

Koltay Tibor

Adatkönyvtáros vagy adattudós

Néhány gondolat a különbségekről és hasonlóságokról*

Egyre gyakrabban találkozunk az adatkönyvtáros és az adattudós elnevezésekkel, ezért érdemes a köztük meglévő különbségek és hasonlóságok mellett azoknak a szolgáltatásoknak a természetével is foglalkoznunk, amelyeket a világ számos táján egyetemi- és szakkönyvtárak nyújtanak a tudományos kutatás során keletkező adatok kezelése kapcsán.

Tárgyszavak: *adat; adatkezelés; információ; döntés-előkészítés; könyvtártudomány*

Bevezetés

A kutatási adatok kezelésével, valamint a velük kapcsolatos tevékenységek széles körében feladatokat végző, többnyire adatkönyvtárosnak (data librarian) nevezett egyetemi és szakkönyvtárosok munkájával kapcsolatos gyakorlati szakirodalom mellé egyre nagyobb mértékben felzárkóznak az elméleti írások, amelyek az adatok és az információ, az adatkönyvtárosok és az üzleti szférában foglalkoztatott adattudósok közötti különbségekkel és hasonlóságokkal foglalkoznak.

A megváltozott információs környezet

A kutatási adatok nemcsak a tudományos kutatás nyersanyagát képezik, hanem a kutatási tevékenység eredményeit is.¹ A tudományos kutatás egyaránt felhasználja az úgynevezett kis és nagy adatokat. Az előbbieket azért, mert az adatok mennyiségének növekedésével fordítottan arányos az a képességünk, hogy az egyes egyedeket megfigyeljük, ráadásul nehezebb megállapítani, hogy ezek a megfigyelések mit jelentenek, mivel nem feltétlenül tudjuk, hogy miként gyűjtötték, hogyan kezelték vagy miként alakították át őket. Ezért is van, hogy sok kutató inkább kisebb, közlelől is „szemügyre vehető” adatmennyiségekkel szeret dolgozni.² A nagy adatok divatosak, de már viszonylag régóta tudjuk, hogy felhasználásuk számos problémát vet fel³, ezért azt is kimondhatjuk, hogy a nagy adatok nem feltétlenül jó adatok. Emellett a kutatók sokszor nem tudják feltárni az adatokat vagy azok nem is feltárhatóak, illetve nem lehet és (főként) nem szabad őket nyilvánosságra hozni.⁴

A kutatási eredményként létrejött adatok mindenestre csak akkor hasznosíthatók, ha hozzáférhetővé tesszük, szervezzük és hosszú távon megőrizzük őket.⁵ Ennek fontosságát a könyvtárak is felismerték, a tudatosság szintje azonban országoként és szakterületenként is eltérő.

Az adatok kezelése és gondozása újfajta szakemberek közreműködését igényli vagy új típusú elvárásokat támaszt azokkal a szakemberekkel szemben, akik eddig is foglalkoztak adatokkal és információkkal.⁶ A kutatási információkkal a könyvtárakban elsősorban a többnyire *adatkönyvtáros* (data librarian) néven emlegetett szakemberek foglalkoznak. Az üzleti szférában dolgozó szakemberek esetében is többféle elnevezéssel találkozunk, amelyek közül az *adattudós* (data scientist) a leginkább közkedvelt. A köztük meglévő különbségek ellenére azonban vannak átfedések a két szakterület megkövetelte kompetenciák és feladatkörök között, ami azt is jelenti, hogy az adatkönyvtárosoknak és az adattudósoknak van mit tanulniuk egymástól.

Az adatok megítélése

Az adatokat, az információkat, a tudást és a bölcsességet magába foglaló értékpiramis a könyvtár- és információtudomány (a továbbiakban *könyvtártudomány*) egyik alapvető modellje. Az adatok e

* Készült az EFOP-3.6.1-16-2016-00001 „Kutatási kapacitások és szolgáltatások komplex fejlesztése az Eszterházy Károly Egyetemen” projekt támogatásával

piramis legalján, tehát az információ és a tudás szintje alatt helyezkednek el, ezért ahhoz szoktunk hozzá, hogy – egyébként teljesen legitim módon – úgy tekintünk rájuk, mint ennek a hierarchiának olyan elemeire, amelyeknek nincsen jelentése (értelme) vagy értéke, mivel nincs kontextusuk, és nem rendelkeznek interpretációkkal, ráadásul addig nem használhatók fel, amíg nincsenek használható (azaz releváns) formában.⁷

Az adatok vizuális megjelenítése, elemzése, vagy az adatbányászat viszont az információ közegében zajlik, tehát semmiképpen sem alacsony szintű tevékenység, hanem a döntéshozatalt segíti, így megmutatja az adatok fontosságát. Ezt támasztja alá az okos adat (smart data) fogalmának megléte is.⁸

Mivel a számítógépek ma már nagy mennyiségű adatot tudnak tárolni és kezelni, az adatok jelentősége és gyakorlati fontossága is megnőtt. Ennek megfelelően a könyvtárosok egy részének egyre inkább az adatok szakembereivé is kell válnia.

