

naše baze kemijske dokumentacije

j.benkovič
a.kornhauser
m.vrtačnik

UDK 681.3 : 002

Univerza v Ljubljani FNT - VTO kemija in kemijska tehnologija - RCPU

Kemijsko raziskovanje ter prenos raziskovalnih dosežkov v proizvodnjo in izobraževanje sta vedno terjala urejanje kemijske dokumentacije, njeno evaluacijo in diseminacijo. Z eksponencialno rastjo števila kemijskih informacij je postala uporaba računalnika na tem področju neogibna.

V Sloveniji imamo v okviru Računalniškega centra za programirano učenje (RCPU) pri Fakulteti za naravoslovje in tehnologijo dve bazi kemijske dokumentacije. Prva je iz uvoza, Chemical Abstracts Condensates (Chemical Abstracts Service- Columbus, Ohio, USA), druga pa lastna baza podatkov o alkaloidih.

Chemical Abstracts Service objavlja tekoče izvlečke iz publikacij kot Chemical Abstracts, ki so v kondenzirani obliki prenešeni na računalnik. To bazo v okviru projekta Raziskovalne skupnosti Slovenije uvaja zgoraj navedena institucija tako, da je opravila prenos na naš računalniški sistem (Republiški računski center), nudi stalno pomoč uporabnikom pri oblikovanju profilov, izvaja operativna dela, vodi stalno analizo in evaluacijo dosežkov, skrbi za nenehno optimizacijo profilov in uvajanje novih oblik te baze. V prispevku je opisan celotni proces te dejavnosti.

Med uporabniki baze CAC so fakultete in visoke šole, tehnične knjižnice, raziskovalne organizacije ter številne OZD iz industrije. Podan je podroben pregled števila uporabnikov po institucijah in strokah.

RCPU pa razvija tudi lastno bazo kemijskih informacij na področju alkaloidov, ki je pomembno tako za kemijsko teorijo kot kemijsko-farmaceutsko industrijo. Pri tem uvaja v delo uporabnike v industriji in raziskovalnih organizacijah, pa tudi študente kemije in biologije. Prispevek podaja pregled problematike pri gradnji lastne baze, postopke za njeno gradnjo, uporabo in evaluacijo. Daje primere tega dela in podatke o obsegu doslej zgrajene baze dokumentacije. Vabi k sodelovanju.

CHEMICAL DATA BASE IN SLOVENIA. Chemical research and the transfer of its achievements into industry and education have always required the processing of chemical documentation, its evaluation and dissemination. Due to the exponential growth of chemical information the use of computer has become inevitable.

In Slovenia, the Computer Centre for Programmed Learning (RCPU) at the Faculty of Natural Sciences and Technology processes two data bases of chemical information. The first one is imported Chemical Abstracts Condensates (Chemical Abstracts Service - Columbus, Ohio, USA), and the other the Alkaloid Data Base, developed in the Centre.

Chemical Abstracts Service publishes current abstracts of publications in the form of Chemical Abstracts which are transferred in the condensed form onto computer. Within a project sponsored by the Research Community of Slovenia this base has been adapted to the available computer system (Republic Computer Centre). RCPU offers assistance to the users in the formation of profiles, and carries out the operational work, continuous analysis and evaluation of results. The paper presents the processes of this work.

Faculties and colleges, technical libraries, research institutions and numerous enterprises are the users of CAC data base. A detailed survey of users is presented.

RCPU also develops its own data base of chemical information in the field of alkaloids, important for chemical theory and for chemical-pharmaceutical industry. The Centre introduces into development of this data base the users from industry and institutes, as well as chemistry students. The paper provides a survey of problems involved in the formation of such a base, procedures used in its formation, application and evaluation. Examples are given for the use of this data base and the survey of results accomplished so far is presented. The wish for cooperation is expressed.

UVOD

Narava dela v kemijskem raziskovanju, pri prenosu raziskovalnih dosežkov v prakso in v kemijsko izobraževanje je vedno terjala urejanje dokumentacije, njeno evaluacijo in njeno diseminacijo.

Število kemijskih informacij narašča eksponentialno. Če vzamemo kot primer Chemical Abstracts, to lepo ilustrira obdobje, potrebno za milijon abstraktov: za prvi milijon je bilo potrebno 32 let, za drugega le še 18, za tretjega 8, za četrtega 4 leta in 8 mesecev, za petega komaj tri leta.

