

# LEVÉLTÁRI KÖZLEMÉNYEK

KILENCVENHATODIK ÉVFOLYAM  
2025



Budapest, 2026

LEVÉLTÁRI KÖZLEMÉNYEK

Archival Publications Revue  
des Archives

A Magyar Nemzeti Levéltár Országos Levéltára  
Közleményei

Journal of the National Archives of Hungary  
Revue des Archives Nationales de Hongrie

Kilencvenhatodik évfolyam

*Főszerkesztő*  
Szabó Csaba

*Szerkesztőség*  
Fiziker Róbert (felelős szerkesztő),  
Hegedűs István, Kulcsár Krisztina, Miklós Dániel

*Szerkesztőbizottság*  
Csetz Balázs, Mikó Zsuzsanna,  
Németh László Sándor, Rácz György, Reisz T. Csaba

Székhely (postacím):  
1014 Budapest, Bécsi kapu tér 2–4.  
(1250 Budapest, Pf. 3.)  
<https://leveltarikozlemenyek.mnl.gov.hu/index.php/lk>  
[mnl.gov.hu/mnl/ol/leveltari\\_kozlemenyek](https://mnl.gov.hu/mnl/ol/leveltari_kozlemenyek)  
Email: [leveltari.kozlemenyek@mnl.gov.hu](mailto:leveltari.kozlemenyek@mnl.gov.hu)

A borítóképhez felhasznált fotó:  
A budapesti Tudományos Akadémia homlokzati rajza (1861) – részlet  
Forrás: HU-MNL-OL-T 15-14.

HU ISSN 0024–1512  
DOI: 10.71065/LK.2025  
A kiadásért felel a Magyar Nemzeti Levéltár főigazgatója  
A borítót tervezte: Nagy Attila  
Műszaki szerkesztő: Máté István  
Nyomdai kivitelezés: OOK-Press Kft., Veszprém  
Felelős vezető: Szathmáry Attila  
Megjelent 21,5 (B/5) ív terjedelemben

## TARTALOM

### A Magyar Tudomány Éve

|                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------|----|
| Szabó Csaba: Levéltári utópia                                                                                   | 9  |
| Kazareczky Noémi: A Drágffy család margójára – gazdálkodás, innováció, tudomány                                 | 25 |
| Kelemen Dávid: Falukutató mozgalom, monográfiák és szociográfiák az 1930-as és '40-es évek Sopron vármegyéjében | 53 |
| Bakó Zsigmond: A Magyar Nemzeti Levéltár 2025. évi berlini tanulmányútja                                        | 75 |

### Levéltártan

|                                                                                                                                                                                                  |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Lise Jaillant: Hogyan tehetjük a born-digital és a digitizált levéltári dokumentumokhoz való hozzáférést könnyebbé? Az akadályozó tényezők azonosítása és a megoldások (fordította: Túrke Gábor) | 119 |
| Reinold Schmücker: Eredetiben megőrzés vagy digitalizálás? (fordította: Szabó Attila András)                                                                                                     | 145 |
| Izsák Gergely – Závoczki Adrienn: Egy levéltári segédlet új útjai                                                                                                                                | 153 |
| Ugo Taraborrelli: „Ne quid rerum dictae Poenitentiariae lapsu temporis pereat”. A Sacra Poenitentiaria Apostolica levéltárának rövid bemutatása (fordította: Rácz György)                        | 161 |

### Műhely

|                                                                                                                               |     |
|-------------------------------------------------------------------------------------------------------------------------------|-----|
| Tóth Árpád: Három londoni vonatkozású evangélikus polgári végrendelet Pozsonyból                                              | 177 |
| Deák Ádám: Az 1847. évi vármegyei anyautasítások összehasonlító elemzése                                                      | 215 |
| Csik Tamás: Etetés és itatás a csongrádi és a dési választásokon 1901-ben. „Fertőzött” szavazatok és politikai közösségépítés | 259 |
| Lenkey Tibor: A másik Iványi. Ifj. dr. Iványi-Grünwald Béla                                                                   | 291 |

### Irodalom

|                                                                                                                                                                                                                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Ordasi Ágnes: „Na ja, a Birodalmat érteni kell!” (Ress Imre: Nemzet, birodalom, identitás. HUN-REN Bölcsészettudományi Központ Történettudományi Intézet, Budapest, 2024)                                                                                                                                                                | 305 |
| Fiziker Róbert: Egy polgár életútja (Borbás Barna: Az első gól. Borbás Gáspár élete. Válasz Online Kft., [Budapest], 2023)                                                                                                                                                                                                               | 316 |
| Reisz T. Csaba: Teljessé vált Budapest várostörténeti atlasza (Simon Katalin – Sudár Balázs – Ternováczy Bálint: Pest. A kezdetektől 1848-ig / From the beginning to 1848. HUN-REN Bölcsészettudományi Kutatóközpont – Budapest Főváros Levéltára. Bp., 2025 [Magyar Várostörténeti Atlasz, 10. Hungarian Atlas of Historic Towns, 10.]) | 319 |

|                                                                                                                                                                    |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Bessenyei Vanda: A diplomácia sűrűjében (Miklós Dániel: Csehszlovák<br>–magyar politikai kapcsolatok 1938–1945 között. Magyar Nemzeti<br>Levéltár, Budapest, 2024) | 326 |
| <b>Pro Archivo</b>                                                                                                                                                 | 331 |
| <b>In memoriam</b>                                                                                                                                                 | 335 |
| <b>Angol nyelvű összefoglalók</b>                                                                                                                                  | 337 |

## TABLE OF CONTENTS

### **The Year of Hungarian Science**

|                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------|----|
| Csaba Szabó: Archival Utopia                                                                                    | 9  |
| Noémi Kazareczky: Farming, Innovation, Science – Notes on the Drágffy Family                                    | 25 |
| Dávid Kelemen: Village Research Movement, Monographs, and Sociographies in Sopron County in the 1930s and 1940s | 53 |
| Zsigmond Bakó: A Study Trip to Berlin by the National Archives of Hungary, 2025                                 | 75 |

### **Archival Studies**

|                                                                                                                                                                        |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Lise Jaillant: How Can We Make It Easier to Access Born-Digital and Digitized Archival Documents? Identifying Obstacles and Solutions<br>(Translated by: Gábor Türke)  | 119 |
| Reinold Schmücker: Preservation in Its Original Form or Digitization?<br>(Translated by: Attila András Szabó)                                                          | 145 |
| Gergely Izsák – Adrienn Závoczki: New Approaches to an Archival Finding Aid                                                                                            | 153 |
| Ugo Taraborrelli: „Ne quid rerum dictae Poenitentiariae lapsu temporis pereat”. Discovering the Archives of the Apostolic Penitentiary<br>(Translated by: György Rácz) | 161 |

### **Workshop**

|                                                                                                                                        |     |
|----------------------------------------------------------------------------------------------------------------------------------------|-----|
| Árpád Tóth: Three Lutheran Civil Wills From Pressburg Relating to London                                                               | 177 |
| Ádám Deák: Comparative Analysis of the County Instructions of 1847                                                                     | 215 |
| Tamás Csik: “Feeding and Watering” During the Elections in Dés and Csongrád in 1901. “Infected” Votes and Political Community Building | 259 |
| Tibor Lenkey: The Other Iványi. Dr. Béla Iványi-Grünwald jr.                                                                           | 291 |

### **Book Reviews**

|                                                                                                                                                                                                                                                                                                                                                  |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Ágnes Ordasi: „Well, You Have to Understand the Empire!” (Ress Imre: Nemzet, birodalom, identitás. HUN-REN Bölcsészettudományi Központ Történettudományi Intézet, Budapest, 2024)                                                                                                                                                                | 305 |
| Róbert Fiziker: The Life of a Citizen (Borbás Barna: Az első gól. Borbás Gáspár élete. Válasz Online Kft., [Budapest], 2023)                                                                                                                                                                                                                     | 316 |
| Csaba Reisz T.: Budapest’s Historical Atlas Comes to Completion<br>(Simon Katalin – Sudár Balázs – Ternováczi Bálint: Pest. A kezdetektől 1848-ig / From the beginning to 1848. HUN-REN Bölcsészettudományi Kutatóközpont – Budapest Főváros Levéltára. Bp., 2025 [Magyar Várostartörténeti Atlasz, 10. Hungarian Atlas of Historic Towns, 10.]) | 319 |

|                                                                                                                                                                         |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Vanda Bessenyei: In the Thicket of Diplomacy (Miklós Dániel: Csehszlovák–<br>magyar politikai kapcsolatok 1938–1945 között. Magyar Nemzeti Levéltár,<br>Budapest, 2024) | 326 |
| <b>Pro Archivo</b>                                                                                                                                                      | 331 |
| <b>In Memoriam</b>                                                                                                                                                      | 335 |
| <b>Abstracts of the Papers</b>                                                                                                                                          | 337 |

# HOGYAN TEHETJÜK A BORN-DIGITAL ÉS A DIGITIZÁLT LEVÉLTÁRI DOKUMENTUMOKHOZ VALÓ HOZZÁFÉRÉST KÖNNYEBBÉ?

## Az akadályozó tényezők azonosítása és a megoldások

<https://doi.org/10.71065/LK.2025.5>

*Levéltári Közlemények* 96 (2025) 119–143.

### Bevezetés

Az adatokhoz való hozzáférést manapság kulcsfontosságú tényezőnek tekintjük. Ennek ellenére a levéltárakban őrzött digitális adatállomány oroszlánrésze adatvédelmi okokból, szerzői jogi védettség következtében, esetleg technikai problémákból kifolyólag hozzáférhetetlen. Az emailek és más born-digital gyűjtemények legtöbbször nincsenek katalogizálva, lehetetlen bennük keresni, ennél fogva használhatatlanok.<sup>1</sup> A digitalizálásukat megelőzően papíralapon keletkezett dokumentumok esetében a jogvédettség jelentős akadályt jelenthet a hozzáférhetőség szempontjából. Mint irodalmár és mint digitális bölcsész kiterjedt tapasztalattal rendelkezem a digitális gyűjteményekhez való hozzáférésre irányuló próbálkozások (és gyakran azok kudarca) terén. „Hagyományos” tudományos kutatóként szerettem a képeztésemet és éveket töltöttem levéltárakban régi levelek és más papír alapú dokumentumok tanulmányozásával. A közgyűjteményekben elhelyezett emailekhez, PDF és Word dokumentumokhoz való hozzáférés sokkal nehezebb. Még szerencsésnek számítok, hogy korlátozott számban kutathattam a Harry Random Centerben található Ian McEwan-gyűjteményben, valamint a manchesteri Carcanet sajtóarchívumban őrzött emaileket. A born-digital adatokhoz való hozzáférés azonban továbbra is a kivétel körébe sorolható, nem a szabályhoz.

A digitális archívumokhoz való hozzáférés problémájának megoldásához elkerülhetetlen a különböző szakterületek, tudományágak közti együttműködés. Korunk nagy kihívásai – a globális felmelegedéstől kezdve a társadalmi egyenlőt-

\* Lise Jaillant a Digitális Kulturális Örökség professzora (School of Social Sciences and Humanities, Loughborough University, Egyesült Királyság). A cikk eredeti nyelven megjelent a *Archival Science* 22/3. Springer, 2022. számában, online közlése a <https://doi.org/10.1007/s10502-022-09390-7> oldalon található. Jelen cikk nyílt hozzáférésű a Creative Commons Attribution 4.0 International License alapján. Az engedélyt tartalmazó dokumentum megtekinthető a <http://creativecommons.org/licenses/by/4.0/> linken.

