

Digitális bölcsészet a könyvtár- és információtudományban 4. rész

Történeti kutatásokhoz kapcsolódó európai adatbázisok*

A digitális bölcsészet, valamint a könyvtár- és információtudomány szoros kapcsolatban álló diszciplínák. Tanulmány sorozatunk negyedik részében a European Association for Digital Humanities (EADH) által közzétett multidiszciplináris projektekből válogatva a történettudományhoz kötődő adatbázisokat (1914-1918-online, International Encyclopedia of the First World War, ALCIDE – Analysis of Language and Content In a Digital Environment, CENDARI – Collaborative European Digital Archive Infrastructure, COURAGE – Cultural Opposition: Understanding the CultuRal HeritAGE of Dissent in the Former Socialist Countries, eHeritage – Expanding the Research and Innovation Capacity in Cultural Heritage Virtual Reality Applications, Europeana Newspapers, Forced Labor 1939–1945. Memory and History, PELAGIOS – Pelagios: Enable Linked Ancient Geodata In Open Systems, Fasti Online) elemezzük. Vizsgálatunk eredményeinek közzétételével elsősorban a magyarországi könyvtári tájékoztatást és a könyvtárosképzést szeretnénk támogatni, valamint a hazai tartalomszolgáltatások színvonalának emelését ösztönözni.

Tárgyszavak: bölcsészettudomány; digitális technika; információmenedzsment; innováció; adatbázis; tartalomszolgáltatás; nemzetközi együttműködés; régészet; történelem

Digitális bölcsészet és történettudomány

Közleményszériánk jelen részében a *European Association for Digital Humanities (EADH)* projektjei közül¹ a történeti témájúakkal foglalkozunk, melyekből kilenc kezdeményezést veszünk górcső alá:

1. *1914-1918-online. International Encyclopedia of the First World War,*²
2. *ALCIDE – Analysis of Language and Content In a Digital Environment,*³
3. *CENDARI – Collaborative European Digital Archive Infrastructure,*⁴
4. *COURAGE – Cultural Opposition: Understanding the CultuRal HeritAGE of Dissent in the Former Socialist Countries,*⁵
5. *eHeritage – Expanding the Research and Innovation Capacity in Cultural Heritage Virtual Reality Applications,*⁶
6. *Europeana Newspapers,*⁷
7. *Forced Labor 1939-1945. Memory and History,*⁸
8. *PELAGIOS – Pelagios: Enable Linked Ancient Geodata In Open Systems,*⁹
9. *Fasti Online.*¹⁰

A klasszika-filológiai adatbázisok után az európai digitális bölcsészeti jó gyakorlatok komparatív bemutatására irányuló választásunk azért esett a *történelemtudományra*, mert szélesebb érdeklődésre számot tartó diszciplínáról van szó, továbbá a könyvtárak jelentős része rendelkezik különböző mélységű történeti gyűjtőkörrel, így a *szakértői tájékoztatásban* is alkalmazhatók munkánk eredményei.

Elemzésünk módszertana megegyezik a korábban alkalmazottal, azaz szólunk a portálok mindegyikére, egy (kisebb) csoportjára, vagy csak egyikére érvényes tulajdonságokról is. Ezután összevetjük a gyűjtemények profiljait, továbbá azt, hogy mely szervezet(ek) támogatják, illetve melyek a partnerintézményeik (ha vannak). Végül pedig – könyvtár- és információtudományi megközelítésünkben eredetileg eltérő terjedelmi keretek közt tárgyalva – kiemeljük a projektek egyedi vonásait, keresési, böngészési, rendezési, megjelenítési stb. lehetőségeit.

* A tanulmány az Emberi Erőforrások Minisztériuma (EMMI) támogatásával készült.

A kutatási projektek átfogó összehasonlítása

A *nyílt hozzáférésű*, az adatok megosztását és felhasználását célzó projektek közös jellemzője, hogy *kollaboráción alapuló* informatikai eszközökkel segítik a kutatói együttműködések elmélyítését, *új összefüggések* feltárását. Az 1914-1918-online portál például az első világháborúval foglalkozó szakértők közös munkája révén jött létre, és máig is várnak tanulmányokat, amelyeket online publikálnak a témakörben. A PELAGIOS fogadja további gyűjtemények csatlakozását, hogy minél láthatóbbá tegyék egymás eredményeit, valamint összekapcsolják azokat más, hasonló tematikájú adatbázisokkal. Szintén a PELAGIOS-nál a Recogito online felületen a saját magunk által összeállított szövegekből és képekből lehetőség nyílik egyedi munkaterület létrehozására, ahová meghívhatunk partnereket is a közös kidolgozásra. De a rendszerváltás előtti kulturális ellenállást bemutató COURAGE is arra buzdítja a hasonló kollektívával rendelkezőket, hogy csatlakozzanak a hálózatukhoz.

A portálok felhasználnak olyan haladó technológiákat, mint a *Named Entity Recognition (NER)*, másnéven *Named Entity Extraction*,¹¹ amely gépi eszközökkel azonosítja a személyneveket, helyszíneket, egyéb elnevezéseket, majd hozzákapcsolja azokat az alapértelmezett vagy beállított szótárhoz, enciklopédiához. A konnektió létrejötte után a linkre kattintva bővebb információt tudhatunk meg az adott elemről. Ilyet használ a PELAGIOS Recogito felülete, az ALCIDE, a CENDARI és a European Newspapers. Utóbbinál egy kimondottan ezt támogató nyílt forráskódú programot is fejlesztettek, és szabadon elérhetővé tették a *GitHub* platformján.¹² A *GitHub-on*¹³ – mint programokat, számítógépes kódokat megosztó és publikáló, verziókövetést biztosító tárhelyen – történő közzétételt más tartalomszolgáltatók (pl. PELAGIOS) is alkalmazzák.

1. táblázat

Az EADH portálján regisztrált történeti tematikájú európai digitális bölcsészeti projektek áttekintése

A portál neve és profilja	Támogató(k) és partnerintézmény(ek)
<p>1914-1918-online. International Encyclopedia of the First World War</p> <ul style="list-style-type: none"> • angol nyelvű virtuális referenz szolgáltatás az első világháborúról • 50 országra kiterjedő, nyilvános kutatásokat hálózatba szervező projekt • innovatív navigációs és keresési lehetőségeket nyújt • több mint 1600 közleményt tartalmaz 	<p>Támogatók:</p> <ul style="list-style-type: none"> • German Research Foundation (DFG) • Freie Universität Berlin <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • German Historical Institute • Deutsches Historisches Institut Paris (DHIP IHA) • CENDARI • Europeana • University of Birmingham

A kapcsolatok megteremtéséhez gyakran igénybe vesznek *Linked Open Data* megoldásokat: a Fasti Online a Pleiades helységnévtárat alkalmazza a régészeti helyszínek beazonosítására (Pleiades URI), de ilyen módszerrel dolgozik a PELAGIOS, az 1914-1948-online, a European Newspapers, a CENDARI és az ALCIDE. A többi a saját rendszerén, adatbázisán belüli hivatkozásokat használ.

