

- A könyvtári szolgáltatások továbbfejlesztése egész Európára kiterjedően, amelyek megfelelően szakosítva hozzáférhetővé teszik az ismeretek tárházát, elősegítve a polgárok műszaki, gazdasági, kulturális és szociális fejlődését.
- Munkapiaci adatszolgáltató hálózat megszervezése, amely európai méretekben tartja nyilván a munkaerő-szükségletet és -kínálatot, csökkenti a munkanélküliséget.
- Új hozzáférési módszerek kialakítása, amelyek például egyes e célra kialakított helyeken, (hitel)kártyák felhasználásával nyilvánosan teszik lehetővé egyes elektronikus szolgáltatások elérését.

Az alapszolgáltatások területén

- Transzeurópai kapcsolt távközlési hálózat kialakítása, amely lehetővé teszi az e-mail és a fájlátvitel széles körű lebonyolítását, valamint az adatbankok és videoszolgáltatások online elérését, megfelelő személyazonosítással, adatvédelemmel és információbiztonsággal, akár nyilvános állomásokról is.
- Az alapszolgáltatások progresszív továbbfejlesztése *multimédia-szolgáltatásokká*, amelyek felölelik a multimédia-levelezés, a nagy sebességű fájlátvitel, az igény szerinti videoműsor-szolgáltatások stb. lebonyolítását. Támogatni kell egyes szolgáltatások, például a gépi fordítás, a beszéd- és grafikafelismerés multimédia-szolgáltatásokba való integrálását.

- Általános és biztonságos *elektronikus személyazonosító* rendszer kialakítása válik szükségessé az elektronikus úton elérhető szolgáltatások megsokasodásával.

Hálózati területen

- *Transzeurópai ISDN-hálózat* kiépítése, amely képes lesz kapcsolt digitális rendszerben az új, szélessávú szolgáltatások lebonyolítására is. A jelenleg szigetszerűen kialakuló ISDN-hálózatok európai méretű integrálása.
- Az *ATM és más szélessávú rendszerek* kereskedelmi alapú meghonosítása, amelyek majd lehetővé teszik a különböző célú fix, mobil és műholdas szolgáltatások egy közös, szélessávú, nagy sebességű hálózatba való integrálását.

/Trans-European telecommunications networks. = CORDIS Focus, 40. sz. 1995. június 16. p. 19-20./

(Reich György)

Bevezetés a Szabványos Általános Feltáró Nyelvbe

Háttér

Az *SGML (Standard Generalized Markup Language)* a különböző szövegkezelési eljárásokból született meg, az egymással nem kompatibilis szövegformátumokból fakadó elektronikus információkezelési problémák megoldása céljából. Ahhoz, hogy az információátadás a heterogén hardver- és szoftverkönyezeteken belül és a különféle szervezetek között világszinten is sikeresen megvalósítható legyen, a dokumentumszerkezeteket egységes, szabványos módon kell meghatározni. A túlságosan egyszerű szerkezet nem felel meg. Egyrészt korlátozza a szerző, illetve a dokumentumalkotó munkáját, másrészt pedig egy-egy dokumentumot nem minden olvasó keres ugyanabban a formátumban. A szabványos dokumentumszerkezet létrehozásánál mind a szerzők, mind az olvasók igényeit figyelembe kell venni. A struktúrának közvetítenie kell a dokumentum készítőjének elképzeléseit az olvasónak, illetve az olvasó számára lehetővé kell tennie a különféle információkeresési módszerek használatát, például a tartalom vagy a szerkezet alapján történő keresést.

