



# A kutatási adatok és a könyvtár

KOLTAY Tibor

## Bevezetés

Aki nyomon követi a könyvtártudomány szakirodalmát, mostanában azt látja, hogy szinte hirtelen megjelennek meg cikkek a kutatási adatokról, azok megosztásáról és gondozásáról, a „nagy adatokról” és számos kapcsolódó problémáról, ide értve természetesen azt is, hogy milyen szerepet játszhat az adatok kezelésében a könyvtár és a könyvtáros. Igaz, *Christine Borgman* már 2007-ben úgy látta, hogy az adatok megosztása és maguk az adatok a publikált cikkekhez hasonló tudástöket jelentenek.<sup>1</sup> Ugyanabban az évben *Z. Karvalics László* hívta fel a figyelmet az adat-intenzív tudományra.<sup>2</sup> Hozzáteszem, hogy régebbi keletkezésű szakirodalmat is bőven találunk. Nem kétséges tehát, hogy az egyre inkább digitális eszközöket alkalmazó kutatás nyomán keletkező adatárdatot kezelni kell, tehát hozzáférhetővé kell tenni, szervezni kell és hosszú távon meg kell őrizni.<sup>3</sup>

A kérdéskör rendkívül összetett, bár nyilvánvaló, hogy a legtöbb kutató munkájában számos területen alapvető az adatokra alapozott gon-

dolkodás, ami összekapcsolódik annak képességével, hogy az adatokat absztrakt fogalmakra fordítsuk le.<sup>4</sup>

*Borgman* 2012-ben nagy ívű áttekintést nyújtott a kutatási adatok megosztásával kapcsolatos kérdésekről, és megállapította, hogy az adatokban gazdag világgal szemben óriási a várákozás, legyen szó új gyógyszerek felfedezéséről, a világ időjárásának, történelmünknek vagy kultúránknak jobb megismeréséről. Az adatok mennyisége számos tudományterületen új modellek és új tudományágak létrejöttét (mint például asztroinformatika, bioinformatika és digitális bölcsészet) eredményezte. Ahhoz, hogy mindez megvalósítható legyen, a kutatóknak meg kell osztaniuk másokkal az általuk létrehozott adatokat, ráadásul értelmezhető és újrafelhasználható módon kell ezt tenniük. Annak ellenére, hogy a kutatásokat finanszírozó szervezetek nyomást gyakorolnak a kutatókra az adatok megosztása érdekében, ráadásul az adatok megosztása növelheti a hivatkozások mértékét, valószínű, hogy kevés adat megosztására kerül sor.<sup>5</sup>

## Mi az adat?

Kutatási adatokról beszélve elkerülhetetlen, hogy valamennyire meghatározzuk, hogy mit is értünk adat alatt. Az adatok olyan szimbólumok, amelyek tárgyak, események és környezetük tulajdonságait reprezentálják. Az adatok megfigyelések eredményeként jönnek létre, meghatározásainak közös vonása, hogy elsősorban bizonyos jellemzők hiányával operálnak a következő módon:

- az adatoknak nincsen jelentése (értelme) vagy értéke, mivel nincs kontextusuk, és nem rendelkeznek interpretációkkal;
- az adatok diszkrét, objektív tények vagy megfigyelések, amelyek szervezetlenek és feldolgozatlanok, továbbá nem közvetítenek specifikus jelentést;
- az adatok dolgok, események, tevékenységek és tranzakciók elemi és rögzített leírásai.<sup>6</sup>

A tudományos adatok bármely, megfigyelésen, kísérleten vagy valamely hipotézis megvizsgálásán alapuló, szisztematikus kutatás végtermékei. Ha ezeket összegyűjtjük és kontextusba helyezzük, továbbá szakszerűen értelmezzük, új tudás jön létre. A természettudományok, a társadalomtudományok és a humán tudományok kutatói egyaránt létrehozhatnak adatokat.<sup>7</sup> A tudományos adatok lehetnek numerikusak, szövegalapúak, hangok, álló- és mozgóképek, továbbá a formátumuk is eltérő lehet.<sup>8</sup>

Az adatok forrásai változatosak lehetnek. A fizikai és az élettudományokban megfigyelések, kísérletek vagy modellek eredményeznek adatokat. A társadalomtudományokban a kutatók maguk is létrehozhatnak vagy a nyilvánosan hozzáférhető iratokból is nyerhetnek adatokat. Az adatfogalom a humán tudományok területén a legkevésbé fejlett. A digitális bölcsészet megjelenése azonban ahhoz vezetett, hogy ott is elterjedtebbé vált ennek a fogalomnak a használata.

Hogy mit használunk fel adatként, azt nagymértékben az adott szakmai érdekközösség határozza meg, viszont egy-egy kutató akár több ilyen közösség tagja is lehet, továbbá ezeknek a közösségeknek eltérőek lehetnek az adat fogalmá-

ról alkotott elképzelései és a velük kapcsolatos gyakorlata. Ezeknek az érdekközösségeknek a határai nem egyértelműek és nem is tartósak.

Az adatok fontos jellemzői között szerepel gyűjtésük céljának specifikus volta. Ez a feltáró kutatásoktól az obszervatóriumokig terjed. Az előbbieket jól körülhatárolt kérdésekre keresik a választ, gyakran egy-egy specifikus helyen, többnyire meghatározott jelenség kapcsán. Ilyen vizsgálatok folyhatnak laboratóriumokban, terepen, vagy ezek kombinációjában. Az obszervatóriumok értéke pedig abban van, hogy hosszú időn át, szisztematikusan végzik el azonos megfigyelések sorát.

Maga a kutatás, többek között, lehet empirikus vagy elméleti jellegű. Ezeknek a kutatásoknak a terminológiája is eltérhet egymástól. Az empirikus kutatás során bizonyos változók nem ellenőrizhetők. Ilyen például az atmoszféra vagy a világegyetem vizsgálata, amelyek során a kutatók nem a „való világból” vett adatokat tanulmányoznak, hanem elméleti modelleken végzik a kísérleteket. A világmindenség megfigyelése egyedi helyen és időben történik, tehát nem rekonstruálható, míg a kísérletek és modellek megismételhetők. Az adatok abban is különböznek egymástól, hogy mennyire munkaigényes a feldolgozásuk, továbbá az adatgyűjtés hatóköre is eltérő lehet. Az utóbbi tekintetben az egyik véglet a konkrét eseményeket és jelenségeket leíró kutatás, míg a másik oldalon az egész rendszereket modellező vizsgálat áll. A köztes fázist a felmérések képviselik, például azért is, mert az online eszközök segítségével akár több száz egyetem hallgatóit is megkérdezhetjük. Ezekben könnyebben megoldható az adatok bizalmas kezelése, mint az interjúkban, így eredményeiket könnyebb megosztani másokkal. Ezzel lehetségessé válhat, hogy helyi vizsgálatok adatait össze tudjuk hasonlítani egymással, ami viszont nem mindig kivitelezhető vagy kívánatos.<sup>9</sup>

Az adatok minősége egy sokdimenziós jellemző. Rése a bizalom, amelynek mértéke számos szubjektív tényezőtől függ. Ezek közé tartozik, hogy az adatokat autentikusnak ítéljük-e meg, elfogadhatónak találjuk-e felhasználásukat vagy alkalmazásukat. Hasonló tényező azoknak a jó

híre, akik az adatokat létrehozták. A minőség megítélését az adatok értékelőinek az elfogultsága és előítéletei befolyásolják. Annak ellenére, hogy igen összetett kérdés, a bizalom problémáját is megemlíthetjük, különösen azért, mert felülírhat más szempontokat. A hitelesség ebben a kontextusban annak durva mértéke, hogy az adatok mennyire a „jó tudományt” tükrözik, és olyan kérdéseket állít-e középpontba, mint az adatgyűjtés eszközeinek megbízhatósága, az elméleti alapok megfelelő volta, az adatok teljessége, pontossága és érvényessége, valamint az adatok ontológiai következetessége. Annak érdekében, hogy a hitelességet meg tudjuk ítélni, az adatoknak érthetőnek kell lenniük.

