

ELEKTRONIKUS SZÁMITÓGÉPEKSEL DOLGOZÓ VEGYIPARI DOKUMENTÁCIÓS OSZTÁLY AZ ORÉAL KUTATÓKÖZPONTJÁBAN

Georges Sag

Az IBM France 1965 szeptemberében Le Vesinetben megtartott "A számítógépek szerepe az információ visszakeresésében" témájú szemináriumán először volt alkalmunk nyilvánosságra hozni gépesítési kísérletünk kezdeti eredményeit. Az Association Française des Documentalistes et des Bibliothécaires Spécialisés Informations c.lapjának a franciaországi dokumentáció automatizálásával foglalkozó különszámában (Documentaliste 1966, 107-14) R.Charle műszaki vezérigazgatóval kifejtettük gépesítési módszerünk alapelveit és módszereit. (A Könyvtártudományi Szakkönyvtárban 4-2407 számon található).*

Jelen tanulmányunkban összefoglaljuk fent említett két közleményünket, de főleg gépesítési kísérletünk továbbfejlesztésének tapasztalatait kívánjuk ismertetni. Hangsúlyozzuk, hogy valamely gazdag dokumentációs anyag gépesítése, annak kezdeti szakaszában jelentős költségtöbblettel jár és anyagi előnyökre csak a teljes régi anyag feldolgozása után számíthatunk.

Az Oréal cég, 40 gyárüzemében -Párizstól Tokióig- mintegy 1200 kozmetikai és háztartási vegyiterméket gyárt, évi 170 millió dollárt meghaladó értékben. A központi kutató laboratórium (Aulnay sous Bois) ma több mint 100 vegyészt foglalkoztat, de e sorok írója még 1953-ban szervezte meg a cég dokumentációs osztályát, amikor a kutatói létszám sokkal szerényebb volt.

Az osztály tagozatai a következők:

1. Gyűjtemények:

- a/ 1000 kötetes könyvtár, amely magában foglalja az alapvető műveket (Beilstein, Organic Reactions, Organic Syntheses, Teil-

* Charle, Roger - Sag, Georges: Essai d'automatisation de la documentation cosmétologique (Laboratoires Oréal). - Documentaliste. 1966. 1. trim. [2.] no. spec. 107-111. p. Illusztr. Bibliogr. 114. p. 5 tétel.

neimer, Uilman, Houben-Weyl stb.)

- b/ Folyóirat gyűjtemény (mintegy 120 periodikum), számos sorozat visszamenőleg az 1920-1930-as évekig.
- c/ Minden országra kiterjedő, mintegy 10 000 darabból álló szabadalmi leírás gyűjtemény
- d/ Mikrofilm gyűjtemény, mintegy 10 000 mikrofilmcsikkal, Minnesota Mining and Mfg Co. leolvasó-nagyító géppel
- e/ Katalógus és prospektus gyűjtemény, több mint ezer dokumentummal
- f/ Központi katalógus-karton terem a hozzátartozó sokszorosító berendezéssel (Colos)

Ez a részleg 3 személyt foglalkoztat. Egy segédkönyvtárosnőt, aki egyben az osztályvezető titkárnője, egy gépirónőt és egy alkalmazottat, aki a sokszorosítási munkákat és a kartonok osztályozását végzi. A fényképészeti laboratórium, a xerográfikus (Rank) és fénymásoló (blue-print) berendezés, valamint az Ofset-sokszorosító a dokumentációs osztály rendelkezésére áll, de attól független.

2. Feltárási részleg:

Minden kutatólaboratóriumi csoportvezetőnek (5 főmérnök) egy dokumentalista áll rendelkezésére, aki a beérkező dokumentumok feltárását végzi az illetékes kutatócsoport érdeklődési körének megfelelően. Ezeket 21x27 cm-es leírólapokon rögzítik (alkoholos sokszorosítás, 2 gépirónó). A kivonatok a leg-rövidebb időn belül az illetékes kutatók, a tudományos igazgató, valamint a dokumentációs osztály vezetője elé kerülnek.

3. Központi szerv:

A feltárási részleg leírásait és az üzem érdeklődési körébe tartozó új szabadalmi leírásokat a központi szerv vezetőjének elbírálása szerint két csoportba soroljuk:

- a/ a dokumentum a kutatási programot közvetlenül vagy közvetve, távolról vagy közelről érinti;
- b/ a dokumentum csak műszaki, laboratóriumi vagy analitikai számadatot tartalmaz. A leírás csak a kutatók általános ismereteinek gazdagítását célozza, vagy pedig az üzem adottságait tekintetbe véve, a leírás érdektelen.