Változó könyvtárosi szerepek

A LIBER (az Európai Tudományos Könyvtárak Egyesülete) tagjai közül 22 ország könyvtárainak *Carol Tenopir* által irányított felmérése** azt mutatja, hogy igazgatóik felismerték a kutatási adatok kezelésének fontosságát, ezért keresik az intézményeik számára legmegfelelőbb megoldásokat.⁹ Egy korábbi, ugyancsak Tenopir vezetésével elvégzett vizsgálathoz¹⁰ hasonlóan ezek az európai könyvtárak is számos szolgáltatást nyújtanak. Ilyen szolgáltatás a kutatók igényeinek megfelelő, már létező adatállományok felderítése, hiszen az adatok az adatkönyvtárosok gondjaira vannak bízva, és ők azok, akik használatra és elemzések elvégzéséhez rendelkezésre bocsátják ezeket az adatokat.¹¹ Az ő feladatuk tehát az adatok reprezentációja, szervezése és terjesztése, és ők azok, akik a kutatási adatok kezeléséhez kötődő szolgáltatások hátterében álló technológiákat kialakítják.¹² Ezt megkönnyíti, hogy számos feladatot már meglevő ismereteik és készségeik birtokában tudnak megoldani.¹³

Viszonylag gyakori szolgáltatásuk a segítségnyújtás adatkezelési tervek elkészítéséhez.¹⁴ A kutatást finanszírozó szervezetek ugyanis gyakran előírják, hogy egy-egy kutatáshoz készüljön adatkezelési terv. Amikor a könyvtárosok segítenek ezek megírásában, hasonló készségekre van szükségük, mint amikor (hagyományos) tanácsok-

kal szolgálnak a felhasználóknak. Az adatkezelési terv esetében azonban ismerniük kell a finanszírozók előírásait és a helyi adatkezelési folyamatokat.

Számításba jöhet még a szöveg- és adatbányászat is, amelynek kapcsán főként a licencek, a felhasználóképzés és az újonnan keletkező (derivált) adatok kurátori gondozása területén játszhat szerepet a könyvtár. Emellett bibliometriai és alternatív metrikák (altmetrics) alapján történő számítások népszerűsítéséről is szó lehet, sőt ezeket el is végezhetik az adatkönyvtárosok.¹⁵

Az adatok esetében is van referenz interjú, amely általában több kérdésből áll, mint hagyományos (és jól ismert) formája. Lehetséges ugyanakkor, hogy az interjúk eredményeképpen a könyvtárosok hasznos, de nem végleges választ tudnak csak adni a felhasználónak.¹⁶ Az ilyen referenz interjúk során, a hagyományos tájékoztatáshoz hasonlóan fontos, hogy az Amerikai Könyvtáregyesület Tájékoztatási és Olvasószolgálati Szervezetének (Reference and User Services Association of the American Library Association, RUSA) irányelvei közül érvényesüljön az, hogy el kell ismerniük a felhasználó által az információs interakcióba hozott tudást.¹⁷

Az adatkönyvtárosok az adatokra való hivatkozás területén is a kutatók rendelkezésére állhatnak. A digitális források jövőbeni használhatóságához pedig azzal járulhatnak hozzá, hogy dokumentálják, honnan származik az adott adatállomány, miért fontos és hogyan kell használni. Ennek része, hogy azonosíthatóságuk érdekében metaadatokkal látják el az adatállományokat. Bár az adatok esetében ez a folyamat sokkal összetettebb annál, mint amit szöveges dokumentumok esetében ismerünk, ez a feladat ismerős számukra.¹⁸

Az adatkönyvtárosok a repozitóriumokban elhelyezendő adatállományok kiválasztásában is segítenek, amihez jól jönnek gyűjteményszervezési ismereteik és tapasztalataik.¹⁹

Az információkeresési készségek mindig is a könyvtárosok fontos kompetenciái közé tartoztak. Az adatok visszakeresése azonban nem olyan