Az érthetőség értékeléshez nélkülözhetetlen, hogy az adatokat leíró dokumentáció, a metaadatok vagy az adatok eredetére vonatkozó információk formájában elegendő kontextus álljon rendelkezésre, valamint az, hogy az adatok használhatóak legyenek.

A használhatóság megköveteli, hogy az adatok megtalálhatók és hozzáférhetőek legyenek, a fájlformátumok megfelelőek legyenek, az adatok minőségét megítélő egyének megfelelő eszközökkel rendelkezzenek az eléréshez, és biztosítva legyen az adatok megfelelő mértékű integritása.

Az integritás feltételezi, hogy az adatok a bitek szintjén bizonyíthatóan azonosak legyenek egy korábbi, elfogadott és ellenőrzött állapottal.<sup>10</sup> Az integritás az adatok minőségével szembeni elvárás, tehát az a bizonyosság, hogy az adatok teljeseek és hiánytalanok, konzisztensek és helyesek mind intellektuálisan, mind technikai szempontból. Az integritást a létrehozás és a használat bármely fázisában veszélyeztethetik emberi hibák. Mivel az adatok javítása mindig költséges, a legjobb gyakorlat az, ha kezdettől fogva helyes adatokkal dolgozunk.

Bizonyos tudományterületeken nem elégséges a nyers adatokat megosztanunk, a használhatósághoz szükség lehet az adatok létrehozásához használt, elemzéséhez és megosztásához használható eszközök és módszerek leírására is.<sup>11</sup>

Az adatfogalom értelmezése azért is eltérő, mert bizonyos megfigyelések eredményeit adatok-

nak tekintheti az, aki gyűjtötte őket – azonban a befogadó nem így fogja fel. Fordított esetben a kutató birtokában lehetnek anyagok anélkül, hogy tudatosulna benne, milyen értéket képviselnek azok adatként.<sup>12</sup>

## A „nagy adatok”

A digitális adatok egy része nemcsak „nagy” lett, hanem strukturálatlan is. Áramlásuk gyors, így elárasztanak bennünket. *Steve Smith* szerint kezdik betölteni azt a szerepet, amelyet korábban a média játszott.<sup>13</sup> Azt nevezzük nagy adatnak, amelynek a feldolgozása meghaladja a hagyományos adatbázis-rendszerek és szoftverek képességeit, mivel gyűjtése és feldolgozása több időt vesz igénybe, mint amennyit tolerálni tudunk.<sup>14</sup> Mérete azonban „mozgó célpont” abban az értelemben, hogy egy-egy adatállomány esetében (jelenleg) néhány tucat terabájttól több petabájtig terjedhet.<sup>15</sup> A mai értelemben vett nagy adatok esetében sokkal inkább arról van szó, hogy a számítástechnika hatékonyságának maximalizálása folytán lehetséges a nagy adatállományok összegyűjtése, elemzése, összekapcsolása és összevetése. Kereshetünk bennük, és keresztutalásokat alakíthatunk ki közöttük.<sup>16</sup>

Nem lényegtelen az sem, hogy a digitális eszközök széles körű elterjedtsége folytán korunk fogyasztói óriási mennyiségű adatot állítanak elő.<sup>17</sup> Éppen ezért mondható el, hogy a közösségi média új utakat nyitott a társadalmi és a kulturális folyamatok tanulmányozásában. Emberek százainak, sőt ezreinek elképzeléseit, véleményét, gondolatait és érzéseit tanulmányozhatjuk segítségével.<sup>18</sup>

*Boyd* és *Crawford* számos problémára hívja fel a figyelmünket. Rámutatnak arra, hogy más társadalmi-technikai jelenségekhez hasonlóan a nagy adatok jelensége egyszerre vált ki az utópiára és a disztópiára (anti-utópiára) épülő retorikát. Megszületett ugyanis az a mítosz, hogy ezek az adatállományok a tudás és az intelligencia az eddiginél fejlettebb formáját kínálják, olyan meglátásokat téve lehetővé, amelyek eddig nem voltak elképzelhetőek. Mindezt az igazság, az objektivitás és a pontosság aurája veszi körül.

A másik oldalon viszont sokan úgy tekintenek a nagy adatokra, mint ami lehetővé teszi a magánélet titkosságának megsértését, a szabadságjogok csorbitását, az állam és a cégek által gyakorolt ellenőrzés megnövekedését.

Ahogy az egyéb társadalmi-technikai jelenség-gel is megtörténik, a remények és a félelmek gyakran akadályozzák, hogy meglássuk a kisebb és nehezen megfogható változásokat. Mindenesetre kritikus szemmel kell annak a kérdéseit vizsgálnunk, hogy mit jelentenek ezek az adatok, ki fér hozzájuk, miként elemzi őket és milyen célból.

Boyd és Crawford azt is kiemeli, hogy a nagy adatok felhasználása új utat nyit a humán tudományok számára ahhoz, hogy a kvantitatív és objektív módszertannal dolgozó tudomány státuszát követeljék maguknak. Tény, hogy a nagy adatok számos társadalmi jelenséget tesznek számszerűsíthetővé, azonban továbbra is szubjektívek maradnak, ráadásul, amit számszerűsítenek, az nem feltétlenül kerül közelebb az objektív igazsághoz. Figyelmeztetnek arra is, hogy az internetről vett nagy adatállományok gyakran megbízhatatlanok, ezért jó, ha ismerjük adataink jellemzőit és korlátait. Lehet, hogy egy adatállomány mérete milliós nagyságrendű, ami azonban nem jelenti azt, hogy reprezentatív. Hogy statisztikailag érvényes következtetéseket vonjunk le egy-egy adatállományból, tudnunk kell, honnan származik, és melyek a gyengeségei. Tudatában kell lennünk azoknak a tényezőknek, amelyek az értelmezést torzítják. A nagy adatok ezen kívül arra hajlamosítanak bennünket, hogy ott is összefüggéseket lássunk, ahol valójában nincsenek. A több adat tehát nem mindig jobb adat, a nagy adat pedig nem azonos az adatok teljességével.

Gondolnunk kell arra is, hogy a nagy adatokkal kapcsolatos etikai kérdésekről keveset tudunk. Az a tény, hogy valami elérhető, nem feltétlenül teszi felhasználását etikussá. Az adatközlők névtelensége például hamar elillanhat, ami különösen összetett kérdéseket vet fel a közösségi oldalokról vett adatok esetében. Felmerül az a kérdés is, hogy engedélyt kell-e kérnünk felhasználásukhoz az érintettektől, továbbá mit

jelent az egyén számára, ha tudtán kívül vagy az eredetitől eltérő kontextusba helyezve elemzik tevékenységét?<sup>19</sup>

## Az adatok megosztása

Az adatok megosztása viszonylag kevés tudományterületre korlátozódik, és a gyakorlat még ezeken a területeken sem következetes. Az Amerikai Egyesült Államokban a *National Science Foundation (NSF) Science and Technology Center* (a Nemzeti Természettudományi Alapítvány Tudományos és Technológiai Központja) adatokkal kapcsolatos gyakorlatának vizsgálata azt mutatta, hogy az adatokat létrehozó teameken túl kevés adat kerül be valamilyen tágabb körforgásba, sőt kevesen is kérik ezeket az adatokat, aminek számos oka lehet. A kutatóknak például többnyire nincs meg az ez irányú szakértelme, nincsenek eszközei hozzá, valamint ösztönzést sem éreznek arra, hogy másokkal megosszák kutatási adataikat. Bizonyos adatok – a már részben említett – etikai és ismeretelméleti okokból nem oszthatók meg. Néhány esetben pedig nem világos, hogy melyek az adott kutatási projekthez tartozó adatok.