A Nemzetközi Szabványügyi Szervezet (International Standard Organization = ISO) Szöveg és Iroda Rendszerek Albizottsága (Subcommittee of Text and Office Systems) munkája eredményeként 1986-ban született meg az ISO 8879, a Szabványos Általános Feltáró Nyelv. Az ISO 8879 szabvány meghatározza, hogy egy dokumentumot hogyan oszthatunk fel két részre: egyrészt arra, amely a szöveget (vagy más információs formát) tartalmazza, másrészt arra a részre, amely megadja a dokumentum struktúráját anélkül, hogy bármilyen szövegkezelő eljárással kapcsolatban lenne. Az SGML függetleníti az információt a közzététel és a továbbítás eszköztől, lehetővé teszi a dokumentumszerkezet kódolását és gépről gépre történő ellenőrzött átvitelét. Az információ kezelése, aktualizálása és fenntartása egy semleges SGML adatbázis keretén belül történik, szállítható papírhordozón, CD-ROM-on vagy hálózaton. Röviden a következő fő jellemzőkkel határozható meg:

Szabványos

- az SGML nemzetközi szabvány (ISO 8879).

Általános

- semleges formában tárolja a dokumentumokat;
- alkalmazáspecifikus azonosítók helyett generikus azonosítót használ;
- független a rendszerektől, eszközöktől, nyelvektől és alkalmazásoktól.

Feltáró

- feltáró jelöléseket kapcsol a dokumentumhoz, amelyek megmutatják annak szerkezetét;
- megadja a feltáró jelölés módjának szabályait;
- a feltáró jelölés nincs szabványosítva.

Nyelv

- metanyelv – szintaxisa van, de szemantikája nincs;
- reprezentációs nyelv;
- bármely dokumentum leírására használható.

Az SGML céljai

Egy dokumentum létrehozása három fő komponensből áll: a szerző üzenetéből, a közzététel és a továbbítás eszközéből, valamint az olvasó/felhasználó nézetéből, véleményéből. Az SGML létrehozásakor több kritériumot tartottak szem előtt. A legfontosabb a függetlenség feltétele. A dokumentumot kezelő és szerkezetét jelölő nyelvnek *rendszer-, eszköz-, nyelv- és alkalmazássemlegesnek* kell lennie. A nyelvnek bármely dokumentum-előállító és/vagy -eladó rendszeréhez illeszkednie kell; a dokumentumok nem kötődhetnek egy meghatározott hardvertípushoz, képernyőhöz vagy nyomtatóhoz; a szabvány legyen jól használható különböző nyelveken; végül az SGML-nek a nagyon egyszerű és nagyon bonyolult dokumentumok szerkezetének leírására is alkalmasnak kell lennie.

Az SGML-dokumentumok felépítése

Az SGML-dokumentumok azonosítható *elemekből* állnak. Minden elem szöveget vagy további *alelemek* tartalmaz. Többszintű felépítés esetén a legalsó elem mindig szöveg vagy más formátumú információ. Az SGML-dokumentum SGML-deklarációból, egy dokumentum típusú definícióból (DTD) és a dokumentumtartalomtól áll.

Az *SGML-deklaráció* meghatározza a dokumentumleírás karakterkészletét, a konkrét szintaxist, az opcionális jellemzőket és a kapacitás feltételeit. A *DTD* azokat a szabályokat tartalmazza, amelyek meghatározzák a dokumentumosztályok struktúráját. Ez mindazokat a szabályokat jelenti, amelyek az adott dokumentumtípusra vonatkoznak. Tartalmazza az elemek nevét, attribútumait, és azoknak az elemekhez való viszonyát. Egy dokumentum típusú definíció lehet nagyon rövid és egyszerű, de méretét és összetettségének mértékét semmi nem korlátozza. Az SGML-elemző programok (parser) azt tesztelik, hogy a dokumen-

tum megfelel-e a DTD-ben leírt szabályoknak. A *dokumentumtartalom* értelemszerűen magát a dokumentumot jelöli.

Az SGML az egyes elemek jelölésére *címkéket* használ. A címkét egy-egy elem elején (címké) és végén (/címké) is feltüntetik. A címké tartalmaz egy generikus azonosítót, amely azonosítja az elem típusát. A 1. ábra egy könyvelem generikus azonosítóját mutatja.

```

<könyv>
  <fejezet>
    <cím>...adat...</cím>
    <bekezdés>...adat...</bekezdés>
  </fejezet>
  <fejezet>
    <cím>...adat...</cím>
    <bekezdés>...adat...</bekezdés>
    <bekezdés>...adat...</bekezdés>
  </fejezet>
</könyv>

```

1. ábra Egy könyvelem generikus azonosítói

Az SGML lehetővé teszi, hogy egy jelölő címke a jelölésen túl többtinformációt, a szövegelemre vonatkozó valamilyen tulajdonságot, jelzőt tartalmazzon, például, hogy az adott dokumentum egy tervezet.

