

**A TUDOMÁNY, AZ OKTATÁS ÉS A
KÖZGYŰJTEMÉNYI KISZOLGÁLÁS ÚJ
INFORMATIKAI SZINERGIÁI**

**NETWORKSHOP 2026
35. Országos Informatikai Konferencia**

**2026. március 31–április 2.
Debreceni Egyetem, Debrecen**

Szerkesztette: Tick József, Kokas Károly, Holl András

**HUNGARNET Egyesület
Budapest, 2026**



HUN-REN
Magyar Kutatási Hálózat

NETWORKSHOP

Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Korrektúra: Danyi Melinda

Angol nyelvi lektor: Lukács Katalin

Networkshop 2026 konferencia előadásainak közleményei

Debreceni Egyetem, Debrecen

2026. március 31–április 2.

ISBN 978-615-6792-29-7

DOI: <https://doi.org/10.31915/NWS.2026>

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével

Budapest

2026

Borítókép: [freepik.com](https://www.freepik.com)

AZ ARK (ARCHIVAL RESOURCE KEY) AZONOSÍTÓ BEVEZETÉSE ÉS KÖZZÉTÉTELE A MAGYAR NEMZETI LEVÉLTÁR "ADATBÁZISOK ONLINE" FELÜLETÉN

Debreczeni László, Magyar Nemzeti Levéltár Digitális Szolgáltatásfejlesztési Osztály
debreczeni.laszlo@mnl.gov.hu

Simon András, Magyar Nemzeti Levéltár Digitális Szolgáltatásfejlesztési Osztály
simon.andras@mnl.gov.hu

Köszönetnyilvánítás

A szerzők szeretnék köszönetüket kifejezni a „Digitális Irodalmi Akadémia köteteinek és hierarchikus egységeinek ARK azonosítóval való ellátása” munkaanyag szerzőinek a dokumentumban közölt értékes elemzésért és információkért.¹

Absztrakt

A Magyar Nemzeti Levéltár nyilvános webes felületein egyre több digitalizált levéltári irat elérhetőségét biztosítja, emellett egyre nagyobb mennyiségben tesz megtekinthetővé földrajzi- és személynévtér rekordokat is. Ezek közvetlen, felhasználóbarát módon történő hivatkozhatóságát biztosítani kell. A hivatkozás módjának a világban az egyik elterjedt módja az ARK (Archival Resource Key) azonosító használata, melyet ma már számos levéltár, mindenekelőtt a Francia Nemzeti Levéltár (ANF – Archives Nationales de France) is használ. Az azonosító a nemzeti levéltárban való bevezetésével kapcsolatban számos kérdést kellett eldönteni (saját adatbázisból származó vagy generált rekordazonosítók használatba vétele, „okos” vagy „buta” azonosítók alkalmazása, a csatolt digitális tartalmakkal nem rendelkező metaadatrekordok hivatkozásának a felvállalása, hivatkozás megoldása a teljes dokumentumok mellett azok részeire is, feloldás saját vagy külső eszközzel, ARK azonosító módosíthatósága és törölhetősége). Az előadás az intézmény által meghozott döntések ismertetése mellett bepillantást enged a döntési folyamatba és bemutatja a megszületett konkrét megoldásokat is.

Kulcsszavak: Permalink, ARK azonosító, levéltári nyilvántartás, webes szolgáltatások

¹ Készítők: Horváth Ádám, Mohay Anikó, Radics Péter, Bánki Zsolt, Szabó Mihály, Csáki Zoltán, Virág Gabriella

Abstract

The Hungarian National Archives provides a great amount of digitized archival documents on its public web interfaces and also makes geographical and personal namespace records available in increasing quantities. Their direct, user-friendly references must be ensured. One of the possibilities is to use the ARK (Archival Resource Key) identifier used by many archives, especially the French National Archives (ANF – Archives Nationales de France). Several issues had to be decided in connection with the introduction of the identifier in the national archives (the use of record identifiers from its own database or generated ones, the use of opaque or non-opaque identifiers, the acceptance of referencing archival catalogue records without attached digital content, the resolution of references to parts of documents in addition to entire documents, resolution with its own or external tool, the modifiability and deletion of the ARK identifier). In addition to describing the decisions made by the institution, the paper provides an insight into the decision-making process and presents the specific solutions that were created.