Néhány projektnél fontos szerepet tölt be az *annotálás*, amelynek során a gép által automatikusan nem felismert adatelemekről tudjuk megmondani azok típusát (például személynév vagy helység-név), készíthetünk róluk részletes leírást és egyéb, kutatáshoz fontos megjegyzéseket fűzhetünk az adott szövegekhez. A CENDARI által használt *Access to Memories (AtoM)*¹⁴ annotáló alkalmazás igen hasonló a PELAGIOS Recogito felülete által kínált lehetőségekhez. Főbb különbség, hogy az AtoM különállóan is letölthető és felhasználható, míg a Recogito az implementált online felületen vehető igénybe.

A *vizualizáció* is fontos funkcióként jelentkezik. A CENDARI például összesíti és megjeleníti az adatokat rögzítésük után, valamint az ALCIDE is tartalmaz adatprezentálási eszközöket.

Több projektnél kulcsszerepet játszanak a földrajzi szempontok szemléltetésére szolgáló *interaktív térképek*, így a PELAGIOS-nál, a Fasti Online-nál, az 1914-1918-online-nál és a Forced Labor 1939-1945 adatbázisnál is.

A programokat szervező legaktívabb szereplőként egyértelműen a *Freie Universität Berlin*¹⁵ kell kiemelnünk, támogatóként pedig az *Európai Unió* különböző pályázati forrásait (pl. Horizon 2020).

A *1. táblázatban* az EADH oldalán megjelenő bemutatkozó projektadatok alapján szemléltetjük az egyes adatbázisok főbb jellemzőjét és intézményi hátterét.

<p>ALCIDE</p> <ul style="list-style-type: none"> • történelmi források és irodalmi művek webalapú, nagytömegű elemzésére szolgáló platform egyelőre három mintakorpusszal (Kennedy, Nixon és Marinetti beszédei) • haladó szövegfeldolgozási technológiákat és intuitív vizualizációt használ fel a korpuszok szoros (close) és távoli (distant) olvasásához • dokumentumgyűjtemények közötti tallózást és interaktív elemzést kínál • lexikai, szemantikai, geográfiai és időbeli dimenziók szerinti vizsgálati lehetőség 	<p>Támogatók:</p> <ul style="list-style-type: none"> • Fondazione Cariplo • Fondazione Caritro <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Digital Humanities Research Group at Fondazione Bruno Kessler • SIG (Italian – German Historical Institute) at Fondazione Bruno Kessler
<p>CENDARI</p> <ul style="list-style-type: none"> • intézményi és nemzeti korlátok nélküli virtuális kutatási infrastruktúrát biztosít a történelmi forrásokhoz • átfogó célja a történettudományi kutatások európai szintű katalizálása a digitális technológiákat kihasználva • az első világháborúra és a középkorra vonatkozó nemzetközi történelemtudományi forrásokban kereshetünk: több mint ezer intézet anyagában és több mint egymillió adatcsomagban 	<p>Támogató:</p> <ul style="list-style-type: none"> • European Commission under the 7th Framework Programme <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Freie Universität Berlin • University of Birmingham • Trinity College Dublin • Czech National Library • Università di Cassino • The European Library • Consortium of European Research Libraries (CERL) • King's College, London • Institut national de recherche en informatique et en automatique (INRIA) • University of Stuttgart • Goettingen State and University Library • Mathematical Institute of the Serbian Academy of Sciences and Arts • Fondazione Ezio Franceschini • Società Internazionale per lo Studio del Medioevo Latino
<p>COURAGE</p> <ul style="list-style-type: none"> • a kulturális ellenállás emlékeit gyűjti az egykori szocialista államokból • célja a gyűjtemények egyes országok társadalmában betöltött funkciójának bemutatása, a kollekciók és az eredmények mind nemzeti, mind nemzetközi szinten történő prezentálásának elemzése • az első ilyen jellegű összefoglaló adatbázis online és offline forrásokból, köz- és magángyűjteményekből felépítve • időhatára a kommunista uralom kezdetétől a Vasfüggöny leomlásáig terjed • magyar projektvezetés • jelenleg 419 gyűjteményt tartalmaz, 816 személyről találunk leírást és 136 történelmi eseményről olvashatunk • honlapjuk 16 nyelven érhető el 	<p>Támogató:</p> <ul style="list-style-type: none"> • European Union's Horizon 2020 <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Research Centre for the Humanities, Hungarian Academy of Sciences (MTA BTK) (coordinator) • Institute of Philosophy and Sociology, Polish Academy of Sciences (IFIS PAN) • Trinity College Dublin (TCD) • Institute for East and Southeast European Studies, University of Regensburg (IOS Regensburg) • Centre for Social Sciences, Hungarian Academy of Sciences (MTA TK) • Lithuanian Institute of History (LII) • Charles University in Prague (CUNI) • University of Bucharest (UB) • Croatian Institute of History Comenius University in Bratislava (HIP) • The University of Oxford (UOXF) • Institute for Computer Science and Control, Hungarian Academy of Sciences (MTA SZTAKI)

<p>eHeritage</p> <ul style="list-style-type: none"> • a Brassói Egyetem koordinálásával a projektben résztvevők virtuális valóságra fókuszáló kutatásait segíti • kulturális örökséggel kapcsolatos programok megvalósítása, tudástranszfer a történelem és a művészetek területén 	<p>Támogató:</p> <ul style="list-style-type: none"> • European Union's Horizon 2020 <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • University of Transilvania of Braşov • Jožef Stefan Institute (JSI), Ljubljana • Scuola Superiore Sant'Anna, Pisa
<p>Europeana Newspapers</p> <ul style="list-style-type: none"> • 23 európai könyvtár 12 millió oldalnyi archív folyóiratához, valamint több mint 20 millió újságoldalhoz kapcsolódó metaadathoz biztosít hozzáférést • a dokumentumok a Europeana platformján keresztül vagy a The European Library felületén lévő dedikált újságnézegető alkalmazással kutathatók • megtekinthető 5 millió nyilvános (public domain) hírlap • Open Source eszközök publikálása a GitHub-on 	<p>Támogató:</p> <ul style="list-style-type: none"> • European Commission Competitiveness and Innovation Programme <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Staatsbibliothek zu Berlin • Koninklijke Bibliotheek • National Library of Estonia • Österreichische Nationalbibliothek • National Library of Finland • Staats- und Universitätsbibliothek Hamburg • Bibliothèque nationale de France • National Library of Poland • University of Salford • CCS Content Conversion Specialists GmbH • Stichting LIBER • National Library of Latvia • National Library of Turkey • University Library of Belgrade • University of Innsbruck • Landesbibliothek Dr. Friedrich Tessen • The British Library • Europeana Foundation
<p>Forced Labor 1939-1945</p> <ul style="list-style-type: none"> • az adatbázis annak a több mint 20 millió embernek kíván emléket állítani, akik kényszermunkát végeztek a náci Németországban • 590 élettörténet-interjú tartalmaz a koncentrációs táborok foglyaival • audio- és videointerjúk, valamint kísérő- és oktatási anyagok a gyűjteményről, illetve a kényszermunkáról • németországi iskolák számára 2016-ban összeállított tananyaga: Learning with interviews 	<p>Támogató:</p> <ul style="list-style-type: none"> • Foundation "Remembrance, Responsibility and Future" <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Foundation "Remembrance, Responsibility and Future" • Freie Universität Berlin • German Historical Museum (DHM)
<p>PELAGIOS</p> <ul style="list-style-type: none"> • célja a Linked Open Data bevezetésének és elterjedésének elősegítése a helyszínekre vonatkozó online források esetében • megtaláljuk benne az ókori római helyneveket és a moderneket is, utóbbit az egész világra vonatkozóan • szabadon lehet csatlakozni a vonatkozó forrásadatbázisokkal • a GitHub-on publikál a Recogito keretében például szemantikus annotációs programokat és geográfiai beépülő alkalmazásokat 	<p>Támogatók:</p> <ul style="list-style-type: none"> • Jisc • The Andrew W. Mellon Foundation <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • University of Southampton • The Open University • Austrian Institute of Technology