Tudi število kemijskih spojin silovito narašča. Chemical Abstracts Registry kaže tedenski porast okrog šest tisoč spojin. Tako tudi na ožjem področju kemije ni več mogoče spremljati informacij v celoti.

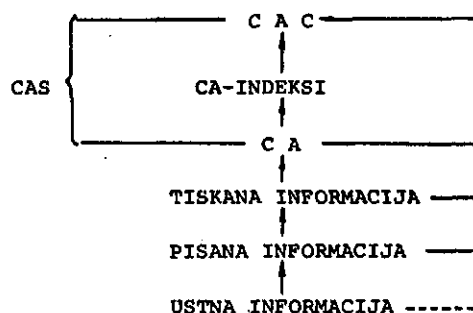
Za vsako eksponentialno rast je značilno, da vodi v eksplozijo, če se ne umiri oz. stagnira. Danes že govorimo o eksploziji kemijskih informacij. Umiritev seveda ni mogoča, možen je le prehod na novo kvaliteto. To novo kvaliteto omogoča uporaba računalnika za vsa tri področja: za proizvodnjo, evaluacijo in diseminacijo informacij.

V okviru Računalniškega centra za programirano učenje (v nadaljnjem besedilu RCPU) pri VTO kemija in kemijska tehnologija Fakultete za naravoslovje in tehnologijo Univerze v Ljubljani (center je že bistveno presegel področje svojega imena, vendar ga je zaradi ljubezni do tradicije ohranil....) delujeta dve bazi, zanimivi za kemika, tehnologa, farmacevta, biologa, medicinca in vrsto drugih, ki jim je kemija osnova pri delu. Prva je baza CAC, druga pa alkaloidna baza, ki jo razvijamo sami.

BAZA PODATKOV CHEMICAL ABSTRACTS CONDENSATES (CAC)

Kemijski informacijski servis (Chemical Abstracts Service-CAS) objavlja tekoče izvlečke iz publikacij v seriji Chemical Abstracts ter indekse, ki omogočajo retrospektivo. Med pomembne novejšje dosežke pa sodi računalniška baza podatkov Chemical Abstracts Condensates (CAC).

Shema kemijske dokumentacije



Baza Chemical Abstracts je razdeljena na 80 sekcij. Prvih 34 sekcij izhaja v lihih tednih in prinaša podatke o naslednjih področjih: biokemija (20 sekcij) in organska kemija (14 sekcij). V sodih tednih pa izhajajo sekcije, ki pokrivajo naslednja področja: makromolekularna kemija (12 sekcij), uporabna kemija in kemijsko inženirstvo (18 sekcij), fizikalna in analizna kemija (16 sekcij). Baza je pisana v standardnem formatu, izbira dokumentov pa je zasnovana na definiranju kemijskih področij. V bazo so vključene lastnosti elementov in spojin, metode njihove analize in sinteze, fiziološki učinki, uporaba.

PRIPRAVA RAČUNALNIŠKE OBDELAVE

Računalniške obdelave smo ves čas razvijali v dveh smereh: za selektivno diseminacijo informacij (SDI) ter za retrospektivo.

Za SDI smo uporabili modificirani programski paket PRETEXT (Programme de Recherche dans le Texte), ki so ga razvili na Institut Français du Pétrol. Ta program ima širok spekter nastavitvenih parametrov za oblikovanje izpisa: sortiranje in utežitev odgovorov, izbira formata in števila kopij, omejitev izpisanih dokumentov. Številni so tudi ključi za iskanje dokumentov: CODEN, vrsta dokumenta, jezik, sekcija, patentna klasifikacija, avtor, institucija, tekst.

