

Hogyan alakítják át a közgyűjtemények a szolgáltatásaikat a digitális bölcsészet eszközeinek és módszereinek használatával?

Mi a feladata egy digitális bölcsésznek a könyvtári környezetben, hogyan integrálódik a digitális bölcsészeti munka a könyvtárak napi tevékenységeibe, és milyen speciális készségekre van szükség ehhez a területhez?



A válasz elején mindenképp szólni kell egy olyan tudományterületről, a humáninformatikáról, melyet magam is művelek, s amely ezernyi szálal kapcsolódik a digitális bölcsészethez, határaiik sok esetben igen képlékenyek, sőt, önkényesek. Napjaink könyvtári munkatársa digitális bölcsész vagy humáninformatikus is lehet, akár egyszerre mindkettő. Az egyik esetben végezhet webarchiválást, digitális szövegkiadásokat készíthet elő, az intézmény által birtokolt digitális vagy digitalizált szövegtárakon különféle, "távlati" elemzéseket futtathat, látványos vizualizációkkal teheti fogyaszthatóbbá a nagy adathalmazokat, vagy épp művek és címeik hatalmas, hiteles és adatgazdag tudástárát alkothatja meg. A humáninformatikus kevésbé van otthon az irodalom- vagy a nyelvtudomány területén, ami a digitális bölcsész "alapkészletének" része. Munkájának jelentős része az intézményi metaadatok körül forog: karbantartásukkal, szabványosságukkal, átalakításukkal és közzétételükkel foglalkozik. Tudását gyűjteménykezelő rendszerek és más elektronikus közgyűjteményi szolgáltatások, például discovery-felületek üzemeltetésében is jól használhatja. Tájékozottsága kiterjed a legkülönbözőbb közgyűjteményi metaadat-formátumokra, konverziós lehetőségeikre és a hazai, illetve nemzetközi (!) szoftvermegoldásokra.

Az informatikai technológia és kisebb részben az elméleti háttér módosulása, változása napjainkban is folyamatosan zajlik. Az entitásalapú forrásleírás, a webarchiválás, a szemantikusweb-technológia vagy éppen a mesterséges intelligencia alkalmazási megoldásai egyre nagyobb teret kérnek, amelyek régi gyakorlataink felülírását, vagy bevált módszereink finomítását is elhozhatják.

Hubay Miklós
MNMKK
Petőfi Irodalmi Múzeum

A digitális bölcsész feladata a könyvtári környezetben rendkívül szerteágazó, átfogóan talán a következőképpen fogalmazható meg: digitális eszközök és módszerek integrálása a könyvtári gyakorlatokba, digitális tartalmak létrehozása és kezelése. Ide tartozik a digitális adatbázisok, archívumok kialakítása, folyamatos bővítése, a munkafolyamatok minél magasabb fokú gépi automatizálása, online szolgáltatások, platformok fejlesztése, tömeges digitalizáció támogatása, kutatástámogatás, ismeretterjesztés, oktatás digitális bölcsészeti eszközök használatával. A közgyűjteményekben már nemcsak digitalizált, hanem eleve digitálisan született (born digital) tartalmakkal is számolnunk kell (kéziratok, fotók stb.), sőt ide kapcsolhatjuk a webarchiválás monumentális feladatát is.

Hogy mit kell egy digitális bölcsésznek mindezekhez tudnia? Ismernie és használnia kell a területhez kapcsolódó legújabb informatikai eszközöket, követnie kell a nap mint nap változó és fejlődő elméleteket, tendenciákat, s szintén alapvető fontosságú, hogy erős humán tudományos alapokkal rendelkezzen. A digitális bölcsész tehát egyfajta keverék kutya, akinek feladata, hogy a hagyományos bölcsészet tiszteletben tartása mellett újabb, modern eszközöket, eljárásokat, lehetőségeket fedezzen fel és vezessen be a bölcsészettudományokba, ezzel nemcsak az eddig megválaszolatlan kérdésekre feleletet adva, hanem egyben új kérdéseket is feltéve. Emellett a bölcsészet néha megfoghatatlan és szabadon szárnyaló világot összeillessze az informatika szigorúan következetes menetelésével, kihasználva mindkettő előnyeit. S mindez már túlmutat a bölcsészeten: valójában ez a kultúra, a művészet, a tudomány, az információ új utakra terelése, az emberrel való merőben új viszonyának megmutatkozása.

Mihály Eszter
MNMKK OSZK
Digitális Bölcsészeti Központ

A közgyűjteményi tartalomszolgáltatásba megérkezett a digitális bölcsészet diszciplinája. Ahhoz, hogy az átalakulás perspektíváját értsük, pillantsunk rá az idáig vezető evolúcióra. A számítógépes eljárások hajnala az analóg nyilvántartások adatbázisba rögzítése, később a katalógusok elérhetővé tétele következett a felhasználók számára offline környezetben, és az internet elterjedésével online publikáltuk is azokat. A fejlődés következő lépését az integrált szolgáltatások jelentették (pl. discovery). A tartalomszolgáltatás mérföldkövei a metaadatok mellett a digitalizált objektumok közzététele képként, majd az elektronikus könyvtárak, ahol teljes szövegű tartalomszolgáltatás valósult meg, de még kézzel rögzített vagy hagyományos OCR eljárással felismertetett módon rögzítve.