Keywords: Permalink, ARK identifier, archival catalogues, web services

Permalink a levéltári nyilvános felületeken

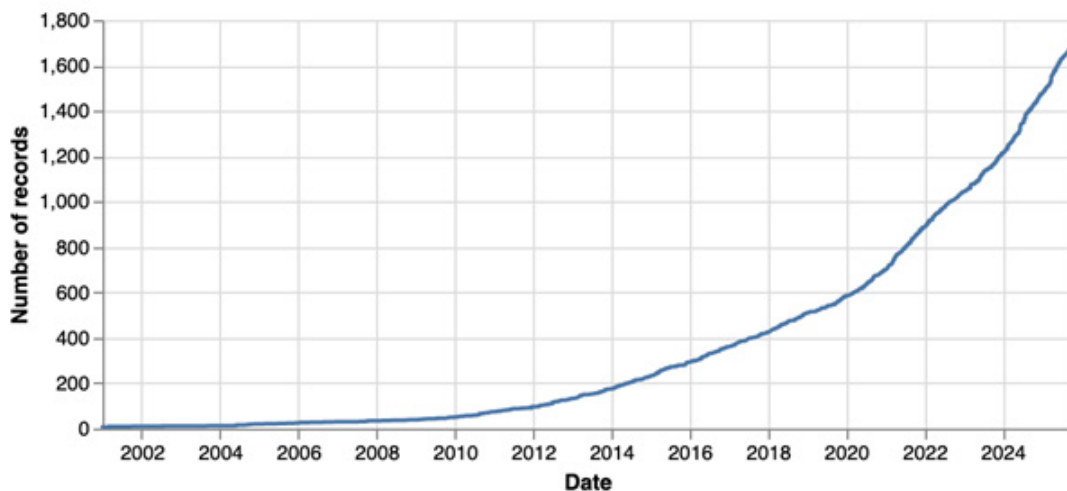
Egy URL link átlagos élettartama nem nagy. Nem kell sok időnek eltelnie, hogy egy korábban elmentett linkre kattintva a "404 Not Found" üzenet fogadjon bennünket. A perzisztens linkek és az ezekről a gyűjteményi tartalomra mutató, a link feloldását biztosító rezolverek alkalmazásával közelebb kerülhetünk ahhoz, hogy felhasználóink tartósan és közvetlenül elérjék és hivatkozassák az általunk nyilvánosságra hozott dokumentumokat, akkor is, ha az intézményünk gyűjteményi struktúrája vagy nyilvános felülete lényegesen megváltozott.²

Ezek alkalmazása a levéltárakban kiemelt fontosságú, mert gyűjteményeik nagy többségben nem publikált dokumentumokat tartalmaznak, ezért azok nem rendelkeznek DOI-val sem. A levéltári iratokat nehezebb is azonosítani, mint egy könyvet, folyóiratcikket vagy akár a képeket. Sok esetben címmel és szerzővel sem rendelkeznek. Másrészt, vélelmezhetően tartósan hozzáférhetők lesznek az iratórző adatbázisában, és mint primer források sohasem avulnak el. A dokumentum hozzáférhetőségének biztosítását a Magyar Nemzeti Levéltár – mint roppant stabil intézmény – biztosítani tudja.

² <https://arks.org/about/ark-features/>

Az ARK azonosító

A lehetséges permalinkmegoldások közül a Magyar Nemzeti Levéltár Digitális Szolgáltatásfejlesztési Osztályán figyelmünk az ARK azonosítóra (Archival Resource Key) irányult. Az ARK ugyanis egyre szélesebb körben elterjedt, mindenekelőtt a Francia Nemzeti Levéltár (ANF – Archives Nationales de France) is használja.



