

29. SZOKOLOV, A. V. – MANKEVICS A. I.: Szocial'naja informatika i bibliotecsno-bibliograficeszkije diszciplinu. Leningrad, Leningradskij Goszudarsztvennij Insztitut Kulturü imeni N. K. Krupszkoj, 1974. p. 5–9.
30. ENGELBERT, H.: Zu einer Theorie der gesellschaftlichen (sozialen) Information. Ilmenau, Technische Hochschule, 1976. p. 45–54.

/ENGELBERT, H.: Wissenschaftliche Informations-tätigkeit und wissenschaftlich-technischer Fortschritt = Informatik, 25. köt. 5. sz. 1978. p. 41–47./

(Vermes Mária)



A tudományos–műszaki irodalom referálása: elméleti alapok

Az automatikus referálással foglalkozó művek a különböző alkalmazott módszereket az elsődleges dokumentumban lévő mondatok „szűrésére” használt eljárások szerint osztályozzák [1–4].

V. E. Berzon [1] a szöveg automatikus tömörítésének következő módszereit különbözteti meg: 1. statisztikai, 2. pozíciós, 3. deskriptoros, 4. megkérdézéses, 5. asszociációs, 6. szemantikai, 7. szintaktikai, 8. a hasonló szövegek szerkezetének vizsgálatára épülő módszerek.

I. P. Szevbo [2] a statisztikai és szintaktikai módszerek mellett az információs–logikai nyelvek alkalmazására épülő eljárást emeli ki.

Az amerikai kutatásokat áttekintő munkájában V. P. Leonov [3] statisztikai, logikai–matematikai és nyelvészeti módszereket említi.

Ugyanezeket a szövegrészeket értékelő módszereket R. E. Wyllis [4] statisztikai, szemantikai–szintaktikai és szöveges osztályokba sorolja.

A fentiekből is látható, hogy az automatikus referálási módszereknek nincs egységes osztályozási rendszere. E hiány kiküszöbölésére a cikk párhuzamot von a referáló szakember munkája és a folyamat modellje között.

A szakértő általi referálás mechanizmusát kísérleti szinten két mű vizsgálja [5, 6]. G. Sz. Zsdanova [5] az elsődleges és a másodlagos dokumentumok tartalmát hasonlította össze, míg N. A. Helmuth [6] a megkérdézéses módszert alkalmazta. Valószínűleg ez az oka annak, hogy a két mű kiegészíti egymást: amit az egyik szerző kimondatlanul is feltételez, a másik le is írja. Így pl. Helmuth a referálás folyamatát nem osztja szakaszokra, Zsdanova pedig 3 szakaszt különböztet meg:

1. Az első szakasz a referálandó dokumentum végigolvasása, amely az elsődleges dokumentum tartalmának interpretálását eredményezi a közleményben kifejtett gondolatok egymásutánja vagy a cikk tartalmáról alkotott általános elképzelés formájában.

2. A második szakasz műveleteit a referáló egyidőben végzi ugyan, de célszerű megkülönböztetni:

a dokumentum tartalmát tükröző fő elemek kiválasztását és

a másodrendű, kevésbé fontos információk kiszűrését.

3. Végül a harmadik szakaszban a dokumentum tartalmát tükröző fő elemek kerülnek közlésre a leg-tömörebb formában.

Zsdanova megemlíti a referálási folyamatnak a gyakorlatban alkalmazott két változatát is.

Egyik esetben az elsődleges dokumentum tartalma a közleményben kifejtett gondolatok egymásutánja vagy a cikk tartalmáról alkotott általános elképzelés formájában jelenik meg. Ebben az esetben az elsődleges dokumentum szövegrészeinek felhasználási aránya majdnem zérus.

A másik esetben az elsődleges dokumentum tartalmát kiválasztott mondatok, gondolati egységek és szövegrészek tükrözik. Az elsődleges dokumentum szövegével való egybeesés különböző fokú lehet. A referátum leírása, a folyamat harmadik szakasza ennek a kiválasztásnak az alapján történik. Ebben az esetben a referálási folyamat első és második szakasza egybefolyik, ezért a kvázireferáláshoz hasonlít, melynek során a referátumban szereplő mondatok sem lexikailag, sem szintaktikailag nem térnek el az elsődleges dokumentum mondataitól.