Az *adatmegosztás* szónak számos jelentése van. Meghatározhatjuk úgy, mint az adatok közreadását abból a célból, hogy mások is használhassák őket. Ennek a formája lehet a kérésre történő magánjellegű csere, és egészen addig az igényig terjedhet, hogy az adatok nyilvános adatgyűjteményekbe kerüljenek. Megosztásnak minősül az is, ha egy nyilvános weboldalon helyezük el az adatokat, vagy mellékletként egy folyóirat rendelkezésére bocsátjuk őket. A hasznosság, a megbízhatóság mértéke és az adatok értékes volta azonban sokféle lehet. Bizonyos adatok jellemzője a gazdag strukturáltság és gondozottság, míg mások lehetnek minimális dokumentációval ellátott nyers fájlok. Hasonlóképpen a célközönség is egy szűk szakterület kutatóitól a nagyközönségig terjedhet.<sup>20</sup>

Miközben a kutatók körében kulturális természetű hagyománya van annak, hogy inkább önmagukra és megbízható kollégáikra támaszkodjanak, mint valamilyen központi szolgáltatásra,<sup>21</sup>

a kutatásokat finanszírozó ügynökségek kezdik megkövetelni az adatok közreadását, bár ennek mértéke és a végrehajtás kötelező jellege változó. A *National Institutes of Health*, az Egyesült Államok Nemzeti Egészségvédelmi Intézete például 2003 óta megköveteli, hogy az 500 ezer dollár feletti támogatásokhoz adatkezelési terv készüljön. Az NSF 2001 óta előírja az adatok megosztását az általa finanszírozott kutatások esetében, azonban nem volt következetes abban, hogy ezt meg is követelje. 2010-ben viszont bejelentették, hogy a jövőben minden pályázatnak tartalmaznia kell egy kétoldalas adatkezelési tervet. Egyes folyóiratok is elvárják, hogy a szerzők náluk helyezték el a cikkekhez kötődő adatokat és más kutatási dokumentációt. Az erre vonatkozó irányelvek szigorodtak is az idők folyamán.<sup>22</sup> Nagy-Britanniában hasonló a helyzet. A legtöbb finanszírozó szervezetnek vannak elvárásai, amelyek azonban eltérő természetűek.<sup>23</sup>

A kutatók ösztönzése hatékonyabb lenne, ha valamilyen ellenszolgáltatást kapnának, például úgy, hogy az adatok publikálása és a rájuk történő hivatkozás a tudományos teljesítmény elismerésének része lenne.<sup>24</sup> A motiválás mellett érvényesíthető irányelvekre is szükség volna.<sup>25</sup> Az adatok megosztásának a célja nemcsak az lehet, hogy a közpénzekből finanszírozott kutatások eredményeit hozzáférhetővé tegyünk a nyilvánosság számára. Célba vehetjük a hatékonyabb kutatást is. Mindezek mögött az a – vonzó, de túlzott – állítás rejlik, hogy a számítástechnika segítségével meghaladható a tudomány eddigi eszköztára, tehát immár a „negyedik paradigma” meglétéről beszélhetünk.

Az adatoknak – ahogyan a tudományos kutatásnak is – megvan a maga *életciklusa*. Az utóbbit egyaránt leírták már lineáris és több részműveletről álló ciklikus formában. A kutatás valamilyen ötlettel vagy hipotézissel kezdődik, majd valamilyen publikációval végződik. A kutató a kutatás minden fázisában használ, gyűjt és létrehoz adatokat, ezért az adatok életciklusa és vele gondozásuk nem tér el jelentősen a kutatás életciklusától.<sup>26</sup>

Nem téveszthetjük szem elől, hogy bizonyos

adatok megosztható formában állnak rendelkezésre, míg mások nem. Bizonyos adatok értékét felismeri az adott közösség, míg másokét nem. Egyes kutatók minden adatukat bármikor hajlandók megosztani, míg mások soha semmilyen adatukat nem akarják megosztani. A leggyakrabban azonban az, hogy bizonyos adataikat néha hajlandóak megosztani.<sup>27</sup>

Ahhoz, hogy az adatok hatékonyan legyenek megoszthatók, számos követelménynek kell megfelelniük. Konzisztens mutatók segítségével a jövőben is megtalálhatóaknak kell lenniük. Ehhez őket meg kell őriznünk, és hosszú távon hozzáférhetővé kell tennünk. Minőségüknek lehetővé kell tennie, hogy használhatók legyenek.<sup>28</sup>

Az adatok megosztásához tudnunk kell a következőket anélkül, hogy a kérdésekre egyértelmű, igen vagy nem válaszokat adhatnánk:

- ☉ létezik-e az adatkészlet,
- ☉ hol található,
- ☉ sérült vagy elavult-e,
- ☉ van-e engedélyünk a használatára,
- ☉ eléggé szabványos-e ahhoz, hogy elfogadható mértékű erőfeszítéssel használni tudjuk,
- ☉ elég világos-e, hogy mit reprezentál,
- ☉ eredete és hibaarányai ismertek és elfogadhatók-e, tehát megbízhatónak tekinthető-e,
- ☉ célszerű-e felhasználnunk céljaink elérésére.

Ezek a kérdések láncolatot alkotnak, tehát ha létezik az adatkészlet, lehet, hogy nem férhetünk hozzá. Ha hozzáférünk, előfordulhat, hogy nem használható. Ha használható, nincs a használatára engedélyünk, és így tovább. Bármelyik probléma magakadályozhatja a felhasználást. Amikor pedig elhárítjuk az egyik akadályt, egy másik még mindig lehetetlenné teheti a felhasználást.<sup>29</sup>

Az adatok megosztását jól segíthetik a *nyílt adatok*. Az adatok ebben az értelemben újrafelhasználható kutatási eredményeket jelentenek, és a kutatások finanszírozóinak közelmúltban tett lépései azt mutatják, hogy ez a nyitottság hamarosan kötelező feladattá válik.<sup>30</sup>

## Az adatgondozás

Annak érdekében, hogy az adatok megtalálhatók, elérhetők és használhatók legyenek, ki kell alakítanunk a tárolás és a visszakeresés megfelelő technikai és szervezeti infrastruktúráit. Bár a nagy adatállományok archiválásának megvan a technológiája, ez nem egyszerű feladat, mert adatainkat könnyen hozzáférhetővé és jól leírhatóvá kell tenni.<sup>31</sup>

A Charles Bailey által összeállított bibliográfia arról tanúskodik, hogy az adatok (kuratori) gondozásának kérdésével számtalan (angol nyelvű) írás foglalkozik.<sup>32</sup>

Ha a digitális gondozás meghatározását<sup>33</sup> az adatokra alkalmazzuk, akkor az adatgondozást úgy definiálhatjuk mint azon eljárások összességét, amelyek ahhoz szükségesek, hogy digitális kutatási adatokat egész életciklusukon át és a jövőben is, a jelen és az elkövetkező felhasználói generációk számára fenntartsunk. Az adatgondozásnak biztosítania kell a tudományos eredmények reprodukálhatóságát, újrafelhasználását, továbbá hozzáadott értéket kell létrehozni.<sup>34</sup>

A gondozásnak a digitális források jövőbeni használhatóságát olyan módon kell biztosítania, hogy a kutatók tudhassák, hogy micsoda és honnan származik az adott objektum, miért fontos, és hogyan kell használni. Az ehhez szükséges kontextust dokumentációk vagy metaadatok, valamint ezek keveréke adja meg.

