skončení projektu, kdy její správu převezme organizace neziskového charakteru.

Neméně důležitým krokem je příprava nových instrukcí autorům pro zasílání spektroskopických a doprovodných dat elektronickou poštou spolu s vlastní publikací.

Pro řešení vzniklých problémů projekt počítá s dvěma poradními orgány.

Jeden poradní orgán tvoří vydavatelé a jeho úkolem je garantovat, že projekt bude odpovídat současným trendům a technickému vývoji zvláště v oblasti elektronických publikací. To je podstatné, má-li se projekt vyhnout rozhodnutím, jejichž výsledkem by byla investice do oblastí, které budou v krátké době nedůležité.

Druhým poradním orgánem je konzultační komise koncových uživatelů, jejíž primární povinností bude radit v otázkách přijatelnosti technologických řešení. V informační technologii se u takto velkých projektů vždy objevuje nebezpečí, že jejich účastníci chtějí využívat nejnovější či nejelegantnější řešení, koncoví uživatelé však často nejsou schopni nejnovější trendy ve vývoji hardware či software sledovat.

Bezprostřední pomoc při ukládání dat budou poskytovat školení pracovníci na telefonních linkách, z nichž jedna bude na Vysoké škole chemicko-technologické v Praze.

Autoři dat musí rychle vidět prospěch ze své práce v tom, že uložená data jsou pro ně také využitelná. To je klíčový úkol řešitelů projektu. Projekt podpořili jak akademičtí, tak průmysloví spektroskopici. Pro úspěch projektu a vytvoření kvalitní databáze je třeba, aby projekt podpořila široká obec spektroskopiků. To bude představovat pro každého jednotlivce či pracovní tým určitou práci navíc, ale toto úsilí by mělo být mnohonásobně vyváženo prospěchem, které vytvoření databáze přinese v následujících letech. Věříme, že i s Vaší podporou se tento záměr podaří.

Dodatek

Během přípravy rukopisu k tisku byl projekt mezinárodní banky spektroskopických dat oživen. Systém pro vstup dat a server archivu IS-DB byly zpřístupněny autorům vědeckých publikací dne 19. května 2003 na Achemě. EuroSpec již žije! Chcete-li zaslat spektra do této databáze, zaregistrujte se na webové stránce http://www.is-db.org.

L. Volková and K. Volka *(Department of Analytical Chemistry, Institute of Chemical Technology, Prague):* **International Spectroscopic Data Bank Project EUROSPEC**

A short introduction to the International Spectroscopic Data Bank (IS-DB) is given. This database was established in May 2003 . It is up to spectroscopists around the world to feed the system with good reference spectroscopic data. Register and deposit data at http://www.is-db.org.