<p>Fasti Online</p> <ul style="list-style-type: none"> • 2000-től épített, archeológiai ásatásokra vonatkozó adatbázis • több mint 3700 helyszínt tartalmaz 14 országból • egy ásatási terület többször történő feltárását (kronologikusan) is bemutatja • a Fasti Online Documents & Research, Archaeological Conservation (FOLD&R Italia) című kapcsolódó folyóiratcsaládban az ásatásokról készült illusztrált riportok olvashatók 	<p>Támogatók:</p> <ul style="list-style-type: none"> • International Association of Classical Archaeology (AIAC) • Center for the Study of Ancient Italy of the University of Texas at Austin (CSAI) <p>Partnerintézmények:</p> <ul style="list-style-type: none"> • Packard Humanities Institute • Center for the Study of Ancient Italy at the University of Texas at Austin • Baron Lorne Thyssen • Italian Ministry for Cultural Heritage and Activities • Society for the Promotion of Roman Studies • Classical Association • Oxford Journal of Archaeology • Advanced Research Infrastructure for Archaeological Dataset Networking in Europe (ARIADNE) • Ukrainian Studies Fund of Harvard University
--	--

A tartalomszolgáltatási projektek sajátosságai

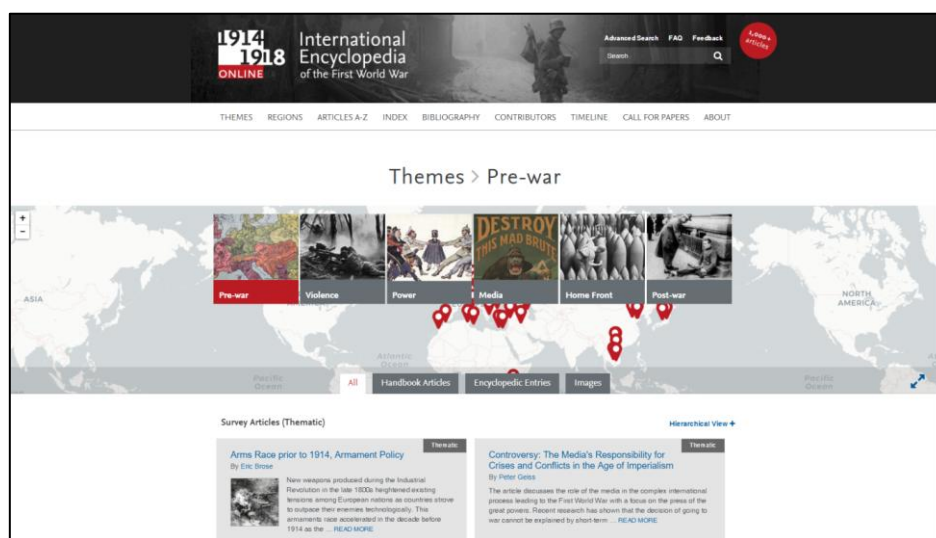
1914-1918-online. International Encyclopedia of the First World War¹⁶

Az első világháború történetét komplexen feldolgozó, több mint 1000 szerzővel dolgozó, kezdetben egydimenziós publikációs adatbázis 2016-tól az interaktív munkakörnyezetet biztosító *Open Encyclopedia System (OES)*¹⁷ használatára tért át, melynek kivitelezését a *Center for Digital Systems*,¹⁸ a *Center for Modern Greece*¹⁹ és a *Bavarian State Library*²⁰ végezte. Az első két szervezet a Freie Universität Berlinhez tartozik, a harmadik pedig a Bajor Állami Könyvtár, így egyértelműen jelentős német közreműködéssel valósult meg a fejlesztés, de a projekt intézményi és sze-

mélyi működtetése is Németországban összpontosul.

A közlemények tudományos színvonalát garantáló *szerkesztőbizottság* 22 ország 90 neves történészből áll, akik arra törekednek, hogy ne csak az európai színterek, hanem a kolóniákkal kapcsolatos kérdések is terítékre kerüljenek.

A szolgáltatás *Semantic MediaWiki (SMW)*²¹ technológiára épülő nem lineáris elérésű, a *kapcsolatok vizualizációja időre, témára és régióra* orientáltan valósul meg. A feldolgozás kézikönyvfejezetekre, enciklopédiatételekre, valamint képekre fókuszál (1. ábra). Az új és aktualizált tartalmakra folyamatosan felhívják a figyelmet.



1. ábra 1914-1918-online. International Encyclopedia of the First World War. Téma szerinti böngészés

A nyitólap felső menüsorából elérhető főbb *böngészési opciók*:

- témák (pl. háború előtt, média, hátszág);
- régiók (pl. Közép-Európa, Afrika);
- tanulmányok (cím, szerző, régió és téma szerinti rendezésben);
- mutatók (név, helyszín és tárgy);
- bibliográfia;
- közreműködők (szerzők, szerkesztőbizottság, külső bírálók);
- idővonal.

A *haladó keresés* számos szűkítési lehetőséggel bír, itt jelennek meg olyan paraméterek, mint a személyek, intézmények, események vagy az egyes tárgykörök (pl. kultúra, gazdaság, társadalom) hierarchikus bontásban.

Ahogy a 2. ábrán látható, a böngészés vagy keresés során nyert tanulmányokat *nyomtathatjuk és letölthetjük* különböző formátumokban (PDF, EPUB, Kindle). A közleményeket kiegészítik a kapcsolódó térképek, képek, az összefüggésbe hozható cikkekre mutató URL-ek és egyéb források (ide integrálták a CENDARI adatbázis vonatkozó tételeinek hivatkozásait is²²). A szövegben a neveket, országneveket, fogalmakat linkekkel látták el, amelyekkel bővebben informálódhatunk a háttéradatokról. Minden cikket letölthetünk vagy megjeleníthetünk RDF/XML formátumban is. A tanulmányok visszakereshetőségét DOI-k (Digital Object Identifier)²³ garantálják. A közleményeknél feltüntetik az utolsó frissítés dátumát, így nyomon követhetők a módosítások. A *metaadatok* jól körülhatárolható egységet képeznek (3. ábra).




The screenshot shows the website '1914-1918 ONLINE'. The main article is titled 'Western Front' by Jonathan Krause. It features a 'Table of Contents' with sections like 'Introduction', 'The War of Movement', '1915: The Early Trench Battles', etc. On the right, there is a 'Map' showing Belgium and Luxembourg with red location pins, and an 'Images' section with a photo of 'American Expeditionary Forces...'. Below the map are links for 'Related Articles', 'External Links', 'Resources', and 'RDF/XML'. The top navigation bar includes 'THEMES', 'REGIONS', 'ARTICLES A-Z', 'INDEX', 'BIBLIOGRAPHY', 'CONTRIBUTORS', 'TIMELINE', 'CALL FOR PAPERS', and 'ABOUT'. A search bar and a '1,000+ articles' badge are also visible.

