



2.6 TARTALOMKEZELÉS

A tartalomkezelés témaköre alatt a különböző típusú információk (szöveg, kép, hang, mozgókép, stb.) egyéni és csoportos előállítását, fejlesztését, tárolását, rendszerezését, és visszakeresését értjük. Elsősorban nem a produktumot, a különböző technológiák által létrehozott tényleges tartalmat, hanem az ahhoz vezető – vagy azt megőrző – eljárásokat, a tartalom-előállítás, tartalomtárolás, és -elérés formáit elemezzük.

Az internet fejlődésével (még nagyobb) teret nyernek a tartalom-előállítás új, kollektív fajtái (BLOGOK, WIKIK, fotóalbumok, fórumok, stb.). Az „Alkalmazási eszközök” fejezetben tárgyalt adat, szöveg- és ismeretbányászati megoldások szintén elősegítik, gyorsabbá és hatékonyabbá teszik a tartalomgenerálást és -kezelést. A valóságot és a virtualitást különböző mértékben keverő (multimédiás) technológiák a jelenleginél lényegesen interaktívabbá alakítják a médiatartalmat; egyre gyakoribb a reális és a virtuális közötti átjárás.

A Szemantikus Web ajánlásait követő „intelligens” keresők jelentősen hozzájárulnak a rendezettebb adatáramláshoz, az információhoz való könnyebb és pontosabb hozzáféréshez.

Részterületek fejlődése

2.6.1 Multimédia tartalmak előállítása és megosztása

A mobil kommunikáció, a digitális műsorszórás és a hálózati infrastruktúra konvergenciája az egyéni és a társadalmi igényekhez alkalmazkodó, azokra reagáló MULTIMÉDIA-tartalom előállítását teszi szükségessé. A modernebb multimédiás tartalom (mozgás, illat, stb.) gyártása és kezelése kontextus-alapú megközelítéseket, nagyobb interaktivitást, speciális hardvert és szoftvert igényel. Elterjednek a jelenleg kísérleti fázisban lévő tartalomtípusok, alkotói módszerek: nem elsősorban a felhasználó, hanem mesterséges intelligenciával működő karakterek által aktívan alakított játékok, interaktív történetmesélés, stb.

A MULTIMÉDIA-forradalom egyik legfontosabb, a közeljövőben még erőteljesebben érzékelhető következménye, hogy a különböző munkákat már nem speciális terjesztési médiumokra készítik (például egybeépített CD- és mp3-lejátszón, mobiltelefonon, digitális kamerán, rádión, stb. egyaránt hallgatunk zenét). A médiumok közötti, multimédiás információ-megosztásban és keresésben olyan eljárások kerülnek előtérbe, mint a



képeket, szöveget, numerikus és audió adatokat egyetlen konzisztens reprezentációba integráló „ismeretbányászat”.

A nemcsak multimédiás, hanem az összes online és elektronikus tartalomra vonatkozó mennyiségi növekedés kettős irányba mutat: egyrészt potenciálisan „mindentudást” eredményezhet, de legalábbis mindennek a tárolását és visszakereshetőségét, amely azonban a megfelelő információ megkeresésének és kiszűrésének a nehézségével, jár, másrészt a túl sok tartalom „információs környezetszennyezést”, „adatszemet-áradatot” generál. Eltolódnak a mennyiség és a minőség arányai, egyre gyakrabban merül fel a kompetencia és a minőségbiztosítás kérdése.

2.6.2 Interaktív médiatartalom

A tartalom-előállítást nagymértékű interaktivitás (pl. VR, AR) technikák használata) jellemzi, amelyben – többek között – az adat- és szövegbányászat is fontos szerephez jut.

A virtuális világgal kiegészülő valódi világ, a BŐVÍTETT VALÓSÁG (AUGMENTED REALITY) legnagyobb kihívása az időnek és a térnek SZENZOROKKAL, pozíció-detektorokkal és képfelismerés segítségével történő szinkronizálása, stb. Egyszerre kell pontosnak, és gyorsnak lennie. A való és a virtuális világ képeinek olyan szinten kell egymáshoz kapcsolódniuk, hogy a néző ne tudjon különbséget tenni közöttük. A vizuális, audió és tapintásalapú információ vagy a környezet részeként jelenítődik meg, vagy mobil, rajtunk hordott eszközökön keresztül jut el majd hozzánk. Egyelőre azonban távol vagyunk ettől a szinttől – a Gartner szerint talán csak a következő évtized közepére, vagy még később lesz kivitelezhető.

2.6.3 A kollektív tartalom-előállítás eszközei

A tartalom-előállítás és -kezelés kollektív formái (főként a BLOGOK és a WIKIK) egyre fontosabbak lesznek. Nemcsak a dokumentumok készítését alakítják át, hanem kihatnak a döntéshozatalra is (például a vállalatoknál). Az egyéni szerzők által jegyzett „befejezett”, „zárt” munkák mellett előtérbe kerül a csoportosan, akár naplószerűen is létrehozott, permanensen változó „nyílt” tartalom. Kialakul a blogosphere (blogoszféra, blogvilág).

A weblogoknak is nevezett (de a szerverek naplóállományával nem azonos) BLOGOK világhálón publikált periodikus írások (fordított kronológiai sorrendű) gyűjteménye: egyéni és vállalati naplók, meghatározott témára fókuszáló közösségek munkái, politikai kampányszövegek, médiaprogramok, stb.