Helmuth nem a referátumkészítés módszereiről beszél, hanem azokról az intellektuális folyamatokról, amelyek eredményeként a kvázireferátum létrejön. 390 könyvtári és információs intézményben dolgozó szakembernek kiküldött kérdőíven a következő kérdést tette fel: „milyen intellektuális folyamatok alkotják a kvázireferálást?” A kérdőívekre mindössze 57 válasz érkezett:

ugyanazok a folyamatok, mint a referálást és indexelést – 9 válasz;

a dokumentumban lévő főbb gondolatok áttekintése – 16 válasz;

a dokumentumban kifejtett gondolatok szelektív értékelése – 11 válasz;

az adott terület szakértője kell a referáláshoz – 19 válasz (mivel a referáló ismeri a felhasználók igényeit és érdekeit);

egyéb – 2 válasz.

W. Ashworth [7] referálási útmutatójában 5 szakaszból álló folyamat betartását ajánlja a referálóknak.

1. a cikk elolvasása, hogy a tartalomról világos és összefüggő elképzelése legyen;

2. a tartalom leglényegesebb elemeinek kijelölése;

3. az egyes tartalmi elemek fontosságának értékelése;

4. a kiválasztott tartalmi elemek logikus elrendezése, mivel a referátumnak logikailag összefüggő információk kell tartalmaznia;

5. a referátum megszerkesztése és ellenőrzése.

Ez a felosztás nagyjából megfelel a *Zsdanova*-féle felosztásnak. *Ashworth* azt is hangsúlyozza, hogy a fő tartalmi elemek kiválasztásakor azt is figyelembe kell venni, hogy a felhasználónak ezek mennyire fontosak.

Hasonló gondolatokat fejtenek ki *B. H. Weil* és szerzőtársai [8]. Ha a referáló nem teljes mértékben ismeri a felhasználók igényeit, kívánságait és céljait, nehézségei támadnak a referálás során.

Bár a fenti gondolatok magától értetődőnek látszanak, az automatikus referálási módszerek osztályozásakor mégsem vették őket eddig figyelembe. Ezért azoknak az ismereteknek az elemzése, amelyekre a referálóknak a másodlagos dokumentum készítésekor szüksége van, nem teljesen nyilvánvaló következtetésekhez vezet.

J. A. Srejder [9, 10] a „külső világról alkotott elképzelések” modelljét, vagyis a teauruszt bizonyos ismeretek rendszereként határozza meg. Véleménye szerint az információcsere csak egymáshoz eléggé közel álló (egyeztetett) teauruszokkal (ismeretanyagot kifejező szókészlet) lehetséges, különben nem jön létre a kölcsönös megértés. Ezért van az, hogy a referálóknak a vonatkozó ágazat szakértőjének kell lennie azaz ismernie kell az adott területen elért eredményeket és a felhasználói kör igényeit is. Ezenkívül világosan kell látnia munkájának célját és ismernie kell a referátummal szemben támasztott követelményeket. Minél jobban

ismeri a referáló ezeket a követelményeket (referálási útmutató), minél nagyobb a referálói gyakorlata, annál jobb lesz a referátum, abban az értelemben, hogy a tudományos információ teljesebb és pontosabb felhasználására vezet.