** Ciprus, Franciaország, Horvátország, Luxemburg, Magyarország, Málta, Portugália, Románia, Szerbia, Szlovákia és Törökország könyvtárigazgatóit nem kérdezték.

magától értetődő tevékenység, mint amit folyóirat-cikkek visszakeresése kapcsán ismerünk. A fájlok átlagos hossza többszöröse lehet a tudományos publikációk szokásos méretének, ráadásul az adatállományok különböző típusú vagy formátumú fájlokból állhatnak, továbbá sokszor csak letöltés után olvashatók és használhatóak.²⁰

Az adatkönyvtárosok fontos feladata az is, hogy az információs műveltséggel közeli rokonságban álló adatuműveltség (adat-írástudás) alapjait megismer- tessék az oktatókkal és a kutatókkal, továbbá az egyetemi hallgatókat arra neveljék, hogy az adatok kezelése tekintetében is műveltek legyenek. Ehhez maguknak is el kell sajátítaniuk azokat a készségeket és képességeket, amelyek a kutatási adatok eléréséhez, megértéséhez, interpretálásához, kezeléséhez, kritikai értékeléséhez, valamint etikus használatához szükségesek.²¹ Az adatuműveltség oktatásához az adatkönyvtárosoknak pedagógiai felkészültséggel is rendelkezniük kell, ami a felső- oktatásban dolgozó könyvtárosokkal szemben egyre gyakrabban támasztott követelmény. Nem kétséges az sem, hogy az adattudósoknak is magas szinten kell(ene) elsajátítaniuk az adatuműveltség készségeit.

Összességében és általánosságban az adatkönyvtárosok új könyvtári szolgáltatásokat hoznak létre, figyelve a heterogén és kontextusfüggő kutatási adatokra, és a kutatás minden szakaszában segítik a kutatók munkáját.²²

Az adatkönyvtárosi munkakörök többfélék lehetnek. Egyik típusuk megkövetelheti valamely (nem könyvtári) szakterület beható ismeretét, ami egyúttal azt is jelenti, hogy az adatkönyvtárosok munkája jól körülhatárolt felhasználói csoportokhoz kötődik. Szükség van azonban olyan adatkönyvtárosokra is, akik széles körű ismeretekkel rendelkeznek arról, hogy miként használják fel különböző felhasználói csoportok a kutatási adatokat, ezért különböző szakterületeken dolgozó oktatókkal, kutatókkal és hallgatókkal állnak kapcsolatban.²³

Közös érdeklődés és érdekek

Az adattudomány rövidtávú célokat tűz ki, amelyek eléréséhez az adattudósoknak ismernie kell a dizájn és az információépítészet elveit, számítástechnikai és statisztikai módszereket kell közös nevezőre hozniuk, továbbá adatelemzésekre épülő új termékeket és szolgáltatásokat kell létrehozniuk.²⁴

Az egyetemi és szakkönyvtárak nem preferálják sem ezt a megközelítést, sem az alkalmazott eszközöket, azonban hosszú távon fontossá válhat számukra, hogy elemzéseket és statisztikai szolgáltatásokat is nyújtsanak.²⁵

Látható tehát, hogy az adatkönyvtárosok és adattudósok munkakörnyezete, kultúrája és feladatköre eltér egymástól. Mindazonáltal az adattudomány jól használható, új módszereket adhat a könyvtárosok kezébe. Hogy ezekkel éljenek, meg kell ismerniük az adattudomány által használt új gyakorlatokat. Ezzel már most találkozhatunk, hiszen az adatkönyvtárosok esetenként közvetlen módon is segítik a kutatók tevékenységét azzal, hogy adatokat elemeznek és vizuálisan jelenítik meg őket.²⁶ Ettől függetlenül, az adatkönyvtárosoknak nem kell programozókká vagy statisztikusokká válniuk, viszont érdeklődniük kell a programnyelvek és a programozás logikája iránt.²⁷

Ha azt nézzük, hogy mekkora figyelmet szentel az adattudomány az adatok minőségének, eltérő nézetekkel találkozunk. *Wang* szerint az adattudomány egyelőre főként az adatok mennyiségével törődik.²⁸ *Cao* viszont úgy látja, hogy a két területet és szakmát éppen a minőség iránti kitüntetett figyelem köti össze egymással. Igaz, ő még azt a kreativitást és intelligenciát is közös elemnek tekinti, amelyek korunk, a nem szabványosított, személyre szabott és gyakran még nem automatizált információs környezetben szükségesek a döntéshozatalt elősegítő indikátorok és bizonyítékok felfedezéséhez.²⁹