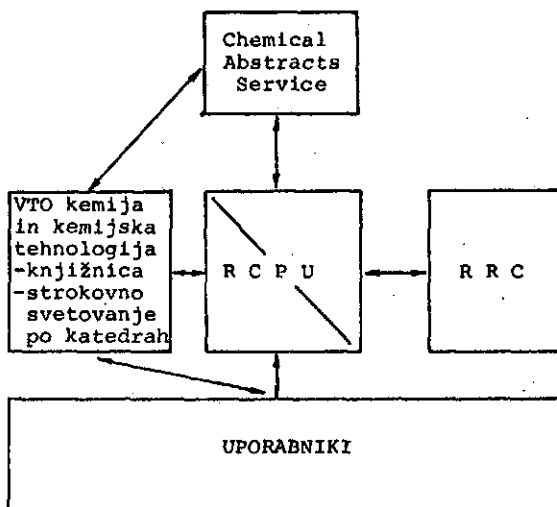
V okviru retrospektive smo zaradi omejenega prostora na diskih tvorili doslej bazo podatkov le iz nekaj zvezkov Chemical Abstracts. Ob tem smo izpisali slovar besed s frekvenco in slovar okrajšanih besed, ki sta nepogrešljiva pri formuliranju profilov. Princip retrospektivnega iskanja s pomočjo kodiranih ključnih besed pa se je izkazal kot neprimeren, ker je frekvenca ključnih besed premajhna.

Pri transformaciji baze pa je prišlo do zastoja zaradi doslej neugotovljenih napak v sistemu PRETEXT - verjetno zaradi novega operacijskega sistema na računalniku v Republiškem računskem centru. Pri tem se je tudi tu pokazala velika pomanjkljivost uvoženega softwara, kadar ni poskrbljeno za permanentno vzdrževanje. Po nekaj neuspešnih poskusih, da bi s pomočjo tujega avtorja odkrili in odpravili napako, smo se odločili za izdelavo lastnega programskega paketa, ki smo ga razvili iz obstoječega sistema DORS (Documentation Retrieval System) - ta je instaliran v Republiškem računskem centru. Novi razviti paket CACDORS je prevzel prej navedene možnosti PRETEXTA, ima pa prednost, da za izvajanje programov porabi manj računalniškega časa, ker vanj nismo zajeli ključev, ki zelo redko pridejo v poštev. Vanj smo vključili le avtorja, vrsto dokumenta, sekcijo in tekst - vse z možnimi levimi in desnimi okrajšavami. Izpis smo oblikovali za format A4, ki omogoča arhiviranje rezultatov v običajnih mapah.

Novi paket za SDI je v evaluaciji. Dosedanji rezultati so ugodni. Prav zato že načrtujemo tvorbo retrospektivne baze na večjih diskovnih pomnilnikih, organizirane na osnovi posameznih besed.

V drugi polovici letošnjega leta se bosta dosednji bazi CAC in CASIA (Chemical Abstracts Subject Index Alert) združili v novo bazo CA SEARCH. Do tedaj bo treba prirediti programe za nove, povečane možnosti iskanja podatkov.

ORGANIZACIJA SPECIALIZIRANEGA INDOK CENTRA ZA KEMIJO



NEPOSREDNO DELO Z UPORABNIKI PRI OBLIKOVANJU PROFILA

Oblikovanje profila, to je vprašanja uporabnika, ki ga moramo prirediti za procesiranje po bazi podatkov, poteka v več stopnjah. To delo je treba skrbno opraviti, kajti če vprašanje za poizvedbo ni precizno oblikovano, nastopa v izhodnih dokumentih veliko "šuma". Lahko pa je tudi odzivnost baze premajhna - tedaj govorimo o "tišini". "Šum" so vsi neustrezni dokumenti, ki jih je računalnik zajel zaradi premalo precizno postavljenega vprašanja. "Tišina" pa so dokumenti, ki obstajajo v bazi podatkov, a jih računalnik ni zajel v izpis, ker je bil profil slabo nastavljen.

Za iskanje informacij smo se odločili v obliki naravnega teksta - to je, iskanje poteka po celotnem tekstu sekundarnega dokumenta (CAC). Tekst so v tem primeru popolni bibliografski podatki, polje deskriptorjev in rezime, če ga baza vsebuje. Oblikovanje zahtevkov na osnovi naravnega teksta terja od dokumentalista solidno znanje, saj mora obvladati angleško terminologijo na izbranem strokovnem področju.

Stopnje oblikovanja profila:

1. Vprašanje uporabnika

Pri oblikovanju vprašanja ima uporabnik dve možnosti:

- a) jedrnato opredeli vprašanje po vsebini,
- b) sam napiše nekaj ključnih besed in jih poveže v ustrezno logično zvezo.