Ezt haladják meg a digitális bölcsészeti eljárásokkal feldolgozott és publikált szolgáltatások. Gyűjteményeink közzétételének új tárháza nyílt meg nagy tömegű szövegtárak intelligens szolgáltatásokban való megjelenésével. Az adatfeldolgozás során ebbe a körbe tartoznak az automatikus kézírás-felismerés (HTR), a digitális filológia tudományos forrásközlései, az adatvizualizációk, a korábban is használt optikai karakterfelismerés (OCR) jelentős minőségjavítása MI modellekkel, a metaadat-kinyerés automatikus eljárásokkal (NER) és a szintén MI alapú entitásjegyzés. Kutatástámogatási területen is nagy lehetőségeink vannak az adatgazdagításra különböző szabály- és modellalapú eljárások alkalmazásával. És még előttünk áll a Retrieval-augmented generation (RAG) – természetes nyelvi kutatástámogatási eszköz bevezetése közgyűjteményi területen is.

Bánki Zsolt
Magyar Nemzeti Levéltár

A digitális médium társadalmi penetrációjának radikális növekedése, amely újabb és újabb hullámok formájában ér el bennünket, amelyeket már-már a lökéshullámok intenzitásával élünk meg, a digitális bölcsészetnek és a gyűjteményi praxissoknak a közeledésével jár együtt. A célok ugyanakkor eltérőek, még akkor is, ha az eszközök hasonlóak: az előbbi a módszertanok – gyakran homályban maradó – működésének, hatásainak feltárásában, reflexív elemzésében és továbbfejlesztésében érdekelt (ezt takarja (!) az "unblackboxing" új keletű fogalma), az utóbbi ezzel szemben a gyűjteményi anyagok széles körű, tömeges digitális közzétételében érdekelt. Ha a digitális tudós célja a reflektált, elemző "távols olvasás", addig a közgyűjteményi szakember célja a távols olvasásra alkalmas anyagok magas minőségű előállítás. Látunk persze példákat arra, hogy digitális könyvtárak egyes digitális bölcsészeti eszközöket integrálnak a szolgáltatásukba (ilyen kísérlet a Gale Digital Scholar Lab), de a kutató jó esetben tartózkodik az ilyen, rejtett algoritmusokara épülő eszközök felhasználásától. Vannak ugyanakkor olyan területek, ahol a legkomplexebb gyakorlatok gyűjteményi közegből származnak, ilyen terület az OSZK-ban is magas szinten művelt web-archiválás és a digitálisan született, anyagok, hagyatékok feldolgozásának komplex folyamatrendszer. De visszatérve a legutóbbi "lökéshullámhoz": a generatív MI megjelenése a digitális tudóst és a közgyűjteményi szakembert egyaránt új helyzet elé állítja. A fekete doboz jóval mélyebb, mint a korábbi digitális módszertanok esetében, reflexiója ennek megfelelően sokkal nehezebb, azt gondolom, kreatív és kritikai felhasználása mindkét területen az egyik legfontosabb készség és tudás lesz a közeli jövőben.

Palkó Gábor
ELTE BTK
Digitális Bölcsészet Tanszék



A közgyűjtemények digitális átállásával egyre inkább elvárás, hogy a memóriaintézmények anyagaikat digitális formában tegyék közzé és adatbázisokat használjanak és építsenek. Ezáltal a digitális filológia, amely digitális tudományos szövegkiadások készítésére vállalkozik, elsősorban kutatók számára, és a közgyűjteményi munka, ahol a gyűjteményi anyagok tömeges digitalizálása a cél, átfedésbe kerül. A digitális filológia területén az elmúlt évtizedekben a jelölőnyelvi átíratok készítése, azon belül a Text Encoding Initiative (TEI) ajánlás felhasználása egyeduralmukodóvá vált. A TEI-t eredetileg bölcsész kutatók bölcsész kutatások támogatására és sztenderdizálására hozták létre, de, nemzetközi jó gyakorlatokat követve (pl. a Deutsches Textarchiv), már Magyarországon is léteznek közgyűjtemények által kiadott forráskiadások jelölőnyelvi átíratban – ebben például az OSZK Digitális Bölcsészeti Központja élén jár. Ha dokumentumokat, legyenek azok kéziratok, kódexek vagy akár modern kiadások, jelölőnyelvi átíratban teszünk közzé, lehetővé válik azok egységes, intézményeken és szolgáltatásokon átívelő kereshetősége és összevetése, illetve a dokumentumok tudományos célú, számítógéppel történő feldolgozása. Ugyanakkor kevés példát látunk még rá, hogy a szöveget nyers adatként tegyék közzé, ami elengedhetetlen a gépi feldolgozás, a "távols olvasás" gyakorlásához. A másik terület, amelyben a közgyűjteményeknek kiemelt feladatuk van, az az adatok közzététele a szemantikus webben (ittthon a Petőfi Irodalmi Múzeumban láthatunk erre példát). A szemantikus web, az adatok nyílt és egységes szintaxisú hálózata lehetővé teszi a különböző adatterekben való együttes keresést, illetve digitális kulturális örökség objektumok szemantikus leírását.

Fellegi Zsófia

HUN-REN Bölcsészettudományi Központ
Irodalomtudományi Intézet



Összeállította:
Szüts Etele