2. ábra 1914-1918-online. International Encyclopedia of the First World War. Részlet egy tanulmányból

Bergen, Leo Van: Medicine and Medical Service , in: 1914-1918-online. International Encyclopedia of the First World War, ed. by Ute Daniel, Peter Gatrell, Oliver Janz, Heather Jones, Jennifer Keene, Alan Kramer, and Bill Nasson, issued by Freie Universität Berlin, Berlin 2014-10-08. DOI: 10.15463/ie1418.10221.

License

This text is licensed under: [CC by-NC-ND 3.0 Germany - Attribution, Non-commercial, No Derivative Works](https://creativecommons.org/licenses/by-nc-nd/3.0/de/).

Metadata	
Subjects	Science and technology > Medicine Warfare and the military > Trauma and medical care > Medical services, military
Author Keywords	medical service ; healthcare ; wounds ; mental illness military medicine
GND Subject Headings	Weltkrieg [1914-1918] ; Medizinische Versorgung 
LC Subject Headings	World War, 1914-1918--Medical care 
Rameau Subject Headings	Soins médicaux ; Guerre mondiale (1914-1918) 
Key Persons	Andreas Latzko (1876-1943) ; Clovis Vincent (1879-1947) ; Emil Kraepelin (1856-1926) ; Ernst Toller (1883-1939) ; Fritz Wittels (1880-1950) ; George Grosz (1893-1959) ; Henriette Rémi (1885-1978) ; Julius Wagner-Jauregg (1857-1940) ; Kurt Singer (1886-1962) ; Leonhard Frank (1882-1961) ; Lewis Yealland (1884-1954) ; Louis Barthas (1879-1952) ; Oskar Maria Graf (1894-1967) ; Sigmund Freud (1856-1939) ; Theodor Lessing (1872-1936) ; nurses
Regional Section(s)	Central Europe > Austria-Hungary > Austria Central Europe > Germany Western Europe > Belgium Western Europe > France Western Europe > Great Britain and Ireland > Great Britain
Thematic Section(s)	Violence
Title	Medicine and Medical Service
Author(s)	Bergen, Leo Van

3. ábra 1914-1918-online. International Encyclopedia of the First World War. Egy tanulmány metaadatai

Az enciklopédiából praktikus és látványosan, sokszempontúan és tudományos igényességgel tájékozódhatunk az első világháború eseményeiről és háttérinformációiról. A szerkesztők továbbra is várják a témába vágó írásokat, az erre vonatkozó feltételeket a *Call for papers* menüpont alatt lehet megtekinteni.

ALCIDE²⁴

A történeti kutatások támogatására létrejött olasz ALCIDE a *politikai kommunikáció* elemzésére törekszik modern IT-eszközök bevonásával. A trentói *Istituto Storico Italo-Germanico (ISIG)*²⁵ berkeiben 2014-ben elkészült alkalmazás prototípusa *Alcide De Gasperi*²⁶ írásait és beszédeit tartalmazta. Az ALCIDE (Analysis of Language and Content In a Digital Environment) *három szövegtörzsmintát* tárol:

- kettőt angolul – *Nixon* (282 dokumentum) és *Kennedy* (598 tétel) 1960-as választási kampányban elhangzott beszédeit;
- egyet pedig olaszul – *Marinetti* 1909–1921 közötti írásait (49 tétel).

Az eredeti dokumentumokat XML-be konvertálták, majd a *TextPro*²⁷ program *Natural Language Processing (NLP)*²⁸ moduljaival történt a *szövegfeldolgozás* az alábbiak szerint:

- feldarabolás (Tokenisation),
- mondatokra tagolás (Sentence Splitting),
- morfológiai elemzés (Morphological Analysis),
- szófaji egyértelműsítés (Part-of-Speech – PoS),
- lemmatizáció (Lemmatisation),
- szó szerkezetek felismerése (Multiword Recognition),
- kulcsszavak kivonása (Keywords Extraction),
- szintaktikai alapú fráziselemzés (Chunking),
- névelemek felismerése (Named Entity Recognition – NER).

E módszerek lehetővé teszik a szövegek *lexikai, szemantikai, geográfiai* és *időbeli dimenziók* szerinti vizsgálatát, továbbá a korpuszok *szoros (close) olvasása* mellett a *távoli (distant) tanulmányozást* is. Szoros olvasás alatt a hagyományos olvasást értjük, míg a távoli alatt a beállított gépi elemzési metódus eredményeiből következtetések levonását. Példaként említhetjük, hogy a szógyakoróság, illetve a kulcsszavak felismerése automatizált eszközök nélkül – kizárólag szoros olvasással – nehezen (rendkívül időigényesen) kivitelezhető ekkora szövegmennyiségben.²⁹

A MySQL adatbáziskezelő rendszerben tárolt dokumentumokból készíthető *diagramok* megjelenítését a *Highcharts*³⁰ biztosítja, míg az egyedi és interaktívabb *adatvizualizációt* a *Data-Driven Documents (D3.js)*³¹ Az *interaktív térképeket* a *Leaflet library*³² felhasználásával alakították ki. A szövegeken – regisztráció után – szabadon tesztelhetők az elemzési lehetőségek.

A Kennedy-gyűjteményt példaként kiválasztva megjelenik a dokumentumok száma Highcharts eszközön intervallumra korlátozhatóan, majd a *Persons* szempontra kattintva ábrázolódnak a beszédekben leggyakrabban elhangzott személyek nevei (4. ábra). A *Keywords* szerinti kimutatásban pedig a kulcsszavak statisztikája rajzolódik ki (5. ábra). Érdekes funkció még az elhangzott helynevek térképen való ábrázolása (6. ábra).

