Z. KARVALICS LÁSZLÓ

Csapajev naplója, avagy a szintetikus dokumentumok és az archívumi világ

Hozzászólás „A Szintetikus dokumentumok és a jövő archívumai” vitához

A Vörös Hadsereg frontkiadójának lektora lerúgja a csizmáját, bebújik az ágyba, és felüti a hős Csapajev vaskos naplóját, amit aznap kapott meg kiadáselőkészítésre. Nagy ákombákom betűkkel indul a szöveg: *„Ezerkilencszáztizennyolc október 3-án még alig virradt, én máris felkaptam jó Mirza lovam hátára, hogy elinduljak a több mint húsz versztányira lévő hadosztályparancsnokságra”*. A lektor, ahogy szokta, unott mozdulattal az utolsó oldalra ugrik, a bejegyzések végéhez: *„és szerencsésen meg is érkeztem”*. Döbbenten kezdi lapozgatni a kéziratot, hogy ugyan mi is történhetett közben, és 150 oldalon keresztül ezt látja leírva: *„gyí, gyí, gyí, gyí….”*

A szintetikus dokumentumok kapcsán nem véletlenül jutott az eszembe a sok kiváló Csapajev-vicc közül éppen ez, amit az Oreate AI segítségével vizuálisan is átélhetővé tettem (azzal a kis csavarral, hogy vágtázás közben nem napló, hanem egy szintetikus dokumentum születik, s azzal a megjegyzéssel, hogy lám, a bölcs MI a magyar szerző angol nyelvű kérését magától, kontextushelyesen, orosz nyelven, bár több csúnya helyesírási hibával teljesítette).



Oreate AI által generált kép

Most tételezzük fel, hogy a kép életre kel, megjelenik, valaki felfigyel rá, megosztja, online vándorútra indul, és előbb-utóbb elszakad a szerzőségemtől, kielégítve Székely Iván munkadefinícióját a szintetikus dokumentumokról.[1]

Az első kérdés az, hogy vajon *létezik-e olyan felhasználási-kutatói igény, amely archiválási kényszert* teremthet a képpel mint dokumentummal kapcsolatban?

Az affordanciák hatalma

Próbáljunk tehát azonosítani potenciális érdeklődésközösségeket.

Alapvetően egy Csapajev-vicchez kapcsolódó, ahhoz kiegészítő jelentést másodlagos humorforrásként rendelő „tartalomformáról" van szó. Vele kapcsolatban érintettek lehetnek például a *Csapajev-viccek gyűjtői* (vannak ilyenek, különösen Oroszországban számosan, de ismerek magyar lelkitársakat is), akik egy (talán) több mint száz éve gyarapodó korpuszt megfrissítő friss adaléként tekinthetnek rá.



Csapajev egy korabeli képen (Majláth 2025)[2]

Csakhogy a Csapajev-motívum kulturálisan sokkal mélyebb kisvilágokhoz vezet. Itt most nincs idő kitérni arra, hogyan teremtette meg az Ural folyóba fulladt vörösparancsnok kultuszát egykori komisszárja, Dmitrij Furmanov egy róla írt re-

[1] A mesterséges intelligencia (MI) által, autonóm módon, közvetlen humán kezdeményezés vagy közreműködés nélkül született adatok, iratok, képek, videók, tudományos vagy irodalmi művek, hírek stb. (összefoglalóan: dokumentumok), amelyeknek a valósághoz, *humán entitásokhoz való viszonya külső szemlélő számára nem ellenőrizhető, nem visszafejthető* (Idézet forrása EBBEN A LAPSZÁMBAN). Ha ez a definíció el is fogadható, az nem tűnik szerencsésnek, hogy a szintetikus dokumentumokról homogén kategóriaként beszélünk. Ez olyan, mint amikor azt mondjuk, hogy *tévézés*, és közben a sportközvetítéstől kezdve a szappanoperán át a természetfilmekig iszonyú sok lehetséges tartalom van, és mindegyikre más attribútumok lesznek érvényesek. Vagy amikor azt mondjuk, hogy a gyerek *kütyüzik*: a tanulástól a játékon keresztül ez is nagyon sok tevékenységet takarhat. Ha a szintetikus dokumentumoknak vannak fő osztályai, és ezeknek különböző funkciói, akkor nevezzük meg ezeket a területeket, és egyesével keressünk választ az archiválási igény kérdésére.

