

(5) Kvalitatív szempontból, a kivonatok szerzői stílus-sajátosságai folytán, az eredmények nem mindig egyértelműek. Egyre nő a kivonatok tartalmi és formai egységesítésének fontossága és szükségessége!

A fenti megközelítéssel olyan információábrázolásra törekednek, amely végső soron mondatfunkció – szignifikáns szókapcsolat-párok formájában valósulna meg. A kulcsszavazást az ilyen információábrázolás speciális esetének lehetne tekinteni, amikor csak a „téma” van megadva. Ami egy ilyen alapú információkereső rendszer felépítését illeti, további kutatásokat igényel (a) a dokumentumok asszociatív tartalmi feltárása – tehát nem egyszerűen a rokon tartalmú dokumentumokból való információfeltárás, hanem egyes dokumentumokban levő tartalmi asszociációt reprezentáló mondatok egymásnak megfeleltetésével, (b) az automatikus tezauszépítés és (c) az automatikus kivonatolás kidolgozása.

/MAEDA, T. – MOMOUCHI, Y. – SAWAMURA, H.:
An automatic method for extracting significant phrases in scientific or technical documents. = Information Processing and Management, 16. köt. 3. sz. 1980. p. 119–127./

(Szöllősy Éva)



Fogalmi kapcsolatok szemléltetése tezauszokban

Az AIDOS–OS/ES programrendszer, valamint a megfelelő normatív dokumentumok és irányelvek alapján, továbbá a már meglévő tezauszok alkalmazása során nyert tapasztalatok felhasználásával állították össze az NDK-ban „Kisfeszültségű megszakítástechnika” tezauszát. A szerkesztési eljárásokat és módszereket részletesen leírták és szabályokban rögzítették. A tezausz szisztematikus részénél olyan megoldást alkalmaztak, amely lehetővé teszi valamennyi – az indexelés és visszakeresés szempontjából lényeges – fogalmi kapcsolat megkülönböztetését és áttekinthető ábrázolását.

1. Elméleti alapok

Előjáróban a következő alapelveket szögezték le:

A deskriptorok és nem-deszkriptorok – szavak és szócsoportok alakjában – fogalmakat fejeznek ki. Ugyanez vonatkozik a deskriptorok és a több szóból álló kifejezések közötti kapcsolatokra is. (Ez utóbbiak nemcsak mondatok, hanem egyéb értelmes szerkezetek is lehetnek.)

A fogalmak (szemantikai megközelítésben) szavakat és szócsoportokat jelentenek, és (az ismeretelmélet tükrözési szabályainak megfelelően) az osztályok, az egyének gondolati visszatükröződései.

A fogalmi kapcsolatok így, egyrészt a szavak és szócsoportok jelentései közötti kapcsolatok; másrészt (ismeretelméleti szempontból) az osztályok gondolati visszatükröződései közötti kapcsolatok. A tezauszban ez az egyes deskriptorok közötti és az egyes nem-deszkriptorok közötti, valamint a deskriptorok és nem-deszkriptorok közötti kapcsolatokban jut kifejezésre.

Az eddigi tezauszoknál alkalmazott eljárás módokkal ellentétben, az új tezauszban kizárólag az úgynevezett *egyértelmű ekvivalencia elve érvényesül.* Ennek előfeltétele azonban, hogy az AIDOS–OS/ES programrendszer inverz adatai, a deskriptorok hivatkozási száma mellett, a nem-deszkriptorokat is tartalmazzák. A prekombinált, csoportosított megnevezések mellé pedig felvették a tezauszba mindazokat a szerkezeteket is, amelyek – megadott kritériumok szerint – a deskriptorok és nem-deszkriptorok kiválasztásához és meghatározásához szükségesek. Ezáltal, egyrészt a kifejezhetőség vált pontosabbá: a tezauszban nem szereplő fogalmakat deskriptor-kapcsolásokkal (posztkombinált megnevezésekkel) lehet kifejezni; más részről így biztosították, hogy *a tezausz a referátumok – későbbre tervezett – gépi indexelésénél is alkalmazható legyen.*

2. A fogalmi kapcsolatok szemléltetése és a gépi adatfeldolgozás

A tezausz szerkesztése közben arra törekedtek, hogy a fogalmi kapcsolatokat és, ha ezek anyagi jelenséget tükröznek, akkor a fizikai jelenségek, a gyártmányok, a gyártási eljárások stb. közötti kapcsolatokat, a tezausz felépítése áttekinthetően szemléltesse.