Prvi pristop je za računalniškega dokumentalista boljši, saj mu daje več prostosti pri oblikovanju zasnove profila ob uporabi ustreznih indeksov, ki jih izdaja CAS. Pokazalo se je namreč, da zgolj iz ključnih besed, ki so vrh tega pogosto še okrajšane, dokumentalist zelo težko razbere ključno vsebino za iskanje dokumentov. Omejen je v svojem razvijanju profila. Izkušnje kažejo tudi, da uporabniki po večini uporabljajo le Boolove operatorje, kar pogosto osiromaši vsebino izpisanih dokumentov.

2. Identifikacija konceptov

Izraz "koncept" je v uporabi za skupine spojin, postopke, vrste materialov, načine uporabe in podobne sklope pojmov s sorodnim pomenom, ki nastopajo v vprašanju uporabnika. Dokumentalist mora te koncepte prepoznati v vprašanju uporabnika, da jih lahko razvija.

3. Razvijanje konceptov

Ta faza oblikovanja profila je ključnega pomena, saj je treba za dane izraze (koncepte) poiskati čim večje število možnih sinonimov, da bi se izognili šumom in tišini. Pri tem delu si pomagamo s CA indeksi. Posebej ustreza INDEX GUIDE. V pomoč pa so nam tudi slovarji besed, ki smo jih pripravili v centru z namenom, da bi se čim boljše seznanili z besediščem baze CAC in si olajšali krajšanje besed na levi in desni strani.

4. Oblikovanje profila kot logičnega izraza

Posamezne termine za določene koncepte izrazov, ki smo jih okrajšali na levi in desni strani, povežemo z uporabo Boolovih in sintaktičnih operatorjev ter oklepajev v logični izraz, ki je s tem prirejen za procesiranje po bazi podatkov.

5. Evaluacija izhodnih dokumentov

Po procesiranju sledi ovrednotenje izhodnih dokumentov, pri katerem morata sodelovati uporabnik in dokumentalist. Zadnji lahko oceni dokumente le glede na prisotnost zaželenih terminov in njihovo povezavo, ne pa tudi dejanske vsebine dokumenta. To je delo strokovnjaka-uporabnika, ki mora pri svoji oceni upoštevati predvsem ustreznost izpisa glede na vsebino primarnega dokumenta. Le stalno sodelovanje uporabnika in dokumentalista omogoča sprotno popraviljanje profila, vnašanje novih terminov, opuščanje neustreznih in s tem optimiziranje profila.

V našem delu zagotavljamo tako povezavo v osebnih stikih, pismeno in po telefonu. Kvaliteta prvih seveda daleč prekaša ostala dva načina.

POMEN IN UPORABA OPERATORJEV

Pri oblikovanju profila uporabljamo Boolove in sintaktične operatorje. S prvimi - AND, NOT, OR - določimo prisotnost ali odsotnost izbranega termina v informaciji CAC. S sintaktičnimi operatorji pa lahko povežemo v logične zveze besede v istem izrazu (stavku, polju). V polju teksta je beseda definirana kot vrsta alfanumeričnih znakov med dvema separatorjema. Razdalja med izrazoma v istem stavku je določena s številom separatorjev med izrazoma. Z uporabo sintaktičnih operatorjev lahko določimo maksimalno število separatorjev med izrazoma in vrstni red izrazov v polju naslova in/ali teksta.

Navajamo primer:

Operator:

AVEC	AAnn- asociacija asimetrična nn separatorjev
	ASnn- asociacija simetrična nn separatorjev
IGNOR	IAAnn- ignoranca asimetrična nn separatorjev
	ISnn- ignoranca simetrična nn separatorjev

Kot nn lahko nastopajo cela števila od 00 (izraza sta v tem primeru dela iste besede) do 31 (to pomeni, da izraza nastopata v istem polju teksta ali naslova).

"CARBON" AA01 "DIOXID"

Besedi CARBON in DIOXID sta ločeni z enim samim separatorjem, vrstni red pa mora biti tak kot je postavljen v zapisu. Članki, pri katerih bi bilo zaporedje teh dveh besed zamenjano, ne bi nastopali v izpisih CAC.

V logičnem izrazu lahko uporabimo več sintaktičnih in Boolovih operatorjev. V takem primeru moramo posamezne logične podenote ločiti med seboj z oklepaji. Pri kombinaciji operatorjev pa je treba upoštevati mejne pogoje dopustnih kombinacij. Pri tem velja:

ISnn (B NOT C)

Sintaktičnemu operatorju IGNOR ne sme slediti Boolov operator NOT.