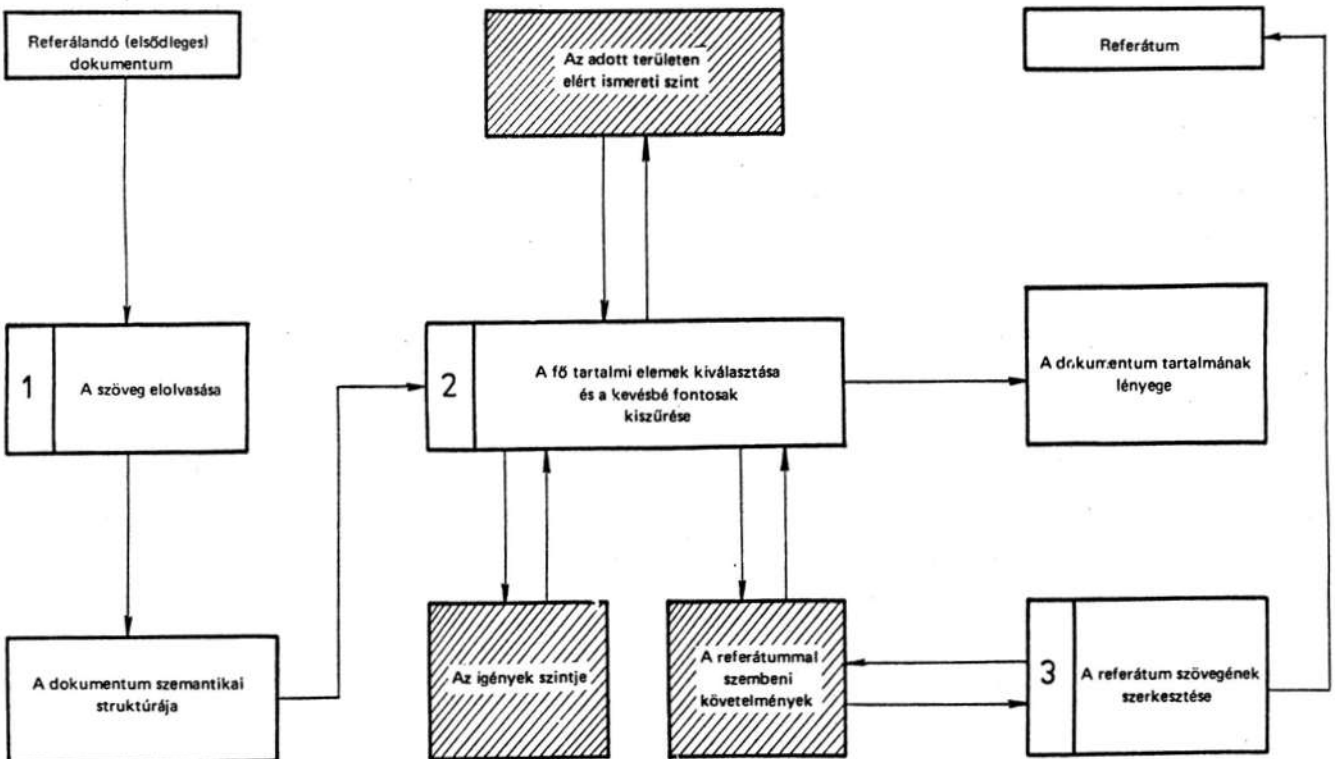
Összegezve az elmondottakat: a dokumentum elemző-szintetizáló feldolgozása során a referáló ismerete a következőkből áll:

- a konkrét tematikai területen elért eredmények ismerete;
- a felhasználók igényeinek ismerete;
- a referátummal szembeni követelmények ismerete.

Feltételezve, hogy ezek az ismeretek elegendőek a referálási folyamat leírásához, a folyamat *Zsdanova* által leírt menete a fenti ismeretekkel kiegészítve az 1. ábrán bemutatott modellhez vezet.

Az ábrán 1–3 jelöli a referáló által végzett műveleteket. A referálandó dokumentum elolvasásának (1. művelet) eredményeként világossá válik a szemantikai struktúra, azaz megjelenik az *elsődleges dokumentum tartalma*.

A szemantikai struktúra további elemzésének (2. művelet) eredménye a *fő tartalmi elemek kiválasztása és a kevésbé lényeges információk kihagyása*. E művelet során van szüksége a referálóknak a világról meglévő ismereteire. Arra vonatkozó döntése, hogy a dokumentumból mely információkat tart fontosnak és melyeket másodrendűnek, lényegesen függ az ismereteit tükröző



1. ábra A referálási folyamat műveletei

szókészletben és a referálandó szövegben lévő információk összehasonlításának eredményétől.

Igy szűrődik ki az elsődleges dokumentum tartalmának lényege, amely a szintézis szakaszában (3. művelet) *referátum formájába kerül*.