Az a/ csoportba tartozó dokumentumok tartalmát gépesítésre előkészítjük - kódoljuk. Erre a célra kizárólag az eredeti dokumentumokat használjuk fel és sohasem azok leírásait. A leírásokat, esetleges kiegészítés vagy egyszerűsítés után, katalógus-kártyákra sokszorosítjuk, a központi katalógus-terem, valamint

az érdekelt részlegek (műszaki vezérigazgató, szabadalmi osztály karton-katalógusa, végtermékeket formuláló laboratóriumok stb.) részére. A b/ csoportba tartozó leírásokból kódolás nélkül készítünk katalógus-kartonokat (2 gépirónó) az a/ csoportban felsorolt érdekelték részére. Jelenleg a központi szerv egy vezetőt és 3 dokumentalistát foglalkoztat, ami figyelembe véve a nagy mennyiségű régi anyag feldolgozásának szükségességét, nem elég a munka elvégzésére.

A katalóguskartonokat a szokásos 3 csoportba soroljuk:

- szerzői katalógus
- tárgyyszó-katalógus (deszkriptorok)
- ország-és számrendi szabadalmi katalógus

A tárgyyszó-katalógus mintegy 2000, betűrendbe sorolt tárgyyszót (deszkriptort) foglal magában, melyet az osztály első 5 működési éve alatt állított össze. A deszkriptorok kidolgozásának problematikáját - különös tekintettel a gépesítési szempontokra - az ADBS műszaki és tudományos csoportja részére tartott előadásunkban ismertettük. (Ennek szövege a Könyvtártudományi Szakkönyvtárban a: érdeklődők rendelkezésére áll a 4-2311/1 számon)*

Miután gyűjteményeink az elmúlt évek folyamán, évente mintegy 7000 dokumentummal gyarapodtak, már néhány esztendeje nyilvánvalóvá vált, hogy a tárgyyszó katalógusban, a katalógus kartonok felhalmozódása miatt, gyakorlatilag nehézkes a feladott kérdések felkutatása, sőt rendkívül nehéz megállapítani azt is, ha valamely téma nem szerepel a katalógusban.

A mai ipari élet követelményei mellett egy tradicionális 150 000 kartonos katalógus üzemeltetése korszerűtlen. Ennek a ténynek felismerése ellenére, nem lehetett átmenet nélkül, a kizárólagos gépi dokumentációra átállítani az osztályt, egyrészt a nagymennyiségű feldolgozatlan régi anyag, másrészt a b/ csoportba tartozó új dokumentumok miatt.

A régi anyag és a b/ csoportba tartozó új anyag feltárásának elégtelenségét azzal jellemezhetjük, hogy az átlagos deszkriptor szám /dokumentumonként csupán 1,5.

Gépesítésünk kísérleti szakaszában 1963-ban mintegy 1000 dokumentumot kódoltunk a hajfestés és a hullámosítás témaköréből. Ehhez a munkához 1600 különböző deszkriptort alkalmaztunk. Az átlagos deszkriptor szám 10 felett volt dokumentumonként, ami az ismert franciaországi gépesítési kísérleteknél jóval mélyrehatóbb analízist mutat. Az első nálunk ismert kísérletet a vegyiparban a Compagnie St-Gobain végezte (Levéry F. et al.: Une expérience de sélection

* Sag, Georges: Essai d'automatisation de la documentation cosmetologique. Seminaire IBM, role de l'ordinateur dans la recherche de documentation. Le Vesinet 27-28 septembre 1965. 5. p.

automatique de documentation, Saint-Gobain-IBM-France, Paris, 1960, 52 pp.), 8,5 átlagos deskriptor számmal. Az egyéb iparágakból ismert gépesítési kísérletekben az analízis intenzitása még alacsonyabb volt, a Compagnie Générale de Télégraphie sans Fils (Gamb, J., IBM-France szeminárium, Le Vesinet, 1965.szeptember) 7 deskriptort használt dokumentumonként, a Compagnie Générale d'Electricité (Lehmann, G. - Guedeney, F.: Mémoires soc. ingénieurs civils France, 1965, mai, 12 pp.) pedig csak 3,5 deskriptort. Jelenleg az átlagos deskriptor számunk az egész memorizált anyag után számítva 12 felett van.

Az osztályunkon végzett gépesítés lehetővé tette:

1. az analízis igen jelentős elmélyítését, és azt, hogy minimális idő alatt válaszoljunk olyan kérdésekre, hogy pl. milyen kozmetikai célokra alkalmaztak eddig X nyersanyagot, vagy pedig Y nyersanyag szerepel-e szabadalmi leírásokban Z alkalmazási területen;
2. a témafigyelést;
3. a gépi memorizálás melléktermékeként szerzői, tárgyi és szabadalmi szám szerinti index rendszeres kiadását;

Végül pedig a gépesítés kizárja az elkallódott vagy rendezetlen katalógus-cédulákat.

Gépesítési eljárásunkban eddig - hetenként 2 órán át - az IBM-1460 számítógépet használtuk. Ez az időtartam természetesen csak a lyukkártyák visszakeresési, illetve az index rendezési és kinyomási idejét jelenti; a lyukkártyaműhely igénybevétele havonta mintegy 100 munkaórát jelent és egy program-speciális hosszú hónapokat töltött problémáink megoldásával. Gépparkunk most az eddiginél jóval nagyobb kapacitású IBM-360 számítógéppel bővült.