[2] Látható, hogy az AI képének Csapajev-arca erősen hasonlít az eredetire. A későbbiek szempontjából fontos adalék, hogy a „hivatalos" portréfotó mint dokumentum nem archívumból, hanem egy cikk illusztrációjaként került elő (Majláth 2025). Ez nem zárja ki, hogy a fotó eredetije valamikor egy muzeális gyűjteményből kelt vándorútra.

gényben. A lényeg: a nem különösebben eszes, faragatlan, de hithű harcos gyorsan folklorizálódott, 1934-ben film készült róla, amit egy év alatt 30 millióan néztek meg. Népdalok ihletője lett,[3] s mivel szinte fikciós hőssé vált, alkalmas volt arra, hogy rajta nevetve vezessék le a feszültséget akár a legsötétebb sztálini időkben is.

Ám itt még nincs vége. Az egész Csapajev-mitologéma új életre kelt Viktor Pelevin besorolhatatlan műfajú *Az agyag géppuska* című regényében[4], és ez is mind semmi a Jelcin-érában megjelent eszement kalandjátékhoz képest (Red Comrades Save the Galaxy),[5] amelynek nyolc folytatása jelent meg, de túl későn törte át a nyelvi korlátokat ahhoz, hogy a világ más pontjain is popkulturális ikonná váljon.[6]

Ám nemcsak a kép „csapajevsége", hanem „mémsége" is lehetne kutatói érdeklődés tárgya. A közösségi média, a hétköznapi ikonológia, a tömegkultúra búvárai előszeretettel nyúlnak a mémekhez, akár illusztratív, akár meghatározott jelenségek elemzésének szándékával.

S természetesen a mém megszületését lehetővé tévő technológiának magának is akadhatnak búvárai, akik az MI-képgenerálás használatával, elterjedésével, a felhasználói szokásokkal foglalkozva igénylik a vizsgálódásra kiválasztott tartalomtípus egyfajta sokaságát.

Végül – csakis logikai lehetőségként – az alkotó személye is érdekessé válhat, ha nagy kisugárzó hatású gondolkodóként (luminary) nemcsak a tudományos vagy irodalmi életműve, hanem annak rekvizitumai is biográfiai jellegű megközelítés tárgyává válnak.

Ötféle lehetőséget mérlegeltem, ötféle érdeklődésközösséget írtam körül. Akár ebben a tárgykörben is találhatnánk még további lehetőségeket, de azt szerettem volna bemutatni, hogy minden tartalomforma (és ezek dokumentumreprezentációja) egy sokelemű *affordanciatér*ben helyezkedik el. A potenciális kapcsolódások más elemekhez (és az ezeken keresztül meghatározható érdeklődésközösségek világa) csak részlegesen tárható fel a tartalmak születésekor.

Egy pillanatra képzeljük el, még a preszintetikus dokumentumok világában, hogy a 19. század elején vagyunk, és egy szobát föl kell szabadítani egy skandináv levéltárban, amelyik tele van hajónaplókkal, amiket a 17–18. század balti- meg északi-

[3] Hogy mást ne mondjunk, az óbudai Mókus (majd később: Miklós) utcai zenetagozatos Általános Iskola hetedikes diákjaként, 1974 őszén a legendás Fodor tanár úr énekóráján mi még harsogtuk a Csapajev-indulót („száz dal zengi vörös hadak dicsőségét / Csapaj hősi népét).

[4] Cserna-Szabó András egyszerre nevezi ultraskizofrén őrületnek és fontos műnek, s nehéz szó nélkül elmenni amellett, hogy melyik gondolatot illeszti elénk afféle központi üzenetként: *Csak egyetlen szabadság van: amikor az ember megszabadul attól, amit az értelem épít. Ezt a szabadságot hívják úgy, hogy „nem tudom"*. https://litera.hu/irodalom/netnaplo/meg-ez-a-nincs-sincs.html Ez a maxima ugyanis, bármilyen furcsán hangzik is, nem más, mint amire épp most szeretnék a mérnökök a nagy nyelvi modelleket megtanítani.

[5] https://ipon.hu/magazin/cikk/red-comrades-2-for-the-great-justice-reloaded A cselekményről elég annyit mondani, hogy Csapajev és tisztiszolgája sikerrel akadályozzák meg a Föld lerohanására készülő idegeneket ördögi tervük végrehajtásában, de Csapajev nem tud megszabadulni a fejébe helyezett chiptől és félig ember, félig gép terminátorként kénytelen élni.