<sup>1</sup> Born-digital dokumentumokon azokat tartalmakat értjük, amelyek eleve digitális formában keletkeztek és valamilyen fizikai formátumból lettek utólag digitalizálva.

lenségeken át – nem oldhatóak meg egyetlen tudományterületen belül. Ugyanez áll a kutatók elől elzárt „sötét” iratállományokra is. Nem várhatjuk el a levéltárosoktól, vagy akár a digitális bölcsészeketől, hogy maguktól találják fel a csodaszert, amely varázsütésre hozzáférhetőbbé teszi a digitális dokumentumokat. Ehelyett együttműködést kell kialakítanunk olyan tudományágak közt, amelyek egyébként meglehetősen ritkán kerülnek kapcsolatba egymással.

Különösen az informatikusokkal való együttműködés lehet gyümölcsöző. A mesterséges intelligencia és a gépi tanulás<sup>2</sup> felhasználható arra, hogy iratállományokhoz hozzáférést biztosítsunk, könnyebben fellelhetővé és elérhetőbbé tegyük őket – például a szenitív adatokat tartalmazó és nem tartalmazó anyagok különválogatása által, vagy tárgyszavak és más metaadatok automatikus rögzítésével a fellelhetőség javítása érdekében. Kiemelten fontos azonban, hogy kerüljük az elfogultságot az adatok kiválasztása és feldolgozása során, amelynek következtében diszkrimináció érhet egyes csoportokat és kihatással lenne a kollektív emlékezetre. Mindez megköveteli, hogy elmélyedjünk az algoritmusokban és ne kezeljük a MI-t egyfajta „fekete doboz”-ként. A „megmagyarázható” MI alkalmazása alapvetően fontos, hiszen az emberek így érteni fogják, hogy a gép hogyan jutott az adott eredményre és hozta az adott döntést. Az olyan hálózatok, mint az AURA (Egyesült Királyság/Ír Köztársaság levéltárai és MI) és az AEOLIAN (Egyesült Királyság/USA: MI a kulturális szervezetek számára),<sup>3</sup> összehozzák a digitális bölcsészeket, az informatikusokat, a levéltárosokat, a könyvtárosokat és más szakmabelieket.

Digitális bölcsészekből álló csapatunk jelen cikk megírásához számos interjút készített az Egyesült Királyság, Írország, valamint az USA könyvtári és levéltári szakembereivel. A born-digital és a digitizált iratanyagokhoz való hozzáférés fő akadályozó tényezőit az iratörzők nézőpontjából vizsgáltuk. A szűk keresztmetszetet adó okok közül a legfőbbek jól ismertek (köztük az adatvédelmi szabályozás és a szerzői jog), de az interjúk a hozzáférés további akadályait tárták fel – például a kockázatkerülő hozzáállást, melynek következtében számos intézmény hírnevét védve rendkívül szigorúan értelmezi a jogszabályokat. Interjúalanyaink gyakran elégedetlenek voltak ezen intézményi hozzáállás miatt, és az irathozzáférés és a jogi szabályozás közti harmonikusabb egyensúly megteremtésének szükségességét hangsúlyozták. Az akadályok megnevezése mellett, az interjúalanyok lehetséges megoldásokat is felvetettek a digitális iratállományok hozzáférhetőbbé tételére a gyűjteményhasználók számára. Mindazonáltal többen sajnálkozásukat fejezték ki,

<sup>2</sup> A Mesterséges intelligencia (MI) egy nagyszabású elképzelés, olyan intelligens gépek megalkotását célozza, amelyek képesek az emberi gondolkodást és viselkedést utánozni. Gépi tanulás – Machine Learning (ML) az MI egyfajta alkalmazása, ha úgy vesszük, leágazása, amely lehetővé teszi a gépek számára, hogy kívülről történő programozásuk helyett a rendelkezésükre álló adatok felhasználásával maguk fejlesszék képességeiket. A gyakorlati életben az „MI” és „ML” kifejezéseket gyakran azonos értelemben, ugyanarra a dologra alkalmazzák.

<sup>3</sup> [www.aura-network.net](http://www.aura-network.net) és [www.aeolian-network.net](http://www.aeolian-network.net).

hogy a tudósok és a többi gyűjteményhasználó is ritkán mutat érdeklődést a digitális kor levéltárait érintő problémák iránt és rendszerint meglehetősen régimódi kutatási módszerekkel dolgoznak, papír alapú iratgyűjteményeket használva, mellőzve az újabb keletű, korszerűbb forrásoknak a munkájukba való bevonását.

Jelen cikk első szakasza röviden áttekinti a born-digital és digitizált iratgyűjteményekhez való hozzáférés tárgyában rendelkezésre álló kutatásokat. A következő szakasz az interjúk lezárultát követően általunk alkalmazott módszereket és megközelítéseket taglalja. Szó esik majd az eredményekről is, valamint felhívjuk a figyelmet arra, hogy a kutatók nem rendelkeznek beleszólással a digitális iratgyűjteményeket érintő vitákba. Ajánlásokat teszünk, hogy a tudomány képviselői és más felhasználók is „odaülhessenek az asztalhoz” és hozzájáruljanak a digitális levéltári anyagokhoz való hozzáférésre vonatkozó szabályrendszer alakításához.

Összegezve, jelen írás a következőkre tér ki:

- A levéltárosokkal, könyvtárosokkal és a közművelődés területén dolgozó más szakemberekkel lefolytatott interjúkat alapul véve azonosítjuk a digitizált és born-digital gyűjteményeknek a gyűjteményhasználók számára történő hozzáférhetőbbé tétele útjában álló fő akadályokat.
- Felvázoljuk, hogy a különböző, közművelődés területén működő szervezeteknél a gyűjtemények széles körére nézve milyen a hozzáférés jelenlegi szintje, beleértve, ha egyáltalán nem lehetséges a hozzáférés, vagy korlátozott hozzáférésről van szó (például, ha a gyűjteményhasználónak az őrzés helyére szükséges utaznia a dokumentumok kutatása céljából).
- Lehetséges megoldásokat javasolunk a hozzáférés problémakörére – beleértve a mesterséges intelligencia etikus módon történő alkalmazását az eddig elzárt „sötét” iratállományoknak a gyűjteményhasználók számára elérhetővé tételére.
- Javaslatot teszünk egy globális felhasználói közösség létrehozására, amely közreműködhetne a digitális gyűjtemények hozzáférhetőségére vonatkozó döntések meghozatalában.

## **Kapcsolódó művek**

Az elmúlt három évtizedben számos tudományos cikk és más munka jelent meg a digitális megőrzés témájában. Ezek gyakran a levéltártudomány területéről származnak, míg más területek (köztük a digitális bölcsészet) ritkán kapcsolódnak be ezekbe a vitákba. A gyűjteményhasználók ezen viszonylagos hallgatása kihat a keretekre, amelyek közt a digitális gyűjteményekkel kapcsolatos egyeztetések folynak. És valóban a fókusz többnyire a megőrzésen van, míg a digitális gyűjteményekhez való hozzáférés kérdését jóval ritkábban tárgyalják. Egyik interjúalanyunk megjegyezte: *„Nagyon régi meggyőződésem, hogy a megőrzés előtérbe helyezése anélkül, hogy azon is gondolkodnánk, hogyan biztosítjuk a hozzáférést... már*

*kudarcc rögtön az indulásnál, ami abból ered, ahogy az egész kérdést az elejétől fogva megközelítettük*” ([anon.] 2021a).

A levéltáros közösség az 1990-es évek közepén látott hozzá kidolgozni a digitális iratanyag – köztük a megőrzésre szánt emailek – megőrzésére irányuló stratégiákat.<sup>4</sup> A fő hangsúly akkoriban a hanyagság és a technikai avulás következtében elveszni látszó veszélyeztetett anyagokon volt. A „digitális sötét kor” bekövetkezése miatti növekvő aggodalmak 2002-ben a Digital Preservation Coalition (Digitális Megőrzési Szövetség – DPC) létrehozásához vezettek, amelyet több, az Egyesült Királyságban és Írországon működő ügynökség együttműködése keretében alapítottak meg. A 2000-es évek végén a levéltárosok és a tudósok közti együttműködések olyan nyílt forrású digitális könyvtár eszközöket hozott létre tartalomgondozás céljára, mint a BitCurator.<sup>5</sup> Az együttműködésnek ezen példái azonban kivételek maradtak, ráadásul ritkán érintették a „sötét” iratanyagokhoz való hozzáférés problematikáját.

Az Email Munkacsoport munkáját követően a DPC emailek megőrzésével kapcsolatos jelentése ((Prom 2019) rámutatott, hogy a megőrzés többé nem számít akkora kihívásnak, mint azelőtt. A fókusz most a megőrzésről a maradandó értékű anyagok felbecslésére és kiválogatására helyeződött. A RATOM project egy Natural Language Processing (NLP – Természetes Nyelvek Feldolgozása – a ford.) alkalmazást fejleszt az email archívumok felmérésére és feldolgozására.<sup>6</sup> A Brit Nemzeti Levéltár szintén alkalmaz mesterséges intelligenciát a kormányzati dokumentumok felmérése és történeti dokumentumként való megőrzésre történő kiválogatása céljából.

Nincs értelme olyan dokumentumok megőrzésének, melyeket nem lehet akár azonnal, akár egy későbbi alkalommal használatba venni. Az állami levéltárakkal kapcsolatos legutóbbi cikkében Sara Martínez-Cardama és Ana Pacios megemlíti, hogy a „hozzáférés” és a „használat” fogalmát 14, illetve 12 intézmény jelölte meg mint prioritást. A gyakorlatban azonban a fő akadályok – mint a személyes adatok védelme, szerzői jog és technikai problémák – rendkívül megnehezítik a digitális kulturális adatokhoz való hozzáférést. Az email archívumok esetében például a személyes adatokkal kapcsolatos aggályok gyakran a teljes gyűjtemény elzárásához vezetnek.<sup>7</sup> Amikor a brit író, Ian McEwan átadta a hagyatékát a texasi Harry Ransom Center számára, belefoglalta 17 évet felölelő, 1997 és 2014 között folytatott emailezését is. Az iratjegyzék röviden utal McEwan emailezésére, ami „*egyelőre*

<sup>4</sup> Lásd például Bearman (1994), Cook (1994), Ross és Gow (1999).

<sup>5</sup> Lásd Kirschenbaum és mások 2010. Ez a jelentés a sarokkövévé vált annak az együttműködésnek, amely fejlesztette a digitális levéltárosok által ma már széleskörben használt BitCurator rendszert.

<sup>6</sup> <https://ratom.web.unc.edu/about/> (utolsó letöltés: 2021. november 17.).

<sup>7</sup> Lásd Jaillant L. (2019) és Johnston J. A., és mások (2019) az email archívumokhoz való hozzáférés tárgyában.

*feldolgozatlan és kutathatósága nem biztosított*”.<sup>8</sup> Számos gyűjtemény rejtve marad és a kutatóknak még csak értesülésük sincs bizonyos digitális dokumentumok létezéséről. Az író Will Self emailézése a British Libraryban például nem szerepel a gyűjteményt leíró jegyzékben és a kutatók számára sem helyben, sem interneten nem elérhető. A tény, hogy nagyon kevés tudományos kutató rendelkezik hozzáféréssel az archivált emailgyűjteményekhez, egyfajta hézagot eredményezett a téma tudományos feldolgozottságában. Az adatvédelmi és szerzői jogi aggályok a közlési szakaszban is nehézségekhez vezethetnek. A 2020. évi „Levéltárak, hozzáférés és MI” témájában szervezett konferencián tartott előadásában Sophie Baldock elmondta, hogy doktorandusz korában a British Libraryben kizárólagos hozzáféréssel rendelkezett a Wendy Cope brit költő emailjeit tartalmazó gyűjteményhez (Baldock 2020). Ennek ellenére nehéznek bizonyult kutatását publikálni, mivel az emailkezés sűrűn tartalmazott személyes, illetve szenzitív adatot. A publikációs engedélyek megszerzéséhez szükséges az emailek szerzőjének a hozzájárulása, ráadásul harmadik személyek is szerepelnek az emailekben. Nem meglepő, ha a kutatók gyakran inkább a könnyebben hozzáférhető gyűjteményekre koncentrálnak, melyek adatait könnyebb hivatkozni a tudományos munkákban.