A modellekkel szemben azt a követelményt támasztjuk, hogy megmagyarázzák a meglévő jelenségeket és segítséget nyújtsanak a még nem ismert tények feltáráshoz. A fenti modell segítségével könnyen megmagyarázható, hogy ugyanazt a publikációt különböző referálók miért referálják eltérő módon. A referátumok különbözőségének oka ugyanis az, hogy *a referálók eltérő ismeretanyaggal rendelkeznek*. A modell segítségével ismeretlen dolgokat is fel lehet tárni, ha az automatikus referálással foglalkozó munkák osztályozására használják.

Az ábra alapján az *automatikus referálási módszerek* a következő csoportokra oszthatók:

1. azok a módszerek, amelyekkel a referálandó dokumentum szemantikai struktúrájának leírása alapján *kiválasztják a fő tartalmi elemeket* és kiszűrik a kevésbé fontos információt;

2. azok a módszerek, amelyekkel *a referátummal szembeni követelményeket modellezik*, és ezt a modellt vagy a fő tartalmi elemek kiválasztásához és a kevésbé fontos információk kiszűréséhez, vagy a referátum szövegének szintetizálásához használják;

3. azok a módszerek, amelyekkel *a konkrét tematikai terület ismereti szintjét modellezik*.

Az egyes csoportokba tartozó módszerek rövid jellemzése során legszembetűnőbb, hogy a legtöbb tanulmány az 1. csoportba tartozik, a 2.-ba már kevesebb és a 3. csoportban mindössze egyetlen szöveget szerző által írt művet [12] sikerült találni.

Az 1. csoportban lévők közül is a legkidolgozottabbak a szöveg kutatók munkái [1, 2, 13–19]. A leegyszerűbb módszerek a statisztikai eljárást alkalmazzák a szövegszerkezet elemzésére [20–25], a bonyolultabbak, amelyek a leegyszerűsített szabványosított mondatok kiválasztásához a szintaktikai elemzés algoritmusát használják [26–30], egyelőre nem vezettek eredményhez.

A 2. csoportba tartozó módszerek az indexelés és a referálás formalizálására irányuló kutatásokra [31–37] épülnek.

A referálási modell alapján lehetséges a módszereknek egy negyedik csoportja is, amelyekkel az *információs igényeket modellezik*. Az irodalomban azonban nem sikerült ebben az irányban folytatott automatikus szövegtömörítésről szóló munkát találni. A gyakorlatban azonban ez egy megszokott eljárás, eredménye a meghatározott felhasználói kategóriára orientált referátum.

