



Nem szövegalapú információszolgáltatás a hannoveri TIB-ben. Kihívások, tervek

A tudományos kutatás egyre összetettebb és egyre több adattal dolgozik: nagyon adatintenzívvé vált. Az információtechnológia hatására a tudomány is változik. Az adatáradat befolyásolja mind a kísérletező, mind az elméleti tudományokat, mind a számítógép-tudományokat. A tudomány paradigmái négyre bővültek: ezer évvel ezelőtt a tudomány a tapasztalásra épített, az utóbbi néhány száz évben elméleti jellegű volt, modelleket alkotott és általánosított. Az utóbbi néhány évtizedben összetett jelenségeket szimulált a számítógép segítségével. A jelenlegi e-tudomány adat-tömegekkel foglalkozik, amelyeket műszerekkel rögzít vagy szimulátorokkal hoz létre. Az adatokat szoftverekkel dolgozza fel, a tudást számítógépekben tárolja, és a tudós adatbázisokat, fájlokat elemző adatkezelő rendszerek segítségével. Az adatintenzív tudomány a tudományos kutatás negyedik paradigmája.

A tudományos kommunikáció is alapvetően megváltozik: nyilvános digitális könyvtárak veszik át a hagyományos könyvtáraktól a kiadványok őrzésének szerepét: a költségek és a korlátlan, állandó elérhetőség igénye miatt, valamint azért, mert a kísérleti adatokat együtt kell tartani az adatokról szóló dokumentumokkal.

A tudományos információ több a kiadott könyvnél vagy megjelent cikknél: a könyvtáraknak az adatok felé is nyitniuk kell.

A hannoveri TIB (*Technische Informationsbibliothek* = *Műszaki Könyvtár*) GetInfo szolgáltatása ezt az új megközelítést ülteti át a gyakorlatba.

A GetInfo tudományos és műszaki portál egyszerre nyújt hozzáférést nagy szakterületi adatbázisokhoz, kiadók kínálatához és könyvtári katalógusokhoz. Közvetítésével a teljes szöveget is meg lehet rendelni, valamint jelenleg az egyetlen olyan

könyvtári portál Európában, amely tudományos adathalmazokhoz vezet el.

Ma már 1500 feletti adathalmaz érhető el a hozzá tartozó tudományos kiadványok részeként. Ha a felhasználó eljut egy adathalmaz állandó azonosítójához, a kattintás után először egy oldallal találja magát szemben, amelyen az adathalmaz metaadatait tanulmányozhatja, mielőtt megabájnja adattömegeket töltene le. Hasonlóan működik, mint a DOI használat a tudományos folyóiratokban. A DOI feloldása először a kiadó oldalára, a cikk metaadataihoz vezet el.

A nem szöveges dokumentumok körében a közelmúltban generált multimédia-tartalmak száma évről-évre nő. Például a 3D-s modellek használata a mérnöki gyakorlatban egyre fontosabbá válik.

Tartalomalapú indexelés

A nem szöveges dokumentumokhoz általában nem tartozik leíró adat, ha előzőleg nem katalogizálták adatbázisban. Manapság a digitális könyvtárakban a nem szöveges dokumentumokat indexelik és szöveges metaadatokkal elérhetővé teszik. Ezeket azonban kézi munkával költséges kinyerni a tartalomból, és a szerzők ritkán motiváltak abban, hogy dokumentumaikhoz metaadatokat adjanak meg.

A TIB célja PROBADO elnevezésű, több német digitális könyvtárat átfogó projektjében, hogy ezen túllépjen, a kézi katalogizálási munkát minimálisra csökkentse, és ahol csak lehet, automatikus metaadatokat generáljon a nem szöveges dokumentumokból. A PROBADO nem tisztán kutatási projekt, a kifejlesztett rendszereket hosszú távon szeretnék használni a könyvtárakban.

A jelenlegi 3D-s tartalomra alapuló indexelő eljárások vagy a globális térrel vagy az építészettel kapcsolatosak.