Az emailekkel ellentétben a digitális gyűjtemények esetén nem egyszer könnyebbnek ítélik a hozzáférést. A korábban hozzáférhető gyűjtemények azonban adatvédelmi és szerzői jogi okokból később lezárásra kerülhetnek. Az egyének személyes adataik védelméhez való jogát hangoztatja Michelle Moravec feminista kutatási módszerek és digitális archívumok témájában írt cikkében, amelyben elsősorban a *Spare Rib* című (Oldalborda – a ford.) feminista magazin példányainak a British Library által 2013-ban végrehajtott digitalizációjával foglalkozott. „Beleegyezésüket adták-e ehhez azok a személyek, akiknek a munkái ezekben az anyagokban megjelennek?” – kérdezte Moravec (186).<sup>9</sup>

A British Library 2020 decemberében bejelentette, hogy amint az Egyesült Királyság kilép az EU-ból, meg fogja szüntetni a *Spare Rib*-hez való hozzáférést. Valóban, az anyag az EU árva művekre vonatkozó direktívája alapján volt hozzáférhető. A Brexittel összefüggésben nézve, a szerzői jogi keretek megváltozása vezetett ahhoz a döntéshez, hogy *Spare Rib* magazin teljes sorozatához megszűntették a hozzáférést.<sup>10</sup>

A levéltárosok egyre több figyelmet kezdenek fordítani a hozzáférés kérdéskörére. A Born-Digital Archives Working Group (egy többnyire amerikai levéltárosokból összeállt munkacsoport) 2020-ban megjelentette a „*Levels of Born-Digital*

<sup>8</sup> <https://norman.hrc.utexas.edu/fasearch/findingAid.cfm?eadid=01073> (utolsó letöltés: 2021. november 17.).

<sup>9</sup> A digitális gyűjteményekkel kapcsolatos adatvédelmi aggályok vonatkozásában további információért lásd LeClere (2018), Hart és mások (2019).

<sup>10</sup> Lásd <https://journalarchives.jisc.ac.uk/britishlibrary/sparerib> (utolsó letöltés: 2021. november 17.).

*Access*” – A born-digital dokumentumokhoz való hozzáférés fokozatai – című dolgozatát, amely egy viszonyítási pontokból és gyakorlati útmutatókból álló összeállítás a born-digital iratanyagokhoz való hozzáférés megkönnyítésére. „A hozzáférési szinteket szakmabeliek alakították ki, a szakmabeliekre gondolva”, közölte a beszámoló. Azok részére szánták felhasználás céljára, illetve hivatkozási alapul, akik rendelkeznek a digitális megőrzés eszközeire és elveire vonatkozó alapismeretekkel. (Peltzman és mások 2020). A digitális levéltárosok által digitális levéltárosok számára készült beszámoló több hozzáférési szintet vázol fel. Az 1-es szint esetében a kutatóknak a megőrzés helyszínére kell utazniuk, ahol az anyagokat nyilvános hozzáférésű számítógépeken kutathatják. A spektrum túlsó végén, a dokumentumokat távolról lehet elérni és kifinomult eszközökkel lehet elemzésüket folytatni. Másképp kifejezve, a kutatók az otthoni íróasztaluk mellől tekinthetnek bele az archívumokba, és számítógépes eljárásokat alkalmazhatnak a digitális dokumentumokon.

Bár értékes az eddig végzett munka, sürgetően szükség szerű a tudósok és más gyűjteményhasználók mélyebb bevonása. A digitális gyűjtemények különböző felhasználóinak nem egyformák a szükségletei. A magasabb szintű számítástechnikai készségekkel rendelkező kutatók kisebbséget képviselnek a bölcsészettudományok művelői és a nem tudományos kutatók közt. A levéltárosok által a kutatók részére készített sorvezetők helyett közösen kialakított stratégiákra lenne szükség a digitális iratanyagok könnyebben hozzáférhetővé tételére.

## **Módszerek, szemléletmód és az eredmények**

A hozzáférhetetlen iratanyagok nyilvánvaló befolyással vannak a kutatókra, ennek ellenére a digitális megőrzéssel és hozzáféréssel kapcsolatos irodalmat levéltárosok írják és nem a levéltárhasználók. Ebben a cikkben ezt a téma tudományos feldolgozottságában jelentkező hézagot interdiszciplináris nézőpontból tárgyaljuk. Ameddig a digitális bölcsészek ritkán szembesültek a zárt digitális archívumok problémájával, a megközelítésünk az volt, hogy a kutatókat közvetlen kapcsolatba hozzuk levéltárosokkal és más könyvtári, levéltári és múzeumi szakemberekkel. Intézetünk Etikai Tanácsának beleegyezését megszerezve, három fős digitális bölcsészekből álló csapatunk 26 levéltáros, könyvtáros és muzeológus szakemberrel összesen 21 interjút folytatott le. Az egyes interjúk hossza 30 perctől 1 óráig terjedt, az online kivitelezés lehetővé tette, hogy az Egyesült Királyságban, Írországban és az USA-ban élő szakembereket egyaránt elérhessük. Az interjúalanyok többsége engedélyezte, hogy nevét feltüntetve szerepeljen az eredményként létrejött cikkben és egyéb kutatási kiadványokban.

Az interjúalanyok kiválasztása a meglévő személyes kapcsolataink felhasználásával történt. A megkeresett személyek közül, akik nem tudtak rendelkezésünkre állni, gyakran ajánlották munkatársaikat és ez a „hólabdázós” eljárás további interjúkat eredményezett. Az interjúalanyok meglévő kapcsolati hálókön keresztül tör-

ténő beszervezése az eredmények torzulásához vezethet, különösen, ha az összes ismerős hasonló háttérrel rendelkezik. Ezért figyeltünk arra, hogy a lehető legszélesebb metszetet alkalmazzuk a nemek (12 nő és 14 férfi), a pályafutás (pályakezdőtől a már nyugállományba vonult szakemberekig) és a földrajzi hely tekintetében. Kérdeztünk szakembereket a Skót Nemzeti Könyvtártól, a Walesi Nemzeti Könyvtártól, az Ír Nemzeti Könyvtártól és az Ivy League-hoz (Borostyán Liga, az USA nyolc elitegyetemét tömörítő sportszövetség – a ford.) tartozó intézményektől az USA keleti partvidékén (Yale, Harvard). Interjúalanyaink egy része jelentős, nagyvárosi kulturális intézményekben dolgozott, mint a British Library, vagy a Brit Nemzeti Levéltár, de kisebb intézményeket is szerepeltettünk, mint például a Seven Stories (a gyermekkönyvekre specializált múzeum Newcastle-upon-Tyne-ban) vagy az Ír Népzenei Archívum Dublinban.

A mintavétel tekintetében a diverzitásra való minden törekvésünk ellenére elismerjük, hogy a nagy kulturális intézmények túl vannak reprezentálva. Ez annak a ténynek köszönhető, hogy ezek az intézmények rendelkeznek azokkal a pénzügyi forrásokkal és a megfelelően képezett személyzettel, amelyek lehetővé teszik számukra, hogy folyamatosan kezeljék a born-digital és a digitizált anyagaikat. A digitalizációs programokban való részvétel vagy a born-digital forrásokkal kapcsolatos stratégiai szerepvállalás rendszerint kívül esik a kisebb gyűjtemények lehetőségein. Rachel Hosker (levéltárvezető és a Különleges Gyűjtemények helyettes igazgatója, Kutatási Gyűjteményi Központ, Edinburgh-i Egyetem) ismerte fel ezt a megosztást a viszonylag kiváltságos helyzetben levő és kevésbé privilegizált intézmények közt. *„Megvannak a képességeink, de a legtöbb helyen nincsenek programfejlesztők, akikkel együtt lehetne dolgozni”*, mutatott rá Hosker (Hosker 2021).

Az interjúk felépítése félig strukturált: egy kérdéssort előre megküldtünk a résztvevőknek, és további kérdéseket tettünk fel az interjúk során. Bár a kérdéseinket az interjúalanyokra szabtuk, összeállítottunk egy minta kérdéssort, amely a következőképpen nézett ki:

#### *Háttér és tapasztalat*

1. Tudná részletezni a szakmai múltját? Mivel foglalkozott, mielőtt [foglalkozás megnevezése] lett?
2. Ismertetné az intézményében őrzött gyűjteményeket?
3. (Megközelítőleg) mennyi kutató látogatja évente az Önök intézményét? A kutatók többnyire tudományos körökből, vagy a szélesebb nyilvánosságból kerülnek ki?
4. Leírná, hogy az Önök intézménye hogyan reagált a COVID járvány jelentette kihívásokra?

### *Digitizált gyűjtemények*

5. Részletezné az Önök intézményében folyamatban levő digitalizációs projekteket? Jelenleg mennyi gyűjteményük érhető el digitális formában?
6. Minden digitális gyűjteményük elérhető távolról is? Ha nem, milyen hozzáférési korlátozások vannak?
7. Szélesebb nézőpontból vizsgálva, milyen a kihívásokat, illetve milyen távlatokat lát az Egyesült Királyság és más területek digitizált iratgyűjteményei előtt?

### *Born-digital gyűjtemények*

8. Mint tudja, a born-digital gyűjtemények számos okból kifolyólag ritkán csak kivételesen hozzáférhetőek a gyűjteményhasználók számára. Ismertetné az Önök intézményében fennálló helyzetet az ilyen gyűjteményekhez való hozzáférés vonatkozásában?
9. Hogyan tehetnénk a born-digital gyűjteményeket elérhetőbbé a kutatók számára?
10. Van valamilyen együttműködésük más intézményekkel – az Egyesült Királyságban, vagy akár máshol – a born-digital projektekkal és gyűjteményekkel összefüggésben?
11. Milyen a kihívásokat, illetve milyen távlatokat lát born-digital iratgyűjtemények vonatkozásában az Egyesült Királyságban és más térségekben?

### *Egyéb*

12. Van más olyan dolog a témában, amit még szeretne megosztani velünk?

Az interjúalanyoknak a fenti kérdésekre adott válaszai segítettek számunkra azonosítani azokat az általános problémákat, amelyek akadályozzák a digitizált és a born-digital gyűjteményekhez való hozzáférést; segítettek továbbá megállapítani az egyes intézményekben a kiválasztott gyűjteményekhez (például webarchívumok, vagy emailarchívumok) való hozzáférés jelenlegi szintjét; lehetséges megoldásokat kipróbálni a hozzáférést érintő problémákra, beleértve az olyan fejlett technológiák, mint a MI és a ML, használatát is. Az eredményekre alapozva amellett érvelünk, hogy az elzárt kulturális értékek nyilvánossá tétele érdekében több összefogásra van szükség a levéltárosok és az gyűjteményhasználók között.

### **A digitizált gyűjteményekhez való hozzáférést akadályozó tényezők**

A digitizált gyűjtemények kisebbséget képeznek a kulturális intézmények teljes dokumentumállományában, nem csupán azért, mert a kiemelkedő minőségű digi-

tális másolatok készítése magas költségekkel jár és technikai szakértelmet igényel, hanem elsősorban azért, mert nem is minden gyűjtemény digitalizálható.