A kapcsolatok ábrázolásának *gépi adatfeldolgozással történő megvalósítása esetén három lehetőség kínálkozik:*

- utaló jelzetek használata;
- a deskriptoroknak és nem-deszkriptoroknak cikkekbe vagy szakcsoportokba való rendezése;
- a deskriptorok és nem-deszkriptorok – különböző részekre tagolt – hierarchikus beosztása és jobbra történő elcsúsztatása.

A betürendes tezauszok esetében kizárólag az elsőként említett megoldást alkalmazzák. Ebből következik, hogy a betürendes tezauszban nem valósítható meg optimális hatásfokkal valamennyi fogalmi kapcsolat átfogó kifejezése; különösen akkor nem, ha sok hierarchikus kapcsolatból tevődnek össze.

A cél olyan tezausz-forma létrehozása volt, amely lehetővé teszi *valamennyi leírandó fogalmi kapcsolat jól áttekinthető ábrázolását viszonylag kevés leírási elem*

alkalmazásával. A teaurusz főrésze szisztematikus felépítésű. A kötelező bevezetést követően valamennyi fogalmi kapcsolat teljes leírását közli és permutált betűrendes mutatót is tartalmaz. Az utaló jelzeteket az ekvivalencia és az egylépcsős hierarchia követelményeinek megfelelően alkalmazzák. A többlépcsős hierarchia kifejezésére elősegíti a különböző asszociatív fogalmi kapcsolatokat megkülönböztetésére kétszámjegyű utaló jelzeteket vezetnek be.

A szakcsoportokat úgy határozta meg és nevezte el, hogy a jelzeteknek azoknak a fogalmaknak a nem-fogalmát jelölik, amelyeket a vonatkozó szakcsoport első hierarchiaszintjének deszkriptorai fejeznek ki. A deszkriptorokhoz tartozó deszkriptorokat egy-egy hozzájuk csatlakozó számmal jelölik meg. Ez megadja azt a szakcsoportot és hierarchiaszintet, amelyben a vonatkozó deszkriptor mint címdeszkriptor szerepel.

3. Az új formájú teaurusz jellemzése

Minden teaurusznak rendelkeznie kell egy *szisztematikus* és egy *betűrendes* résszel, ahol a főrész minden egyes deszkriptor összes adatát tartalmazza. A főrész lehet betűrendes, szisztematikus vagy a kettő kombinációja. Míg a betűrendes résznek deszkriptor- és nem-deszkriptor-cikkekből kell állnia, addig a szisztematikus (hierarchikus) részbe kizárólag a deszkriptorok vehetők fel.

Az újonnan kialakított teaurusz főrésze a *szisztematikus és betűrendes forma keveréke*. (Az 1. ábra mintát közöl a teaurusz szisztematikus részéből; a 2. ábra pedig a hozzá tartozó permutált betűrendes mutatóval ismerteti meg.) Az új formájú teauruszban olyan struktúra létrehozása volt szükséges, amely a benne megnevezett jelenségek és fogalmak objektív kapcsolatait híven tükrözi. Ennek megvalósításához az alábbi módszerek kínálkoznak:

a teaurusz *felépítése induktív úton történjék*: vagyis a hierarchiát a legspecifikusabb fogalmak nem-faj kapcsolataiból kiindulva, „alulról felfelé”, a legspecifikusabb fajfogalomtól a legáltalánosabb nem-fogalom irányába haladva, alakítsák ki;

a leírandó jelenségek és fogalmak közötti valós kapcsolatok meghatározásánál messzemenően alkalmazzák az ontológiai és onomasziológiai vizsgálati módszereket;

kerüljék az „idealizáló” fogalomalkotásokat és osztályozásokat, amelyekkel a szakmai kommunikációban nem lehet mire menni.