Irodalom

1. BERZON, V. E.: Avtomaticheskoe szmüszlovoe szvertüvanie na osnove analiza szverhfrazovüh szvjazej tekszta. Sztatistika recsi i avtomaticheskij analiz tekszta. Leningrad, Nauka, 1974. p. 147–162.
2. SZEVBÓ, I. P.: Sztruktura szvjaznogo tekszta i avtomatizacija referirovanija. Moskva, Nauka, 1969. 135 p.
3. LEONOV, V. P.: O metodah avtomaticheskogo referirovanija (SZSA, 1958–1974) = Naucsno-Tehnicoszkaja Informacija, 2. sor. 6. sz. 1975. p. 16–20.
4. WYLLYS, R. E.: Extracting and abstracting by computer. Borko, H.: Automated language processing. New York–London–Sydney, Wiley, 1967. p. 127–129.
5. ZSDANOVA, G. SZ.: O naucsüh osnovah metodiki referirovanija = Naucsno-Tehnicoszkaja Informacija, 1967. 2. sz. p. 12–16.
6. HELMUTH, N. A.: The use of extracts in information services = Journal of the American Society for Information Science, 22. köt. 6. sz. 1971. p. 382–389.
7. ASHWORTH, W.: Producing and using abstracts. Special libraries and information work. 4. kiad. London, 1975. p. 124–152.
8. WEIL, B. H. – ZAREMBER, I. – OWEN, H.: Technical abstracting fundamentals. 1. Introduction = Journal of Chemical Documentation, 3. köt. 1. sz. 1963. p. 86–89.
9. SREJDER, Ju. A.: Ob odnoj modeli szemanticheskij teorii informacii. Problemu kibernetiki. 13. sz. Moskva, Nauka, 1965. p. 233–240.
10. SREJDER, Ju. A.: Tezaurusz v informatike i teoreticheskij szemantika = Naucsno-Tehnicoszkaja Informacija, 2. sor. 3. sz. 1971. p. 21–24.
11. URSZUL, A. D.: Prakticheskij aszpekt naucsnoj informacii (Metodologicheskije problemü) = Naucsno-Tehnicoszkaja Informacija, 2. sor. 8. sz. 1968. p. 3–8.
12. SZTJAZSKIN, N. I. – PEVZNER, B. R.: Ob odnom metode szpceferirovanija. Problemu logiki. Moskva, 1963. p. 25–37.
13. SZKOROHOD'KO, É. F.: Avtomaticheskaja kompreszija tekszta na osnove analiza ego szemanticheskij sztrukturu. Problemu prikladnoj lingvisztiki. Teziszü mezsvozovszkoj konferencii. 2. rész. Moskva, 1969. p. 290–294.
14. SZKOROHOD'KO, É. F.: Adaptive method of automatic abstracting and indexing. Information processing 71. North-Holland, 1972. p. 1179–1182.
15. SZKOROHOD'KO, É. F.: Analiz teksztov sz primeneniem recsevüh szemanticheskijh szetej. Matematicszkaja lingvisztika. 1. köt. Kiev, Kievskij universzitet, 1973. p. 97–103.
16. SZKOROHOD'KO, É. F.: Szemanticheskije szvjazi v lekszike i teksztah = Voproszü Informacionnoj Teorii i Praktiki, 1974. 23. sz.
17. SZKOROHOD'KO, É. F.: Lingviszticheskije problemü obrabotki teksztov v avtomatizirovannüh informacionno-poiskovüh szisztemah = Voproszü Informacionnoj Teorii i Praktiki, 1974. 25. sz. p. 5–120.