Először is a technológiai avulás következtében nehéz megfelelő eszközöket előkeríteni a régi formátumokban tárolt dokumentumok olvasására. Például a VHS vagy Betamax szabványú mágnesszalagokhoz való lejátszó készülékek már egyre ritkábbak és Stuart Lewis (Digitális Gyűjtemények társigazgatója, Skót Nemzeti Könyvtár) szerint az a veszély fenyeget, hogy három vagy négy éven belül teljesen el is tűnhetnek. A lejátszók nélkül „*még ha az adathordozó jó állapotban is lenne, nem leszünk képesek digitalizálni*” (Lewis és mások 2021). A könyvtár nemrégiben szerzett egy üzleti tervet, lehetővé téve 10 000 mágnesszalag digitalizálását, ami jelentősen megnöveli majd a tárolt adatmennyiséget, az audiovizuális fájloknak a dokumentumképekhez viszonyított nagyobb mérete következtében.

Másodszor, a digitalizációs programok magas prioritásúnak számító anyagok kiválogatásában, egyben a kevésbé fontosnak, vagy értékesnek tartott dokumentumok kizárásában bíznak. „*Nagy vita folyik arról, ... mely közösségekhez tartozó és milyen iratok kerüljenek digitalizálásra, ki legyen megjelenítve és ki nem a digitalizációs stratégiával kapcsolatban hozott döntésekben*” – mondta Anthea Seles (főtitkár, Nemzetközi Levéltári Tanács [ICA]) (Seles 2021). Például, a George Floyd halálát követő események nyomán Dorothy Berry (Digitális Gyűjtemények programigazgatója a Houghton Könyvtárnál a Harvard Egyetemen) felismerte, hogy az afroamerikai kulturális vagyoni vonatkozásában csak nagyon kismértékű digitalizáció történt. Ezt követően végigirányított egy 2000 dokumentum digitalizációját megvalósító projektet ([anon.] 2021b).<sup>11</sup>

A szerzői jogok védelme a harmadik ok, amiért bizonyos gyűjteményeket nem digitalizálnak, vagy tesznek elérhetővé, amennyiben léteznek digitális példányok. Treasa Harkin (könyvtáros, Ír Népzenei Archívum [ITMA]) rávilágított, hogy az ITMA az évek során rendkívül erős bizalmat épített ki a népzenei közösséggel: „*ez részben arra a tényre épül, hogy biztos vagy benne, hogy amit beadsz az archívumba, az biztonságban van és nem fog egyszer csak megjelenni egy weboldalon, vagy máshol az interneten*” (Harkin és O'Connor 2021). Az archívum megpróbál egyensúlyozni, hogy e bizalom elvesztése nélkül növelje a digitizált anyagok elérhetőségét a gyűjteményhasználók számára. Ez a jogbirtokosokkal az engedélyek megszerzése érdekében folytatott tárgyalásokat, valamint a különböző hozzáférési szintek felülvizsgálatát (például kizárólag loginnévvel és jelszóval rendelkező regisztrált felhasználók számára biztosítanak online hozzáférést) igényli.

Az engedélyek megszerzését rendkívül megnehezíti, ha a jogbirtokosok vagy örököseik nyom nélkül eltűnnek. Stuart Lewis megjegyzi, hogy a Skót Nemzeti Könyvtár nem rendelkezik feltétlenül az állományok digitalizálásához az azokat adományozó, vagy letétbe helyező személyek engedélyével. „*Huszonöt évvel ezelőtt*

<sup>11</sup> Lásd még <https://news.harvard.edu/gazette/story/2020/07/houghtons-2020-21-digitization-focus-blackamerican-history/> (utolsó letöltés: 2021. november 19.).

meg sem fordult volna a fejünkben, hogy megkérdezzük: Lehetne ezt kérem digitalizálnunk és online közzélnünk?” (Lewis és mások 2021). A szerzői jogi hozzájárulások visszamenőleges megszerzésének nehézségei megmagyarázzák, hogy a Skót Nemzeti Könyvtár miért nem kezeli első helyen a kéziratos anyagok digitalizációját, és miért fókuszál inkább a kevésbé problematikus dokumentumokra – mint például a köztulajdonnak számító tételek, amelyek a Skót Nemzeti Könyvtár Data Foundry nevű weboldalán elérhető digitizált tartalomnak a nagyobbik részét alkotják.<sup>12</sup>

A digitális kort megelőzően őrzésbe vett anyagokat érintő szerzői jogi problémák más archívumokat is érintettek, köztük a Yale Egyetem Könyvtárának Fortunoff Videóarchívumát. Amikor ez a holokauszt-túlélők visszaemlékezéseit magába foglaló gyűjtemény keletkezett, az interjúalanyok hozzájárulásukat adták a felvételek elkészítéséhez, de ahhoz nem, hogy azok az interneten bárhol elérhetők legyenek. Ez korlátozza az olyan kutatóknak a gyűjteményhez való hozzáférését, mint Abby Gondek (Morgenthau ösztöndíjas, Franklin D. Roosevelt Könyvtár és Roosevelt Intézet): *„Rengeteg a túlélővallomás, de nem hallgathatod őket végig – csak az ízelítőket lehet meghallgatni, hacsak nem mész el magába az archívumba, vagy ha vannak ilyenek, a partnerintézményekben, amelyekkel kapcsolatban állnak, és akkor meghallgathatod a felvételeket ezeken a helyeken”* (Gondek 2021).

Az adatvédelem a negyedik nagy akadály, amely a szenzitív adatokat tartalmazó digitizált anyagok elérhetősége javításának az útjában áll. Az Egyesült Királyság és az EU közgyűjteményi intézményeinek meg kell felelniük az Általános Adatvédelmi Rendeletnek (GDPR) az Európai Unióra és a 2018. évi adatvédelmi rendelkezésnek (Data Protection Regulation 2018) Nagy-Britanniára vonatkozólag. A szenzitív adatok kikerülésének megelőzése kiemelt kérdés a londoni Wellcome Gyűjtemény számára, amely az egészségüggyel és gyógyászattal kapcsolatos bizalmas dokumentumokat – köztük egyének kórtörténetét – őrző intézet. A gyűjtemény más bizalmas dokumentumokat is tartalmaz, például üzleti titkokat a szervezeti anyagok közt. Alapvetően a Wellcome nem digitalizál tíz évesnél nem régebbi anyagokat. Más dokumentumok esetén a levéltárosok kockázatmérlegelő megközelítést alkalmaznak: Veszik egy adott sorozat egy részletét, és előbb azon végeznek felmérést. Amennyiben a bennük levő adatok rendkívül szenzitívnek bizonyulnak, akkor még több anyagot vizsgálnak át abból a sorozatból. Egyes esetekben jóval alaposabb, akár darabszintű átvizsgálás válik szükségessé, a kulturális szempontból érzékeny, sértő nyelvezetet használó, vagy egyszerűen csak kellemetlen látványt nyújtó (mint például a sebészeti beavatkozást ábrázoló képek) dokumentumok beazonosítása érdekében. Ezek a dokumentumok digitalizálhatóak és online elérhetővé tehetők, de a felhasználókat előzetesen figyelmeztetik a megtekinteni szándékozott tartalommal kapcsolatban ([anon.] 2021a).

<sup>12</sup> <https://data.nls.uk/data/digitised-collections/> (utolsó letöltés: 2021. november 19.).

Számottevő méretű gyűjtemények esetén a szenzitivitási vizsgálat sziszifuszi munkává válhat. Bármikor előfordulhat, hogy egy korábban szenzitív adatokat nem tartalmazónak vélt gyűjteményről kiderül, mégis tartalmaz problémás anyagokat. A Skót Nemzeti Könyvtár (NLS) biztonsági ellenőrzések sorozatát folytatja, melyek a digitalizációt azon gyűjteményekre korlátozzák, amelyeket adatvédelmi és szerzői jogi szempontokból átvizsgáltak és további feltételeknek is megfeleltek. *„Az egyik feltétele annak, hogy valami digitalizálásra mehessen, az, hogy hogy véglegesen rendszerbe legyen téve, tehát rendezve van, rendelkezik leírással és foliózták, vagyis van egy hivatkozási rendszerünk [és] vissza tudunk keresni egy bizonyos foliót egy adott kontextuson belül”* – jegyezte meg Stephen Rigden (digitális levéltáros az NLS-nél) (Lewis és mások 2021).

A *„Több Termék, Kevesebb Munkafolyamat”* mozgalom (Greene és Meissner 2005) nyomán egyesek felvethetik, hogy ezek a feltételek túlzóak, és mindig az a jobbik eset, ha digitalizálunk és hozzáférhetővé tesszük a dokumentumokat akkor is, ha azok nincsenek teljesen rendezve és még továbbra is tartalmaznak problematikus információt. Egy gyűjtemény sosem lesz tökéletesen rendezett, és nincs olyan, aki garantálhatja, hogy az összes szenzitív anyagot kiszedték belőle. Mint Dennis Meissner és Mark Greene állítják 2010-ben megjelent cikkükben: *„Nagyobb erkölcsi kárunk származhat az anyagok visszatartásából. Bármely kár, ami szenzitív, de szerződésileg nem tiltott anyagok nyilvánosságra hozatalából származhat, elhalványul ehhez képest.”* (204)

### **Born-digital gyűjtemények: a keletkezéstől a hozzáférésig**

A digitalizálásukat megelőzően papíralapon keletkezett gyűjteményekkel ellentétben a born-digital dokumentumok ritkán kerülnek önmagukban átvételre, ehelyett rendszerint hibrid gyűjtemények részeként érkeznek a megőrzést végző intézményekbe. Egyelőre nincs piaca a born-digital dokumentumoknak. Mint egy interjúalanyunk mondta: *„Ha valaki eredeti példányokat ajánl, és aztán másvalaki jön és ezen eredetiek fénymásolatait kínálja, nem a fénymásolatot veszem meg, hiába azonos a tartalmuk”* ([anon.] 2021b). A fénymásolatokhoz hasonlóan a born-digital dokumentumok is elérhetővé teszik a tartalmat és könnyűszerrel duplikálhatóak. Egy gyűjtemény pénzben mért értékét továbbra is a dokumentum eredetiségéből eredő aura és az alkotójához való köthetősége befolyásolja.

Az egyik ok, amiért a könyvgyűjtemények előnyben részesítik az eredeti anyagokat, az, hogy ezek jól mutatnak a kiállításokon. A gyermekek és szüleik érdeklődésének felkeltése eredeti dokumentumok bemutatásával a Seven Stories Múzeum stratégiájának központi eleme. Krie McKie gyűjteményigazgató ezt mondta:

*„Itt őrizzük Enid Blyton hagyatékát, vagy legalábbis egy részét és rengeteg, az Ötösfogat könyvekhez, a Noddyhoz, satöbbihez készült gépirattal rendelkezünk. Igazán egyszerű ezekből egyet megmutatni a látogatócsoportnak és közölni, hogy ez egy személyesen Enid Blyton által gépelt kézirat és ... árad belőle valami eredetiség...*

*De egy born-digital anyaggal, hát nem tudom, hogy keltenéd fel ugyanazt a fokú lelkesedést, vagy érdeklődést egy Word-fájl iránt”* (McKie 2021. május 28.).

Egyes intézmények tényleges born-digital fájlok bemutatása helyett tárgyi emlékeket állítanak ki, például eredeti számítógépeket. A Houghton Library a Harvardon például a jövőbeli kiállításon Jamaica Kincaid író nő egyik laptopját tervezi kiállítani ([anon.] 2021b). Az eszközt, amit az író érintett munkavégzése során, olyan aura hatja át, aminek a born-digital fájlok, úgy tűnik, híján vannak.

A napjainkban megszerzett vagy adományozott gyűjtemények még többnyire a digitális forradalom előtt keletkeztek. A szerzők újabb generációjának eljövételével valószínű, hogy a jövőbeli gyűjtemények legnagyobb részét a born-digital dokumentumok fogják képezni. Egyes intézmények előre készülnek erre a pillanatra, és aktívan keresik a kapcsolatot az olyan szerzőkkel, akik egész pályafutásuk során számítógépet használtak írógép és más eszközök helyett.