A fejlődés újabb állomását jelentik a BLOGOK és a videó összekapcsolásával létrejött, jelenleg azonban még gyerekcipőben járó videóBLOGok (vlogok).

A WIKIK hiperlinkekkel összekapcsolt weboldalakat kezelő egyszerű, szöveges együttműködés-alapú rendszerek. Lehetővé teszik, hogy a felhasználó megváltoztassa a mások által létrehozott oldalakat, kommentárokat. Több irányba fejlődnek tovább, például a portálmegoldások felé, és megjelennek a bejegyzéseket címkéző szemantikus WIKIK is. Várhatóan 2006 és 2009 között érnek el egy kiforrott állapotot. (Gartner)

2.6.4 Szemantikus Web

Az indexelésben terjednek az MI-módszereket használó automatikus és adaptív tanulórendszerek, kísérleti stádiumból 2009 és 2014 között széles körben elterjedt technológia lesz a Szemantikus Web. (Gartner) A jelen kaotikus adatáramlását egy szabályozottabb, irányítottabb hálózat váltja fel. Az adatok még gondosabban definiált reprezentációja lehetővé teszi, hogy a felhasználók automatikusan, emberi beavatkozás nélkül hasonlítsák össze, osszák meg a különböző honlapokról származó információkat.

A fejlesztés első lépése során nyelveket dolgoztak ki. Például a névszolgáltatásra (naming) URL-eket, a szintaxisra XML-t használó RDF-et (Resource Description Framework). A második, jelenlegi fázis a metaadat (metadata) standardok (pl. WSDL, OWL) létrehozására, kezdeti alkalmazásaira fókuszál. A harmadik, legérdekesebb szakasz az úgynevezett mag körüli szabályalapú rendszer megteremtése lesz. Ennek során komoly kihívás az adatok és a hozzáférési eszközeik közötti együttműködési képesség (interoperability) megoldása.

A Szemantikus Web fogalmához szorosan kapcsolódnak az ontológiák, amelyek szabványosításában döntő lépés volt, a W3C: Web Ontology Language (OWL) ajánlás. Egy ontológia egy domainről (szakterületről vagy annak sajátos nézőpontjáról) szóló viszonylag kimerítő és szabatos fogalmi séma megfogalmazására irányuló kísérlet eredménye; első közelítésben olyan hierarchikus vagy hálós adatszerkezet, amelyben helyet kap az összes odatartozó fogalom, azok relációi és a szabályok. A bennefoglalás viszony mellett azonban a fogalmakat még más szemantikai viszonyokkal is jellemezni lehet, olyanokkal, amelyek az egyes fogalmak összekapcsolódási módját adják meg.

2010-re az ontológiatervezés a szoftverfejlesztés részévé válik, viszont egyelőre csak a leginkább tudás-intenzív területekre (gyógyszeripar, egészségügy, államigazgatás, stb.) korlátozódik. (Gartner)



2.6.5 Intelligens keresők

A komplex, esetenként specializált KERESŐMOTOROKNAK kulcsszavak helyett speciális, pontosan és célzottan megfogalmazott kérdéseket is megadhatunk. Új típusú keresők jelennek meg: a keresők keresői, az úgynevezett metakeresők, illetve a szövegalapúak mellett a hang-, kép- és mozgókép-alapúak is. A keresők képesek lesznek munkájukat a felhasználó személyére szabva, annak szokásaihoz, földrajzi környezetéhez igazodva, egyfajta személyi asszisztensként végezni. Mivel a tartalom elérése során a jelentés és a kontextus lesznek a legfőbb szempontok, lokalizált, fókuszáltabb találatok várhatók.

2.6.6 Archiválás

A tudomány, a kultúra, sok forrását csak digitális formában állítják elő, és terjesztik, illetve a web az egyik legfontosabb információforrás; így az on- és offline digitális örökség megőrzése sürgős feladat. Az archiválás lehet szelektív vagy teljes, ill. statikus vagy dinamikus. Súlyos problémát jelent, hogy a jelenlegi adattároló médiumok és eszközök gyors elavulása (ld. az „Alapok” fejezetben) ellenére kell biztosítani az archivált anyagok jövőbeni elérhetőségét.

Az integrált dokumentumarchiváló és -visszakereső rendszereket főként a nagy mennyiségű statikus dokumentumhoz való hozzáféréshez, ill. tárolásukhoz használják. A rendszerek 2007-re vállalati tartalomkezelő architektúrába (ECM) integrálódnak. Az elektronikus leveleket szintén az ECM-rendszerek archiválják (Gartner).

2.6.7 Digitális könyvtárak

Megkezdődött, és várhatóan hosszú évekig eltart még a világ könyvállományának digitalizálása, online elérhetővé tétele. A munka főbb fázisai: az oldalak beszkenelése/lefotózása, a kép feldolgozása, gépi olvasása, a kép szövegfájlá alakítása, word-dokumentumok készítése a szövegfájlokból, az elektronikus formájú könyv tárolása.

A 2005-ben indult EU-s TEL-ME-MOR projekt támogatja a 10 új tagállam nemzeti könyvtárát az Európai Nemzeti Könyvtárak Igazgatóinak Konferenciája kezdeményezésére létrejött TEL-ben (The European Library/Az Európai Könyvtár) való teljes jogú tagság elnyerésében. Ezáltal a 2004-ben csatlakozott EU tagállamok nemzeti könyvtárai egységes hozzáférést nyernének a fő európai nemzeti könyvtárak elektronikus forrásaihoz, valamint más könyvtári szolgáltatásokhoz.