[6] Ám egy fontos, témánk szempontjából is kiemelkedő jelentőségű kérdéshez jól elvezet. Abban feltehetően nincs vita, hogy a játék maga, minden folytatásával együtt, *digitális játék-archívumokba* kívánkozik. De vajon *az egyedi lejátszások* közül érdemes archiválni bármit? Vajon mi különbözteti meg lényegileg a kalandjáték *egy* lejátszását a sok közül az LLM-nek feltett *egy* kérdésre adott *egy* választól?

meg más tengeri hajóútjaikról a kapitányok följegyeztek. Döntést kell hoznunk, hogy szelektáljuk, kidobjuk vagy megőrzésre érdemesnek gondoljuk az unalmas hajónaplókat? Azt, hogy a hajónaplók elsődleges klímatörténeti források lehetnek, csak 200 évvel később tudtuk meg, amikor számunkra problémává vált, hogy kellenek megbízható, idősoros, dátumhoz lehorgonyozható adatok. Másképp: ez az affordancia csak jóval később nyilvánult meg.[7]

A nagy kérdés az tehát, hogy létezik-e archívumi módszer, technológia, megközelítésmód, filozófia, amellyel a megőrzési igény jó előre kimondható, akár még nem ismert affordanciákra való tekintettel is? Úgy vélem, ehhez találhatunk megfelelő kiindulópontokat.

A kollektív episztémikus többlet keresése

A hajónaplók esetében azt a kérdést kellett volna egykor feltenni, hogy a dokumentum mint kimenet szolgálhat-e megismerő aktusban bemenetként. Egyszerűbben: létezik-e olyan információs többlet, tudástöbblet, amely kizárólag innen érhető el, egyedinek mondható, túlmutat létrehozója individuális cselekvési terén és bármilyen csekély mértékben is, de jelenthet-e nyersanyagot kumulatív természetű kultúránk bármely feldolgozó rendszere számára? Másképp: hiánya jelentene-e erőforrásvesztést az affordanciakapcsolt episztémikus közösségek számára.

S mivel az adott időben, adott útvonalon az adott hajók egyedüliként közlekedtek, rögzített megfigyeléseik máshogyan el nem érhető adalékok, amelyek mindenképp megőrzésre érdemesek.

De vajon mi mondható el a különböző mesterségesintelligencia-platformok kimeneteivel kapcsolatban?

Amit a grafikák, karikatúrák, fotorealisztikus képek, nagy festők stílusában készült alkotások, a Studio Ghibli vagy a Disney formavilágát hűen előállító képgeneráló alkalmazások „kiköpnek”, nem mutatnak túl saját magukon, nem hoznak létre megőrzendő értéket. Hiszen az a nyersanyag, amiből építkeznek, az hordozza a kollektív episztémikus többletet. A generált képek meg nem őrzése semmilyen erőforrásveszteséget nem okoz – és ugyanez igaz természetesen a zenekészítő-szövegíró programokra, meg a dolgozatokat, publikációs kritériumoknak megfelelő szövegeket előállító MI-eszközökre.[8] Egy egyedi lekérdezés mögött van egyedi episztémikus érték, mert az LLM hozzásegíti a felhasználót egy új ismerethez. De ez az ismeret jelen van, visszakereshető módon, az LLM bázisául szolgáló szövegtérben. Bármely lekérdezés eredménye eldobható – annál is inkább, mert bármikor újragenerálható.

Ráadásul, ha az archívumok lelkiismeret-furdalás nélkül negligálják a szintetikus dokumentumokat, akkor is kétszeresen lehetnek nyugodtak. Egyrészt maga a

[7] Az affordanciaprobléma tipikus helyzete a megfigyelőkameráé. Ha nincs „esemény”, akkor meghatározott időszakonként a (szintetikus) állomány törlődik. Ha azonban történik valami (lopás, tűz stb.), akkor az időablakon belül még kinyerhető a fontos információ. A törlés egyetlen kockázata, hogy valami az időablakon túl válik jelentéstelivé az egykori felvételből. De ez pontosan olyan helyzet, mint saját humán percepciós kapacitásunk limitáltsága.

[8] Úgy is mondhatnánk: nagyjából olyanok, mint Csapajev naplójában a sok „gyí”.

