

A tárgy ismeretének szerepe a kémiai irodalom keresésében – buktatók a tapasztalatlan keresők előtt

Az irodalomkeresés nagyon nehéz lehet és sok buktatóval is járhat. A következtetlen indexelés a nomenklatúra, a nyelv és a transliteráció problémáit már vizsgálták és nyilvánvaló, hogy az irodalomkereséshez szükség van a szakterület ismeretére; tudatában kell lenni a bibliográfiai adatbázis és a keresőrendszer korlátainak.

Az vitatható, hogy az irodalomkeresőnek kell-e a kémiai szaktudás a használható eredmények eléréséhez, viszont nagyon kétséges, tud-e tapasztalatlan személy online keresést végezni mások számára. Egy vegyészeti osztályon interaktív irodalomkeresést kívántak bevezetni, de az érdekeltek tapasztalatlanságukban nem tudták, hogy egy vegyületre vonatkozó keresés a BRS rendszeren még elérhető ChemCon adatbázisban általában kiegészítésre szorul a nyomtatott Chemical Abstracts (CA) vegyi anyag indexben vagy a teljes CA Search adatbázisban végzett további kereséssel.

A ChemCon adatbázis erősen korlátozott tartalmú a CA Search adatbázishoz képest, mert csak a szerzők nevét, a címbeli szavakat és a CA füzetekben használt kulcsszavakat (Keywords Index) tartalmazza. A CA Search adatbázis viszont tartalmazza a CA kötetindexeiben (General Subject Index, ill. Chemical Substance Index) található sokkal részletesebb index-kifejezéseket is.

A különböző adatbázis-szolgáltató központok a gazdaságos kémiai irodalomkeresés nehézségeit igyekeznek csökkenteni. Tanfolyamaik azonban nem foglalkoznak kellően az indexelési irányelvek és a nomenklatúra problémáival, sem azzal, hogy a kémiai irodalomkeresőnek legyen valamelyes szerkesztési ismerete.

A CA-nak például az az elve, hogy csak olyan vegyületeket indexel, amelyekre "új" információ érkezik; hiányzik a szabadalmak és disszertációk mély-indexelése. Ilyen és egyéb okokból a teljes CA Search adatbázis (Chem-Con & CASIA) lekérdezése a tökéletesen logikus kereső kérdésmegfogalmazás ellenére is még mindig nélkülözheti a teljességet.

Oppenheim több tanulmány alapján kijelenti, hogy a CA a kémiai szabadalmak tekintetében gyengén informál. Ugyanő másutt rámutat, hogy a CA bizonyos vegyületekről új információt közlő 56 angol szabadalom közül 55-öt kivonatol, de a tárgymutató és a képletmutató használatával ezeknek csak 49%-a volt megtalálható. Freeman és Oppenheim a kémiai kivonat készítő szolgálatokkal kapcsolatban megállapítja, hogy a szabadalmak beépítése,

indexelése és időbeli késedelme tekintetében nagyon különböznek.

Oppenheim és Sutherland a kohászati szabadalmakkal kapcsolatban kiderítette, hogy a CA késedelme kétszerese a World Patent Indexének (WPI); az USA-szabadalmak és az egyéb szabadalmak között a kiválasztás nem azonos mérték szerint történik. Másutt arra mutattak rá, hogy a szabadalmak gyakran az egyedüli információforrást jelentik, több információt adnak és gyakran sokkal korábban informálnak, mint az egyéb források. Oppenheim és van den Ende azt találták, hogy a mikrobiológiai szabadalmakat a fontosabbak közül csak három kivonatoló szolgálat dolgozza fel. A CA mennyiségben és frissességben felülmúlta a másik kettőt, de messze elmaradt a WPI mögött.

Kaye szerint a CA hiányosságai a következők: a nem-francia bejelentők francia szabadalmait mellőzi; nem arányosan fedi le a szabadalmakat a származási országtól függően; a nyomtatott CA indexekben következtetlen a terminológia, különösen egy új vegyészeti ág első publikációjának feldolgozásakor.

Davison szerint valamennyi kivonatoló-indexelő szolgáltatás hatásossága rossz. Hivatkozik más olyan írásokra is, amelyek a tudományos indexek hiányosságairól szólnak és vizsgálja a spektroszkópia tárgykörét felölelő, manuális visszakeresésre szolgáló indexek hatásosságát. Szerinte spektroszkópiai információkat tartalmazó 6–7 cikk közül csak kb. egynek a címe utal erre. Számítógéppel kapcsolt tömegspektrometriára vonatkozó irodalomkeresés során egyetlen forrás sem – köztük a CA sem – adott 40%-nál nagyobb teljességet, ha pedig csak a kézenfekvő címszavak szerint folyt a keresés, akkor a CA-ban a találati szám 11%-ra esett le.

Loev a CA-val kapcsolatban főleg a nomenklatúra problémát veti fel (a CA a dietil-éterre 1,1'-oxybisethane-ként utal). Bár vannak online nomenklatúra fájlok, drágák és használatuk előbb a nyomtatott CA Index-Útmutatókat (Index Guide) és más forrásokat kell áttanulmányozni. A nomenklatúra változások miatt átfogó keresés alkalmával a különböző időperiódusokra vonatkozó valamennyi Index-Útmutatót meg kell nézni. Loev a szerkezeti képleteket tartalmazó képletindexet részesíti előnyben a CA vegyi anyag indexével szemben.