Amikor az adatkönyvtárosok és az adattudósok közötti hasonlóságokat és különbözőségeket nézzük, érdemes megfogadnunk *Lyn Robinson* tanácsát. Ő ugyanis arra figyelmeztet bennünket, hogy lövészárkokat ásunk, és nehezen hozzáférhető adatsílokat építünk, ha kizárólag kvantitatív módszereket használva csak az adatokra fókuszálunk. Ilyenkor ugyanis csak a mintázatokra és a szintaxisra figyelünk, viszont szem elől tévesztjük a jelentést. A szemantika bevonása kvalitatív módszertant kíván, amelynek feltétele, hogy figyeljünk az információra is. Ha ezt nem tesszük meg, túlságosan szűkre szabjuk az adatkönyvtárosok és az adattudósok szakmai profiljait, ami egyik szakmának sem tesz jót.³⁰

Mindezt alátámasztja, hogy a könyvtártudomány és az adattudomány egyaránt a rögzített információ teljes kommunikációs láncát tanulmányozza. Mindkét területre jellemző, hogy egy-egy lépessel

feljebb lép az előbbieken említett hierarchiában. Az adattudomány célja ugyanis az, hogy a nyers adatokból hasznosítható, a döntéshozatalt segítő tudást állítson elő. Ehhez az információn át vezet az út. A könyvtártudomány (bár fő tárgya az információ) a tudás hatékony szervezésére és felhasználására törekszik. Természetesen az ilyen, vagy hasonló alapokon nyugvó elképzelések fő pillére az a megközelítés, hogy az adatból meghatározható az információ, az információból a tudás, ráadásul úgy, hogy minderről azt gondoljuk, hogy az ellenkező irányban is működik.³¹

Következtetések

Egyáltalán nem „rargon aluli” és lebecsülendő feladat, ha azt tűzzük ki célul az egyetemi és szakkönyvtárosok elé, hogy segítsék a kutatókat az adminisztratív teendők ellátásában³², viszont ezeknek a könyvtáraknak arra is törekednie kell, hogy valódi együttműködések alakítsanak ki könyvtárosaik és a kutatók között. Korunk adatintenzív tudománya erre minden eddiginél több lehetőséget kínál, amelyeket viszont csak akkor tudják a könyvtárak kihasználni, ha megváltozik a könyvtárosok önazonossága és tisztán látják, hogy melyek a kutatási adatokkal kapcsolatos technikai és társadalmi, valamint etikai kérdések.³³ A változások lényeges feltétele az is, hogy a szolgáltatásokhoz, amelyeket a könyvtárak kínálnak, rugalmas szervezeti struktúrák és munkaköri leírások illeszkedjenek.³⁴

Hivatkozások

- 1 PRYOR, G.: Managing research data. London: Facet, 2012.
- 2 BORGMAN, Ch.: Big Data, Little Data, No Data: Scholarship in the Networked World. Cambridge, MA: MIT Press, 2015.
- 3 BOYD, D – CRAWFORD, K: Az adatrengeteg kínos kérdései: Vitaindító egy kulturális, műszaki és tudományos jelenségről. = Információs társadalom, 12. évf., 2012. 2. sz. 7–23.p.
- 4 BORGMAN, i.m.
- 5 CARLSON, J., FOSMIRE, M., MILLER, C. C., NELSON, M. S.: Determining data information literacy needs: A study of students and research faculty. = portal: Libraries and the Academy, vol. 11, 2011, no. 2, 629-657.p.
- 6 COX, A.: Academic librarianship as a data profession. Information Today Europe/ILI365, 2018

<http://www.infotoday.eu/Articles/Editorial/Featured-Articles/Academic-librarianship-as-a-data-profession-125376.aspx?PageNum=1> [Utolsó elérés: 2018. augusztus 28.]