(A AND B) ASnn (C AND D)

Levi operand sintaktičnega operatorja ne sme vsebovati Boolovega operatorja AND.

PREGLED UPORABNIKOV SDI V SLOVENIJI

Številni avtorji, ki poročajo o uvajanju računalniško obdelane kemijske dokumentacije, trdijo, da se te bistveno bolj poslužujejo uporabniki iz industrije kot iz "akademske" sfere. Temu naj bi bila vzrok večja izpostavljenost industrije na tržišču.

Naša statistika je sicer še skromna, saj gre za prvo fazo uvajanja CAC, vendar kaže približno ravnotežje med obema sferama.

Vrsta OZD uporabnikov	Število uporabnikov
fakultete	
visoke šole	26
inštituti	12
gospodarske OZD	38
skupaj	76
število profilov	188

Zanimivo sliko kaže tudi ožja stroka uporabnikov, kar je razvidno iz naslednje tabele.

Ožja stroka uporabnikov	Število
Visokošolski kemiki	
- učitelji	11
- študenti	15
Raziskovalci IJS, KIBK	12, skupaj 38

Industrija:

- elektrotehnična	19
- smole, barve, laki	7
- farmacevtska, kozmetična	3
- bazična	3
- prehrabena	2
- lesna	2
- gumarska	1
- strojna	1, skupaj 38

Vendar so to šele začetni rezultati, prve lastovke. Bazo CAC, s katero razpolagamo, je treba čim bolj približati uporabnikom. V ta namen skušamo zlasti:

- intenzivirati stike z uporabniki,
- v sodelovanju z uporabniki stalno izpopolnjevati profile,
- organizirati uvajalne tečaje za nove uporabnike,
- vključiti program uvajanja CAC v redni študij kemije.

RAZVIJANJE LASTNE BAZE KEMIJSKE DOKUMENTACIJE

Za uspešni razvoj lastne baze kemijske dokumentacije je treba najprej definirati - in jasno omejiti! - področje kemije, na katerem želimo imeti takšno bazo. To je lahko zelo zahtevna naloga, na primer na področju kalkulacij v kvantni mehaniki, v organski sintezi, kemijskem inženirstvu itd. Večkrat pa je področje že samo po sebi jasno definirano in omejeno, npr. IR spektra heterocikličnih spojin. Praviloma je manj težav s preciziranjem področja v industriji, kjer večina že dobro pozna svoj cilj. To je lahko na primer proizvodnja fenol-formaldehidnih smol za določene potrebe elektroindustrije in podobno.

V vsakem primeru velja: čim bolj jasno je opredeljeno področje, tem bolj uspešno bo razvijanje dokumentacijske baze. Zato ne kaže preveč hiteti na začetku. Preden se take baze lotimo, je treba razpolagati vsaj z osnovno in urejeno dokumentacijo na karticah, pa tudi

z jasnimi pojmi, kaj pravzaprav želimo zasledovati - denimo: avtorje, citate literature, določene ključne besede ali deskriptorje. Obenem je treba takoj tudi pogledati, kaj lahko k takemu razvijanju pripomorejo komercialno dosegljive baze, npr. baza IR spektrov in podobne.

Še en kriterij je treba upoštevati: ceno. Za enkratne naloge, na primer za sintezo določene spojine, se ne splača razvijati dokumentacijskega sistema. Vlaganja so smoterna le, če gre za bolj ali manj permanentni sistem, ki ga bomo razvijali skozi leta. Le v takem primeru je možno razviti tudi retrospektivo, ki je - čeprav v povojih, eden najboljših dosežkov dokumentacijskega sistema z uporabo računalnika. Kajti šele retrospektiva zagotavlja primerjavo, ta pa odpira pot za marsikateri zaključek, ki ob nepovezanih podatkih ne bi bil mogoč.

