

Holl András

Vízió a tudományos (szakkönyvtári) informatikai infrastruktúráról

Az egyre erősebben digitalizálódó világban jelentős változások mennek végbe. Amint az anyagi világ mind nagyobb mértékben jelenik meg a világhálón – digitális nyilvántartások, digitális kópiák, digitális folyamatszabályozás és megfigyelés, digitális kommunikáció, digitális adatok, illetve a személyes terek virtuális kiterjesztései formájában – a korábban élesen kirajzolódó határok elmosódnak. Nem csupán az országok, de a szakmák, intézmények, hatáskörök, tudományterületek és sok más tényező tekintetében is. A Web 1.0 helyet adott a 2.0-ás és a 3.0-ás koncepcióknak, a felhasználók teljes élete, tevékenysége egyre inkább a webre helyeződik át. Mit jelent ez a tudomány területén, ezen belül is, a tudományos szakkönyvtárak, a közgyűjtemények számára?

Tárgyszavak: tudományos könyvtár; hálózat; tartalomszolgáltatás; nyílt hozzáférés; hitelesség

A számítógépes hálózati alkalmazások immár nem elszigetelt eszközök, hanem egymással kommunikáló, munkafolyamatok végigvitelét segítő, sőt, végigvivő eszközrendszerkévé alakulnak. A tudományos kutatás, oktatás, közgyűjtemények kezelésének a feladatai is összefonódnak. A kutatás – és a vele összefonódó oktatás, ismeretterjesztés, innováció, szellemi vagyonkezelés feladatai is összekapcsolva, zökkenőmentesen kell működjenek. A kutató tájékozódik, kísérleteket, megfigyeléseket végez, modelleket alkot, cikket ír, oktat, népszerűsít és gondoskodik a folyamatban keletkezett adatokról és publikációkról. És persze közben pályázik, értékkel, jelent. Ezeknek a folyamatoknak a termékei – bibliográfiai adatok, teljes szövegű művek, adatok, számítógépes programok, pályázati anyagok, jelentések, elszámolások, előadások stb. – mind ugyanezen folyamatok közül egy következőnek bemenetei is egyúttal. Mindegyik kooperációban keletkezik, minden esetben szükség van az átláthatóságra, mindent meg kell osztani, közzé kell tenni. Eközben az információt felhasználó aktorok egyre nagyobb mértékben gépek, nem közvetlenül emberek.

A fentiek következtében minden eddiginél nagyobb szerepe van a szabványoknak, a nyitottságnak, a folyamatkezelésnek, a kommunikációnak (mind ember és ember, mind gép és gép között – az ember-gép kapcsolatról meg nem feledkezve).

A digitális áradat rengeteg adattal, információval önt el bennünket, ezért az adatok és információk értéke csökken. Az adatból és információból csak úgy lehet felhasználható tudás, ha az informatikai eszközök segítséget nyújtanak a szükséges adatok és információk megtalálásában, feldolgozásában, értelmezésében, még az értékelésében is. Intelligens rendszereket kell létrehozni, amelyek szemantikus keresésre, információkiemelésre, elemzésre képesek.

Milyen következtetéseket kell ebből levonnunk? Nem szabad elszigetelt rendszereket építsünk. Még jobban kell figyelni a szabványokra, protokollokra, olyan rendszereket kell építeni, támogatni, amelyek be tudnak kapcsolódni az alkalmazások hálójába, más alkalmazásoktól át tudják venni, más alkalmazásoknak át tudják adni az adatokat. Alkalmazásokat, amelyek segítik az együtt dolgozó intézményesített vagy virtuális közösségek együttes munkáját. Oda kell figyelni a kutatói, oktatói, közgyűjteményi folyamatok ciklusaira. Nem elég tartalmakat létrehozni a weben, de ezeknek a megosztásáról, nyilvántartásáról, archiválásáról is gondoskodni kell, még hozzá nem utólag, hanem a projekt tervezésekor. A tartalmaknak minden eddiginél dinamikusabbnak kell legyenek, a szolgáltatásoknak adatátvételle és adatátadásra is képesnek kell lennie. Jól definiált felületekre és protokollokra van szükség.