1. ábra: NAAN azonosítóval ellátott intézmények száma³

Alkalmazása nem jelent többletköltséget a levéltár számára. Egyszerű, logikus szerkezet, emellett hatékony hierarchia és verziókezelés jellemzi, és belesimul az intézményi DNS-rendszerbe.

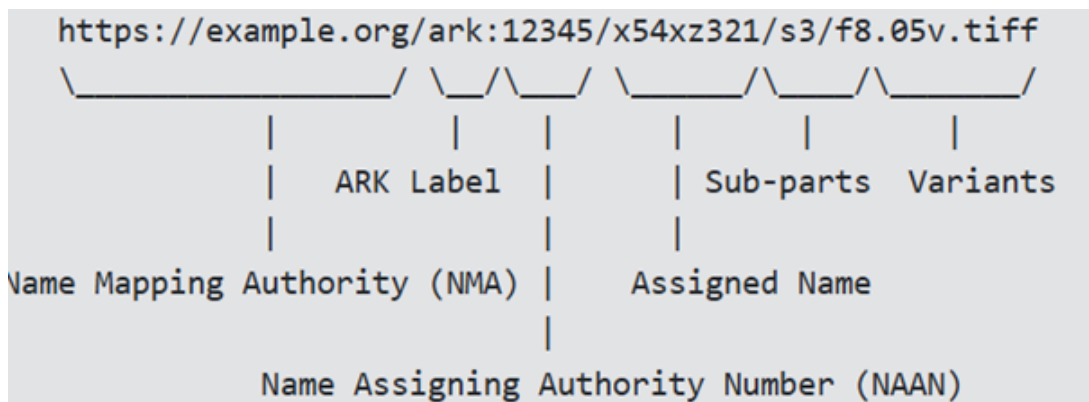
Az azonosító bevezetéséhez első lépésként Name Assigning Authority Number (NAAN) azonosítót kellett igényelnünk az <https://arks.org> weboldalon.

A bevezetéskor döntést kellett hozni néhány alapvető kérdésben:

- Mely domainnévhez akarjuk hozzárendelni a NAAN azonosítót?
- Mely gyűjteményi egységeket és mely adatbázisrekordokat kívánjuk ezzel az azonosítóval ellátni?
- Milyen legyen az általunk generált ARK azonosító szerkezete és hogyan legyenek ezek előállítva?
- Hogyan biztosítjuk az állandóságot, és mi legyen a sorsa a mégis megszűnt azonosítóknak?
- Támasztunk-e elvárásokat a hivatkozott objektumok metaadataira vonatkozóan?

3 <https://arks.org/community/>

A döntések meghozatalához az ARK szerkezetével kellett mindenekelőtt megismerkednünk.



2. ábra: az ARK azonosító szerkezete⁴

Name Mapping Authority: Az intézmény neve az URL-címben.

ARK Label: Ez mindig „ark:”

Name Assigning Authority Number (NAAN) azonosító

Assigned Name: Az objektumhoz rendelt egyedi azonosító.

Sub-parts: a dokumentum alábontásának követése az azonosítóban.

Variants: lehetséges változatok megjelenítése.

Az azonosító képzése a nemzeti levéltárban

Az azonosító bevezetése során első körben az AOL-felületen (Adatbázisok Online) szolgáltatott dokumentumokkal és azon belül két alapkategóriával foglalkozunk:

- Metaadatrekord, amely online felületen elérhető digitális objektumra mutat.
- Metaadatrekord, amely a nyilvános felületen el nem érhető digitális objektumra mutat.
- Metaadatrekord, amely analóg dokumentumra mutat.
- Metaadatrekord, amely névtételelem (személy, testület, földrajzi) azonosítójára mutat.

Az ARK azonosítók előállítása során az adatbázisban már meglévő azonosítók (belső rekordazonosítók, raktári jelzetek, leltári számok) is felhasználásra kerülhetnének, de ez az azonosító által tartalmazott stringben a felhasználható jelek kötött száma miatt nem túl praktikus.

4 <https://arks.org/>

Számjegyek: 0123456789

Betűk, de kizárólag a mássalhangzók (kis és nagybetűk különbsége számít), leszámítva az „l” mássalhangzót.