Ko smo opredelili ožje področje za razvijanje lastne dokumentacijske baze, je treba specializirati kemika, ki bo delal na njej. Normalno se kemik, najsibo v proizvodnji, raziskovanju ali izobraževanju, ukvarja z molekulami oz. z naravnimi sistemi ter pri tem uporablja že uveljavljene simbole za spremembe snovi in energije. Kemijski dokumentalist pa se mora bolj ali manj na novo ukvarjati z besedami in številkami, s teorijo kodiranja, problemi semantike in razvijanjem sistema notacije. Še več: kemik-dokumentalist naj bi se ukvarjal tudi z raziskavami kemijskih informacij per se - npr. s poskusi napovedi neke nove sinteze na osnovi izpeljave sintez homolognih spojin, ali denimo z napovedmi novega alkaloida v določeni drogi, kadar kombinacije strukturnih elementov za znane alkaloidé kažejo na prisotnost oz. možnost obstoja še neizoliranega alkaloida. Take hipoteze lahko bistveno pripomorejo k raziskavam v kemijskem laboratoriju.

Kemik-dokumentalist rabi tudi dobršno mero potrpljenja. Upoštevati je treba, da se večina kemikov zanima prav toliko za programiranje, kodiranje in Boolovo logiko kot za podrobno zgradbo spektrofotometra. Kar želijo, je dobiti odgovor na svoje vprašanje na čim bolj enostaven način, čim hitreje in čim ceneje. Računalniški dokumentalist na področju kemije mora zato strpno pomagati pri pripravi in reviziji profilov, kar terjaja prehodno skupno obravnavo problemov kot osnove profilov, in evaluacijo izpisov. Na začetku se namreč ob slabo izdelanem

profilu kaj rado dogaja, da ima obsežni izpis kup neustreznih podatkov; ključni pa manjkajo.

VIRI

Za izgradnjo lastne kemijske dokumentacijske baze imamo na razpolago niz virov. Mednje sodijo najprej strokovne revije. Pri tem se je treba nasloniti na najboljše, kajti te lahko izbirajo izmed številnih ponujenih prispevkov, članki pa so tudi opremljeni s skrbno navedenimi referencami.

Vendar preteče od oddaje članka v tisk do objave v reviji pol do enega leta. Če k temu prištejemo še čas za pripravo članka, pomeni, da objava zaostaja za eksperimentalnim delom dve do tri leta. Zares sveže podatke lahko dobimo le na simpozijih in kongresih ter v neposredni menjavi med raziskovalnimi institucijami in posamezniki.

V znatni meri lahko zmanjšamo zaostajanje, če uporabimo ASCATOPICS ali CHEMICAL TITLES kot "zgodnji opozorilni sistem". Tak sistem vnaprej opozarja na članke, ki so že v tisku. Isto velja za CAC. Kemijski dokumentalist mora zato zajeti te vire.

Še eno področje dejavnosti kemijskega dokumentalista zasluži posebno pozornost: študij relevance oz. ustreznosti sistema kemijske dokumentacije v realnih situacijah, tj. pri reševanju problemov "in vivo", najsibo v raziskovalnem laboratoriju ali v proizvodnji. Pri tem je eden glavnih problemov, kako ob dobro izdelanem profilu, ki zajema veliko večino obstoječih virov, le-te "filtrirati", tj. izločiti manj pomembne, da uporabnik ne bi bil zasut z nepotrebno navlako. Gre torej za snovanje metod selekcije ne le po ključnih besedah, temveč tudi po kvaliteti podatkov. Evaluacija kvalitete podatkov pa - razen upoštevanja njihove svežine in kvalitete vira - še ni bistveno napredovala.

Končno je naloga kemika-dokumentalista še uvajanje mladih, najprej v uporabo, zatem pa še v razvijanje dokumentacijske baze. Ta je še posebej pomembna za uvajanje računalniško podpiranega programiranega izobraževanja in raziskovanja v kemiji, saj le-tega ni mogoče razviti brez splošne in specifične dokumentacije ter sistema preverjanja. V RCPV v ta namen razvijamo:

- a.) uporabo baze CAC,
- b.) lastno alkaloidno dokumentacijsko bazo,
- c.) banko testov za kemijsko izobraževanje s pripravami na zasnovo CGRT-sistema (Computer Generated Repeatable Test System).