Web működik úgy, mint az archívumok archívuma,[9] tehát ha senki nem teljesít archívumi küldetést, a dokumentumok meghatározott hányada akkor is (kereshetően és sok affordanciakapcsolattal) a digitális örökkévalóság része marad. Másrészt az érdeklődésközösségek saját tartalomhasználata nemcsak szaporítja a példányok számát, hanem reflektálja, szövegekkel, másodlagos tartalmakkal vonatkozik rájuk. Adatolja, kiegészíti, kontextusba helyezi őket. Ha a mesterséges intelligencia által generált Žižek–Herzog beszélgetést egy egyetemi professzor tananyagként használja az óráján, és oktatási modulként sikeres, akkor hosszú lábnyomot növeszt, más is felhasználja, írnak róla, hivatkoznak rá. Vagyis: nem egy független memóriaintézmény felelőssége az archiválás, hanem elsősorban a dokumentum létrehozójáé.

Jogi dokumentum született MI-segítséggel? Repozitáld te magad, jogi szereplő, majd döntsd el egy idő után, hogy hogyan használod az állományodat. Az általam ismert „legszintetikusabb" dokumentumtípus a szenzorok millióinak adatait gyűjtő és továbbító cunamidetektáló riasztórendszer – hiszen itt az adattermelésben végképp nincs azonosítható emberi szereplő. De az így képződő hatalmas adattárak kezelése, gondozása szintén nem archívumi feladat: ha a geológusok, oceanográfusok a riasztást nem igénylő jelsorokat más episztémikus többlet elérésére (például a hullámmintázatok monitoringjára) akarják használni, megtalálják a módját, miképp férhessenek hozzá a valahogyan tárolt adatokhoz.

Minden egy irányba mutat: a digitálisan képződött és a szintetikus dokumentumok kapcsán az archívumi világnak nem azt kell keresnie, hogy miképp „nőhet fel" hagyományos intézményként egy új feladathoz, hanem azt, hogy miképp változott meg a szerepeknek, a környezeteknek és a platformoknak az a világa, amelyben az archiválást mint tevékenységet elgondoljuk.

Az új archívumi paradigma – meg a régi

Arkánum. Muemlekem.hu. Fortepan. Egy tartalomdigitalizáló magáncég, egy közösségi épületállapot-monitoring gyakorlat, egy fényképmentő- és digitalizáló, misszióvezérelt mini-magánmanufaktúra. Folytathatnánk a sort, végtelenségig. Magyary Zoltán archívuma az ELTE-n. Hankiss Eleméré Kőszegen. A Wang folyó blogon összegyűlt elképesztő vizuális kultúrtörténeti kollekció. Falvak, kistelepülések, városok saját örökséggyűjteményei.

Közös jellemzőjük, hogy nem hagyományos archívumok. Nem központosított, bürokratizált, állami fenntartású kulturális intézmények[10] – és mégis, küldetésük és munkájuk egy decentralizáltabb és közösségibb archívumi világ felé visz. Sőt, ez az új archívumi tér minden korábbinál személyesebb is. Részben azáltal, hogy nemcsak a már említett luminaryk, hanem bármely individuum érdemesülhet arra, hogy életének strukturálható digitalizált dokumentumai afféle vertikális portálként

[9] "Web as an archive of archives" (Sentilles 2005)

[10] A "régi" hihetetlen erővel áll ellent az újnak. A múzeumok például fogcsikorgatva őrzik a 25-30-40 évvel ezelőtt becsontosodott küldetésnyilatkozataikat. A múzeumi világszervezet 15 éve képtelen megszülni egy picit korszerűbb mission statementeket azzal kapcsolatban, hogy hogy van jelen a múzeum életében a felhasználó meg az állampolgár.

archívumi jelleggel őrződjenek meg. Vajon 30-40-50 digitális év alatt ki mit visz magával folyamatosan, mit gyarapít, mit menteget át egyik laptopról a másikra? Milyen tanulmányokat töltött le, ami mögött esetleg olvasás is van, milyen képeket néz. Milyen tartalmakat állított elő? Ez egyszerre mikroarchivisztika, milliárdnyi kurátorral,[11] másrészt crowdsource, tömeges erőforrásszerzés, kitágítva az archívumi teret, átrendezve a felelősség, munka, erőforrás, szolgáltatás, felhasználás stb. hagyományos logikáját.