THES SYST.						STAND VOM 06/11/78	SEITE 00001
0	1	2	3	4	5	9100 - 9110	
9100 ELEKTRISCHE SCHUTZTECHNIK, RELAIS- UND SCHALTGERÄTE-TECHNIK							
9110 SCHADENSURSACHEN? STÖRUNGEN							
D BERUEHRUNGSSPANNUNG							
O SPANNUNG 3600-2							
O ELEKTRISCHE GROESSE 3600-1							
AZ BERUEHRUNGSSCHUTZ 9210-2							
AR FEHLERSPANNUNG 9110-2							
D BLITZ							
AZ BLITZSCHUTZ 9210-2							
AR AEUSSERE UEBERSPANNUNG 9110-2							
AR DURCHSCHLAG 9110-2							
AR GASENTLADUNG 3400-1							
D ELEKTRISCHE STÖRUNG							
O STÖRUNG 3760-1							
D DURCHSCHLAG							
S -ELEKTRISCHER DURCHSCHLAG							
AZ DURCHSCHLAGSFESTIGKEIT 7350-2							
AZ UEBERSPANNUNGSSCHUTZ 9210-2							
AR GASENTLADUNG 3400-1							
AR UEBERSCHLAG 9110-2							
D FEHLAUFLÖSUNG							
O AUSLÖSUNG 4330-1							
D FEHLERSPANNUNG							
O SPANNUNG 3600-2							
O O ELEKTRISCHE GROESSE 3600-1							
AZ FU-SCHUTZ 9210-3							
AR BERUEHRUNGSSPANNUNG 9110-1							

1. ábra Részlet a teaurusz szisztematikus részéből

PERMUTIERTES REGISTER						STAND VOM 06/11/78	SEITE 00001
AEUSSERE UEBERSPANNUNG 9110							
-ANFANGSKURZSCHLWECHSELSTR 9110							
B IK 2-STRICH 9110							
-ATMOSPHER UEBERSPANNUNG 9110							
B AEUSSERE UEBERSPANNUNG 9110							
BERUEHRUNGSSPANNUNG 9110							
-BETRIEBSFREQU UEBERSPANNUNG 9110							
B INNERE UEBERSPANNUNG 9110							
BLITZ 9110							
DAUERKURZSCHLUSSSTROM 9110							
DOPPELERSCHLUSS 9110							
DURCHSCHLAG 9110							
-DURCHSCHLAG ELEKTRISCHER 9110							
B DURCHSCHLAG 9110							
ELEKTRISCHE STÖRUNG 9110							
-ELEKTRISCHER DURCHSCHLAG 9110							
B DURCHSCHLAG 9110							
-ELEKTRISCHER UEBERSCHLAG 9110							
B UEBERSCHLAG 9110							
ERDSCHLUSS 9110							
-ERDSCHLUSS DOPPELTER 9110							
B DOPPELERSCHLUSS 9110							
-ERDSCHLUSS ROTOR 9110							
B ROTORERDSCHLUSS 9110							
-ERDSCHLUSS STATOR 9110							
B STATORERDSCHLUSS 9110							
ERREGUNGSFAUSFALL 9110							
EXPLOSION 9110							
FEHLAUFLÖSUNG 9110							
FEHLERSPANNUNG 9110							
FEHLERSTROM 9110							
FEHLERHALTUNG 9110							
FREQUENZABWEICHUNG 9110							

2. ábra Részlet a permutált betűrendes mutatóból

4. Az új módszer megvalósításának lehetőségei

Az új formájú teaurusz gépi előállításához még nincsenek meg a szükséges feltételek az NDK-ban. Ehhez az AIDOS-OS/ES programrendszer első lépcsője teljesítőképeségének további bővítése szükséges. A szakemberek szerint a jelenlegi művelési-programmal a teaurusz szisztematikus részében a deszkriptor-cikkek ki nyomtatása lehetséges lenne, azonban a többi követelmény kielégítése csak a rendszer további fejlesztése révén oldható meg.

/GROHMAN, W. - PROTZ, A.: Zur Darstellung von Begriffsbeziehungen in Thesauren = Informatik, 26. köt. 3. sz. 1979. p. 38-42./

(Takáchné Tóth Mária)