18. BONDARENKO, G. V.: Ob ispol'zovanii formal'nüh metodov analiza sztrukturü celogo teksta dlja avtomaticheskogo indeksirovanija i referirovanija = Naucsno-Tehnicoszkaja Informacija, 2. sor. 5. sz. 1976. p. 26–37.
19. SZMIRNOV, G. A.: Nekotorüe oszobennosti szmüszlovoj organizacii tehnologicseszkih tekstv (k probleme kompreszszii teksta) = Naucsno-Tehnicoszkaja Informacija, 2. sor. 4. sz. 1973. p. 18–21.
20. LUHN, H. P.: The automatic creation of literature abstracts = IBM Journal of Research and Development, 2. köt. 2. sz. 1958. p. 159–165.
21. SCHWEISHEIMER, W.: Auto-abstracting mit einer IBM 704 = Elektronische Datenverarbeitung, 1964. 4. sz.
22. PURTO, V. A.: Ob avtomaticheskijom referirovanii na oszнове sztatisticheskogo analiza teksta. Dokladü na konferencii po obrabotke informacii, masinnomu perevodu i avtomaticheskijomu csteniju teksta. 9. sz. Moszkva, 1961.
23. EDMUNDSON, H. P. – WYLLIS, R. E.: Automatic abstracting and indexing – survey and recommendations = Communications of the ACM, 4. köt. 5. sz. 1961. p. 226–234.
24. DOYLE, L. B.: Indexing and abstracting by association = American Documentation, 13. köt. 4. sz. 1962. p. 378–390.
25. AGRAEV, V. A. – BORODIN, V. V. – GLEBSZKIJ, Ju. V.: O nekotörüh metodah avtomaticheskogo referirovanija. Ucsennüe Zapiszki Gor'kovszkogo Universziteta, 63. sz. 1963. p. 73–86.
26. CLIMENSON, W. D. – HARDWICK, N. H. – JACOBSON, S. N.: Automatic syntax analysis in machine indexing and abstracting = American Documentation, 12. köt. 3. sz. 1961. p. 178–183.
27. PLATH, W.: Automatic sentence diagramming. 1961 International conference on machine translation of languages and applied language analysis. 13. szimp. 1. köt. 1962. p. 175–193.
28. EARL, L. L.: Experiments in automatic extracting and indexing = Information Storage and Retrieval, 6. köt. 4. sz. 1970. p. 313–334.
29. NISTOR, E. – ROMAN, E.: Transformations in automatic abstracting = Cahiers de Linguistique Theorique et Appliquée, 7. köt. 1970. p. 143–159.
30. NISTOR, E. – ROMAN, E.: Constructing automatical abstracts from kernel-sentences = Cahiers de Linguistique Theorique et Appliquée, 8. köt. 1971. p. 249–256.
31. POLUSKIN, V. A.: Nekotorüe voproszü informacionnogo analiza dokumentov = Naucsno-Tehnicoszkaja Informacija, 1966. 8. sz. p. 29–32.
32. LABRENT'eva, G. A.: Ob odnom szposzobe csaszticsno formlaizovannogo indeksirovanija szzsatüh tekstv. Trudü tretej Vszeszojuznoj konferencii po informacionno-poiskovüm szisztemam i avtomatizirovannoj obrabotke naucsno-tehnicoszkij informacii. Moszkva, 1967. p. 373–378.
33. OTRADINSZKIJ, V. V.: Metod szzsatija teksta dokumenta na oszнове aszpektnüh deszkriptorov. Dokumentalishtika–69. Materialü pervogo Vszeszojuznogo szimpoziuma po dokumentalishtike. Palanga, 8–12 szentjabrja 1969. Vil'njusz, 1970. 1. köt. p. 49–57.
34. VASZIL'eva, I. I.: K voproszu o vozmozsnoszti primenija informacionnogo jazüka „sztrandartnüh fraz” dlja formalizovannogo referirovanija. Voproszü szoversensztvovanija szisztemü informacionnüh izdanij. Moszkva, 1973. p. 64–72.
35. EDMUNDSON, H. P.: New methods in automatic extracting = Journal of the ACM, 16. köt. 2. sz. 1969. p. 264–285.
36. RUSH, J. E. – SALVADOR, R. – ZAMORA, A.: Automatic abstracting and indexing. 2. Production of indicative abstracts by application of contextual interference and syntactic coherence criteria = Journal of the ASIS, 22. köt. 4. sz. 1971. p. 260–274.
37. MATHIS, B. A. – RUSH, J. E. – YOUNG, C. E.: Improvement of automatic abstracts by the use of structural analysis = Journal of the ASIS, 24. köt. 2. sz. 1973. p. 101–109.

/KARASZEV, SZ. A.: Referirovanie naucsno-tehnicoszkij literaturü: élementü teorii. = Naucsno-Tehnicoszkaja Informacija, 2. sor. 10. sz. 1978. p. 1–4./

(Viszocsekné Péteri Éva)



Tanulmány az ipari információról

A tanulmányok célja

Eddig Franciaországban a tudományos és műszaki információval foglalkozó tanulmányok mindig csupán egészen szük közönség számára íródtak. Ezen akart változtatni a Marseille-Luminy-ben működő ESCAE (Ecole supérieure de commerce et d'administration des entreprises = Vállalatok Kereskedelmi és Ügyviteli Főiskolája) munkatársainak egy csoportja, amikor az INSA (Institut national des sciences appliquées = Alkalmazott Tudományok Országos Intézete) kérésére tanulmányorozatot készített.

Az információ és a vállalat

Az információt éppügy termelési tényezőnek kell tekintenünk, mint bármelyik klasszikus értelemben vett termelési tényezőt. Egy vállalat működése és virágzása lehetetlen anélkül, hogy kellő időben, kellő helyen és kellő formában hozzájusson a számára szükséges információhoz. Ehhez természetesen információszakemberek kellenek. Az ESCAE már 1973 óta ilyen jellegű képzésben részesíti az elsőséves hallgatókat, 1973 áprilisában pedig felmérte a Marseille-i Kereskedelmi és Ipari Kamarához tartozó vállalatok információs és dokumentációs szükségleteit.