- 7 ROWLEY, J.: The wisdom hierarchy: representations of the DIKW hierarchy. = Journal of Information Science, vol. 33, no. 2, 2007, 163-180.p.
- 8 WANG, L.: Twinning data science with information science in schools of library and information science. = Journal of Documentation, vol. 74, no. 6, 2018, 1243-1257.p.
- 9 TENOPIR, C., TALJA, S., HORSTMANN, W., LATE, E. HUGHES, D., POLLOCK, D. SCHMIDT, B., BAIRD, L., SANDUSKY, R. J. and ALLARD, S.: Research Data Services in European Academic Research Libraries. = LIBER Quarterly. vol. 27, 2017, no. 1, 23-44.p.
- 10 TENOPIR, C., POLLOCK, D., ALLARD, S., HUGHES, D.: Research data services in European and North American libraries: Current offerings and plans for the future. Proceedings of the Association for Information Science and Technology, vol. 53, 2016, no 1, 1–6.p.
- 11 LYON, L., MATTERN, E.: Education for real-world data science roles (Part 2): A translational approach to curriculum development. = International Journal of Digital Curation, vol. 11, 2016, no. 6, 13-26.p.
- 12 RICE, R. and SOUTHALL, J.: The data librarian's handbook. London: Facet Publishing, 2016.
- 13 COX, i.m.
- 14 TENOPIR et al. 2017
- 15 COX, A., VERBAAN, E.: Exploring Research Data Management. London: Facet, 2018.
- 16 RICE, SOUTHALL, i.m.
- 17 HULING, N, DALLAS, L. J. KINDER, J. B., WHITLATCH, J. B., AND WOODARD, B., Professional Competencies for Reference and User Services Librarians, 2017. <http://www.ala.org/rusa/resources/guidelines/professional> [Utolsó elérés: 2018. augusztus 28.]
- 18 COX i.m.
- 19 COX, VERBAAN, i.m.
- 20 BUGAJE, M., CHOWDHURY, G. Is Data Retrieval Different from Text Retrieval? An Exploratory Study. In: International Conference on Asian Digital Libraries. Springer, Cham, 2017. pp. 97-103.
- 21 KOLTAY, T.: Data literacy: In search of a name and identity. Journal of Documentation. vol. 71, no. 2, 2015, 401-415.p.
- 22 SEMELER, A. R., PINTO, A. L., ROZADOS, H. B. F.: Data science in data librarianship: Core competen-

- cies of a data librarian. = Journal of Librarianship and Information Science, 2017
<http://journals.sagepub.com/doi/10.1177/0961000617742465> [Utolsó elérés: 2018. augusztus 28.]
- ²³ FEDERER, L.: Defining data librarianship: a survey of competencies, skills, and training. = Journal of the Medical Library Association vol. 106, no. 3, 2018, 294-303.p.
- ²⁴ VOULGARIS, Z.: Data Scientist: The Definitive Guide to Becoming a Data Scientist. Technics Publications. Basking Ridge, NJ, 2014.
- ²⁵ MAXWELL, D., NORTON, H., WU, J.: The Data Science Opportunity: Crafting a Holistic Strategy. = Journal of Library Administration, vol. 58, 2018, no. 2, 111-127.p.
- ²⁶ OGIER, A. L., STAMPER, M. J.: Data Visualization as a Library Service: Embedding Visualization Services in the Library Research Lifecycle. = Journal of eScience Librarianship. vol. 7, 2018, no. 1, e1126
- ²⁷ SEMELER – PINTO – ROZADOS, i.m.
- ²⁸ WANG i.m.
- ²⁹ CAO, L. Data science: nature and pitfalls, = IEEE Intelligent Systems, vol. 31, 2016, no. 5, 66-75.p.
- ³⁰ ROBINSON, L.: Between the deluge and the dark age: Perspectives on data curation. = Alexandria. vol. 26, 2016, no 6, 73-76.p.
- ³¹ WANG i.m.
- ³² LERU Roadmap for Research Data. League of European Research Universities Research Data Working Group, Leuven, 2013
<https://www.leru.org/publications/leru-roadmap-for-research-data> [Utolsó elérés: 2018. augusztus 28.]
- ³³ ROBINSON, L., BAWDEN, D.: "The story of data" A socio-technical approach to education for the data librarian role in the CityLIS library school at City, University of London. = Library Management, vol. 38, 2017, no. 6-7, 2017 312-322.p.
- ³⁴ BURTON, M., LYON, L.: Data science in libraries. = Bulletin of the Association for Information Science and Technology, vol. 43. 2017, no. 4, 33-35.p.

Beérkezett: 2018. IX. 5.-én.



Koltay Tibor

Eszterházy Károly Egyetem
Tudástechnológiai Intézet.
E-mail: Koltay.Tibor@uni-
eszterhazy.hu

E számunk megjelenését önkéntes munkájával segítette:

Berke Barnabásné
Fonyó Istvánné
Hegyközi Ilona
Koltay Tibor
Prokné Palik Mária
Ungváry Rudolf