A kidolgozott program alapelve a kötött, de nyílt nyelvű kódolás, az adatbeviteli űrlapokon szereplő deskriptorok kódszámait a számítógép állapítja meg. A visszakeresési folyamat mágneses lemezek segítségével történik, melyeken minden deskriptor kódszáma mögött fel vannak tüntetve azon dokumentumok sorszámai, amelyekben a kérdéses deskriptor szerepel. A memória ezen legfontosabb eleme mellett a számítógép létrehoz és időközönként kiegészít:

1. két mágneses deskriptor szalagot - az elsőt a deskriptorok és szinonimáik betűrendben, a másodikon kódszámuk sorrendjében szerepelnek;
2. egy "könyvtár-szalagot", amelyen a dokumentumok szerzői, címei, szabadalmi számai, bibliográfiai adatai és az összes jellemző deskriptorok szerepelnek, a dokumentumok sorszámai szerint rendezve;
3. egy "dokumentum-szalagot", amely abban különbözik az előző mágneses szalagtól, hogy ezen deskriptorok nem szerepelnek. A jelzett mág-

neses szalagok, valamint a mágneses lemezek, a "fichier inversé" tartalmát a számítógép kiadó egysége periódikusan kinyomja.

A betűrendes deszkriptor-jegyzéket kódolási szótárként használjuk. A kódszám szerinti deszkriptor-jegyzék az előző kiegészítésére szolgál, amennyiben ez az azonos kódszámú szinonimákat csoportosítja, megkönnyítve a témának legmegfelelőbb deszkriptorok kiválasztását.

A könyvtár-szalag a gépi visszakeresés eredményének gyors ellenőrzését teszi lehetővé. Ezen a kiadványon ugyanis a deszkriptorok, kódolásuknak megfelelően, logikus sorrendben szerepelnek. A deszkriptorok szintaktikus vagy paradigmátikus összefüggésben való kódolását eleve visszautasítottuk, azzal az indokolással, hogy egyrészt a vegytan elég logikus tudomány ahhoz, hogy a deszkriptorok egyszerű juttapozíciója kifejezze a vezérszavak összefüggését, másrészt, mert az ilyenfajta összefüggések kódolása több idővesztéssel járna, mint az az időnyereség, melyet elkészítésük és használatuk biztosítana.

A dokumentum-szalag az adatvisszakeresési válaszok teljes referenciáinak kinyomtatására, a szerzők névszerinti, vagy a szabadalmi leírások ország és sorszám szerinti jegyzékeinek kiadására szolgál.

A m á g n e s e s l e m e z e k ("fichier inversé") tartalmának kinyomtatott jegyzéke bizonyos esetekben lehetővé teszi a számítógép memóriájának használatát a gép igénybevétele nélkül. Továbbá lehetővé teszi nagyértékű dokumentológiai statisztikák leolvasását, mint amilyen például a deszkriptorok előfordulási gyakoriságának megállapítása, vagy esetleges kiküszöbölésüket, ha a deszkriptorokat csak elenyésző számú dokumentum kódolása tartalmazza.

Már gépesítésünk kísérleti szakaszában is észleltük, hogy a deszkriptorok vegyi jelleg szerinti csoportosítása szükséges; ezt a problémát módszerünk továbbfejlesztésének alapvető feladatául tekintjük. A deszkriptorok értelem szerinti csoportosítása magában foglal egy gépesíthető eljárást és egy szakértői munkát követelő feladatot. 1964 októberében a Sorbonne-on beiktatott doktori értekezésünk tervében többek közt vázoltunk egy módszert, a vegyi szakkifejezések gépi kódolására és csoportosítására. A Chemical-Biological Activities Tapes and Search Service (Columbus, Ohio, 11 pp.) füzeté szintén említi annak lehetőségét, hogy az Amerikai Kémiai Társaság kiadványaként megjelenő fenti című mágneses szalagokon a vegyületek neveit szógyökök szerint lehet analizálni, visszakeresni és csoportosítani.

Az összes lehetséges gépi és manuális csoportok kivonása valamely szókincsből természetesen teljesen felesleges munka lenne, viszont az üzemi érdeklődési körébe vágó csoportok összeállítása és e csoportok

tok állandó időszerűsége igen fontos feladata minden gépesített dokumentációs osztálynak. Így például R. Charle műszaki vezérigazgatónkkal a "Manufacturing chemist and aerosol news"-ban, ez év végén megjelent cikkünk, a haj és szőrme festékekről, manuálisan készített csoport-lexikonunk igénybevételével készült.

Befejezésül még megjegyezzük, hogy a gépesített dokumentáció értékmérője az alkalmazott szókincs minősége. Ez a szókincs a kifejezés irodalmi értelmében értelmezendő, amennyiben ez pontosabban jellemez egy szervezési szellemet, mint egy műszaki lexikont.