ZASNOVA ALKALOIDNE BAZE

Alkaloidi ostajajo v ospredju zanimanja proizvajalcev zdravil, saj raziskave odkrivajo nove fiziološke učinke, primerne za uporabo teh substanc v medicini. Z zapleteno zgradbo in še bolj kompleksno biosintezo ter možnostmi sinteze in tvorbe polysintetskih derivatov z modificiranimi fiziološkimi učinki pa pritegnejo tudi fundamentalno usmerjene raziskovalce. Ta povezava obeh aspektov, fundamentalnega in aplikativnega, daje solidno osnovo za študij kemije, zato raziskave drog vzbujajo zanimanja tudi med visokošolskimi učitelji in študenti.

Število drog in pestrost zgradbe ter fizioloških učinkov njihovih alkaloidov sta dovolj velika, da opravičujeta odločitev za izgradnjo baze alkaloidov v naši dokumentaciji. Raziskavam slede vsako leto številne publikacije. Ker je to gospodarsko donosna veja, pa delujejo tudi zakoni tržišča in konkurence. Zato je lastna alkaloidna dokumentacijska baza lahko bistveni prispevek k napredku na tem področju.

To je vodilo našo skupino v prizadevanjih, postopno razviti računalniško obdelavo baze podatkov na področju alkaloidov. V prvi fazi je delo zasnovano na naslednjih nalogah:

- 1.) organiziranje baze virov literature na izbranih področjih alkaloidov,

2. priprava programa za računalniško obdelavo podatkov,
3. priprava sistema za široko uporabo,
4. uvajanje uporabnikov v izkoriščanje te baze podatkov s pomočjo računalnika,
5. razvijanje baze podatkov z dopolnjevanjem z novimi podatki in izboljševanjem sistema na osnovi izkušenj,
6. evaluacija sistema in opredelitev tistih dosežkov, ki bi lahko služili kot prispevek k razvijanju baz na sorodnih področjih.

1.) Organiziranje baze virov literature na področju alkaloidov

V prvi fazi organiziranja baze podatkov za alkaloidne je bilo delo omejeno na tista področja, na katerih potekajo raziskave teh spojin. Postopno pa širimo bazo tudi na druge droge, zanimive za raziskave zaradi fizioloških učinkov. Trenutno imamo obdelano naslednjo alkaloidno bazo, ki je prenešena na računalnik:

Galanthus	111 dokumentov
Glaucium	62
Hibiscus	50
Sedum	58
Secale (ergot)	1089
LSD	353
Valeriana	108
Veratrum	231
Vinca	1185
Securinega	43
Sylibum	97
Creteagus	24

Skupaj	3411 dokumentov
--------	-----------------

2. Viri literature za alkaloidno bazo

Uporabljeni so bili naslednji viri literature, ki so navedeni v tabeli skupaj z značilnostmi, prednostmi in pomanjkljivostmi.

Vir	Opis	Prednosti	Pomanjkljivosti
1. Originalni članki	v bazi so na razpolago v obliki separata	vsi podatki	niso vedno hitro dosegljivi ob izidu
2. Chemical Abstracts	polni podatki o avtorjih, izvleček iz članka, ključne besede		prepočasno prihajanje
3. Chemisches Zentralblatt	isto kot pod 2 do leta 1969	prenehal izhajati	
4. ASCATOPICS	izhaja od nov. 1974 izdaja Institute for Scientific Information (ISI) Philadelphia, vsebuje naslov članka, imena avtorjev, ime revije, število referenc	prihaja tedensko pred C.A., daje naslov	podatke o avtorjih okrajša največ na 7 znakov
5. Current Abstracts of Chemistry	naslov članka, imena in naslov avtorjev, revija, kratek izvleček, formule, + Index Chemicus v vsaki številki, = revije, formule, ključne besede, nove reakcije	izhaja tedensko, grafično označi glavne razisk. metode, ki so bile uporabljene	
6. Current Contents	publikacija ISI, tedenski pregled publikacij po revijah, indeks bistvenih besed iz naslova vsakega članka	omogoča enostavno iskanje	
7. CAC	ime avtorja, naslov članka, naslov revije + letnik, volumen, strani, število referenc, ključne besede, naslov ustanove	hitra organizirana dostopnost,	

Primer iskanja v CAC za Vinca alkaloida

Podatki, ki jih želimo	Zahteva
Podatki o vseh člankih, ki vsebujejo v naslovu besedo Vinca, npr. dihydrovincamine Vinca minor Vinca species	Vinca
Podatki o vinblastinu in njegovih derivatih	Vinbla
Podatki o vindolinu, vindolininu in njunih derivatih	Vindo

Podatki, ki jih želimo	Zahteva
PAZI! Če želimo čim širše o Vinca, bi lahko uporabili le zahtevo: VIN	Vin
Toda tedaj lahko dobimo kot izpis tudi: vinilne spojine vinjak (samo na papirju ...)	