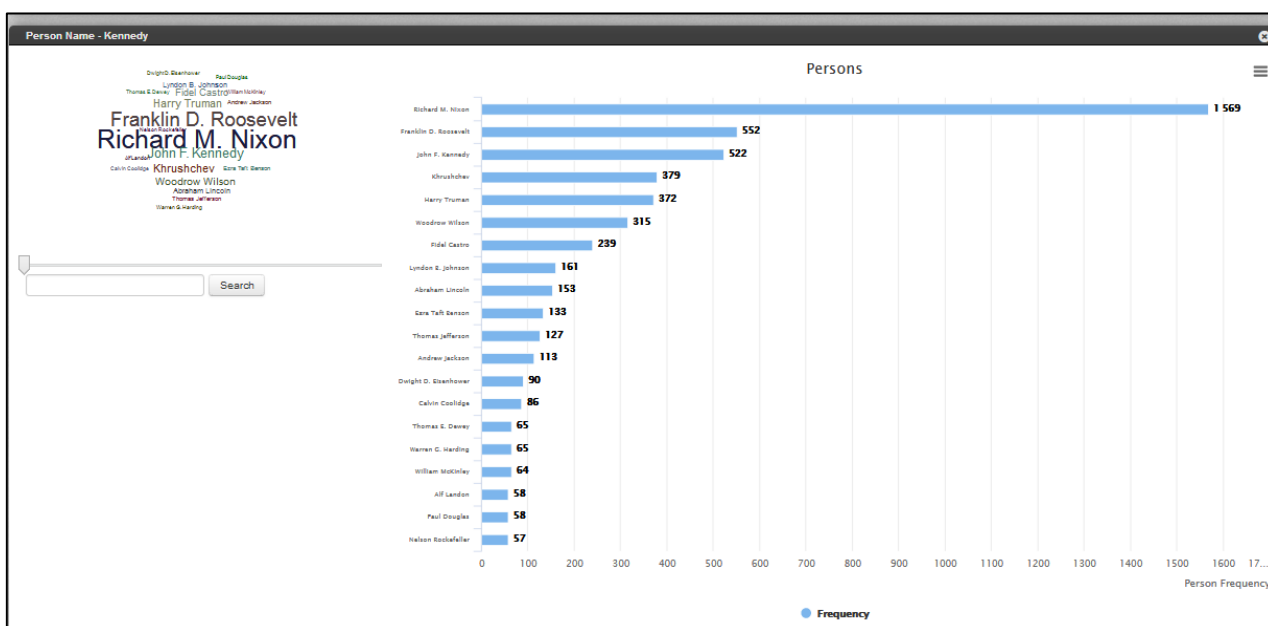
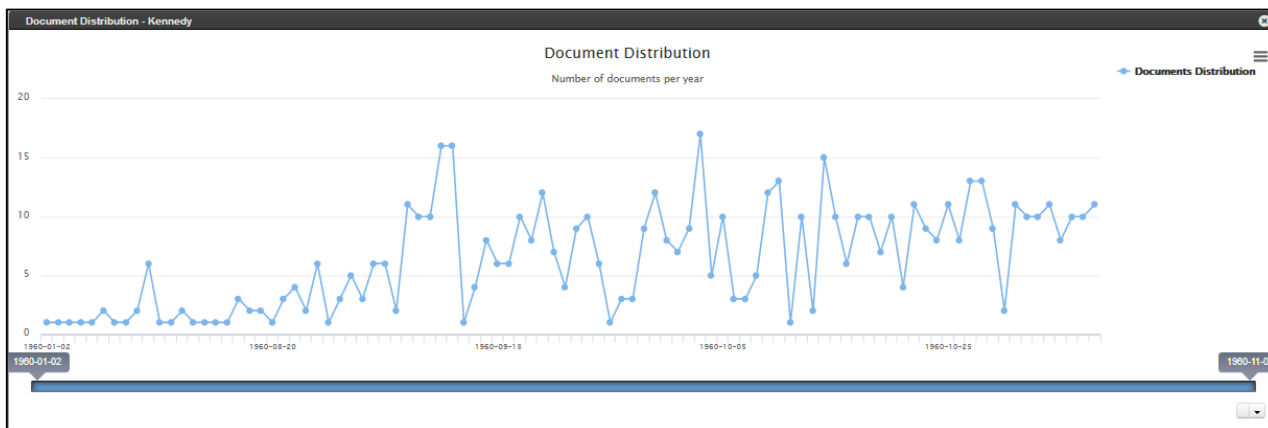
Az ALCIDE kiváló példa arra, hogy megfelelő módon előkészített szövegekből speciális vizualizációs eszközökkel támogatott, a történeti kutatásokat előre vivő *nyelvészeti elemzéseket* is fel lehet építeni.

CENDARI³³

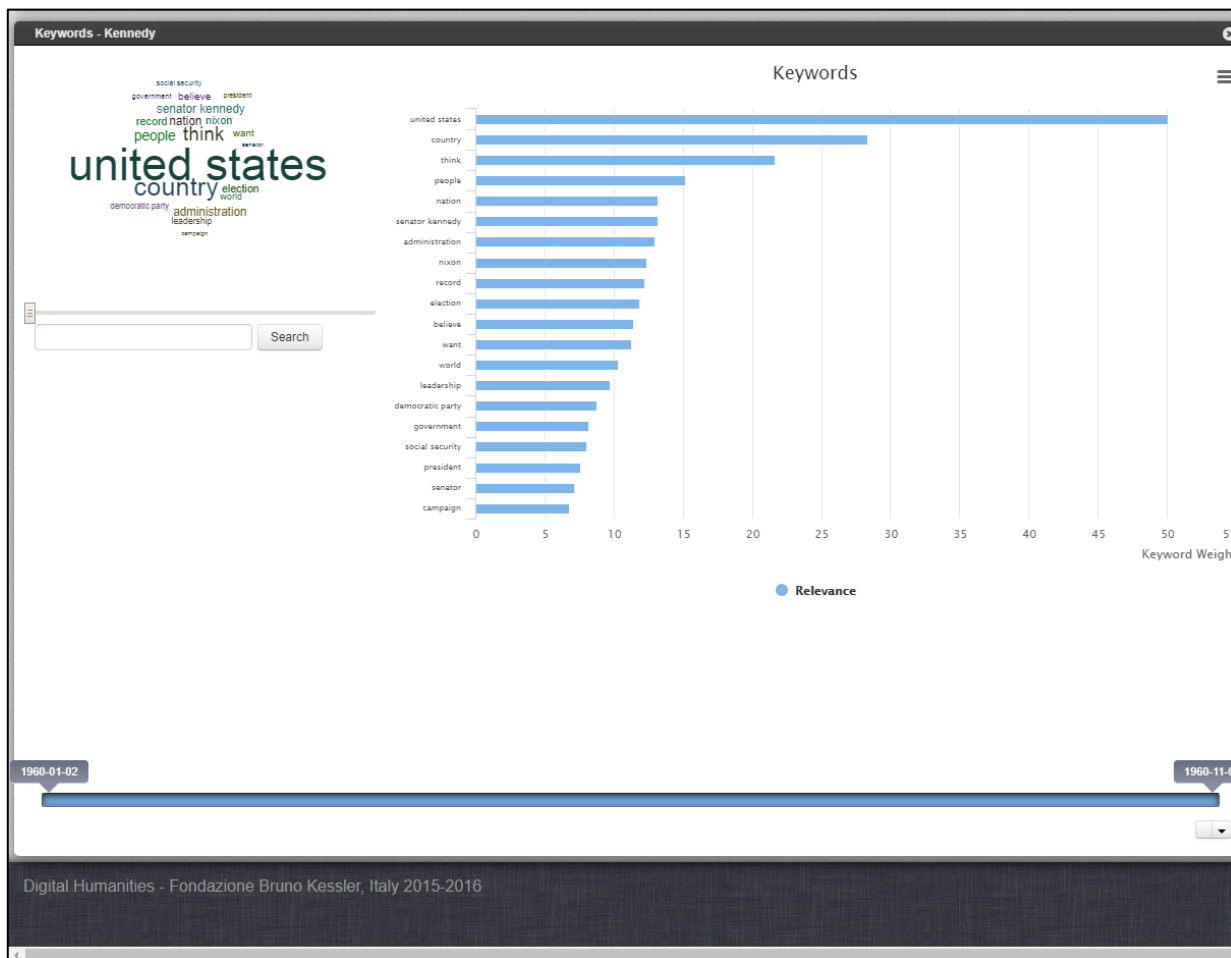
A történeti kutatásokat hálózati megoldásokkal összekapcsoló CENDARI (Collaborative European Digital Archive Infrastructure) projekt három fő pillére:

1. *középkorral és az első világháborúval* foglalkozó, kevésbé ismert archívumokban történő keresés – *Transnational Archival Sources*;
2. jegyzetkészítő alkalmazás – *Working Spaces*;
3. tematikus kutatási útmutatók – *Thematic Research Guidelines* (jelenleg fejlesztés alatt).