Így például az Ír Nemzeti Könyvtár (NLI) 2019-ben bejelentette, hogy a felkaptott szerző, Marian Keyes (született 1963-ban), a *Mercy Close-i rejtély* című regényének kiadásával kapcsolatos digitális anyagokat adományozott a számukra. Ez magában foglalta a borítóterveket, a piszkolatokat és a nyomdai előkészítés előtti példányokat; videókat, amelyeken a szerző a könyv szereplőiről, valamint az írás folyamatáról beszél; továbbá promóciós anyagot, közösségi média bejegyzéseket és a regény kronológiáját.<sup>13</sup> A könyvtár munkatársai ezenfelül fotókat készítettek Keyes dolgozószobájáról – a *Kiterjesztett Gyűjtemény*, egy a British Library által teremtett koncepció szellemiségének megfelelően (Keating és Finegan 2021). A Kiterjesztett Gyűjtemény a kortárs archívumok szerzeményeinek keretét adó új tartalmak létrehozását célozza, különböző módszerekkel – úgymint az írók dolgozószobáit ábrázoló interaktív fotók készítése, a gyűjtemény létrehozójával felvett interjúk készítése és az alkotót körülvevő környezetben tett videóbeszélgetések körutak dokumentálása. A kezdeményezés lehetővé teszi a kutatók számára, hogy további tartalmakat fedezzenek fel, ily módon újra átélhetővé téve az új gyűjteményre először rábukkanó kurátor kiváltságos helyzetét.

Keyes rendkívül aktív az interneten, és az NLI web archivációs programjának részeként már folyamatosan gyűjti a weboldalon és a twitteres hírcsatornáján megjelentetett információkat. Az NLI ezzel a born-digital gyűjteménnyel azt az erős üzenetet közvetíti, hogy érdemes gyűjteni a populáris irodalmat képviselő íróknak-tól származó dokumentumokat. A haladást képviselő intézményként definiálják magukat, akiket nem korlátoznak azok a gyűjteményképzési minták, amelyek hagyományosan a férfi szerzőkre és a papíron keletkezett dokumentumokra fókuszálnak, valamint szigorúan veszik a hierarchiát a magas kultúra és a popkultúra között. „Az igazgatónk a diverzitás és a befogadás nagy bajnoka” – mondta Della Keating (könyvtáros asszisztens, Digitális Gyűjteményi Osztály, NLI) – „és nagy-

<sup>13</sup> <https://www.nli.ie/en/list/latest-news.aspx?article=2a51d71e-584f-4779-a3f0-0f3d023cc179> (utolsó letöltés: 2021. november 19.).

mértékben beépült a gondolkodásmódunkba, egyben a könyvtár szerepével kapcsolatos felfogásunkba, hogy visszatükrözzük a társadalomban folyó változásokat” (Keating and Finegan 2021). A Marian Keyes Archívum részét képezi egy bevezető born-digital projekt sorozatnak – amely magában foglalja a Yes Equality Vizuális Digitális Archívumot (a 2015. évi írószági, a házasság egyenlőségéről szóló népszavazással kapcsolatos).<sup>14</sup>

A Transport for London (TFL) a másik példa az olyan intézményekre, amelyek aktív szerepet játszanak az iratképzőkkel való kapcsolat felvételében, ahelyett, hogy passzívan arra várnának, hogy a gyűjteményt ők ajánlják fel megőrzésre a levéltárba. Tamara Thornhill, a TFL levéltárosa kulcsszerepet játszott a 2012. évi londoni olimpiával kapcsolatos digitális gyűjtemény megalkotásában. A dokumentumok számára a megfelelő kiállítófelület kialakítása érdekében együttműködött az intézmény technikai munkatársaival. Következő lépésként fel kellett venni a kapcsolatot azokkal a cégekkel, amelyek a játékokon szállítmányozást végeztek, és felkérni őket a dokumentumaik átadására az online tárhely számára. Thornhill csapata oral history interjút is készített, amelyeken a dolgozók az olimpia szállítmányozóiként szerzett élményeikről beszéltek. Egyes vezető beosztású dolgozók még mindig „média üzemmódban” voltak, amikor az interjút készítették, és beszámolójukkal olyan benyomást keltettek, mintha minden tökéletesen el lett volna rendezve és semmilyen fennakadás nem fordult elő. Két évvel később újabb interjút készítettek, amelyek már árnyaltabb információval szolgáltak. Az oral history interjúkkal kapcsolatos szervezőmunka mellett Thornhill sokat próbálkozott a céges közösségi oldalakon és a céges intranetfelületeken is, hogy további, a játékokkal kapcsolatos born-digital dokumentumokat kerítsen elő, csak hogy biztos lehessen, *„tényleg a lehető legteljesebb gyűjtemény áll a rendelkezésünkre”* (Thornhill 2021).

Ez a fajta aktív hozzáállás megosztja a levéltáros közösséget. *„A levéltárosok hajlamosak kissé ferde szemmel nézni erre a dologra, mondván: Csak nem történelemhamisítást végeztek? Vagy egyenesen a dokumentumot kívánjátok értelmezéssel ellátni?”* – mondta egy interjúalanyunk, aki az Egyesült Királyság Központi Kormányzatánál dolgozik tudás és információ menedzsment területen ([anon.] 2020). Valóban, a 19. századi *„respect des fonds”* alapelv passzív szerepet ruház a levéltárosra, aki a dokumentumok és azok eredeti elrendezésének megváltoztatása nélkül veszik át a gyűjteményeket azok keletkeztetőjétől. A born-digital gyűjtemények azonban gyakran híján vannak bármiféle rendszernek – könnyedén készíthető róluk másodpéldány, és több eszközön és platformon is képesek aktív beavatkozás nélkül is jelen lenni. Az aktív hozzáállás átvétele a born-digital gyűjtemények kialakításánál ennél fogva tekinthető egyfajta, a jövőbeli kutatók szükségleteire adott pragmatikus válasznak. Csakugyan, a Brexit, vagy a londoni olimpia iránt érdeklődő kutatók könnyebbnek látják majd, ha a dokumentumok mind egy

<sup>14</sup> <https://www.nli.ie/en/born-digital-collection.aspx> (utolsó letöltés: 2021. november 20.).

helyütt pontos jegyzékkel ellátva lesznek megtalálhatóak, mintha több különböző helyen lennének elszórva.

Azonban, mint láttuk, a legtöbb born-digital gyűjtemény nem áll nyitva a kutatók előtt. Ez vonatkozik a NLI-ben a Marian Keyes Archívumra, amely jelenleg feldolgozás alatt áll és a TFL Londoni Olimpia Archívumára is, amelyhez a személyzet néhány tagja rendelkezik hozzáféréssel. Melyek a hozzáférés fő akadályai? A digitizált archívumokhoz hasonlóan a born-digital gyűjtemények is gyakran tartalmaznak szenzitív adatokat. A gyakorlatban sok intézmény inkább a teljes elzárás mellett dönt, minthogy problematikus dokumentumok nyilvánosságra hozatalát kockáztassa. Az Egyesült Államokban, ahol az adatvédelmi jogszabályzás jóval kevésbé szigorú, mint Európában, előfordultak esetek, melyek során born-digital archívumokban megnevezett élő személyek személyiségi jogai sérültek. Interjúalanyaink egyike megemlítette, hogy a Susan Sontag Archívum esete miatt a Kaliforniai Egyetemen (University of California, Los Angeles – UCLA) változott meg a véleménye a levéltári adatvédelem kérdésében:

*„Pályafutásom korai szakaszában nem igazán hittem abban, hogy a levéltár felelőssége lenne az adatok védelmének és a titoktartásnak a biztosítása. És annak következtében, ahogy alkalmazták, gyakran az a benyomásom támadt, hogy a szenzibilitás csak eszköz, hogy a valóban fontos információkat visszatarthassák. Hogy a szenzibilitást rendszerint a hatalom birtokosai alkalmazzák, ezzel akadályozva meg olyan információk kijutását, melyet ők nem kívántak nyilvánosságra hozni”* ([anon.] 2020).

A fordulópontot egy a Los Angeles Review of Books által 2014-ben közölt cikk jelentette, melyet két, az UCLA-n működő doktorandusz írt, akik hozzáféréssel rendelkeztek Sontag archivált emailjeihez, számítógépes tárhelyeihez (Schmidt és Adam 2014). A cikk lerántotta a leplet Sontagnak az Annie Leibovitz fotóssal folytatott viszonyáról, amelyet Sontag egész élete során végig tagadott (2004-ben hunyt el). Az olvasók arról is tudomást szereztek, hogy Sontag számítógépének tárhelye Leibovitz részére nyújtott béranyai szolgáltatással kapcsolatos szenzitív iratokat is tartalmaz. Semmi törvényellenes nem volt a cikkben, interjúalanyunk viszont a magánszféra mélységesen etikátlan megsértését látta benne. Interjúalanyunk úgy vélte, egy olyan korban, amikor a vállalatok hatalmas mennyiségű adatot gyűjtenek be az egyének telefonjairól és számítógépeiről, a könyvtáraknak és a levéltáraknak ellen kellene állniuk ennek a trendnek és tevélegesen védelmezni a magánszférát (Ryan Cordell legutóbbi, *Gépi tanulás és a könyvtárak* című elemzésében [2020] hangoztatott érvelés).

A bizalmas és szenzitív adatokat tartalmazó dokumentumok ugyanúgy érintik a papíron keletkezett iratokból álló gyűjteményeket is, és a levéltárak régóta egy jövőbeli időpont megállapítása útján szabályozzák, mikortól lehet a dokumentumokhoz hozzáférést biztosítani. Még a vállalati irattárak is – amelyeket nem kötelezi a törvény, hogy a gyűjteményeikhez nyilvános hozzáférést biztosítsanak – gyakorta lehetővé teszik a kutatók számára, hogy bizonyos iratokat egy megszabott

időtartam leteltével kutathassanak. A kérdés, hogy nyilvánossá tegyék-e az irattárat vagy egy részét, gyakran heves vitákat kelt a vállalaton belül. Vegyük például a British Telecom Levéltárát Londonban, amely őrzésében egészen 1846-ig visszamenőleg vannak iratai a távközlési vállalatnak. A British Telecom 1984. évi privatizációját megelőzően keletkezett dokumentumok nyilvánosak. A kutatóknak a köziratokról szóló törvény és az információszabadságról szóló törvény értelmében van joguk beléjük tekinteni. Azonban „*még mindig van néhány dolog, amely minősítési jelöléssel rendelkezik és ezért nem hozzáférhető*” – mondta James Elder, aki a BT levéltárának igazgatójaként dolgozott 32 éven át (Elder és Hay 2021). Az 1984 után keletkezett dokumentumok (köztük a born-digital dokumentumok) többnyire el vannak zárva a nyilvánosság elől. Mint David Hay, a BT örökségvédelmi és levéltári szakértője állítja:

„*Továbbra is rendkívül nehéznek bizonyul nyilvánossá tenni, még a privatizációt követő időszak korai szakaszából származó dokumentumokat is, konkrétan az Ön által említett kockázattal összefüggő probléma következtében, amelynél a jogászok és a felső vezetők azt veszik figyelembe, hogy a BT egy óriási, elképesztően szigorúan szabályozott vállalat. Vagyis a szabályalkotót és a kormányzatot kifejezetten érdekli, amit csinálunk. Hatalmas adózási gondjaink vannak, folytonos vitában állunk a HMRC-vel (Her Majesty's Revenue and Customs, a brit adóhatóság – a ford.). Rengeteg a munkahelyi balesettel összefüggő követelés. Miért is akarnánk nyilvánossá tenni bármit, ami valamilyen lehetséges veszélynek tenné ki a BT-t? Sokkal egyszerűbb elzárva tartani, főleg, hogy törvényileg is jogosultak vagyunk rá*” (2021. május 28.).