A kutató szakirodalmat keres – az intelligens keresők segítenek abban, hogy megtalálja a szükséges információt (nem csupán ismert metaadatok, vagy kulcsszavak alapján kell tudnia keresni). A megtalált cikket későbbi felhasználásra elteszi vagy megjelöli, esetleg megosztja munkatársaival, adatait bibliográfiai adatbázisba rögzíti. Kísérletet végez – a berendezéseket vezérlő kódokat más kutatók munkafolyamataiból kölcsönzi, saját munkafolyamatait, kódjait megosztja, közvetlen kollégáival, és a témán dolgozó más, „ismeretlen” kutatókkal kooperálva fejleszti. Az adatokat griden, távoli szuper számítógépen elemzi, modellezi, amelyekhez munkakapcsolatai révén tud hozzáférést kapni. A használt szoftverek nyílt szoftverek, a felhasználói-fejlesztői közösség támogatásával bővülnek. Cikket ír felhőbeli eszközökkel, kollaboratívan. Amit ír, ha elkészül, a kézírata automatikusan az intézményi repozitóriumba kerül, és továbbítódik a kiválasztott folyóiratnak. A megjelent cikk bibliográfiai adatai automatikusan bekerülnek a személyi publikációs listájába, és a kutatás-nyilvántartási rendszerekbe. Ide automatikusan bekerülnek azok az idézetek is, amelyek a publikációra érkeznek. Természetesen a kiadó a közzététel előtt plágiumellenőrzést végez, és azonosított minden társszerzőt. A közzételés díj kifizetése is automatikus: az affiliációk azonosításával a kiadó rendszere felméri, hogy a szerzők intézményei közül melyiknek van egyezsége vele, és az intézményi könyvtárak értesítést kapnak a díjfizetési kötelezettségről. A megjelent cikket bárki – nem csak a szakma – megtalálja, és szabadon elérheti, legyen az ember vagy robot. Ha az eredményből szenzáció lesz, a hatását (tweetek vagy lájkok) érzékelik és elkönyvelik az erre szolgáló rendszerek. A repozitóriumba került vagy a kiadónál elérhető cikket sok más, hasonlóval együtt veszik munkába a szövegbányász rendszerek, hogy a kutatók kérdéseire minél jobb válaszokat adhassanak, olyan összefüggések felderítésében segíthessenek, amelyekre csak óriási méretű információmennyiség együttes elemzésével nyílik lehetőség. A kutató segítséget kér a helyi könyvtárostól az adatai dokumentálásához. Az archiválás csak dokumentálás, metaadatokkal való ellátás és az archív állapot bejelölése – az adatok eddig is a felhőben voltak, ezután is ott maradnak, csak a hozzáférési jogok változnak, és a láthatóság növekszik. Az adatállományok metaadatait egyre kevésbé kell a kutatóknak beírnia, az állományt létrehozó és feldolgozó műszerek és programok gondoskodnak a metaadatmezők nagy részének kitöltéséről. Minden adatfeldolgozási lépés az adatokhoz kapcsolva naplózásra kerül. Miközben ez történik, a kutató

fel sem áll a képernyője előtt – legfeljebb egy teát meginni, lábat mozgatni, hátat kiegyenesíteni. A felgyorsult kutatási ciklusban az egyes lépések – információk felderítése, kódblokkok átvétele, rendszerekhez való hozzáférés megszerzése immár nem évek, hónapok, hetek alatt történik, hanem órák, percek vagy másodpercek alatt.

Ez a vízió informatikai oldalról egyáltalán nem radikális. Minden egyes komponense létezik már, egyesek évtizedek óta. Ami fontos, hogy egyre nagyobb összefüggő „építményeket” kell a lego-elemekből összerakni. További fontos elem a nyitottság még további erősítése. Open Source, Open Access, Open Data – Open Science.

A könyvtári-közgyűjteményi stratégia elemei között a következőkre javasolunk kiemelt figyelmet fordítani:

- 1) Új generációs, felhőalapú, kollaboratív könyvtári-közgyűjteményi platformok bevezetése és támogatása, melyekben különböző intézmények munkatársai közösen végezhetik munkájukat, és közös adatvagyon gyarapíthatnak.
- 2) Közös keresőrendszerek fejlesztése és támogatása, melyek segítségével a felhasználók immár nem szigetszerűen elkülönülő adatbázisokban, hanem intézményi és országhatáron átnyúló forrásokban kereshetnek.
- 3) A források – publikációk, dokumentációk, adatok, kódok – nyílt elérését támogató rendszerek fejlesztésére van szükség.
- 4) A fent említett források feltáráshoz adat- és szövegbányászati eljárások fejlesztése és használata szükséges.
- 5) Olyan rendszerek kialakítását kell támogatni, amelyek segítenek a szabadon hozzáférhető források hitelességének megítélésében – szemantikus webtechnológiák segítségével, az információk provenienciájának: szerzőinek (személyek), forrásainak (intézmények), előzményeinek (dokumentumok), és megalapozottságának (adatok) feltáráásával.

Beérkezett: 2017. VIII. 20-án.



Holl András

az MTA KIK informatikai főigazgató-helyettese.

E-mail: holl.andras@konyvtar.mta.hu