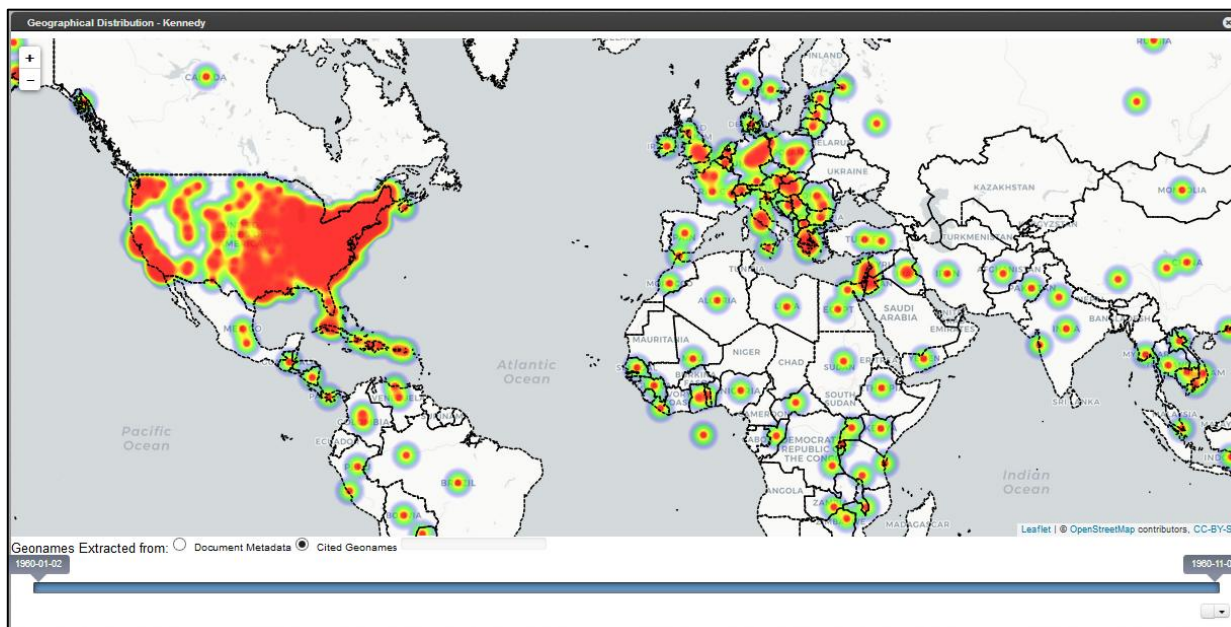
A *Transnational Archival Sources* az *Access to Memories (AtoM)* nyílt forráskódú, archív állományok leírására szolgáló alkalmazásra épül, amely több nyelvet támogat egyszerre és *közös környezetet* biztosít akár több különböző közgyűjteménynek (multi repository). *Nemzetközi szabványokkal* (pl. Dublin Core, Encoded Archival Description – EAD³⁴) kompatibilis, felhasználóbarát eszközöket kínál a gyűjtemények metaadatolására, valamint szerkesztésére. A közel 1000 intézmény 1 millió rekordjában történő *kereséskor* szűkíthetünk nyelvre, archívumra, földrajzi régióra és témára (7. ábra).



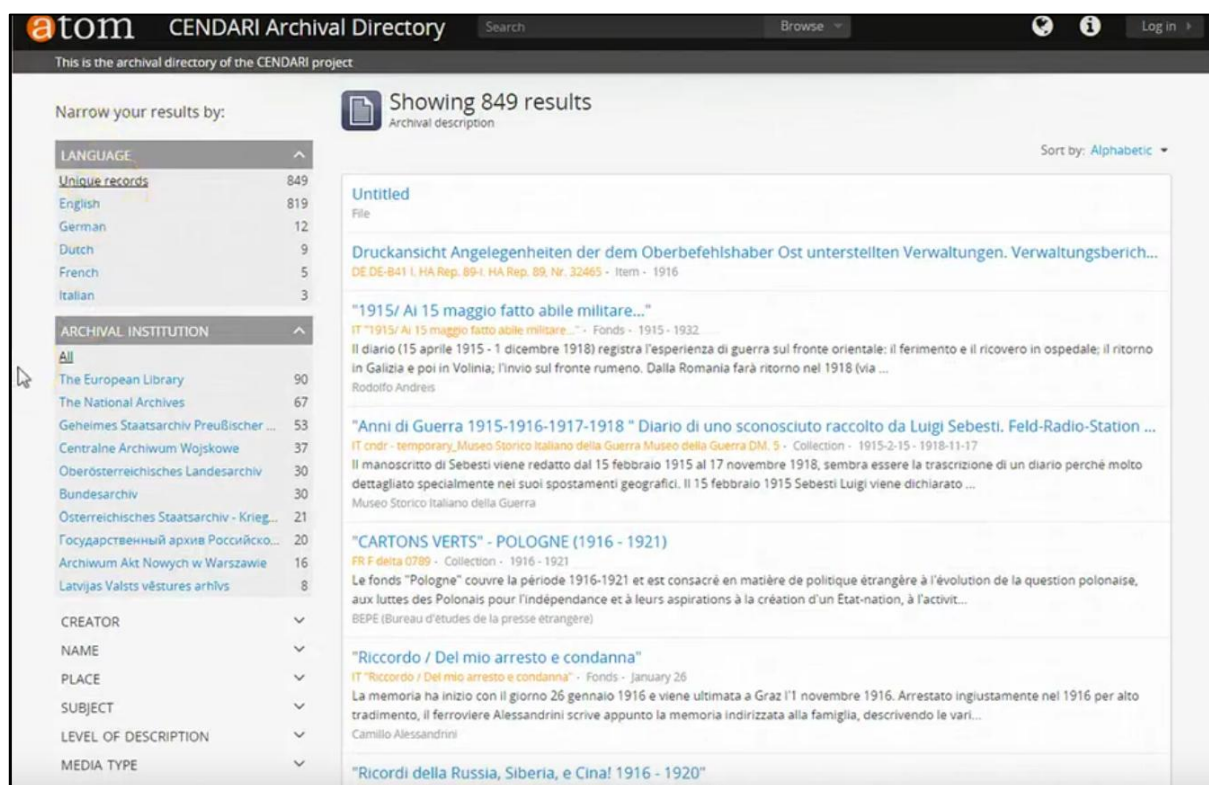
4. ábra ALCIDE. Dokumentumok száma kronologikusan és a beszédekben említett személyek neveinek előfordulási gyakorisága a Kennedy-gyűjteményben