Még ezen rendkívül korlátozó hozzáférési szabályzás keretein belül is vannak tervek egyes anyagok elérhetővé tételére egy még meghatározásra váró időkorlátal. „*Elképzelhetnénk egy olyan felállást, ahol a dokumentumok nyilvánossá tétele egy 30, 40, 50 éves fordulóval, vagy valami ilyesmi alapján haladhat előre*” – mondta Elder, „*és akkor egy applikáción keresztül elérhetővé válnának a kutatók számára, és ez egy előre bejelentett kutatási témán alapuló kutatás alapján és azzal a feltétellel történne, hogy publikálást megelőzően megosszák velünk az eredményeket és lehetőségünk lesz némi ráhatást gyakorolni, azt hiszem ...*” (2021. május 28.). A cenzúráról való félelmében sok kutató ellenezné a kutatási eredményeinek a publikáció előtti megosztását a levéltárral. Talán azonban ezt az árat meg kell fizetni, hogy hozzá lehessen férni egyébként elérhetetlen dokumentumokhoz.

A dokumentumok keletkezése és a nyilvánossá válásuk közötti hosszú időintervallum elméletben bátoríthatná a levéltárosokat a gyűjtemények lehető leginkább felhasználóbarátta tételében. Végso soron az mindig könnyebb, ha a dokumentumok feldolgozására és rendezésére bőséges idő áll rendelkezésre, így jut még esetleg egy stratégia kidolgozására is, hogy a gyűjtemények szenzitív adatokat nem tartalmazó elemeit hamarabb nyilvánossá lehessen tenni. A tudat azonban, hogy a gyűjtemények még 20 évig vagy tovább nem lesznek elérhetőek, kivárára játszó magatartáshoz vezethet. Mint egyik interjúalanyunk mondta: „*Miért fordítanánk*

*most arra sok időt, hogy azon dolgozzunk, miként biztosítsuk a hozzáférést, mikor valójában semennyire nem tudjuk biztosítani azt a hozzáférést jogszerűen?”* ([anon.] 2021a).

Az adatvédelem nem az egyetlen akadályozó tényező, amely korlátozza a kutatók born-digital gyűjteményekhez való hozzáférését. A gyűjtemények mérete szintén nagy gondot jelenthet – különösen kisebb intézmények számára, amelyek korlátozott emberi és pénzügyi erőforrásokkal rendelkeznek hatalmas adatt mennyiségek tárolása és érdemi kezelése tekintetében. A méret gyakran nehezíti a born-digital gyűjtemények megismerését, minthogy a levéltárak megszenvednek a jegyzékelésükkel (Hosker 2021). Szintúgy megnehezíti több, a teljes anyagról készült másolat egyidejű megőrzését, ami pedig fontos alapelv lenne a gyűjtemény fennmaradásának biztosítása érdekében. Jelenleg egyetlen, a British Library által őrzött példány létezik az Egyesült Királyság Web Archívumáról, ezen osztoznak a többi letéti könyvtárral, bár a Skót Nemzeti Könyvtár nemsokára rendelkezik majd egy második példánnyal (Lewis és mások 2021). Teljes gyűjtemények hozzáférhetővé tételére nincs lehetőség, bár kisebb részeket néha a kutatók rendelkezésére bocsátanak. Az Ír Népzenei Archívum terabájtokra rugó anyagot tartalmaz (például fellépésekről készült videófelvételeket). Rengeteg munkával jár ezen felvételek közül néhánynak is az online közzététele, és a másolatok megfelelő minőségéről való gondoskodás (Harkin és O’Connor 2021).

A megőrzést és a hozzáférést biztosító munkafolyamatok közti különbségek a felhasználói élményben is fájó pontokat hoznak létre. Euan Cochrane (digitális megőrzési igazgató, Yale Egyetem Könyvtára) kifejti, hogy a born-digital dokumentumok beszerzése és feldolgozása ömlesztett formában történik. A megőrzési stádiumban a tartalom ellenőrzését végző levéltáros lemezképet készíthet a teljes állományról és eltávolíthatja azt. Ezzel ellentétben az hozzáférést segítő munkafolyamat sokkal tekervényesebb és időigényesebb. *„Ha valaki hozzáférést igényel, [a levéltárosnak] ki kell vennie a fájlokat a lemezképből, megállapítania, mit tartalmaznak, felülvizsgálnia és esetleg még jegyzékelnie is egy keveset, mert még nincsenek eléggé jegyzékelve”* (Cochrane 2021). Röviden, a hozzáférést biztosító munkafolyamatot sokkal nehezebb automatizálni, mint a megőrzésre irányuló folyamatot.

Sok közgyűjtemény számára nagy kihívást jelent egy, a kutatók számára a born-digital anyagokhoz hozzáférést biztosító megfelelő interfész kialakítása, mivel gondoskodni kell arról, hogy a felhasználók ne legyenek képesek a dokumentumokat módosítani, vagy törölni (Hosker 2021). Ahogy a born-digital állományok gondozásának más vonatkozásaiban, a felhasználók digitális fájlokhoz való hozzáféréseinek a felügyelete is magas fokú műszaki képzettséget igényel. Callum McKean (főkurátor, Kortárs Gyűjtemények és Kéziratok, British Library) ezt mondta: *„Nem csupán kurátorokra és műszakiakra van szükség, és persze jó, ha bőségesen áll rendelkezésre idő is, de úgy gondolom, olyan emberek kellene, akik úgymond képesek mozogni e két világ között és tolmácsolni közöttük, hogy hatékonyan végezhesék a feladatot”* (McKean 2021). McKean személy szerint elvégzett egy alkalmazott adat-

tudományi kurzust, hogy a digitális gyűjtemények kezelését segítő programozási eljárásokat tanuljon. Ez a fajta képzettség még szokatlan a közgyűjteményi szakemberek körében. McKean szerint ezen készségek fejlesztése a személyzet körében, vagy ilyen képzettséggel rendelkező személyek felvétele, kulcsfontosságú kellene, hogy legyen az ágazat számára.

Mások, mint például Mark Bell (vezető digitális kutató, Brit Nemzeti Levéltár), úgy vélik, hogy nem életbevágóan fontos, hogy a digitális gyűjteményekkel foglalkozó szakemberek közül mindenki programozói tudással rendelkezzen. Még a levéltári felhasználású mesterséges intelligenciával foglalkozó dolgozóknak sem feltétlenül szükséges programozási nyelvek ismerete. Szükségük van azonban „*az adatkimenetek értelmezésének képességére, legyenek azok akár grafikonok vagy valószínűség-számok, ... illetve numerikusan vagy vizuálisan megjelenített értékek*” (Bell 2021). Összegezve, jelenleg nincs megegyezés abban, milyen formában kellene megvalósulnia ennek a műszaki képzésnek, de az olyan szakemberek, mint McKean és Bell, abban egyetértenek, hogy az ágazat előtt sok tennivaló áll a tudásbázis fejlesztésének tekintetében.

### A digitális gyűjteményekhez való hozzáférés jelenlegi szintjei

Adatvédelmi szabályok, az adatok mennyisége és képzettségbeli hiányosságok: ezek a fő akadályok magyarázzák azt, hogy a born-digital gyűjtemények közül miért olyan sok elérhetetlen a kutatók számára. Sally McInnes (különgyűjteményi és gyűjteménygondozási igazgató, Walesi Nemzeti Könyvtár – NLW) megjegyzi, hogy az NLW-ben őrzött emailarchívumok „*csak úgy vannak a hálózaton valahol*”: „*még nem láttunk hozzá a beazonosításukhoz, de majd fogunk*” (McInnes 2021). Ezek a gyűjtemények teljesen rejtve vannak és megismerhetetlenek a kutatók számára. Ugyanígy jártak a Yale Egyetem Könyvtárának rendkívül szenzitív adatokat tartalmazó emailgyűjteményei is. Cochrane példaként veti fel azt a plasztikai sebészt, aki közvetlenül a halála előtt helyezte letétbe az iratait. „*Ez tartalmazta a postafiókját, az emailfiókját, több százezer emaillel, ... közülük sok páciensekre vonatkozó adatokat tartalmaz. Úgyhogy elképzelhetik, mit lehetne ezzel művelni*” (Cochrane 2021).

Még ha egyes born-digital, illetve digitizált gyűjtemények elérhetőek is, a kutatóknak gyakran az őrzés helyszínére kell menniük, hogy az iratokba betekintést nyerjenek – ami a hozzáférést azokra korlátozza, akik elegendő idővel, pénzzel és testi adottsággal rendelkeznek az utazáshoz. A Seven Storiesnak például létezik egy Chris Fulton illusztrátorhoz köthető gyűjteménye, amely Photoshop fájlokat, PDF-eket, JPEG-eket és más fájlformátumokban levő állományokat tartalmaz. A kutatóknak Newcastle városába kell elutazniuk, hogy ezeket a digitális illusztrációkat vizsgálhassák, részben amiatt, mert az őrzést végző intézménynek nincs kapacitása arra, hogy online is elérhetővé tegye őket (McKie 2021). A Transport for London Levéltárban az anyagok kizárólag helyben érhetőek el. Nem minden anya-

got lehet azonban megismerni. Ahogy Tamara Thornhill megállapította: „Az elektronikus rendszereinkben arra van csupán lehetőségünk, hogy nyílttá, vagy zárttá tegyünk egy állományt, nincs olyan opció, hogy valaminek csak a metaadatait adjuk át” (Thornhill 2021). A gyakorlatban ez azt jelenti, hogy a levéltárosoknak nincs lehetőségük középutas megoldásként csak a metaadatokat megosztani, számos dokumentumot felderíthetetlenül hagyva ezzel. A brit webes archívumokra nézve a nem nyomtatott formában keletkezett könyvtári letétre vonatkozó szabályozásokról szóló, 2013-ban kiadott rendelet szintén előírja, hogy a kutatók olyan anyagok kutatása érdekében is az őrzés helyére utazzanak, amelyek korábban az interneten nyilvánosan elérhetőek voltak. A gyűjtemény legnagyobb része a hat letéti könyvtár összesen kilenc telephelyének valamelyikén tekinthető meg: a British Library (London, St. Pancras és Boston Spa), Walesi Nemzeti Könyvtár (Aberystwyth és Cardiff), Skót Nemzeti Könyvtár (Edinburgh és Glasgow), Bodleiana Könyvtár (Oxford), Cambridge Egyetemi Könyvtár, valamint a dublini Trinity College Könyvtára. Ezenfelül valamivel több mint 19 000 weboldalnak – töredéke az Egyesült Királyság szférájába tartozó, nagyjából valószínűleg 5-10 millió weboldalnak – engedélyezte a tulajdonosa, hogy bárhonnan elérhető legyen a UK Web Archive weboldalon keresztül ([anon.] 2021a).<sup>15</sup>

A helyzet nagyon különbözik ettől az Egyesült Államokban, ahol az Internet Archívum rendszeresen gyűjt és tesz nyilvánosan elérhetővé weboldalakat. A méltányos felhasználásra vonatkozó jogi szabályzás az USA-ban lényegesen liberálisabb, mint a nem nyomtatott formában keletkezett könyvtári letétre vonatkozó szabályozások az Egyesült Királyságban. Míg a brit jogi keretrendszer erősen korlátozó jellegű a hozzáférés tekintetében, ugyanakkor egyértelmű felhatalmazást is ad a letéti könyvtáraknak, hogy olyan anyagokat gyűjtsenek, amelyek egyébként elvesznének. Egyik interjúalanyunk ezt mondta: „A tény, hogy törvényi megbízatásunk és kötelességünk ezeknek a gyűjtése ... jó dolog, mert a British Library és más hasonló könyvtárak hajlamosak hosszú időtávon gondolkodni és azt hiszem, jó esély van rá, hogy létezőnk még 50 év múlva is, 100 év múlva is” ([anon] 2021a). Mások kritikusabban fogják fel a jelenlegi állapotot, amely arra kényszeríti a felhasználót, hogy könyvtári számítógépekről férjen hozzá az Egyesült Királyság web archívumaihoz. William Kilbride (a Digitális Megőrzési Koalíció ügyvezető igazgatója) megjegyezte: „Azok is nagyon kevesen vannak, akik egyáltalán tudják, hogy ez itt van, vagy hogyan tudják kijárni, hogy kutathassák” (Kilbride 2021). Kilbride felfogása szerint nem létezhet megőrzés a hozzáférés biztosítása nélkül, mivel ahhoz alapvetően nélkülözhetetlen, hogy helyesen megértsük a gyűjtemény használó közösség szükségleteit és elvárásait.