Mivel a kutatók – ahogy azt már említettem – nincsenek megfelelően motiválva arra, hogy az általuk előállított adatokat gondozzák, ráadásul eszközeik sincsenek hozzá és idejük sincsen rá, emellett szakértelmük sincs ezen a területen, a kulturális örökségi intézmények gondozzák az adatokat. Ezt azonban csak utólag tehetik meg. A könyvtárakban dolgozó adatkurátoroknak ezért szorosan együtt kell működniük az adatok létrehozóival, hiszen az ő emlékezetükben még frissen él az adatok kontextusa, amelyet egyébként is jobban ismernek, mint a folyamatba utólagosan belépő kurátorok.<sup>35</sup>

A gondozás során szükség lehet arra, hogy döntünk adatok törléséről, ami bizalmas adatok esetében biztonságos megsemmisítésüket jelentheti.

Az adatok gondozásának költségei meg is követelik, hogy időről időre felülvizsgáljuk, mely adatokat őrzünk meg. Ezt kell tennünk akkor is, amikor az adatokat új formára alakítjuk, vagy az elavulástól védendő új hardver- és szoftverkörnyezetbe visszük át.<sup>36</sup>

- ⊙ Az adatgondozáshoz kötődő *irányelvek*nek a következő kérdésekre célszerű válaszolniuk:
- ⊙ Kinek a tulajdonát képezik az adatok?
- ⊙ Milyen követelményeket támasztanak más érdekeltek (finanszírozó szervezetek, kiadók, az adott kutatás együttműködő partnerei)?
- ⊙ Mely adatokat kell megőriznünk? (Nem őrizhetünk meg minden adatot, mivel megőrzésük anyagi erőforrásokat igényel, tehát csak az arra érdemes adatokat tarthatjuk meg: sikertelen kísérletek vagy nyilvánosan elérhető adatok másodlagos elemzéseinek adatait például nem érdemes megőrizni.)
- ⊙ Ki döntsön arról, hogy mely adatokat őrizzünk meg? A döntés a kutató, vagy más kutatók és szakértők kezében van-e? Mely adatállományok esetében valószínű, hogy a jövőben szüksége lesz újrafelhasználásukra?
- ⊙ Milyen hosszan őrizzük meg az adatokat? Milyen metrikákat alkalmazhatunk ennek megállapítására?
- ⊙ Hogyan őrizzük meg az adatokat? Milyen formátumban történjen a megőrzés, és milyen metaadatokkal, valamint más kiegészítő információval kell ellátnunk az adatállományokat?
- ⊙ Vannak-e etikai megfontolások? Hogyan fogjuk kezelni a szellemi tulajdonjogokkal, a magánélet titkosságával, a kényes adatokkal kapcsolatos problémákat?
- ⊙ Hogyan férhetők hozzá az adatok? Az ezzel kapcsolatban felmerülő számos kérdés közül kiemelésre kívánczik, hogy meg kell keresni a hozzáférés metrikáit.
- ⊙ Mennyire legyenek nyíltak az adatok? Ha nincsenek jogi korlátok, az adatok teljesen nyíltak lehetnek. A nyíltság azonban késleltethető is, vagy bizonyos adatok – megfelelő indoklással – sohasem kerülhetnek nyilvánosságra vagy megosztásra.

- 👁 Ki viselje a költségeket? A különböző, magával a kutatással együtt járó és attól független finanszírozási modellek léteznek vagy teremthetők meg ezen a téren.
- 👁 Milyen alternatívái vannak a helyi adatgondozásnak? Nem minden adatot kell saját intézményünkben tárolni, vannak ugyanis nemzetközi, nemzeti és tudományterületi adatközpontok. Lehetséges, hogy az adatokat máshol tesszük elérhetővé, mint ahol hosszú távon őrizzük őket.<sup>37</sup>

### Az adatokra való hivatkozás

Az adatokra való hivatkozás lehetővé teszi a publikált cikkek alapját képező adatok visszakeresését, azonosítását és eredetének igazolását. Egyelőre azonban ritkán és véletlenszerűen találkozunk vele.<sup>38</sup>

Az adatok dokumentálása ma még következtelen és nem követ szabványokat.<sup>39</sup> Az adatokra való hivatkozás területén sincsenek kikristályosodott szabványok. Vannak viszont kezdeményezések. Ilyenek a *Datacite* (<http://www.datacite.org/>) és a *DataVerse Network* <http://thedata.org/>. Az OECD hivatalos kiadványban (white paper) foglalkozik a kérdéskörrel.<sup>40</sup>

Talán a DOI (Digital Object Identifier) metaadat-szabványra épülő szabványok lesznek a leghasznosabbak. Ilyen elképzelés a DataCite metaadat-kezdeményezése.<sup>41</sup> Ausztráliában, a Griffith Egyetemen 2012-ben megkezdett adat-hivatkozási projekt is a DOI-ra épít.<sup>42</sup>

### Az adatok és a könyvtár

A létrejövő adatok pusztá mennyisége, az eszközök, infrastruktúra, a szabványos folyamatok jelenlegi hiánya mellett az adatok gondozásában szakértelemmel rendelkező személyzet sem áll rendelkezésre, ami a digitális alapokon nyugvó kutatás fejlődése számára komoly fenyegetést jelent.<sup>43</sup>

Közben tudnunk kell, hogy bizonytalanság van abban a tekintetben, hogy miként vállaljanak részt az adatok kezeléséből az egyetemek vagy a kutatás finanszírozásában részt vevő szervezetek. A könyvtárak viszont jó helyzetben vannak

ahhoz, hogy kulcsszereplővé váljanak, mivel tapasztalatuk van a válogatásban, a metaadatok használatában, repozitóriumok és más gyűjtemények működtetésében, a megőrzésben, a gondozásban és a hozzáférés biztosításában.<sup>44</sup>

A közelmúltban megtapasztalhattuk, hogy az információközvetítők szerepét technológiai eszközök, mindenekelőtt a keresőgépek vették át. Ez különösen igaz a kutatók információ-használatára. Márpedig, ha a könyvtárosok megint a kutatók munkáját segítő szolgáltatásokat akarnak kínálni, arra jó lehetőséget nyújt a kutatási adatok gondozása.<sup>45</sup>

Egyelőre nem tudjuk, lesz-e a kutatóknak eleendő ideje és hajlandósága, hogy együttműködjenek az egyetemi könyvtárakkal az adatok gondozása terén, pedig éppen ezek a könyvtárak vannak egyedülálló helyzetben ahhoz, hogy megfelelően foglalkozzanak az adatok gondozásával.<sup>46</sup> Számos megoldandó feladat áll előttük: rosszul szervezett adatgyűjtemények, gyenge visszakeresési eszközök, a belső adatállományok hozzáférhetőségének és a megtalálhatóságának hiánya, valamint a szellemi tulajdonnal kapcsolatos kérdések.<sup>47</sup>

A gyakorlati kérdések között pedig ott vannak az adatok a tárolásának, visszakeresésének, leírásának, hosszú távú megőrzésének, az adatkurátorok és a felhasználók oktatásának problémái, valamint a szabályok kidolgozásának szükségessége.<sup>48</sup> A könyvtárosok fontos feladata, hogy előmozdítsák az adatokra való hivatkozás, mint a normatív gyakorlat ügyét bármely szakterületen.<sup>49</sup>