Az alábbi központosítási jelek: = ~ * + @ _ \$. /

A magánhangzók azért vannak mellőzve, hogy elkerülhető legyen, hogy az azonosító értelmes szavakat adjon ki, az l betű az i-hez való hasonlósága miatt marad ki.

Mi a felhasználásra tervezett karakterekből elhagytuk a központosítási jeleket és a nagybetűket, így az ARK azonosítóknban az alábbi karakterek fordulnak majd elő:

0123456789

bcdfghjkmnpqrstvwxyz

A dokumentumok egyedi azonosítója 7 karaktert foglal majd el, a kiadható, összesen 29 féle karakter így több mint 14 milliárd (huszonkilenc a hetediken) lehetőséget jelent.

A későbbiekben tervezzük a – karaktert, melyet az ARK-ban vizuális tagolásra használnak.

Osztályunkon megszületett az a döntés is, hogy – rendelkezvén a szükséges infrastruktúrával – helyi feloldóprogramot (local resolver) fogunk alkalmazni, a webszerver rewriting rules használatával, vagy más, saját fejlesztésű feloldószoftverrel. A táblázatban minden dokumentum egy sor, az ARK azonosítóhoz hozzákeresi az UUID-t, bármilyen változás van a rendszerben, a táblázatot kell módosítani.

Emellett az azonosítók tartalmazni fognak „shoulder-t”, mely az objektumhoz rendelt egyedi azonosító első tagja lehet. Ez a string a hivatkozott rekord jellegére, illetve a szolgáltató tagintézményre nézve fog tartalmazni utalást. Ez az egyediség biztosítását is segíteni fogja. A megbeszéltek szerint tehát a shoulder 4 vagy 5 karakterből áll, és a nagyobb egységeket, tagintézményeket, és azon belül adatbázisokat és adatféléseket jelöli majd. Az azonosító a shoulderrel együtt egy Oracle adatbázisba lesz generálva, mindig az egyes objektumokhoz egy-egy levéltári adatbázis publikálásakor. A generáláskor lesz ellenőrizve, hogy az azonosító nem lett-e már korábban esetleg kiosztva az adatbázison belül.⁵

Az ARK azonosító „Sub-parts” részét képezhetik a „qualifikerek”. Ezek legkézenfekvőbb szerepe egy adott dokumentumon belül az azok részeire (pl. egy több oldalas irat egyes oldalaira) való utalás lehet.

5 A <https://mnl.gov.hu> domain-hez kapott azonosítónk: 52333 lett. Az ehhez tartozó alárendelt domaineik shoulderekkel kerülnek majd azonosításra.

Pl: ark:/12345/x5wf6789/c2/s4.pdf

12345 – NAAN

x5 - shoulder

wf6789 - azonosító

c2/s4.pdf - qualifier⁶

A kvalifikerek alkalmazását – azaz a hivatkozás megoldását a teljes dokumentumok mellett azok részeire is – akkorra halasztottuk, amikor a dokumentumok hivatkozása már széles körben megoldódott, és az egyéb meghozott döntéseket a gyakorlat igazolta.

Az azonosító kialakításakor választani kellett az „okos” non-opaque ARK és a „buta” opaque ARK között. Ez előbbi tartalmazza az azonosítót és egyéb információkat a dokumentum jellegéről, értelmes a felhasználó számára, információt ad számára anélkül is, hogy rákattintana. Mi az utóbbit (opaque) választottuk. Ezt egyszerűbb generálni, rövidebb, a lehetőségek száma nagyobb, az illetéktelen adatgyűjtő programok számára nehezebben kezelhető, és ha a dokumentum bármilyen tekintetben megváltozik, nem kell hozzányúlni. Az egyes azonosítók egyébként – éppen az illetéktelen, automatizált adatgyűjtés elleni védekezésül – nem a stringek ABC-, illetve számsorrendjében lesznek generálva.

A bevezetés mérföldkövei

1. Levéltári irattekintések hivatkozása (akár van már hozzá digitalizált teljes tartalom, akár nincs)
2. Levéltári dokumentumok egyes fejezeteinek, részleteinek azonosítása
3. Névtértekintések (földrajzi, személy és testületi) hivatkozása

Az azonosítók generálása megkezdődött, előállításuk folyamatban van.