A digitizált gyűjtemények esetében az azt végző üzleti vállalkozásokkal kötött szerződések gyakran lehetetlenné teszik, hogy a kulturális intézmények széles kör-

<sup>15</sup> Lásd még <https://www.bl.uk/collection-guides/uk-web-archive> (utolsó letöltés: 2021. november 20.).

ben megosszák a birtokukban levő eredetiktől készült digitális másolatokat.<sup>16</sup> Jellemző, hogy a cégek az általuk készített digitális másolatok elérését fizetőssé teszik. A könyvtárak és levéltárak ingyenes másolatokat kapnak, amelyekhez helyben hozzáférést biztosíthatnak, de nem tehetik ki azokat a weboldalaikra, sem más internetes felületre. *„Lényegében az történik, hogy átruházzuk a saját gyűjteményeinket valaki másra, hogy aztán ne tehesük fel szabadon az internetre”* – mondta Stuart Lewis. A Skót Nemzeti Könyvtár egy másik módszerre tervez áttérni: *„Tulajdonképpen azt próbáljuk velük közölni: Mi végezzük a digitalizációt, és Önök fizetnek nekünk a digitalizáció elvégzéséért, ugyanis gyakorta kiszervezik azt egy harmadik fél számára, és sosem vagyunk igazán elégedettek az általuk végzett munka minőségével. Ezért lényegében igyekszünk megtartani a magunk számára ezt a feladatkört”* (Lewis és mások 2021).

A pandémia különösen rávilágított erre a s problémára a digitizált gyűjteményekhez való hozzáféréssel kapcsolatban. Mint egyik interjúalanyunk rámutatott: digitizált kulturális örökségünk nagy részéhez nem volt szabad hozzáférés, még-hozzá elsősorban épp digitalizációs megállapodásaink következtében. Habár a könyvtáraknak és a levéltáraknak a jogdíjak formájában bevételi forrásuk származik ezekből az üzleti szerződésekből, *„nem érzékelik annak az anyagnak a tényleges gazdasági értékét”*, amit digitalizációs projektek számára biztosítanak ([anon.] 2021a). A felhasználók korlátozva vannak a digitizált anyagokkal végezhető kutatás módjában is. A metaadatok nem szabad hozzáférésűek, ez pedig leszűkíti a kutatók lehetőségeit az elmúlt nagyjából két évtized folyamán készült digitizált anyagok érdemi minőségének megállapítására.

Amikor a pandémia idején a kulturális intézmények bezárásra kerültek, a kutatóknak (és, az akadémiai kutatók esetében, intézeteiknek is) nem volt más választásuk, mint fizetni a fizetőfalakon keresztül kínált digitális másolatokért. Paul Seaward, a Parlamenttörténeti Bizottság igazgatója rámutatott: egyes, a British Libraryban és a Brit Nemzeti Levéltárban őrzött kulcsfontosságú gyűjteményeket a ProQuest digitizált, *„és ezek csak nagyon drágán ... üzleti célú online felületen keresztül érhetőek el, szóval ez probléma”* (Seaward 2021).

## A hozzáféréssel kapcsolatos problémák lehetséges megoldásai

A digitális gyűjteményekhez való hozzáférés problematikájára nem lesz könnyű megoldást találni. Elég csak a nem nyomtatott könyvtári letételre vonatkozó szabályozást példaként felhozni. Számos interjúalany rámutatott, hogy ez a jogszabályi keret nem kielégítő egy olyan korban, amikor az open access vált a fő hajtóerővé. Nem sok értelme van korlátozni a hozzáférést olyan webarchívumokhoz, amelyek

<sup>16</sup> A digitalizációt végző üzleti vállalkozásokkal (különös tekintettel a digitális sajtóarchívumokra) kapcsolatos további információkért lásd Hauswedell és mások (2020), Milne (2002), Smith (2006), Terras (2011) és Gooding (2014).

értékes kutatási eredmények forrásaként szolgálhatnak, amiket aztán a kutatók amúgy is nyílt hozzáférésű folyóiratokban kötelesek publikálni. Még ha a jelen jogi szabályozás szigora enyhülne is, sok kulturális intézmény akkor is inkább a kockázatkerülés mellett maradna. Ez sok olyan kutatót frusztrál majd, akiktől megtagadták, hogy fontos dokumentumokhoz férjenek hozzá. Azon kulturális intézmények, amelyek elszántan őrzik évszázados hírnevüket, különösen nagy kockázatként fogják fel a problémás anyagok kikerülését. *„Olyan szervezet akarunk lenni, amelyben meg lehet bízni. Úgy gondoljuk, már annak számítunk – magas fokú bizalmat élvezünk és meg szeretnénk azt tartani”* – nyilatkozta egy interjúalany ([anon.] 2021a).

Valamennyi digitális gyűjteményt nyílt hozzáférésűvé tenni az interneten mindig is utópia marad, ezért néhány intézmény gyakorlatiasabb megközelítéssel élve, regisztrált felhasználók által igénybevehető virtuális olvasószobákat hozott létre. A Yale Egyetem Könyvtára jelenleg egy online környezet kiépítésén dolgozik, ahol a kutatók a saját számítógépükön futtatott webböngésző segítségével nézhetnek meg korlátozott hozzáférésű tartalmakat. Lehetőségük lesz a saját eszközeiket használni, de az anyagokat nem tudják majd letölteni (Cochrane 2021). A Brit Nemzeti Levéltárnál Mark Bell szintén érdeklődik a különböző hozzáférési szinteket biztosító biztonságos virtuális környezetek iránt. Ehhez számára egy, az Alan Turing Intézettel közösen szervezett webarchívumokra szakosodott adatvizsgáló csoporttal szerzett tapasztalatai szolgáztatják az alapot. Mivel Bell és munkatársai nem szenzitív anyagokkal dolgoztak, az általuk használt virtuális környezet a legalacsonyabb szintű biztonsági besorolással rendelkezett. Ez lehetővé tette számukra, hogy új adathalmazokat vigyenek fel, másoljanak és beillesszenek, és így tovább. Ezek a funkciók fokozatosan kiestek a magasabb biztonsági szinteken. A legmagasabb fokozaton már nem lehetett képernyőmentést készíteni, sem adatot felvinni, sem kivenni. A környezet erősen le volt korlátozva, de még mindig biztosította a felhasználók számára, hogy dolgozhassanak az adatokkal. Bell szerint ez a modell, amely erősen korlátozva a végezhető műveleteket, de hozzáférést biztosít egy számítógéphez, „járható út” a teljesen hozzáférhetetlen adatok problémájának megoldása felé (Bell 2021).

Egyre több intézmény fedezi fel magának a mesterséges intelligenciát és a gépi tanulást, hogy a gyűjteményeit megismerhetőbbé és elérhetőbbé tegye a kutatók számára. A metaadatok automatikus generálása olyan feladat, amelyet át tudnak venni az MI rendszerek. Az Ír Népzenei Archívum ilyenformán dallamok megnevezésére próbálja az MI-t /ML-t használni. Egy, az 1960-as években készült nyitott tekercses magnószalagról másfél óra zenei anyag digitalizálható, ami 50-60 dallamot tartalmaz. Egy többszáz magnótekercset tartalmazó gyűjtemény esetében az MI a dallamokból adathalmazok készítését tenné lehetővé, amelyekhez aztán a kutatóknak hozzáférést lehetne biztosítani (Harkin és O'Connor 2021).

A webarchívumok a gyűjteményeknek egy másik, nagyon minimális metaadatot tartalmazó formáját adják. Ez tartalmazhatja az oldal címét, de semmi mást. A MI /ML képes a weboldalakakat tág kategóriákba sorolni, úgymint politika, vagy kultúra. Ez viszont segíti a kutatókat, akik így az őket érdeklő témákra tudnak koncentrálni

([anon.] 2021a). A mesterséges intelligencia ráadásul alkalmazható ellentmondásos metaadatok azonosítására a történeti gyűjteményekben, köztük olyan festmények esetében, amelyekről előbb fénykép készült és később azt digitalizálták. Antoine Isaac (kutatási és fejlesztési igazgató, Europeana) rámutatott, hogy a cél nem a kellemetlen történelmi események elrejtése: *„ha van egy festményünk, ami rabszolgákat ábrázol, nem leszünk képesek eltitkolni a tényt, hogy létezett rabszolgaság”* (Isaac 2021). A megosztó fogalmakkal kapcsolatos növekvő tudatosság azonban a gyűjteménysmertető írásának új módjaihoz vezethet – egy olyan korban, amikor egyes múzeumok nekiláttak a kiállított műtárgyaik címének módosításához, például, hogy az a rabszolgaságba vetett személy egyéniségét emelje ki.

Az MI/ML egy másik felhasználási módja a fájltypusok meghatározása. Santhilata Venkata (digitális megőrzési specialista, Brit Nemzeti Levéltár) megjegyezte, hogy a munkahelyén gyakran kapnak sérült adatállományú szöveges fájlokból álló dokumentumokat. A mesterséges intelligencia képes automatikusan megvizsgálni a fájl tartalmát és megállapítani a fájl típusát, ami ugyancsak hasznos a digitális megőrzés és hozzáférés tekintetében (Venkata 2021).

A metaadatok javítása és a fájltypus-megállapítás természetesen nagyon fontos, de az MI/ML talán legfontosabb levéltári alkalmazási lehetősége a szenzitív anyagok azonosításának képessége. A szenzitivitási vizsgálat már jelenleg is működő funkciója az ePADD-nak, a Stanford Egyetem Különgyűjteménye és Egyetemi Levéltára által fejlesztett ingyenes nyílt forráskódú szoftvernek, ami támogatja a történeti értékű e-mail archívumok felmérését, feldolgozását, megőrzését, felderítését és továbbítását. Azonban nincs mindenki meggyőződve arról, hogy az ePADD Natural Language Processing funkciója elegendő lenne a problémás és problémamentes anyagok szétválogatására. Callum McKean, aki Wendy Cope brit költő archívumát gondozta, ezt mondta: *„Belefutsz olyan dolgokba, amiket az ePADD problémaként jelöl, mint például a kábítószerhez sorolt bevásárlólistája, mivel gombát vásárolt és ... a szótár szerint a gomba szó kábítószernek minősül, szóval elég sok a téves riasztás”* (McKean 2021). Ellentétképpen az ePADD szokott problémamentes minősítést is kiadni tévesen, mivel Cope igencsak személyes információkat közöl más írókkal kapcsolatban kódolt formában, amin a gép képtelen átlátni.