Az SGML-ben arra is megvan az eszköz, hogy bármilyen adattípus szimbolikus nevét definiáljuk, ezt az eszközt egységnek nevezik (*entity*). Az egységek lehetnek hosszú szövegrészeket helyettesítő jelölések, a klaviatúrán nem található karakterek jelölői, vagy fődokumentumba kívülről behozott és beültetett dokumentumok jelölői.

A dokumentumstruktúrálás szabályait tartalmazó DTD kétféle módon hozható létre: az aktuális dokumentumból kiindulva alulról felfelé történő építkezéssel, vagy már meglévő, általános szerkezetet tartalmazó DTD-nek a kívánalmaknak megfelelő finomításával.

Az első esetben három fő lépést kell követni. Először össze kell gyűjteni az összes dokumentumot, amit egységesen akarunk kezelni. Utána azonosítani kell a mindegyikükre jellemző közös szerkezeti jellemzőket, végül meg kell írni a definíciót, magát a DTD-t, amely minden azonosított elemet és szerkezetet tartalmaz, és amelynek a körbe tartozó minden dokumentum megfelel. Ezt a megközelítést gyakorlatias szempontúnak és tartalomorientáltnak tartják, de néhány hiányosságára is rámutattak már. Például: elsősorban papíralapú dokumentumokra használható; nem biztosítja a dokumentumok hierarchikus jellegéből fakadó tartalomjegyzék felépítését, és az ez alapján vagy valamilyen indexrendszer alapján történő keresést;

figyelmén kívül hagyja a felhasználói igények alakulásának azt a tendenciáját, hogy minél kevesebb, és minél szélesebb körű dokumentumalmazra alkalmazható DTD készüljön.

Az általános szerkezetet hordozó, és a fentről lefelé irányuló finomítással készülő DTD-k előnye, hogy a különböző DTD-k között világos és kifejezett kapcsolat van, és hogy egy DTD egyszerre több DTD-nek a finomításával is létrehozható. A fentről lefelé építkező struktúra esetén a meghatározó szabályrendszerek világos, jól meghatározott kapcsolatban vannak egymással, ún. DTD-hálózatot alkotnak.

A SGML-dokumentumok keresésére több különböző módon van lehetőségünk. A keresés lebonyolítható különböző keresési szerkezetek mentén (pl. „menj a következő címre”, vagy „keresd azokat a paragrafusokat, amelyek az alábbi tulajdonságot tartalmazzák”, vagy „keress meg minden, ehhez a grafikához tartozó referenciát” stb.), a tartalomjegyzék böngészésével, az adott indexben szereplő kulcsszavak megadásával, a kereszterenciák követésével, vagy akár egyes szavaknak a dokumentumszövegben történő keresésével.

A hagyományos dokumentumoktól az integrált dokumentumbázisokig

Az SGML alkalmazása hosszú távra szóló stratégiai döntést igényel, ezért mielőtt dokumentumkezelési feladataink esetén e nyelv mellett döntenénk, mérlegeljük az alábbiakat:

- Valóban szabványra van-e szüksége a szervezet dokumentumkezelésének?
- Kiválasztása esetén milyen szerepet játszana az SGML? Különböző szervezeteknél különböző változatok lehetségesek. Az SGML a különféle információkezelési feladatokhoz közös alapot biztosíthat.
- Milyen típusú dokumentumokat kell szabványosítani?
- Kik a felhasználók, milyen igényeik vannak, melyek az információs források, a tároló és terjesztő eszközök?

Az SGML a legszélesebb értelemben vett, a hagyományos kiadói termékektől a multimédiákig terjedő dokumentumtípusok (újságok, könyvek, szótárak, enciklopédiák, kézikönyvek, hivatali dokumentumok, katalógusok) kezelésére alkalmas.