Vizuális kereső interfész a nem szöveges adatokhoz

A kereső interfész megjelenítése és használhatósága döntő szempont a szolgáltatás elfogadtatásánál. A PROBADO 3D-s modelleket kereső interfészei támogatják a keresést a szöveges metaadatokban: a különböző szűrők szerinti böngészést, minta szerinti lekérdezések feltöltését, 3D-s modellek interaktív felvázolását. A különböző megjelenítések a felhasználó számára lehetővé teszik az adathalmazok feltárását, az eredmény felhasználását új keresésben.

Egy másik, 2010-ben induló TIB projektben nem szöveges adatokon akarnak vizuális kereséseket elvégezni és az eredményeket megjeleníteni.

Tartós azonosítás – nem szöveges dokumentumoknál is

Az eredeti kutatási adatokból a tudományos folyóiratok csak nagyon keveset publikálnak. Az archiválási politikák ellenére ezeket általában privát fájlokban őrzik, nem pedig biztonságos intézményi repozitóriumokban. A kutatási adatok elérésének hiánya gátat jelent a nemzetközi kutatásban, ami az erőfeszítések szükségtelen megismétlését eredményezi és az eredmények bizonyítása nehezzé, majdnem lehetetlenné válik. Óriási összegeket emészt fel a már létező adatok újbóli létrehozása.

A szerzők számára ösztönzőleg hatna, ha az adathalmazaik ugyanolyan fontosságúak lennének, mint az idézhető publikációik.

A TIB az adathalmazokra is kifejlesztette a DOI-t és támogatja használatát. A DOI szorosan kapcsolódik az objektumhoz, és nem ahhoz a helyhez, ahol az objektum éppen megtalálható. 2005-től a TIB kutatási adatok hivatalos DOI regisztrációs ügynöksége. Magukat az adatokat nem a TIB őrzi, hanem minőségbiztosítással rendelkező, tárolásra és elérésre szakosodott megbízható adatközpontok. A TIB már több mint 6000 kutatási adathalmaznak adott DOI-t.

DataCite

A tudományos kutatási adatok elérése a nemzeti tudományos intézmények felelősségi körébe tartozik, ugyanakkor a tudomány globális, a kutatók az egész világon együtt dolgoznak közös projekteken, nemzeti és külföldi információszolgáltatókat is használnak.

A kutatási adatokhoz való hozzáférés kihívására a válasz: globális együttműködés nemzeti képviselőikkel.

A *Globális DOI Ügynökség (Global DOI Registration Agency)* minden kutatónak felkínálja az adathalmaz regisztrálását és katalogizálását. A DataCite konzorciumot 2009. december 1-jén Londonból indították el és sok intézmény és könyvtár a tagja (köztük a TIB). A DataCite átvette az ügynökség szerepét: állandó azonosítót ad adathalmazokhoz, közös kutatási módszereket, legjobb gyakorlatokat és segédleteket hoz létre és támogat. Az intézmények önállóan dolgoznak adatközpontjaikkal és más kutatási adatokat birtokló szervezetekkel. Mivel a tudomány is globális egyéni kutatókkal, a DataCite is globális, és együtt dolgozik nemzeti partnerekkel.

/BRASE, Jan – BLÜMEL, Ina: *Information supply beyond text: non-textual information at the German National Library of Science and Technology (TIB) challenges and planning. = Interlending & Document Supply*, 38. köt. 2. sz. 2010. p. 108–117./

(Burmeister Erzsébet)



Thieme Clinical Collections

The prime resource for essential clinical E-Books

Thieme Clinical Collections contain an extensive assortment of E-Books carefully selected to provide the students, researchers, and clinicians at your institution with thorough coverage of all major medical fields.

This currently expanding resource consists of 314 titles include titles published during the following years:

Collection 1	2000–2003:	65
Collection 2	2004–2006:	87
Collection 3	2007–2009:	111
Collection 4	2010:	51

Thieme Teaching Assistant

One powerful resource.
Two innovative teaching tools.

The **Thieme Teaching Assistant** platform is the ultimate web-based presentation tool, featuring powerful functionality to enrich and deepen the classroom experience.

Anatomy
2,000+ full-color
illustrations

Physiology
792 stunning
color images, charts,
and tables

Ideal for:

- Presentations • Course materials
- Class handouts • Review

For more information and a free trial please contact:
Thieme Institutional Sales
E-mail: eproducts@thieme.de • Tel.: +49-711-8931-407