A könyvtári és a levéltári mesterséges intelligencia, illetve gépi tanulás kapcsán lényeges, hogy a felhasználók hallassák a hangjukat és befolyást gyakoroljanak a gyűjteményekhez való hozzáférés, valamint használatuk módjára. Mint láttuk, a könyvtáros és levéltáros szakemberek az MI/ML-t használhatják metaadatok kinyerésére, vagy a katalogizálási és a karakterizálási műveletek segítésére. A szenzitív és nem szenzitív dokumentumok megkülönböztetése szintén lényeges alkalmazása ennek az új technológiának. A mesterséges intelligenciát és a gépi tanulást azonban a kutatók is használhatják, hogy hatalmas mennyiségben tudjanak dokumentumokat (például webarchívumokat) kielemezni. Ezek a technológiai vívmányok megkívánják az adatokhoz való hozzáférést. Hozzáférés nélkül nem valósulhatnak meg a tudás új formái.

Míg az MI-ben és az ML-ben megvan a képesség, hogy segítséget nyújtson olyan feladatokban, mint a szenzitivitás-vizsgálat és a metaadat-címkézés, a technikai vívmányok önmagukban oldják meg a hozzáférés problémáját. A szerzői jogi és adatvédelmi korlátok képesek megbénítani a digitális szellemi tulajdonhoz való hozzáférést szolgáló rendszereket. Lehetséges, hogy ezek a törvények egyszer odáig fejlődnek majd, hogy elősegítsék a nem különösebben szenzitív, vagy bizalmas dokumentumokhoz való nyílt hozzáférést. Nagy-Britannia nemzeti adatstratégiája hangsúlyozza, hogy az adatnak a „megfelelő módon hozzáférhetőnek, mobilisnak és újra felhasználhatónak kell lennie”.<sup>17</sup>

Ez a stratégia megállapítja, mi az Egyesült Királyság számára a legjobb módszer az adat hatalmának felszabadítására. A brit kormánynak a mesterséges intelligenciát és az adattudományt a jobb adathozzáférés érdekében kulcsprioritásként kezelő álláspontja határozza meg a brit MI Tanács MI Ütemtervének<sup>18</sup> és az UKRI (UK Research and Innovation – a brit tudományos kutatás támogatását szolgáló nemzeti finanszírozási ügynökség) MI Ütemtervének<sup>19</sup> (mindkettőt 2021 elején adták ki) szövegét is. Különösen az UKRI Ütemterve hangsúlyozza ki az „adatokhoz való gyenge hozzáférés” problémáját és a tudományágakon (elsősorban a bölcsészet- és a természettudományok között) átívelő kutatások bővítésének szükségességét.

A „digitális sötét kor” bekövetkezte miatti növekvő aggodalmak az Egyesült Királyságban és Írországbán a Digital Preservation Coalition (Digitális Megőrzési Szövetség – DPC) létrehozását eredményezték két évtizeddel ezelőtt. Azóta ez több mint 100 tagintézményből álló globális szervezetté nőtte ki magát. A figyelmet a megőrzés felől a hozzáférés irányába mozdítva, az általunk javasolt globális felhasználói közösség, levéltárosokkal, könyvtárosokkal és más szakemberekkel együttműködve közösen gyakorlati megoldásokat dolgozhatnak ki a „sötét” archívumok problémájára. Ez a közösség felhívhatná a figyelmet a hozzáférést a jogi szabályozások formájában nehezítő számottevő korlátokra – például a szerzői jogi törvények szigorának enyhítése és a kutatókra vonatkozó jogi mentességek bővítésének érdekében folytatott lobbitevékenységgel. A közösség emellett kampányolhatnak digitalizálásra fordítható további forrásokért és szerzői jog alóli mentesülésért is. Ahogy a Brit Nemzeti Levéltár munkatársa, Mark Bell állítja: *„kifejezetten fontos lesz, hogy közösségként foglalkozzunk mindezen kihívásokkal... Akár az elképzelésekről, vagy magáról a technológiáról, akár a platformokról vagy csupán módszerek vagy adatok megosztásáról van szó, úgy gondolom, széleskörű megosztásra van szükség”* (Bell 2021).

<sup>17</sup> <https://www.gov.uk/government/publications/uk-national-data-strategy> (utolsó letöltés: 2021. november 18.).

<sup>18</sup> <https://www.gov.uk/government/publications/ai-roadmap> (utolsó letöltés: 2021. november 18.).

<sup>19</sup> <https://www.ukri.org/about-us/what-we-do/strategies-and-reviews/ai-review-transforming-our-world-with-ai> (utolsó letöltés: 2021. november 18.).

## **Hivatkozások:**

- [Anon.] (2020): A szerző által 2020. november 20-án készített interjú.
- [Anon.] (2021a): A szerző által 2021. május 27-én készített két interjú.
- [Anon.] (2021b): A szerző által 2021. május 28-án készített interjú.
- Baldock S. (2020): Wendy Cope's email archive at the British Library, presentation at the conference archives, access and AI: working with born-digital and digitised archival collections. [https:// www.poetrysurvival. com/ program me-aracai-conference/](https://www.poetrysurvival.com/programme-aracai-conference/) (utolsó letöltés: 2021. november 17.).
- Bearman D. (1994): Managing electronic mail. *Archives and Manuscripts* 22: 28–50. <https://publications.archivists.org.au/index.php/asa/article/view/8385/8379> (utolsó letöltés: 2021. július 12.).
- Bell M. (2021): A szerző által 2021. május 27-én készített interjú.
- Cochrane E. (2021): A szerző által 2021. május 26-án készített interjú.
- Cook T. (1994): Electronic records, paper minds: the revolution in information management and archives in the post-custodial and post-modernist era. *Archives and Manuscripts* 22: 300–28. <https://publications.archivists.org.au/index.php/asa/article/view/8433/8427> (utolsó letöltés: 2021. július 12.).
- Cordell R. (2020): Machine learning + libraries. Library of Congress, Washington, D.C. <https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf?loclr=blogsig> (utolsó letöltés: 2021. július 14.).
- Elder J., Hay D.: (2021): A szerző által 2021. május 28-án készített interjú.
- Gondek A. (2021): A szerző által 2021. május 20-án készített interjú.
- Gooding PM (2014): Search all about it : a mixed methods study into the impact of large-scale newspaper digitisation. University College London (University of London), Ph.D.
- Greene M., Meissner D. (2005): More product, less process: revamping traditional archival processing. *American Archivist* 68:208–263. <https://doi.org/10.17723/aarc.68.2.c741823776k65863>
- Harkin T., O'Connor L. (2021): A szerző által 2021. május 20-án készített interjú.
- Hart T.R., de Vries D., Mooney C. (2019): Australian law implications on digital preservation. In: Presentation at the IPRES conference.
- Hauswedell T., Nyhan J., Beals M.H., Terras M., Bell E. (2020): Of global reach yet of situated contexts: an examination of the implicit and explicit selection criteria that shape digital archives of historical newspapers. *Archival Science* 20:139–165. <https://doi.org/10.1007/s10502-020-09332-1>
- Hosker R. (2021): A szerző által 2021. május 19-én készített interjú.
- Isaac A. (2021): A szerző által 2021. május 28-án készített interjú.
- Jaillant L. (2019): After the digital revolution: working with emails and born-digital records in literary and publishers' archives. *Archives and Manuscripts* 47:285–304. <https://doi.org/10.1080/01576895.2019.1640555>

- Johnston J.A., Wallace D.A., Punzalan R.L. (2019): Messages sent, and received? Changing perspectives and policies on US federal email as record and the limits of archival accountability. *Archival Science* 19:309–329. <https://doi.org/10.1007/s10502-019-09318-8>
- Keating D., Finegan J. (2021): A szerző által 2021. május 28-án készített interjú.
- Kilbride W. (2021): A szerző által 2021. május 19-én készített interjú.
- Kirschenbaum M., Ovenden R., Redwine G. (2010) Digital forensics and born-digital content in cultural heritage collections. *CLIR*. <https://www.clir.org/pubs/reports/pub149/> (utolsó letöltés: 2021. július 12.).
- LeClere E. (2018): Breaking rules for good? How archivists manage privacy in large-scale digitisation projects. *Archives and Manuscripts* 46:289–308. <https://doi.org/10.1080/01576895.2018.1547653>
- Lewis S., Hamilton G., Rigden S. (2021): A szerző által 2021. május 19-én készített interjú.
- Martínez-Cardama S., Pacios A.R. (2022): National archives' priorities: an international overview. *Arch. Sci.* 22:1–42. <https://doi.org/10.1007/s10502-021-09367-y>
- McInnes S. (2021): A szerző által 2021. május 26-án készített interjú.
- McKean C. (2021): A szerző által 2021. május 28-án készített interjú.
- McKie K. (2021): A szerző által 2021. május 28-án készített interjú.
- Meissner D., Greene M. (2010): More application while less appreciation: the adopters and antagonists of MPLP. *J. Arch. Organ.* 8:174–226. <https://doi.org/10.1080/15332748.2010.554069>
- Milne R. (2002): The research support libraries programme, the distributed national collection, and the importance of cross-sectoral collaboration. In: McKitterick D (ed.) *Do we want to keep our newspapers?* pp. 81–85.
- Moravec M (2017) Feminist research practices and digital archives. *Australian Feminist Studies* 32:186–201. <https://doi.org/10.1080/08164649.2017.1357006>
- Peltzman S., Dietz B., Butler D., Walker P., Farrell J., Arroyo-Ramirez E., Macquarie C., Bolding K., Helms A.M., Venlet J. (2020): Levels of born-digital access. Digital Library Federation (DLF) part of the Council on Library and Information Resources (CLIR), Alexandria, VA. <https://osf.io/r5f78/> (utolsó letöltés: 2021. július 12.).
- Prom C. (2019): *Preserving Email*, 2nd Edition. Digital Preservation Coalition. <https://doi.org/10.7207/twr19-01>
- Ross S., Gow A. (1999): *Digital archaeology: rescuing neglected and damaged data resources: a JISC/NPO study with the Electronic Libraries (eLib) Programme on the preservation of electronic materials*. Library Information Technology Centre, London.

- Schmidt J., Ardam J. (2014): On Excess: Susan Sontag's Born-Digital Archive. Los Angeles Review of Books. <https://lareviewofbooks.org/article/excess-susan-sontags-born-digital-archive/> (utolsó letöltés: 2021. július 14.).
- Seaward P (2021): A szerző által 2021. május 28-án készített interjú.
- Seles A (2021): A szerző által 2021. május 28-án készített interjú.
- Smith N. (2006) Digitising documents for public access. In: MacDonald L. (ed.) Digital heritage: applying digital imaging to cultural heritage. Butterworth-Heinemann, Oxford, pp. 3–32.
- Terras M. (2011) The rise of digitization: an overview. In: Rikowski R. (ed.) Digitisation perspectives. Sense Publishers, Rotterdam, pp. 3–20.
- Thornhill T. (2021): A szerző által 2021. május 27-én készített interjú.
- Venkata S. (2021): A szerző által 2021. május 21-én készített interjú.

**Lise Jaillant** vezető tanár (egyetemi docens) a Loughborough Egyetem Digitális Bölcsészeti Tanszékén. Fő kutatási területe a publikációtörténet és a digitális bölcsészet. Az elmúlt öt évben tapasztalatot szerzett a born-digital archívumok és azok megőrzéssel/hozzáféréssel kapcsolatos problémáinak területén. 2020 óta három, a levéltárak és mesterséges intelligencia témájában futó, külső finanszírozású projekt vezető kutatója: (1) „EyCon (Vizuális MI és a korai háborús témájú fotográfia)”, (2) „AEOLIAN: Mesterséges Intelligencia a Kulturális Szervezetek részére”, (3) „AURA (Az Egyesült Királyság és Írország levéltárai és az MI): összehozni a digitális bölcsészetet az informatikusokkal és a többi szakmával a kulturális javak elérhetővé tétele érdekében”. E nemzetközi projektek célja az, hogy a kutatók digitizált és born-digital archívumokhoz könnyebben hozzáférhessenek és hogy olyan innovatív kutatási módszereket vezessenek be az archív adatok elemzéséhez, mint az AI.

*Fordította: Törke Gábor*