Amerikai és kanadai könyvtárosok munkakörük ellátása során rendszeresen vagy alkalmanként szolgáltatnak kutatási adatokat, ami azonban többségüknek még nem rendszeres feladata. Ugyanakkor meg vannak győződve arról, hogy rendelkeznek a szükséges tudással és készségekkel, hogy ilyen szolgáltatást nyújtsanak a jövőben.<sup>50</sup>

A könyvtárosok új szerepet is betölthetnek „kutatási információs szakember” elnevezéssel. Ebben a szerepkörben együtt dolgoznak a kutatói teamekkel, segítik az adatok kezelését, megőrzését, a keresést és a bibliometriai tev-

kenységet.<sup>51</sup> Az ilyen és hasonló feladatokat ellátó könyvtáros valójában *adatkurátor*, akit a következők jellemeznek:

- ⊗ digitális dokumentumokat választ ki a hosszú távú megőrzésre,
- ⊗ szakértelemmel rendelkezik arra vonatkozóan, hogy mi az egyes digitális entitások rendeltetése egy adott közösség számára, és az milyen hatással van megőrzésére,
- ⊗ ismeri a különböző digitális objektumok szerkezetét, és meghatározza, hogy azok milyen gondozást igényelnek,
- ⊗ ismeri a tárolásra és a megőrzésre vonatkozó irányelveket és gyakorlatot, amelyek biztosítják a digitális dokumentumok megbízhatóságának és elérhetőségének fennmaradását,
- ⊗ ismeri a megfelelő hozzáférést, tárolást és adat-helyreállítást lehetővé tevő információs infrastruktúrával kapcsolatos követelményeket,
- ⊗ az információtechnológiai szakemberekkel együttműködve problémákat diagnosztizál és old meg a digitális objektumok folyamatos hozzáférhetőségének biztosítása érdekében,
- ⊗ figyeli a fájlformátumok, a hardver és a szoftver elavulását és az új formátumok kialakulását,
- ⊗ biztosítja, hogy olyan módszerek kerüljenek alkalmazásra, amelyek segítik a különböző alkalmazások és megőrzési technológiák együttműködési képességét különböző helyeken lévő felhasználók esetén,
- ⊗ megvizsgálja a megőrzendő adatok eredetét, és gondoskodik annak megfelelő dokumentálásáról,
- ⊗ képes folyamatosan felmérni a digitális objektumok hitelességét, integritását és pontosságát.

Vannak kifejezetten menedzsment-jellegű feladatai is:

- ⊗ digitális gondozási projekteket tervez, valósít meg, továbbá figyeli ezeket,
- ⊗ ismeri, valamint a jelenlegi és a potenciális érdekeltek tudomására hozza a digitális gondozás gazdasági előnyeit,

- ⊗ digitális gondozási előírásokat, gyakorlatot és szolgáltatásokat alakít ki, továbbá tudatában van annak, hogy azok milyen hatást gyakorolnak a digitális objektumok létrehozóira és (újra)felhasználóira,
- ⊗ együttműködéseket hoz létre és tart fent különböző érdekeltekkel annak érdekében, hogy a digitális gondozás céljai elérhetőek legyenek,
- ⊗ szervezi a munkatársak oktatását és képzését, valamint minden olyan tevékenység támogatását, amely segíti őket a digitális gondozás új fejleményeinek alkalmazásában,
- ⊗ tudatában van annak, hogy lépést kell tartani a digitális gondozás területén bekövetkezett nemzetközi fejleményekkel, továbbá ismeri azokat a szakmai hálózatokat, amelyek ezt lehetővé teszik,
- ⊗ ismeri és kommunikálni tudja a digitális entitások sérülésének vagy elvesztésének a kockázatait,
- ⊗ szervezi és menedzseli a metaadat-szabványokat, hozzáférési és hitelességi procedúrákat,
- ⊗ ismeri az erre a területre vonatkozó minőségbiztosítási szabványokat, és alapos mérlegelés után dönt ezek alkalmazásáról vagy elvetéséről, jogi szakemberekkel együttműködve betartja a jogi és egyéb előírásokat. Ragaszkodik ezekhez, amikor dönt a megőrzésről, a digitális objektumok használatáról és újrahasznosításáról.<sup>52</sup>

Álláshirdetések elemzésre alapozva *Li Si* és munkatársai a következő feladatokat jelölték ki számára:

- ⊗ konzultációt és tájékoztatási szolgáltatásokat nyújt a tudományos kutatás és adatgondozás területén,
- ⊗ felvilágosítást kér, és összegyűjti a visszacsatolást a kutatók által az adatgondozással szemben támasztott követelményekről,
- ⊗ adatgondozási oktatást és képzést nyújt a felhasználóknak,
- ⊗ metaadat-szabványokat hoz létre, és ellátja ezekkel a kutatási adatokat,



- 👁 segíti a felhasználókat a kutatási adatok gyűjtésében, elemzésében és tárolásában,
- 👁 az adatgondozást elősegítő irányelveket és eljárásokat dolgoz ki,
- 👁 részt vesz az adatgondozással kapcsolatos programokban, kezdeményezésekben és szervezetekben.

Ezek megvalósításához a következő készségeket és képességeket látják szükségesnek:

- 👁 a csapatmunka és az együttműködés képessége;
- 👁 szóbeli és írásbeli kommunikációs készségek;
- 👁 interperszonális készségek;
- 👁 az adatgondozási eszközök és technológiák bizonyított ismerete;
- 👁 az általános metaadat-szabványok ismerete;
- 👁 az adatgondozás trendjeinek ismerete, és hajlandóság, hogy ezekre figyeljen;
- 👁 a kvantitatív kutatási módszerek ismerete;
- 👁 a tudományos adatforrások alapos ismerete.<sup>53</sup>

A LIBER, az Európai Tudományos Könyvtárak Egyesülete ajánlásaiban kifejti, hogy a könyvtárak:

1. Nyújtsanak adatkezelési támogatást, ideértve a pályázók számára készített adatkezelési terveket, a szellemi tulajdonjoggal kapcsolatos tanácsadást és információs anyagokat. Segítsék az oktatókat adatkezelési tervek elkészítésében és az adatkezelésnek a tantervekbe történő integrálásában;
2. Kezdzének metaadatokat és az adatokra vonatkozó szabványokat kialakítani, továbbá szolgáltatassanak metaadatokat a kutatási adatokhoz;
3. Hozzanak létre adat-könyvtárosi állásokat és fejlesszék munkatársaik adat-könyvtárosi készségeit;
4. Aktívan vegyenek részt a kutatási adatokra vonatkozó intézményi irányelvek kidolgozásában, ide értve a források tervezését, ösztönözzék a nyílt adatokra vonatkozó irányelveket és vegyenek át ilyeneket, ahol ez helyénvaló a kutatási adatok életciklusa szempontjából;

5. Legyenek összekötetésben és teremtsenek partnerséget a kutatókkal, kutatócsoportokkal, adat-archívumokkal, adatközpontokkal annak érdekében, hogy elősegítsék az adatok elérése, megtalálása és megosztása együttműködni tudó (interoperábilis) infrastruktúrájának létrejöttét;
6. Támogassák a kutatási adatok életciklusát azzal, hogy tárolási, az adatok megtalálását és az állandó hozzáférést segítő szolgáltatásokat nyújtanak;
7. Segítsék elő az adatokra való hivatkozás ügyét azzal, hogy állandó azonosítókat használnak a kutatási adatokhoz;
8. A rendelkezésre álló infrastruktúra függvényében hozzanak létre intézményi adatkatalógust vagy adat-repozitóriumot;
9. Vegyenek részt szakirányú adatkezelési tevékenységekben;
10. Az intézményi IT egységekkel együttműködve kínálják fel, vagy közvetítsék a dinamikus és a statikus adatok biztonságos tárolását, vagy felkutatják megfelelő felhő-szolgáltatások használatát.