Az SGML-szabvány megfelel mind a hagyományos dokumentumok, mind az integrált dokumentumbázisok információkezelésére. A dokumentumbázis felosztható például az egyes részek aktualizálásának felelőssége alapján. Az integrált dokumentumbázisból származó információk megjeleníthetők és továbbíthatók papírhordozókon, képernyőn, mikrofilmen vagy mikrofilmlapon, optikai vagy mágneses hordozón, vagy akár hálózaton is.

HyTime, az SGML kibővítése

Az információkezelés jövőjét a hipermedia dokumentumok kezelése jelenti, amely rendszerben igen különböző információk állnak egymással kapcsolatban. A ma létező hipermedia dokumentumok meglehetősen szorosan kötődnek az egyes platformokhoz és a megjelölt technikákhoz. A hipermedia dokumentumok szerkezetének szabványos leírása lehetővé teszi olyan alkalmazások létrehozását is, amelyek biztosítják az egyik alkalmazási platformról egy másikra történő átvitelt, az alkalmazások rendszerfüggetlenné válását.

A *HyTime (Hypermedia/Time-based Structuring Language)* az SGML továbbfejlesztése, amely azt definiálja, hogyan lehet a dokumentumszerkezet jelöléseit és a DTD-eket hiperkapcsolatú, időalapú multimédia dokumentumok szerkezetének leírására használni. A HyTime az SGML használatának nemzetközi szabványát adja meg, amely kifejezi a dokumentumokon belüli és közötti belső kapcsolatokat, bármely típusú információs tárgyak között lévő hiperkapcsolatokat, valamint a tárgyak időbeli és térbeli ütemezését.

Az SGML alkalmazása, SGML-termékek

Az SGML vagy a hozzá kapcsolódó projektek végrehajtásában elsősorban Kanada, Franciaország, Németország, Japán, Svájc, Hollandia, az Egyesült Királyság és az Egyesült Államok járnak az élen. Az SGML-t használja pl. az amerikai Védelmi Minisztérium beszerzési és logisztikai rendszere. A technikai-műszaki dokumentumokra vonatkozó SGML-használat a repülőgépiparban, a hajógyártásban, az olaj- és gáziparban, valamint a távközlési és elektronikai iparban kezd terjedni. A kiadók közül az Amerikai Kiadók Szövetsége alkalmazta először. Elektronikus kézirat projektjében három DTD-t készítettek: egyet könyvek, egyet cikkek, és egyet sorozatok számára. Az Európai Szabványügyi Hivatal szabványdokumentumokra, a Nemzetközi Szabványügyi Szervezet szabványokra, míg az Európai Unió Hivatalos Publikációinak Irodája többnyelvű folyóiratokra és kiadványokra alkalmazza. Megemlíthetjük még az USA Energiaügyi Minisztériumát, vagy az angol királynő hivatalos kiadóját, ahol szintén az SGML alapján történik a dokumentumkezelés. A szabványt mind általánosabban használják még a tudományos és kutatói szférában is.

Az SGML-termékek három fő csoportba oszthatók: elemzők (OmniMark, Balise, DynaTag), szerkesztő szoftvercsomagok (pl. SoftQuad Author/Editor) és a kiadói szoftvercsomagok (pl. Interleaf/World View, DynaText).

Az SGML olyan nyelv tehát, amely elválasztja egymástól a dokumentum szerkezetét és tartalmát. Ily módon az alkalmazástól független és újra felhasználható dokumentumokat hozunk létre. Az SGML használatának eldöntése mindig szervezet-specifikus kérdés. A döntés előtt mérlegelni és elemezni kell többek között azt, hogy más szervezetek hogyan használják, melyek az ottani alkalmazások tanulságai, milyen technikai megoldásokkal éltek, hogyan fogjunk hozzá az SGML bevezetéséhez, milyen költségek merülnek fel, milyen jellegű és mértékű képzésre van szükség a szervezeten belül, mit jelent az SGML-alkalmazás fenntartása és működtetése, milyen új technikai megoldások érhetők el az SGML segítségével.