Teendők egy ARK azonosító esetleges megszűnése esetére

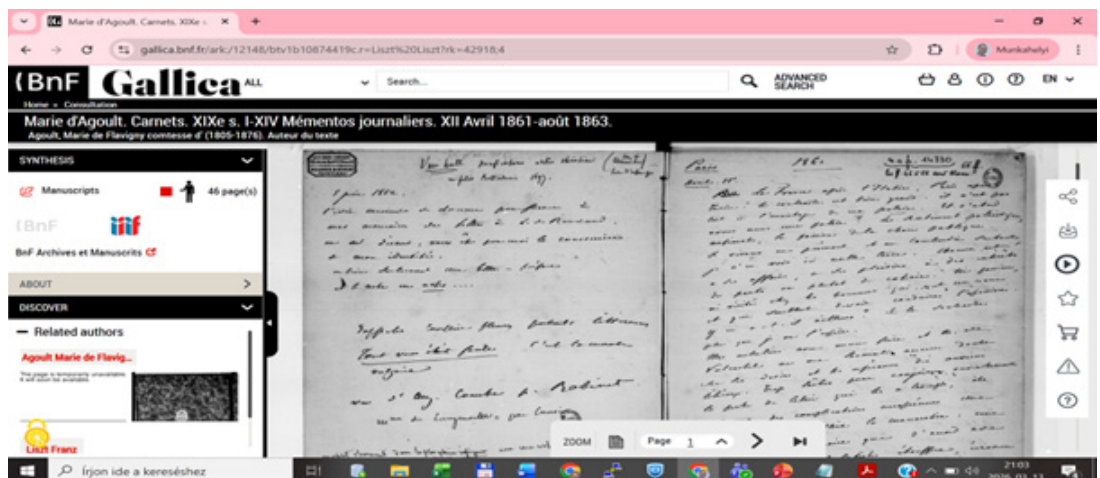
Módosításra, törlésre fel kell készülni. Lehet törölni egy ARK azonosítót, és persze – főként a resolver által használt táblában lehet módosítani a dokumentum elérési útját. Mindenképpen törekedni fogunk rá, hogy a kiadott ARK azonosítók mögött ne legyenek törölt linkek. Ha megszűnik egy dokumentum vagy egy állomány, és töröljük a rá hivatkozó rekordot, akkor egy standard oldalra visz majd az ARK azonosító. Ha olyan strukturális változás következik be, hogy az ARK azonosító módosul, mert a shoulderben a 2–3 vagy a 4. pozíción valami módosul, akkor erre a linkekre való kattintáskor tájékoztató üzenetet küldünk. Ha az adatbázisban történik változás, akkor az eggyel fentebbi levéltári nyilvántartási szintre irányítjuk az azonosítót, és erről tájékoztató üzenetet küldünk a felhasználónak.

6 <https://arks.org/>

Példák a világban:



