

amely ezáltal a komplex struktúrák különféle aspektusait fejezi ki;

Információs rendszer = az a rendszer, amelyben a társadalmi információs folyamat meghatározott oldalait (a szakmai információk keletkezésének, terjesztésének, átalakításának és felhasználásának folyamatát) regisztráljuk, és ennek alapján határozzuk meg egyes elemeit és a közöttük meglévő kapcsolatokat. Az ilyen természetű rendszer dinamikus és célra-orientált egészet képez;

Információs intézményrendszer = a ténylegesen létező intézményeknek az a halmaza, amely részben vagy egészében be van kapcsolva a szervezett információátvitel és információátalakítás társadalmi folyamatába, illetve az a halmaz, amely magába foglalja az intézményrendszer működési produktumainak összességét;

A rendszer közege = az a struktúra, amelyhez képest a rendszer elhatárolódik, s így, bár a rendszerhez képest külsődleges, mégis feltétele a rendszer működésének;

A rendszer környezete = a közeg azon része, amely tartós kölcsönhatásban áll a rendszerrel, és amelyet ismerünk kell ahhoz, hogy megismerhessük a rendszer belső funkcióit és kritériumait;

Dinamikus rendszer = az a rendszer, amely az egyes elemei közötti kölcsönhatások következményeit, továbbá a rendszer magatartásában, külső vagy belső struktúrájában és egyes elemeiben (a környezet egyes elemeiben) bekövetkezett változásokat regisztrálni tudja.

A fenti definíciók alapján könnyen belátható, hogy a rendszer kategóriájának használatát az informatikában is az anyagi struktúrák komplex egészeinek és azok fejlődésének sokoldalú dialektikus elemzése határozza meg.

/KUBÁTOVÁ, V. – TLUSTÝ, V.: *Systémový přístup ke studiu společenských informačních systémů* = *Československá Informatika*, 18. köt. 9. sz. 1976. p. 233–238./

(Futala Tibor)



TÁJÉKOZTATÁSGÉPESÍTÉS – REPROGRÁFIA

A kémiai információ gépesítése

A kémiai a tudományokat művelők igényeiből kiindulva, a kémiai tárgyú tájékoztatás olyan módszertana alakult ki, amely más szakterületeken is alkalmazható. Fejlődése a következőkben foglalható össze.

A kémiai szakirodalom jellegzetességei

Az első, kizárólag kémiai tárgyú folyóirat, az *Annales de Chimie* 1790-ben jelent meg. A kémiai folyóiratok száma a 19. század végére már jelentősen megnőtt, mert egyre több ország indította meg saját szakfolyóiratát, azonkívül megindult a tárgyköri specializálódás és helyet követelt magának az alkalmazott kémia is. A kutató már nem volt képes a primer szakirodalmat átfogni figyelemmel, és ezért megszülettek az olyan felbecsülhetetlen értékű, átfogó jellegű segédkönyvek, mint a szerves kémia területén BELSTEIN, a fizikai tulajdonságok területén LANDÖLT–BÖRNSTEIN munkái, az *International Critical Tables* stb.

Az első kémiai referálólapot Németországban indították 1830-ban. Ezt angol, amerikai, japán, majd szovjet követte. Közülük 1907-ben indult az amerikai *Chemical Abstracts (CA)*, az angol és német referálólappal megszűnése után átvette azok szerepét is. Becslések szerint egyébként is 1965-ben a világ kémiai szakirodalmi termésének több mint 50%-a angol nyelvű volt. Mindezek indokolják, hogy a kémiai szakirodalom legjellem-

zőbb vonását, a mennyiségi növekedést a CA adatain keresztül érzékeltessük:

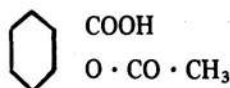
Dekád	A megjelent referátumok száma (kb.)
1917–1926	200 000
1927–1936	500 000
1937–1946	500 000 (benne a 2. világháború)
1947–1956	650 000
1957–1966	1 500 000 (1967-ben becsült adat)
1967–1976	3 500 000

A növekedés folyamatos. A *CA Collective Index* 1907–1916. évi (tíz évet átfogó) kötete még 4823 oldalon jelent meg, az 1972–1976 évi (tehát öt évet átfogó) már 110 700 oldalas.