Antony és Stevens összehasonlította a Ca, a Current Abstracts of Chemistry és az Index Chemicus (CACIC) használhatóságát; a CACIC kötetekben újként indexelt 200 vegyület között 32 olyan akadt,

amely a CA képletindexében vagy vegyi anyag indexében nem szerepelt.

Bawden és Jackson arról számol be, hogy a vegyületek toxicitására és veszélyességére vonatkozó információ a cikkekben gyakran nem kellő súllyal szerepel és ezért nem sikerülhet a keresés a kivonatoló és indexelő szolgálatok révén. Bawden és Brock összehasonlító céllal nyolc próbakeresést végzett különböző forrásokban. A CA kb. 50%-os teljességet biztosított, ami ugyan jó eredmény, de bizonyítja, hogy a CA-t más forrásokkal együttesen kell használni.

Charton különböző fogalmakat keresve a CA-ban megállapította, hogy a korrelációs elemzésről szóló közlemények 50%-a volt visszakereshető, ami megegyezett a Chemical Condensates adatbázissal.

Funk és társai az indexelés következetességét a Medline adatbázisban vizsgálták és megállapították, hogy a manuálisan indexelt adatbázisok közül valószínűleg a Medline a legmagasabb színvonalú, ám ennek ellenére az indexelés következetessége csak kb. 75% volt.

Ezek a megállapítások nem rontják le az egyébként kiváló Medline vagy Ca hitelét, de arra figyelmeztetnek, hogy minden szolgáltatnak megvannak a korlátai. Ahol léteznek konkurens szolgálatok, ott fel kell világosítani a felhasználókat arról, hogy a keresés teljessége a publikációk és/vagy adatbázisok igénybevételének gondosságától függ. Az irodalomkeresést végző könyvtárosnak magas szintű szakismeretet kell szereznie, a gyanútlan felhasználókat pedig figyelmeztetnie kell a kutatással járó problémákra.

Sok esetben a felhasználók nincsenek tisztában azzal, hogy milyen információt veszítettek el. Ez a helyzet tükröződik Jahoda tanulmányában, ahol a ChemCon adatbázissal próbáltak kimerítő keresést végezni anélkül, hogy felhívták volna a felhasználók figyelmét az adatbázis hiányosságaira. Néhány olyan példa, ahol a felhasználók könnyen tévedésbe eshetnek:

1. Az összes kobalt-széndioxid komplex keresését a kobalt és a CO-re vonatkozó nyilvántartási szám (RN=124-38-9) összekapcsolása alapján nehezíti teszi a CA gyakorlata; mely szerint nem indexeli a CO₂-re vonatkozó nyilvántartási számot az online rekordon, amikor a CA nyomtatott vegyi anyag indexben ilyen bejegyzés áll: **Széndioxid**, vegyületek kobalt komplexe (53988-18-4).
2. Különösen zavaró, hogy a CA nem indexel kivonatokat, amikor átfogó bibliográfiára van szükség egy bizonyos szerzővel kapcsolatban. (Szerencsére ezt a hiányt pótolja a Science Citation Index és a SciSearch.)

3. Girard és Moureau közlése szerint a leginkább zavaró három olyan cikk felderítése volt, amelyekben rossz nyilvántartási számot adtak a t-butil-metil-éternek.

Vannak olyan kérdések is, amelyekre nem létezik egyszerű, gazdaságosan elérhető megoldás. A tárgyat és/vagy az adatbázis korlátait nem ismerő tapasztalatlan személy az ilyesfajta keresésekre fizethet ki nagy összegeket és kap nem kielégítő választ.

- ◆ Gázfázisbeli reakciók sebességi állandóinak keresése, pl. NH₃-ra, C₂N₂-re, C₃N-re stb. vonatkozólag, laboratóriumi mérésekre korlátozódva, kizárva, az elméleti számításokat, az asztrofizikai tanulmányokat, a kondenzált fázisokat stb.
- ◆ Hidrogénszulfid-atmoszférában a víz harmatpontjának keresése.
- ◆ Vékony mangánfilmek keresése, amelyek dőpoló szerként kizárják a bizmutot és más elemeket.
- ◆ A fizikai tulajdonságokra irányuló keresés. A Gmelinben, a Beilsteinben és a Landolt-Bornsteinben sokkal több ilyen anyag található, mint ami online adatbázisokban könnyen elérhető.
- ◆ Mellon könyvében a pirokatechin készítésére közölt bonyolult keresési stratégia, amikor az RN=120-80-9P keresőkérdés bevitele is elegendő.

Súlyos a hiba, ha egy egyetemi kutató évek munkája után ébred tudatára, hogy már korábban más kutatta a témáját, vagy a kutató egy vegyi anyagnak már ismertett toxicitásáról nem tudva, végzetes kezelési hibát követ el.

Jogos bizonyos szkepticizmus az online keresést könyvtárakban, információs központokban ilyen irányú szakképzettséggel végző szakemberekkel szemben. A keresés tárgyát ismerő kereső szakember előnyben van az olyan könyvtárosokkal szemben, akiknek nincs komoly ismeretük a technológiáról és az információk közvetítését megalapozó tudományról és nem tudják ismereteiket alkalmazni egy speciális területre.

/ROTH, D.L.: The role of subject expertise in searching the chemical literature ... and pitfalls that await the inexperienced searcher = Database, 8.köt. 1.sz. 1985. p. 43–46./

(Kincses István)