3. Kartoteka podatkov

Podatke iz vseh navedenih virov in iz lastne

raziskovalne dokumentacije smo zajeli v kartoteko. Prva stran kartice vsebuje oznako področja, imena avtorjev, podatke o revijah in C.A. ter oznako, ali imamo original oz. fotokopijo, ali le izvleček iz članka. Sledijo ključne besede ali deskriptorji za vsebinsko zasledovanje člankov. Na hrbtni strani kartice pa nalepimo izvleček iz ustreznega vira.

ZAKLJUČEK

V tem prispevku sta opisani dve bazi kemijske dokumentacije, ki sta v zagonu v RCPU: baza CAC in lastna baza alkaloidne dokumentacije. Tretja baza - banka testov za kemijsko izobraževanje - je tu le omenjena. V njej bo poročano posebej.

Upamo, da ta prikaz ne bo le informiral o poteku dela in o možnostih uporabe in sodelovanja pri nadaljnjem razvijanju obeh baz, temveč tudi pritegnil nove uporabnike k sodelovanju. Tako bo vloženo delo dobilo svoj pravi smisel.

Vse zainteresirane vabimo k sodelovanju. Za neposredne stike z uporabniki skrbita: računalniški dokumentalist Metka Vrtačnik (za bazo CAC) in programer Drago Kardoš (za alkaloidne dokumentacije). Pismo se obračaje na naslov:

RCPU, Vegova 4, p.p. 18/1, 61001 Ljubljana. Po telefonu pa kličite številko (061) 22689.

LITERATURA

1. Arnett, E. McC.; Kent, A. Computer - based Chemical Information, Marcel Dekker, INC., New York, 1973.
2. Association Francaise de Documentation Automatique en Chimie: Manuel de redaction des profils documentaires en chimie, Paris, 1977.
3. Benkovič, J.; Kardoš, D.; Kornhauser, A.; Levovnik, V.; Mešiček, N.; Perpar, M.; Simončič, Š.: Razvijanje računalniške obdelave baze virov literature na področju alkaloidov, Seminar o iskanju kemijskih informacij s pomočjo računalnika, Ljubljana 1977.
4. Benkovič, J.; Kornhauser, A.; Štular, V.: Organizacija specializiranega INDOK centra za področje kemije. (Naloga RSS 784-3732/76 - Poročilo za leto 1976)
5. Benkovič, J.; Kornhauser, A.; Štular, V.; Vrtačnik, M.: Organizacija specializiranega INDOK centra za področje kemije. (Naloga RSS 784-3732/77 - Poročilo za leto 1977)
6. Geist, W.; Ripota, P. Topics in Current Chemistry, 39, Springer-Verlag, 1973.
7. Henley, J. P.: Computer - Based Library and Information Systems, Macdonald, 1972.
8. Klopfenstein, C.E.; Wilkins, C.L.: Computers in Chemical and Biochemical Research, Volume 1, Academic Press, 1972.
9. Klopfenstein, C.E.; Wilkins, C.L.: Computers in Chemical and Biochemical Research, Volume 2, Academic Press, 1974.
10. Li, D.H.: Design and Management of Information Systems, Science Research Associates, 1972.
11. Lucas, H. C. Jr.: Computer Based Information Systems in Organisations, Science Research Associates, 1973.
12. Lynch, M.F.; Harrison, M. J.; Town, W.G.; Ash, J.E.: Computer Handling of Chemical Structure Information, Macdonald, 1971.
13. Moureau, M.; Girard, A.; Delaunay, J.: Search Strategies at the Institut Francais du Petrole Using Non-french Services, Preprint, 1973.
14. Schneider, J.H.; Gechman, M.; Furth, E. S.: Survey of Commercially Available Computer - Readable Bibliographic Data Bases ASIS, Washington, 1973.
15. Vrtačnik, M.: Organizacija in delovanje specializiranega INDOK centra za kemijo, RSS, Komisija za informacijsko - dokumentacijski sistem, Ljubljana, 1978.