A fejlődés azonban nemcsak mennyiségi, hanem a tartalomban is megmutatkozik. A kémiai folyamatok mélyebb megértése egyre újabb szakkifejezéseket szül, és a régieket új tartalommal tölti meg. Mindez különösen megnehezíti az „univerzális” tárgyindex létrehozását és fenntartását.

A legnagyobb probléma azonban a szerves kémiai vegyületeknél jelentkezik. Több mint három millió ilyen vegyület ismeretes, és mindezek neve a szinonimákkal, elméleti és technikai szakkifejezésekkel együtt hatalmas szakszótárat alkot. A kémiai információs rendszerek kulcsproblémája tehát a kémiai nevezéktan.

A vegyész megkülönböztet *triviális* nevet (pl. aszpirin), *tapasztalati képletet* ($C_9H_8O_4$), *szerkezeti képletet*:



szisztematikusan elnevezést (benzolsav, 2-(acetiloxi)-), vagy a WISWESSER-féle lineáris leírást használja: QWR BOV₁, vagy egyszerűen a CA regiszter számot közli: 50-78-2. A különböző elnevezések használatának a maguk helyén azonban különböző előnyei vannak és ezért nem egységesíthetők. A nevezéktan problémakörében előrehaladást hozott a számítógép alkalmazása. Mindezek a feltételek visszatükröződnek a CA szolgáltatásaiban.

A Chemical Abstracts Service (CAS) tevékenysége

Közvetlenül a 2. világháború után a CAS lényegében még a hagyományos információs technológia alapján látta el feladatait. Az 50-es évek közepén azonban az amerikai kormány anyagi támogatásával már megkezdtek a gépesítési program kialakítását és realizálását.

CAS kiadványok közül a három legfontosabb: az *Issues* című folyóirat, a *Volume Indexes*, valamint a tíz, később öt évet átfogó *Collective Indexes*, egymástól lényegében független termelési folyamatban jöttek létre. A bibliográfiai azonosítás, indexelés, vegyület elnevezés, címfordítás, CA fejezetekbe besorolás stb. munkáját különböző személyek végezték.

Az átfogó reform alapelve egyetlen mondatban fogalmazható meg: *egy dokumentumot csak egyetlen elemzésnek vessenek alá, és ennek eredményeként minden CAS szolgáltatás ellátható legyen.* A szerkesztési, tipográfiai információ nagy helyigénye miatt viszont ezt csak a feldolgozás végén, közvetlenül a kiadvány-szerkesztésekor táplálják be a rendszerbe. A dokumentumelemzés eredménye így tehát egy olyan géppel olvasható munkafile, amely nem valamelyik CAS kiadvány előképe, hanem inkább egy hatalmas jegyzetomb, amely minden lényeges információt tartalmaz.

A CAS tevékenységének alapja a *nagymennyiségű dokumentum gyors beszerzése* és belső kezelése, amit hagyományos rendszerben oldanak meg.

Az irodalomelemzési munka sem változott 1907 óta. Az elemzést végző szakember *időfelhasználása azonban lényegesen csökkent*, mert ma már nincs arra szükség, hogy idejének nagy részét helyesírás ellenőrzéssel, vegyület azonosítással stb. töltsse. Idejének közel 100%-át fordíthatja tényleges elemzésre.

A tartalom elemzését megelőzi a bibliográfiai azonosítást szolgáló „fejlec” (cím, szerző, Coden, deskriptorok stb.) kialakítása. Ennek alapján készítették el és adják ki azóta is a CAS, időben első automatizált módon előállított kiadványát, a *Chemical Titles*-t.

A következő, adatrögzítő lépésben az adatokat mágneslemezre írják, kismámítógépek állandó hibaellenőrző

felügyelete mellett. A CAS számára kifejlesztett speciális billentyűzet lehetővé teszi 125 különböző karakter bevitelét.

A számítógép bizonyos hibák esetén nem engedi további adatok bevitelét addig, míg a hibát ki nem javítják. Az ellenőrzést egy főoperátor végzi, aki a kismámítógéphez kapcsolt terminál mellett ülve maga elé vetítheti bármelyik adatelőkészítő munkájának eredményét és esetenként azonnali javításra is lehetőség van. *1972 óta automatikusan szerkesztik az anyagneveket és szerkezeti képleteket.* Ezek ellenőrzése – ha nem is zárhatja ki, de – nagymértékben csökkentheti a hibás bevitel lehetőségét.