Az ajánlás kidolgozói figyelmeztetnek rá, nem szükséges, hogy a könyvtárak az ajánlás összes pontját azonnal teljesítsék, vagy az adatokkal kapcsolatos szolgáltatásokat egyszerre próbálják nyújtani. Helyesebb, ha a bevált gyakorlatot átvevéve mások sikereiből és kudarcaiból okulnak. Az adatok kezelése nem megy egy kaptafára, tehát a különböző könyvtárak szerepe eltérő. Valószínűleg egyetlen könyvtárnak sem kell minden szolgáltatást nyújtania, és bizonyosan nem kell azonnal megtennie ezt.

A könyvtárosok átképzése rendkívül fontos, mivel csak néhány könyvtár engedheti meg magának azt a luxust, hogy az adatok kezeléséhez erre szakosodott új munkatársakat vegyen fel.<sup>54</sup>

### **Az adat-írástudás**

A kutatási adatok talán legtagabb szakmai kontextusát a néha adat-információs írástudásnak nevezett adat-írástudás (data literacy) jelenti. Z. Karvalics László szerint „*a bejövő jelek tömegtermelése olyan nagyságrendekbe emelkedett,*

*ahol számításteljesítménnyel (feldolgozási, átviteli és tárolási kapacitásokkal és mesterséges intelligencia-megoldásokkal) még sikerül követni a mennyiségi kihívást, de valódi elemző és megismerő erőt már nem lehet mellérendelni. A folyamatokból egyre inkább és egyre látványosabban az értelmezésre képes emberi agyak, vagyis a jelekből információt, jelentést és kontextust generálni képes „humán intelligencia” hiányzik.”*<sup>55</sup> Nyilvánvalóan ezért (is) van szükség adat-írástudásra. Mivel a könyvtárosok foglalkoznak az információs műveltséggel, szerepük természetes módon kiterjeszhető az adat-írástudásra.<sup>56</sup>

Az adat-írástudásnak – mint minden hasonló fogalomnak – számos meghatározása van. Legegyszerűbben úgy határozhatjuk meg, mint az adatok megértésének, használatának és kezelésének képességét.<sup>57</sup> Aki ezzel az írástudással (műveltséggel) rendelkezik, tudja, miként kell adatokat keresni, azokat megfelelően szűrni, feldolgozni, létrehozni és szintetizálni.<sup>58</sup> Ennek megfelelően adat-írástudás alatt az adatkörnyezet uralásának és egyszerűsítésének képességét is érhetjük.<sup>59</sup>

Némileg más hangsúlyokat kap, ha azt mondjuk, hogy az adat-írástudás magában foglalja annak ismeretét, mit is értünk adat alatt, miként vonunk le adatokból következtetéseket, továbbá része annak felismerése, mikor használják az adatokat félrevezető vagy nem megfelelő módon. A fogalom egyesíti magában a kutató kettős szerepét, tehát azt hogy egyaránt előállítója és fogyasztója az adatoknak.<sup>60</sup>

Az adat-írástudás oktatása két különböző célt szolgál. Mindkettő az elektronikus adatok tág értelemben vett kezelésének elsajátításán alapul. Az egyik azonban az, hogy hatékony kutató legyen valakiből, a másik pedig, hogy az elektronikus adatkezelés szakemberévé váljon.

Az információs műveltség mindig is magába foglalta az adatok használata iránti érdeklődést, bár az adatoknak információvá alakítására korlátozódott. Mivel azonban a kutatási adatok gondozása sohasem csak a nyers adatokkal foglalkozik, hanem az adatok felhasználásával, újrafelhasználásával, valamint azok kontextusba

helyezésével, ebben az összefüggésben az adat és az információ nem áll távol egymástól.<sup>61</sup> Az adat-írástudás természetesen kevésbé figyel a szakirodalmi jellemzőkre, inkább az adatok gyűjtésének, feldolgozásának, kezelésének, értékelésének funkcionális képességeit állítja középpontba.<sup>62</sup>

Az adat-írástudás oktatásában nehézséget okoz, hogy az egyetemi hallgatók oktatásának az a követelményrendszere,<sup>63</sup> amelyet az ACRL (Association of College and Research Libraries) állított össze, alapvetően az információ fogyasztóira fókuszál, míg az adatmegosztás az adatok létrehozóit (is) érinti. Mindazonáltal az információs műveltség és kompetencia követelményrendszerének számos pontja értelmezhető az adat-írástudás oktatása szempontjából is.<sup>64</sup> Ezek a következők:

*Az információs műveltséggel rendelkező hallgató meg tudja határozni és meg tudja fogalmazni a szükséges információ természetét és mértékét.*

Ez a követelmény alapján véve az adatokra is vonatkozik, amikor megfogalmazódik benne, hogy a hallgató tudja, hogy az információ az elsődleges forrásokból származó nyers adatokból is megszerkeszthető. Az adat előállítója számára az adatok iránti potenciális igény természetének és mértékének azonosítása az adatok hatékony megosztásának, újrafelhasználásának, kezelésének és megőrzésének az alapja.<sup>65</sup>

*Az információs műveltséggel rendelkező hallgató hatékonyan és eredményesen jut hozzá az igényelt információhoz.*

A hallgató használja az ismertebb szakterületi és általános adat-repozitóriumokat. Ismeri az adatok elérésének formátumait, és tudja, hogy milyen szolgáltatások érhetőek el annak érdekében, hogy eredményesen és hatékonyan szerezzen meg adatokat. Az adat-írástudó hallgatónak ismernie kell, hol található adatok, és képesnek kell lennie ezek begyűjtésére, átalakítására, esetleg összeolvasztására, valamint betáplálására elemző vagy megjelenítő eszközökbe, amelyek akár további formátumokat is megkövetelhetnek.

*Az információs műveltséggel rendelkező hallgató kritikusan értékeli az információkat és azok forrásait, a kiválasztott információkat beépíti saját tudásbázisába és értékrendszerébe.*

Amikor az adatokat értékeli, az információsan művelt hallgató ismeri és kritikusan értékeli a forrást. Meg kell határozni, hogy az adatokat szolgáltató kutatócsoport jó hírek örvend-e, továbbá az adat-repozitóriumban van-e minőségellenőrzés. Vizsgálnia kell a relevanciát, a kompatibilitást és a metaadatok minőségét.

*Az információs műveltséggel rendelkező hallgató – egyénileg vagy egy csoport tagjaként – hatékonyan tudja alkalmazni az információt egy meghatározott cél elérésének érdekében.*

Az információ alkalmazása a kommunikálást is jelenti a megfelelő kommunikációs technológiák alkalmazásával, ami magába foglalja az adatok megosztását és gondozását is.

*Az információs műveltséggel rendelkező hallgató ismeri az információ használatát körülvevő gazdasági, jogi és társadalmi problémákat, az információkhoz való hozzáférést és azok alkalmazását etikusán és jogszerűen végzi.*

A hallgatónak nemcsak azt kell meghatározni, hogy mikor és hogyan oszta meg valakivel adatokat, hanem saját adatait is dokumentálnia kell.