5. ábra ALCIDE. Kulcsszavak gyakorisági listája a Kennedy-gyűjteményben



6. ábra ALCIDE. Helynevek ábrázolása a Kennedy-gyűjteményben

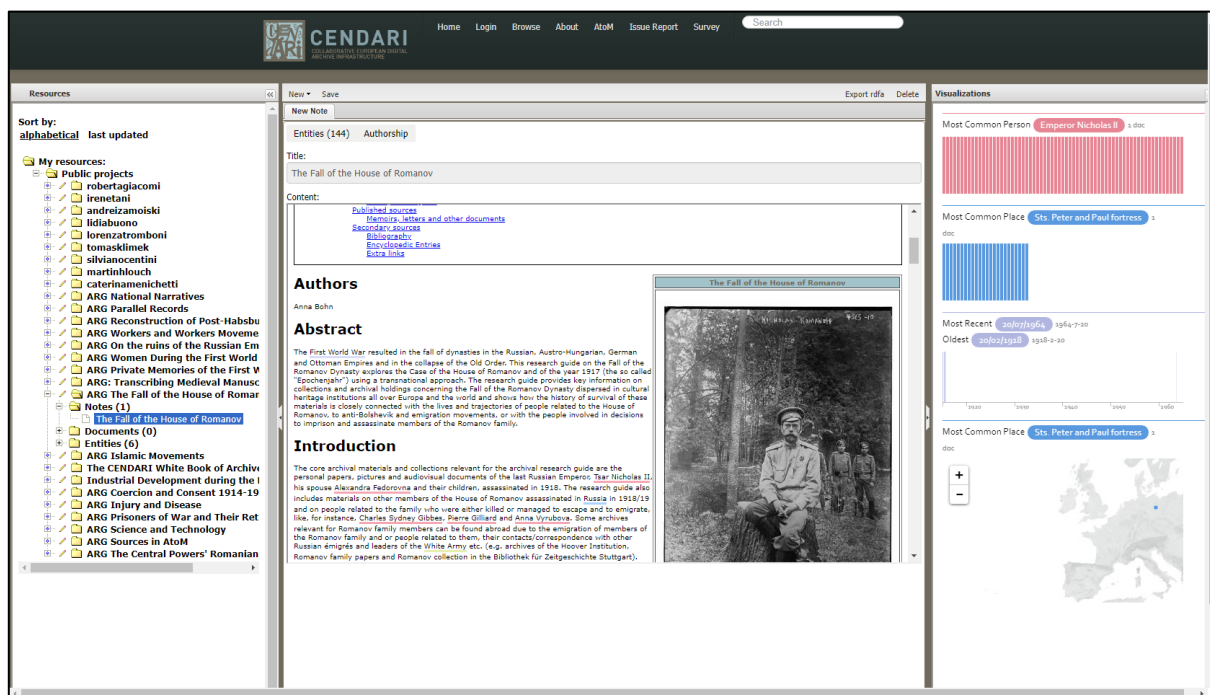


7. ábra Példa a CENDARI találati lista nézetére

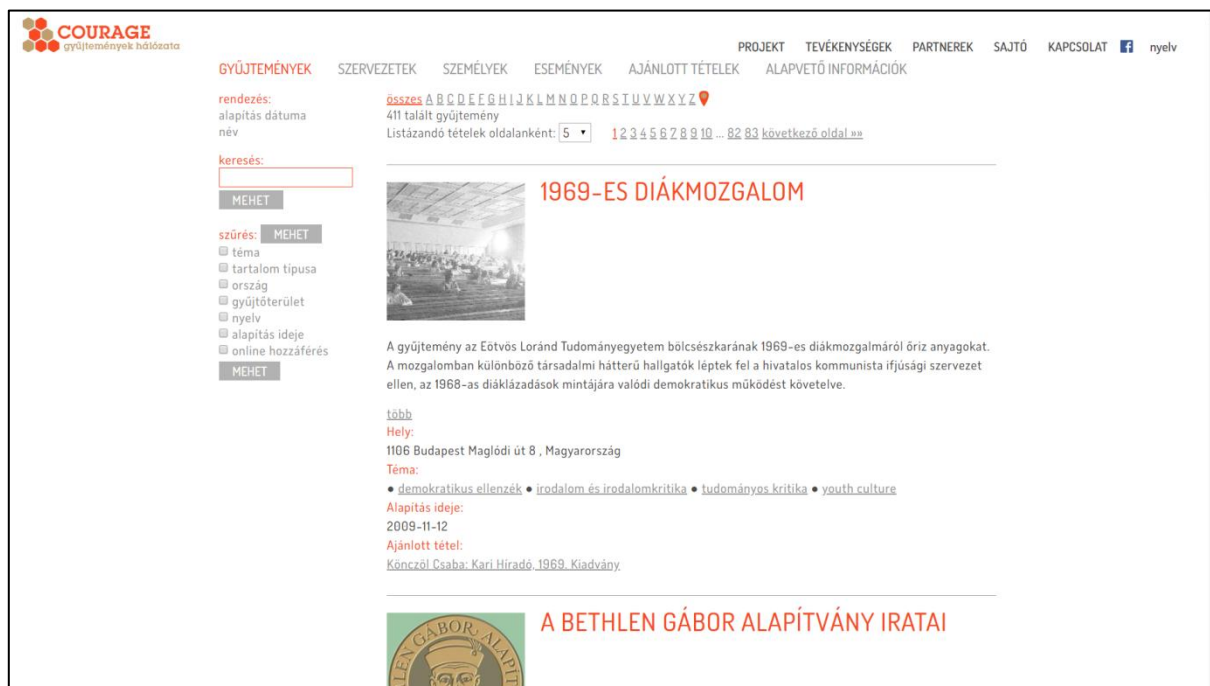
A *Working Spaces* részben a középkorhoz vagy az első világháborúhoz kapcsolódó témákban hozhatunk létre *annotált dokumentumokat, jegyzeteket* (8. ábra). A szövegeket elláthatjuk tag-ekkel, így megjelölhetjük a személyneveket és a helyszíneket, melyekről a beépített vizualizációs eszköz rögtön készíti egy összegzést. A *szemantikus annotációra* is támogatott kontrollált szótárak és ontológiák használatával a rendszer a jelölt nevekhez automatikusan hozzácsatolja a vonatkozó CENDARI *wiki-bejegyzést* (ha létezik, akkor kattintásra meg is jelenik a bővebb információ, ha nem létezik még a tétel, akkor az egérmutató rámozgásakor érzékelhetjük, hogy még nincs ilyen). A szövegeken kívül *táblázatokat* is előállíthatunk, a képek hozzáadásának lehetősége egyelőre nem biztosított. A felület rendelkezik *HTML-szerkesztővel*, a megjelenítéshez egyéni stílust is beállíthatunk. Az *RDF-szerkesztő* funkciója még nem készült el. A létrehozott dokumentumokat nyilvánossá tehetjük, növelve a hozzáférési lehetőségeket. Ilyen módon a téma kutatójaként a források mellett gyűjthetjük a saját megjegyzéseinket is egy, a felületen létrehozott szövegfájlban, amelyeket megoszthatunk a kollégáinkkal.