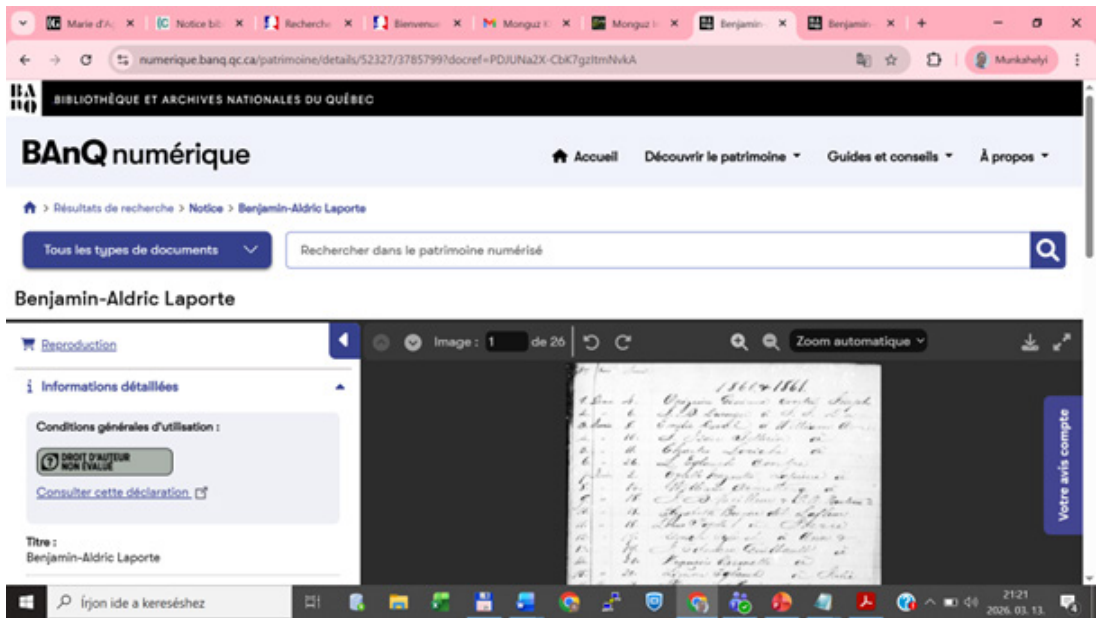
3. ábra: ARK azonosítóval hivatkozott rekord a Francia Nemzeti Levéltár weboldalán⁷



4. ábra: ARK azonosítóval hivatkozott rekord a Gallica weboldalán⁸

7 <https://catalogue.bnf.fr/ark:/12148/cb43509416d>

8 <https://gallica.bnf.fr/ark:/12148/btv1b10874419c.r=Liszt%20Liszt?rk=42918;4>



5. ábra: ARK azonosítóval hivatkozott rekord a Québec állam Nemzeti Levéltárának weboldalán⁹

Saját megoldásunk bemutatása

Jelenlegi permalink a saját weboldalunkon teszt verzióban:

<https://test-zala.natarch.hu/aol/ark:52333/0416rbbd4nws>

A permalink feloldott állapotban:



<https://test-zala.natarch.hu/aol/adatbazis/a-masodik-vilaghaboru-vesztesegi-adatbazisa/adatlap/6FA5D5ED58C44E412DD92202AB2B65C4>

⁹ <https://collections.banq.qc.ca/ark:/52327/3785799>

Ádám József

Így hivatkozzon erre a tartalomra: [ark:52333/0416rbbd4nws](https://aol.mnl.gov.hu/ark:52333/0416rbbd4nws)

HU-MNL-PML-XXXIII.1a.-55.-C-4.-1944-54

Jetzet	HU-MNL-PML-XXXIII.1a.-55.-C-4.-1944-54
Forrás	Gomba halotti anyakönyve
Elhalt neve	Ádám József
Foglalkozás	földműves napszámos
Életkor	34
Vallás	református
Elhalálozás oka	kézigránát robbanás
Elhalálozás dátuma	1944-11-30 du. 4 óra
Elhalálozás helye	Gomba 
Lakhely	Gomba 162. 
Anyja neve	Németh Zsófia
Apja neve	Ádám Ferenc

6. ábra: ARK azonosítóval hivatkozott rekord a nemzeti levéltár weboldalán

A tervezett permalink struktúrája éles felületen így néz majd ki:

<i>shoulder</i>	<i>domain</i>
ark:52333/s1	https://aol.mnl.gov.hu/ark:\${tartalom}
ark:52333/s2	https://nevter.mnl.gov.hu/ark:\${tartalom}

Bár számos intézmény igényelt már mostanra NAAN azonosítót, valójában még kevés felületen alkalmazzák. Magyarországon tudomásunk szerint elsőként fogunk tömegesen ARK azonosítókat előállítani, és saját felületünkön szolgáltatni. Reméljük ezzel a módszertani újítással szolgáltatásunk elnyeri a szakma tetszését, és hozzájárulhatunk az interneten való tömeges, hatékony és hosszú távon megbízható tartalomszolgáltatáshoz.