## Következtetések

Ne felejtsük el, hogy a könyvtárak egy átmenet korai szakaszát élik meg, amikor a kutatási adatok kezelésében betöltendő szerepüket formálják.<sup>66</sup> Közben az adatokkal való „foglalatoskodás” területén nemcsak könyvtárosok vannak jelen, hanem szociológusok, közgazdászok, statisztikusok és informatikusok, hogy csak néhány szakterület szakembereit említsük. Érdemes tehát *Nicholas Weber* tanácsát megfogadnunk: „*Amit a könyvtártudomány tud a hivatkozásról, az információ visszakereséséről és az információs viselkedésről, azt az adat-intenzív paradigmára is használhatja, és alkalmaznia kell.*”<sup>67</sup>

## Irodalom

1. BORGMAN, Ch.: *Scholarship in the digital age: Information, infrastructure, and the internet.*, Cambridge, MA: MIT Press, 2007.
2. Z. KARVALICS László: A cyber-infrastruktúra mint aktuális kihívás, és mint tudományszociológiai probléma = *Magyar Tudomány*, 167. évf. 2007. 4. sz. 475–489. p.
3. CARLSON, J. [et al.]: Determining data information literacy needs: A study of students and research faculty = *Portal: Libraries and the Academy*, vol. 11. 2011. no. 2. p. 629–657.
4. DAVIES, A. – FIDLER, D. – GORBIS, M.: *Future work skills 2020*. Palo Alto, CA: Institute for the Future, 2011. <http://www.iff.org/our-work/global-landscape/work/future-work-skills-2020/> (2014. január 4.)
5. BORGMAN, Ch.: The conundrum of sharing research data = *Journal of the American Society for Information Science and Technology*, vol. 63. 2012. no. 6. p. 1059–1078.
6. ROWLEY, J.: The wisdom hierarchy: representations of the DIKW hierarchy = *Journal of Information Science*, vol. 33. 2007. no. 2. p. 163–180.
7. PRYOR, G.: Why manage research data? In: *Managing research data*. Ed. Graham Pryor. London: Facet, 2012. p. 1–16.
8. KOWALCZYK, S. – SHANKAR, K.: Data sharing in the sciences = *Annual Review of Information Science and Technology*, vol. 45. 2011. no. 1. p. 247–294.
9. BORGMAN, i. m.
10. GIARLO, M.: Quality hubs = *Journal of Librarianship and Scholarly Communication*, vol. 1. 2013. no. 3. 1–10.
11. KOWALCZYK – SHANKAR, i. m.
12. BORGMAN, i. m.
13. SMITH, S.: Is data the new media? = *EContent*, vol. 36. 2013. no. 2. p. 14–19.
14. GORDON-MURNANE, L.: Big data: A big opportunity for librarians = *Online*, September – October 2012, 30–34.
15. MANOVICH, L.: Trending: The promises and the challenges of big social data In: *Debates in the digital humanities*. Ed. M. K. Gold. Minneapolis: Univ. of Minneapolis Pr., 2012. <http://dhdebates.gc.cuny.edu/debates/text/15> [2014. január 4.]
16. BOYD, D. – CRAWFORD, K.: Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon = *Information, Communication & Society*, vol. 15. 2012. no. 5. p. 662–679.

17. GORDON-MURNANE, i. m. Publishing, 2009 <http://dx.doi.org/10.1787/787355886123> [2014. január 4.]
18. MANOVICH, i. m.
19. BOYD – CRAWFORD, i. m.
20. BORGMAN, i. m.
21. PRYOR, i. m.
22. BORGMAN, i. m.
23. PRYOR, i. m.
24. MOONEY, H. – NEWTON, M. P.: The anatomy of a data citation: Discovery, reuse, and credit = *Journal of Librarianship and Scholarly Communication*, vol. 1. 2012. no. 1. p. 1–14.
25. KOWALCZYK – SHANKAR, i. m.
26. PRYOR, i. m.
27. BORGMAN, i. m.
28. KOWALCZYK – SHANKAR, i. m.
29. BUCKLAND, M.: Data management as bibliography = *Bulletin of the American Society for Information Science and Technology*, vol. 37. 2011. no. 6. p. 34–37.
30. LEWIS, D. W.: From stacks to the web: The transformation of academic library collecting = *College & Research Libraries*, vol. 74. 2013. no. 2. p. 159–177.
31. KOWALCZYK – SHANKAR, i. m.
32. BAILEY, Ch.: Digital curation bibliography: preservation and stewardship of scholarly works. 2012. <http://www.digitalscholarship.org/> [2014. január 4.]
33. BEAGRIE, N.: Digital curation for science, digital libraries, and individuals = *International Journal of Data Curation*, vol. 1. 2006. no. 1. p. 3–16.
34. Harvey, R.: *Digital curation: A how-to-do-it manual*. New York: Neil– Schuman Publ., 2010. – Idézi Madrid.
35. GIARLO, i. m.
36. PRYOR, i. m.
37. ERWAY, R.: Starting the conversation: university-wide research data management policy = *Educause Review Online*, December 6, 2013. <http://www.educause.edu/ero/article/starting-conversation-university-wide-research-data-management-policy> [2014. január 4.]
38. MOONEY – NEWTON, i. m.
39. CARLSON et al., i. m.
40. GREEN, T.: *We need publishing standards for datasets and data tables*. OECD Publishing White Papers. Paris: OECD Publishing, 2009 <http://dx.doi.org/10.1787/787355886123> [2014. január 4.]
41. MOONEY– NEWTON, i. m.
42. SIMONS, N. – VISSER, K. – SEARLE, S.: Growing institutional support for data citation = *D-Lib Magazine*, vol. 19. 2013. no. 1. p. 11–12. <http://www.dlib.org/dlib/november13/simons/11simons.html> [2014. január 4.]
43. CARLSON et al., i. m.
44. SHEN, Y. – VARVEL, V. E.: Developing data management services at the Johns Hopkins University = *The Journal of Academic Librarianship*, vol. 39. no. 6. 2013. p. 552–557.
45. PRYOR, i. m.
46. GIARLO, i. m.
47. GORDON–MURNANE, i. m.
48. KOWALCZYK – SHANKAR, i. m.
49. MOONEY – NEWTON, i. m.
50. TENOPIR, C. et al.: Academic librarians and research data services: preparation and attitudes = *IFLA Journal*, vol. 39. 2013. no. 1. p. 70–78.
51. FEDERER, L. The librarian as research informationist: a case study = *Journal of the Medical Library Association*, vol. 101. 2013. no. 4. p. 298–302.
52. MADRID, M. M: A study of digital curator competences: a survey of experts = *International Information & Library Review*, vol. 45, 2013. no. 3–4. p. 149–156.
53. SI, L. et al.: The cultivation of scientific data specialists: Development of LIS education oriented to e-science service requirements = *Library Hi Tech*, vol. 31. 2013. no. 4. p. 700–724.
54. Christensen-Dalsgaard, B. et al: Ten recommendations for libraries to get started with research data management. Final report. LIBER, 2012. <http://www.libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management> [2014. január 4.]
55. Z. KARVALICS László: Az adatsilóktól a tudomány kontrollforradalmáig = *Magyar Tudomány*, 169. évf. 2007. 3. sz. 352–363. p.
56. HAENDEL, M. – VASILEVSKY, N. – WIRZ, J.: Dealing with data: a case study on information and data management literacy = *PLoS Biology*, vol. 10. 2012, no. 5, e1001339 <http://www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.1001339> [2014. január 4.]
57. QIN, J. – D'IGNAZIO, J.: Lessons learned from a two-