A további adatfeldolgozás teljesen gépi úton történik. *Ha a számítógép minden szükséges adatot megkapott, elkezdődik a különböző kiadványok szerkesztése.*

A számítógép által szerkesztett anyag nyomtatását további emberi beavatkozás – és így újabb hibalehetőség – nélkül célszerű megoldani. Ezen a területen az első jelentősebb eredmény az *IBM 2280 film rekorder* alkalmazásához fűződik, amely programozhatóan jelenítette meg katódsugárcsővön a szöveget és ezt automatikusan fényképezte le (percenként mintegy 1000–1200 karaktert). Jelenleg a korszerűbb *APS-4* vette át a szerepét.

A gépesítés nemcsak a hagyományos kiadványok fenntartását volt hivatott biztosítani és nemcsak forgalomba hozható mágnesszalag file-ok keletkeztek „melléktermékként”, hanem az egész *CAS rendszer vált megbízhatóbbá és jóval rugalmasabbá*, kifinomultabbá. Jelentős a haladás az indexelés és a *Registry System*-ben. Mivel időközben megszűnt az amerikai kormánytámogatás, kérdéses, hogy a fejlődés tovább tart-e, azaz a felhasználó hajlandó lesz-e nagyobb anyagi terheket vállalni.

A kémiai anyagok jellemzői

Bár a vegyület nem írott dokumentum, alkalmas a dokumentálásra, legalább négy különböző módon:

1. *deskriptorokkal* jellemezve, mint az egyéb szakirodalomkutatási rendszerekben;
2. olyan *osztályokba sorolva*, amelyek előre adottak, vagy pedig esetiek a vegyület valamilyen tulajdonsága, szövegbeli előfordulása stb. alapján;
3. a csak rájuk jellemző *fizikai-kémiai tulajdonságaik*, biológiai, gyógyászati stb. hatásaik alapján;
4. fenti tulajdonságaik alapján, de *hivatkozva szerkezetükre is.*

1965-ben a CAS elemzői már hetente 8–10 ezer anyaggal foglalkoztak, ezeknek kb. fele új vegyület volt. A feldolgozást a következő módon végezték: a CAS elemzője *a szerkezeti képlet felírásával kezdte.* Ebből egy *mátrixot konstruált*, amelynek minden sora és oszlopa megfelelt egy-egy atomnak, és a megfelelő helyekre

bejelölte a vegyértékeket is. Az így nyert kapcsolat-mátrixot sorról-sorra lelyukasztva, a számítógépbe táplálta. A feldolgozó program értelmezi a mátrixot, ellenőrzi a szerkezetet a tapasztalati képlet alapján, majd megadja az anyag regiszterszámát, ha ilyen van, illetve (új vegyület esetén) új számot ad neki, és felveszi a regiszter-file-ba.

A kapcsolat-mátrixnál lényegesen gyorsabb bevittelt biztosítottak a később beállított kémiai írógépek, amelyek speciális karakterkészlettel és egyéb technikai módosításokkal segítették a szerkezeti képletek bevitelét. Az igen nagy volumen miatt azonban át kellett térni a „harmadik generációs” bevitelre, azaz *fényceruzás display alkalmazására*.

Az előre rendelkezésre álló és kivetíthető szerkezeti képlettörödékekből (fragmens) a fényceruza használatával összeállítható szerkezeti képletet program kódolja át géppel tárolható formába.

A kapcsolat-mátrix segítségével történő tárolás minden előnye mellett fennáll az a hátrány, hogy a vegyész számára oly lényeges gyűrűk, funkciók csoportok nem értelmezhetők azonnal. Az atomról atomra történő összehasonlítás pedig sok gépidőt követel. Ezért a kereső algoritmusokat sokféle módon fejlesztették.

Az egyéb leírási rendszerek közül talán a legelterjedtebb a *Wiswesser Line Notation (WLN)*. A lineáris leírások és így a WLN is eredetileg abból a célból születtek, hogy a szerkezeteket közönséges írógépen is le lehessen írni. Elvük, hogy az atomokat kötéseikkel együtt reprezentálja egy karakter, ahogyan karakter jelöli az egyes gyakori fragmenseket is.