A mikroarchivisztikai rezsim egyúttal nagy erejű szelekciós mechanizmus is. Az marad meg a privát rendszerekben, ami jelentésteli, ami információs értéket hordoz. Elvileg el lehetne jutni egy olyan bizarr végkövetkeztetésig, hogy itt az ideje megoldást keresni: ha nincs technológiai akadálya annak, hogy elvileg minden valaha volt felhasználónak az utolsó laptop állományának az eredeti tartalomrészét (levelezés, saját szövegek, saját dokumentumok) meg lehetne menteni, milyen konstrukcióban és hová lehetne feltölteni, megőrizni – természetesen úgy, hogy a „dokumentumgazdák" erről még életükben rendelkezzenek. S mindez egyfajta „episztémikus garancia" is arra, hogy ha valamilyen affordancia intézményi vakfoltként nem került gyűjtési tárgylemezre, egy ilyen állományban bizonyosan ott lehetnek a keresett dokumentumok.

S miközben a szintetikus dokumentumok helyét keressük, egyetlen pillanatra se felejtsük el, mennyi adósságunk van még a predigitális dokumentumkezelésben, amely viszont igényli a szervezett, strukturált, hagyományos archívumi világ figyelmét és energiáit.

A teljesség igénye nélkül:

– az iratként őrzött állomány teljes mértékű feltárása és digitalizálása
– az intézményi világon kívüli családi- és magánarchívumok, gyűjtemények, naplók, fotókollekciók, a kallódó, másodlagos forgalomban keringő dokumentumállomány megmentése, feltárása, digitalizálása – a kulturális mikroörökséggel (Z. Karvalics 2019) kapcsolatos teendők ellátása, az adattartalmak elvesztésének megakadályozása.[12]

Arra a kérdésre tehát, hogy mit csináljanak az archívumok, az a válaszom, hogy majd akkor kezdjenek nagyban szintetikus dokumentumokkal kapcsolatos problémákat generálni maguknak, hogyha az összes többi ezt megelőző problémájukat megoldották. Miközben a „mesterfoknak" a jövőjéről beszélünk iszonyú sokat, az alapfoknak a meg nem oldott jelenével nem foglalkozunk eleget.[13]

[11] Számos példa családi és személyi mikroarchívumokra itt: https://ranke2.uni.lu/a/webarchives/05-Micro-archiving-family/ és itt: https://www.ehri-project.eu/call-ehri-looking-micro-archives/

[12] Természetesen sok legjobb gyakorlatot ismerünk, ahol memóriaintézmények (a levéltárak mellett könyvtárak és múzeumok) is „beszállnak" a mikroörökség-őrzésbe, segítenek, metaadatolnak, kiállításokat rendeznek és gyűjteményezik azokat, digitalizálnak stb.

[13] Egy kicsit olyan ez egyébként, mondjuk, az AI kontextusában, mint hogy iszonyú sok szó esik az önvezető járművekkel kapcsolatos etikai problémákról, fejlesztésekről, s közben a legbanálisabb szenzorokkal ellátott okosutak segítségével rugalmassá tett áteresztés-optimalizálás nincs megoldva a lámpás kereszteződésekben. Léteznek, de nem vezetjük be a sok-sok megoldást, amely a vezetésbiztonságot támogatja.

Hivatkozások

Majláth Ronald. „A hős, akit saját hazájában kinevettek.” *Moszkvatér*, 2025. szeptember 08. Utolsó hozzáférés: 2026. március 26. https://moszkvater.com/a-hos-akit-sajat-hazajaban-kinevettek/

Oreate AI. Utolsó hozzáférés: 2026. március 26. https://www.oreateai.com

Sentilles, Renée M. „Toiling in the Archives of Cyberspace” In Burton, Antoinette (ed.), *Archive Stories: Facts, Fictions, and the Writing of History.* Duke University Press, 2005: 136–156.

Z. Karvalics László. „Élet a cipősdobozon túl, avagy ki őrzi a kulturális mikroörökséget?” *Liget*, 2019/5: 5–62. Utolsó hozzáférés: 2026. március 26. https://ligetmuhely.com/liget/elet-a-ciposdobozon-tul-avagy-ki-orzi-a-kulturalis-mikroorokseget/

Z. Karvalics László történész, információs társadalom kutató, az MTA doktora, Felsőbbfokú Tanulmányok Intézete, Kőszeg. Kiemelt kutatási témái: digitális kultúra, mesterséges intelligencia, állampolgári tudomány, kulturális mikroörökség.

Így hivatkozzon erre a cikkre:
https://dx.doi.org/10.22503/inftars.XXVI.2026.1.10