- year experience in science data literacy education. In: Proceedings of the 31st Annual IATUL Conference, June 20-24. 2010. 2. <http://docs.lib.purdue.edu/iatul2010/conf/day2/5> [2014. január 4.]
58. JOHNSON, C.A.: The information diet: a case for conscious consumption. Sebastopol, CA: O'Reilly Media, 2012.
59. Z. KARVALICS László: Információs kultúra, információs műveltség - egy fogalomcsalád értelme, terjedelme, tipológiája és története = Információs Társadalom, 12. évf. 2012. 1. sz. 7-43. p.
60. CARLSON et al., i. m.
61. SCHNEIDER, R.: Research data literacy. In: Worldwide commonalities and challenges in information literacy research and practice. European Conference on Information Literacy. Istanbul, Turkey, 2013. Rev. selected papers. Cham: Springer International Publishing, 2013. p. 134-140.
62. QIN – D'IGNAZIO, i. m.
63. American Library Association, Association of College and Research Libraries, Information Literacy Competency Standards for Higher Education. Chicago, IL: ACRL, ALA 2000. <http://www.ala.org/ala/mgrps/divs/acrl/standards/standards.pdf> [2014. január 4.] – Magyarul: Az információs műveltség és kompetencia követelményrendszere a felsőoktatásban In: Információs műveltség és oktatásügy. Nemzetközi szemle. Szerk. Csík Tibor. Bp.: Országos Pedagógiai Könyvtár és Múzeum, 2006. 84-96. p. <http://www.opkm.hu/download/lmOk.pdf> [2014. január 4.]
64. CARLSON et al., i. m.
65. Itt és a következőkben Carlson és munkatársai megállapításait idézve az említett *követelményrendszer* magyar változatának szövegét használok.
66. TENOPIR et al., i. m.
67. WEBER, N.M.: The relevance of research data sharing and reuse studies = Bulletin of the American Society for Information Science and Technology, vol. 39. 2013. no. 6. p. 23-26.

*Beérkezett: 2014. január 7.*

## Veszteségeink

### Vargha Dénes (1928–2014)

2014. március 15-én, 86. életévében elhunyt a híres kunszentmiklósi Vargha-család nyelvész tagja. A könyvtárosok a formális nyelvek, a számítógépes nyelvészet, a tezausz szerkesztés szakértőjeként és négy meghatározó jelentőségű mű szerzőjeként ismerik:

A dokumentáció nyelvészeti kérdései. Szemelvénygyűjtemény 1. (1966), Dokumentáció és nyelvészet.

Az értelmi összefüggések formalizálása (1969), Információkereső tezausz készítésének módszertana (1969), Petőfi S. Jánossal közösen: A tezausz-kérdés jelenlegi helyzete, különös tekintettel a tudományos, műszaki-gazdasági tájékoztatásra (1969), 1976-ban még részt vett a magyar tezausz szabvány kidolgozásában.

Ungváry Rudolf hiteles szavai (<http://katalist-to-feed.blogspot.hu/2014/03/katalist-in-memoriam-vargha-denes.html>) rávilágítanak, mit is köszönhet neki a könyvtári szakma.



### Vajda Erik (1930–2014)

2014. május 14-én hosszú betegség után elhunyt Vajda Erik az OMKDK, majd az Országos Műszaki Könyvtár és Információs Központ (1963–1996) vezető munkatársa. 1986 és 1994 között az UNIDO szakértője, majd az UNISIST tanácsadó bizottságának volt tagja, a Magyar Országos Közös Katalógus (MOKKA) projektszervezője. 1966-tól a Magyar Könyvtárosok Egyesületének tagja, 1994 és 1998 között elnöke volt. Alapítója volt a Műszaki Könyvtáros Szekciónak, amelyet tíz éven át elnökként is vezetett.

Vajda Erikről lapunk következő számában emlékezünk meg.

(A hetvenedik születésnapja alkalmából a Könyv, Könyvtár, Könyvtáros 2000. 7. számában jelent meg vele egy beszélgetés: <http://ki.oszk.hu/3k/2012/09/%e2%80%a6ugye-nem-kell-bizonyitanom-a-konyvtarak-es-konyvtarosok-kulcsszerepet%e2%80%a6/>)

## Őrzők Éjszakája – A könyv elpusztíthatatlan

Múzeumok éjszakája az OSZK-ban  
(2014. június 21-én 18 órától – június 22-én 2 óráig)

Az Országos Széchényi Könyvtár gyűjteményei és műhelyei több mint kétszáz éve őrzői az írott magyar kultúra teljességének. Ezen az éjszakán nemcsak a gyűjtemények izgalmas darabjait láthatja a látogató – amelyek a betűk mellett az alkotóik lelkét is átmentették a sors és a történelem viharain keresztül – hanem bevezetik látogatóikat az **ŐRZŐK**, a nemzeti könyvtár munkatársainak gyűjtő, konzerváló, tudományos tevékenységébe is.

Ízelítőként néhány program:

KönyvTÁRlat III/10 Bűnök és szenvedélyek – Gyűjtő/Szenvedély

451 Fahrenheit – (Ne) játssz a tűzzel! Mentsd a menthető! Melyik művet mentené meg a pusztulástól, s őrizné meg az utókor számára? Játék a Facebookon 2014. június 7-től és a Múzeumok Éjszakáján a helyszínen. A résztvevők között ajándékokat sorsolnak ki éjfélkor – és a legtöbbek által megmentett könyveket is „megjutalmazák”! Keresse a Széchényi-szobrot!

Írjon Gutenberg Könyvébe! – Egész éjszaka „vendégünk” a galaxis teremtője. Találja meg, és üzenjen az utókornak! – Molnár Dániel designer élő jelmezfantáziája.

Játék a papírral – Kreatív kézműves foglalkozás. Készítsen és vigyen haza saját kis könyvet, ismerje meg a könyvrestaurálás, a megőrzés fortélyait.

Látványdigitalizálás, e-könyvtári szolgáltatások – avagy Pizskos Fred útja a szekrénytől a tabletig.

Dicsőség és pusztulás: Földgömbragasztás, térképnymtatás, térképes puzzle.

Információ:

<http://www.oszk.hu/rendezvenyek/muzeumok-ejszakaja-2014>



### Zöld megoldás

A különféle hordozók környezetterhelését vizsgáló kutatásról beszámoló linkre

[http://www.greenfo.hu/hirek/2014/04/28/nyomtatott-papir-vagy-e-hirek-melyik-a-zoldebb?fb\\_action\\_ids=10152424466874603&fb\\_action\\_types=og.likes&fb\\_ref=.U16qX4\\_hOhg.like](http://www.greenfo.hu/hirek/2014/04/28/nyomtatott-papir-vagy-e-hirek-melyik-a-zoldebb?fb_action_ids=10152424466874603&fb_action_types=og.likes&fb_ref=.U16qX4_hOhg.like)  
hívta fel a figyelmünket Gondos Gábor (*Katalist*, 2014. április 30.).

Egy svéd intézet (KTH Center for Sustainable Communications) azt vizsgálta, hogy hírolvasásnál az energiakibocsátás szempontjából melyik az előnyösebb megoldás: a nyomtatott vagy az elektronikus hírolvasás. Szerintük, ha napi 10 percnél kevesebb időt tölt valaki hírolvasással, az elektronikus olvasás a kedvezőbb. Napi 30 perc hírolvasásnál már a nyomtatott olvasás jár kevesebb szén-dioxid kibocsátással.

A hagyományos nyomdatechnikával készült újságok esetén 28/kg/fő/év, míg online sajtótermékek esetén 35 kg/fő/év szén-dioxid kibocsátást számoltak a kutatók. Az intézet hozzáteszi, hogy amennyiben változik az elektromos energia előállításának módja, és a megújuló energia irányába mozdul el, úgy eredményeik is módosulhatnak.