Gépesítés a CAS-on kívül

A CAS ugyan különleges helyet foglal el a kémiai informatikában, de van még néhány egyéb figyelemre-méltó vállalkozás is.

Így például az amerikai szabadalmakat figyelő *Information for Industry IFI – Információt az Iparnak* állít elő kutatható mágnesszalagokat. Ezeket kezdetben csak logikai operátorokkal kereshették, de a logikai kutatás túlságosan merevnek bizonyult. Ez vezetett a súlyozott deskriptorok alkalmazásához.

A CAS mellett talán legjelentősebb (angol nyelvű) információs rendszer az *Institute for Scientific Information (ISI) – Tudományos Tájékoztatói Intézet*, amely a kémiai szakirodalom területén alkalmazza a legkifinomultabb módszereket. A referált folyóiratokat – mintegy 180-at – úgy választották ki, hogy a legfontosabb (legtöbbet hivatkozott) szakmai lapokat keresték meg. 1960 óta kb. 2 millió vegyületről adtak hírt. A két alapszolgáltatás a *Current Abstracts of Chemistry*, referálólappal, hetente indexszel jelenik meg és az *Index Chemicus*, egy negyedévenként, illetve évenként megjelenő kumulált index. A referátumokban a hangsúly a

vegyületekre és reakciókra esik, a képletek gyakran a hely 75%-át is elfoglalják. Céljaiban a CAS regiszterhez hasonló a *Chemical Substructure Index (CSI)*, amelyben *fragmensek kereshetők*. A keresés egy vagy több absztrakt sorszámhoz, és azon belül a vegyület sorszámához vezet, s így igen kényelmes. A CSI a WLN leírást alkalmazza, a kiadványokat számítógép készíti a szerkesztéstől a nyomtatásig.

A frankfurti *International Documentation for Chemistry (IDC – Nemzetközi Kémiai Dokumentáció)* méreteiben szerényebb vállalkozás, de több szempontból igen figyelemre méltó. Lényegében szövetkezeti alapon működő szervezet, a tagok referátumokat készítenek (decentralizáltan), ennek fejében az összes (centralizált) szolgáltatás megilleti őket. A szervezet a modern informatika sokféle eszközét hasznosítja (szabad indexelés, tezaurusz, hierarchia, a szerkezetet illetően: topologikus rekordok, származtatott fragmens indexelés, reakció kódolás stb.).

A hagyományos szerkezetbeviteli módszerek mellett érdekes eljárást is alkalmaznak. *A kémiai szerkezetet egy számozott hálózaton építik fel, amelyet fotocella tapogat le*. A szerkezet tárolás lényegesen mélyebb, kifinomultabb, mint a CAS módszere. A nem szerkezeti jellegű információkezelésben jelentős újítás a *kontextusban való kutatás*, azaz egy adott deskriptort a szövegben elfoglalt szerepe szerint is kerestetni lehet, pl. az aszpirint egyszer mint reakció terméket, máskor mint kiindulási reagenst.

Gépesített információkeresés

Bármilyen jó azonban az információs file, lehet jól vagy rosszul kutatni benne. A CAS politikája régóta arra irányul, hogy az egész világon szakirodalomkutató központok létesüljenek, ahol a helyi adottságoknak megfelelően kutatják végig a CAS szalagokat. A megfelelő angol központban pl. az ún. *felhasználói profil* kialakításával kísérleteznek. Ennek lényege, hogy a felhasználó értékeli az irodalomfigyelés által kapott dokumentumokat, ennek alapján a számítógépi program módosít a felhasználó által megadott keresési utasításokon, iteratív úton *kialakítja a felhasználó profilját*. Ennek az alapvetően *off-line* típusú keresési stratégiának a szerepe a közeljövőben várhatóan csökken az *on-line* kutatás térhódításával.

Az elmúlt izgalmas és gyümölcsöző évtizedet vajon fejlődés vagy megtorpanás követi majd? Ezt végül is a felhasználók fogják eldönteni.

/BATTEN, W. E.: The mechanization of chemical documentation = Journal of Documentation, 32. köt. 3. sz. 1976. p. 207–234./

(Valkó Péter)