- Összegezve:
- az SGML nemzetközi szabvány,
- független a rendszertől, az eszköztől és a nyelvtől,
- bármilyen információs formát tartalmazhat,
- az SGML-ben mindent szabályok határoznak meg,
- az SGML reprezentációs nyelv, bármilyen típusú dokumentum megjelenítésére alkalmas.

/HEIMBÜRGER, A.: *Introduction to Standard Generalized Markup Language (SGML)*. = *Microcomputers for Information Management*, 11. köt. 4. sz. 1994. p. 239–260./

(Kardos Krisztina)

Promenade: kertészeti képek keresése és lehívása a hálózaton

A *Promenade* a Pittsburgi Egyetem és a National Agricultural Library (NAL) közös projektje. Mivel a NAL képi forrásai – más könyvtárakhoz hasonlóan – hagyományos eszközökkel nehezen kereshetők vissza, és csak helyben használhatók, megvizsgálták, hogyan segíthetnek ezen korszerű technika alkalmazásával.

A projekt egyben kísérlet annak eldöntésére, hogy hogyan lehet képekből tematikusan visszakereshető adatbázist készíteni.

Két, egymástól erősen különböző művet választottak. Elsőként a *Curtis Botanical Magazine* első 26 évfolyamából (1787–1807) választottak 1059 botanikai nyomatot. A Curtis-kiadvány kézzel színezett nyomatokat tartalmaz egyrészt olyan növényekről, amelyek Európából és Észak-Amerika keleti részéből származnak, másrészt olyanokról, amelyeket utazók gyűjtöttek, mint pl. *Pitcairn* és *Bligh*. Minden nyomatot szöveges leírás kíséri, a nomenklatúrával, a jellemzőkkel és a termőhely leírásával.

A másik adatbázis még épülőben van, és *Randy Heatly*, a Michigani Állami Egyetem kutatójának diái tartalmazza növényi kártevőkről, és az azonosításukra alkalmas növénykárosodásokról.

A Curtis-gyűjteményről 35 mm-es színes diákat készítettek, amelyeket külső cégnél digitalizáltak Kodak Photo-CD-re való felvitellel.

A képeket kísérő leírások szokatlan betűtípusa miatt optikai karakterfelismerő programmal nem lehetett ezeket bevinni. A billentyűzetről való adatbevitelt szintén elvetették, mert igen munkaigényes, utólagos szakszerkesztést igényelt volna,

továbbá XVIII–XIX. századi lévén, nem felel meg a korszerű botanikai ismereteknek. A kártevők adatbázisát ugyancsak Photo-CD-re vitték, és a leírások ASCII-szöveggé alakították.

A Curtis adatbázisnál a képeket a szövegtől külön kezelték, mivel a felhasználók gyakran csak a képeket nézik meg. Az indexelést a Curtis adatbázisnál rögzített hosszúságú mezők segítségével, a kártevők adatbázisában ugyanilyen módon és teljes szöveg visszakeresésével oldották meg. Az indexelők a képekre vonatkozó, és modern botanikai jellemzőkkel kapcsolatos indexkifejezéseket is adtak. Az indexkifejezéseket összevetették a CAB tezaurusával, és a Royal Horticultural Society *Dictionary of Gardening* kiadványával is.

A két adatbázisból jött létre a *Promenade*, amelyet az *ONTOS* tárgyorientált adatbáziskezelővel kezelnek. Eredetileg házilagosan fejlesztették ki a grafikus felhasználói interfészt az *ONTOS*-szal való párbeszédre, hogy ezzel lehetővé tegyék távoli felhasználók keresését is az adatbázisban. A *Mosaic* megjelenése után ezzel leálltak, mivel a *World Wide Web* és a *Mosaic* keresőprogram sokkal gyorsabb és hatékonyabb párbeszédet tesz lehetővé.

/McLEAN, S.–RASMUSSEN, E. M.–WILLIAMS, J. G.: *Promenade: networked query and retrieval of horticultural images*. = *Online Information '94*. 18th International Online Meeting, 6–8 December 1994, Proceedings. Oxford, etc.: Learned Information, 1994. p. 457–468./

(Koltay Tibor)