COURAGE³⁵

A COURAGE portálja alapértelmezésben angolul töltődik be, de 16 nyelvet választhatunk, köztük a magyart is. A kutatás ugyanis jelentős hazai közreműködéssel valósul meg, a *Magyar Tudományos Akadémia (MTA)* három szervezeti egysége vállal benne komoly szerepet: a *Bölcsészettudományi Központ (BTK)*, a *Társadalomtudományi Központ (TK)*, valamint a *Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI)*. Az MTA BTK vezetésével, tíz ország szakértőinek bevonásával, az *EU Horizon 2020* pénzügyi finanszírozásával épülő *COURAGE Registry* (9. ábra) a *szocialista tömb alternatív kultúrájának* magán- és közgyűjteményi adatait tárja fel, amelyhez tananyag, kézikönyv³⁶ és további disszeminációs szakmai aktivitások (pl. filmfesztivál, kiállítás, játék) is társulnak. A COURAGE – Gyűjtemények hálózata elnevezésű, a Kulturális ellenállás és ellenkultúra, az ellenzéki-ség öröksége az egykori szocialista országokban tematikát valló nemzetközi programról a hazai könyvtári szakirodalomban is tájékozódhatunk,³⁷ ezért részletesebb vizsgálatától eltekintünk.



8. ábra CENDARI. A Working Spaces egy részlete



9. ábra A COURAGE gyűjteményi keresője

eHeritage³⁸

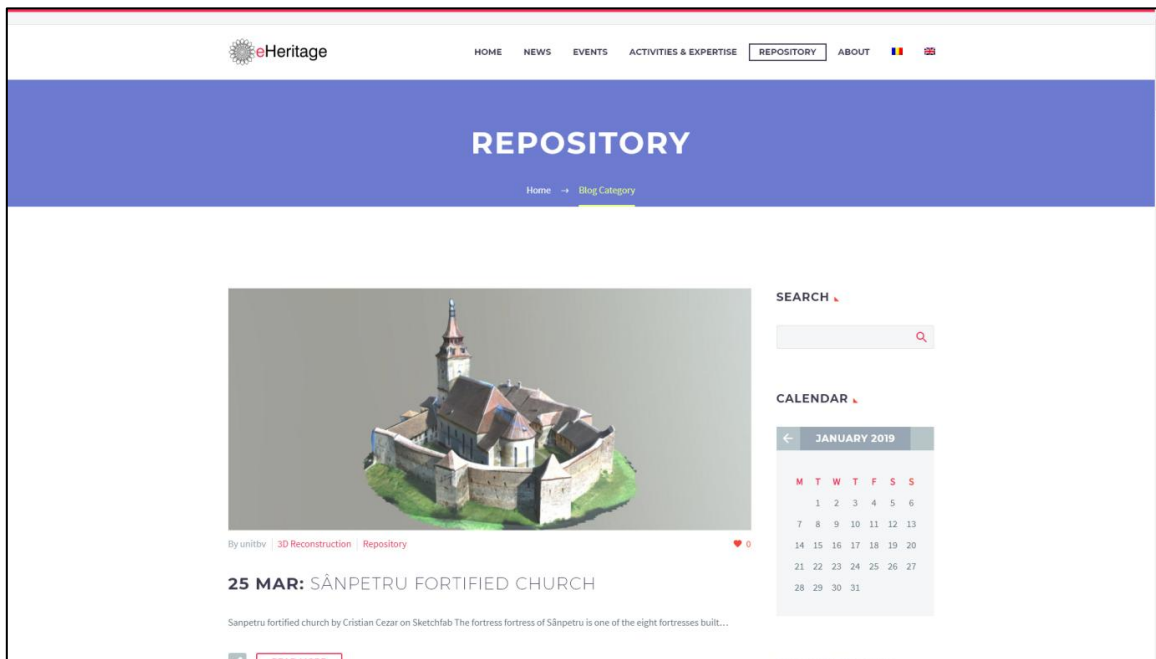
A román, szlovén és olasz kooperációval létrejött eHeritage – Expanding the Research and Innovation Capacity in Cultural Heritage Virtual

Reality Applications projekt a kulturális örökség megőrzését a virtuális valósághoz kötődő technológiákkal kívánja segíteni.

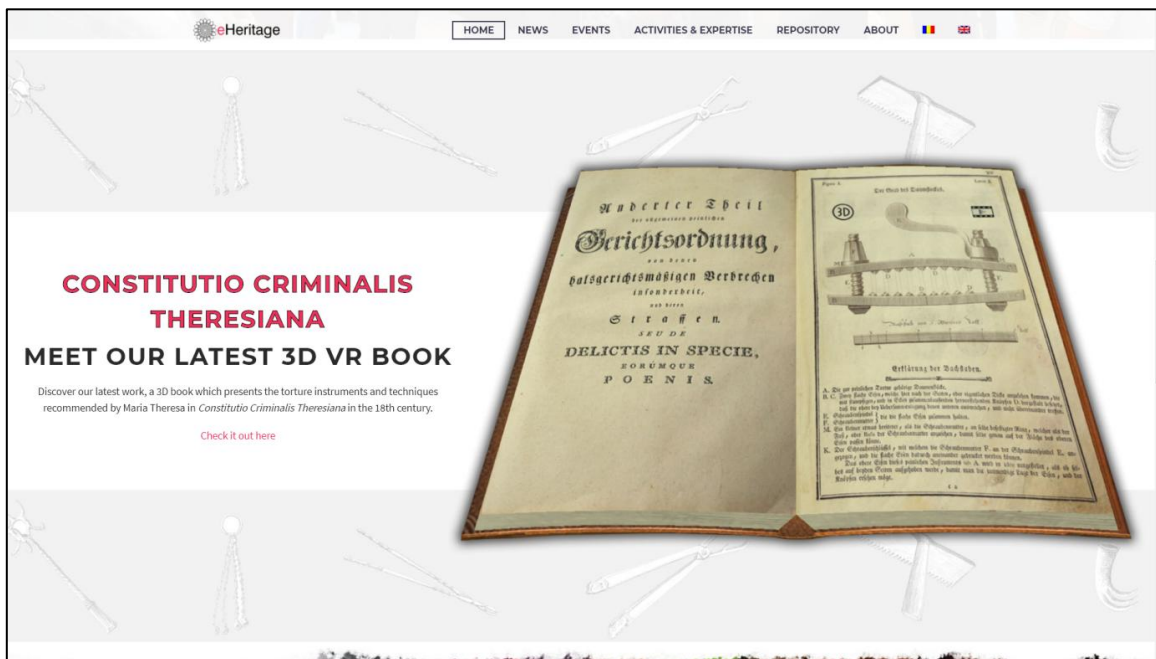
Portáljuk *repozitóriuma* (10. ábra) jelenleg 3D rekonstrukciókat tárol, nevezetesen 9 tételt a Romániában található építészeti emlékekről. A modelleket a *Sketchfab*³⁹ alkalmazással készítették és *OBJ formátumban* szabadon letölthetők.

Legújabb fejlesztésük az 1768-ban *Mária Terézia* által kibocsátott büntetőkódex, a *Constitutio*

Criminalis Theresiana című könyv 3D VR (*Virtual Reality*) technológiával készült rekonstrukciója, melyet a *Brassó Megyei Történeti Múzeumban*⁴⁰ őrzött nem nyilvános státuszú példányról készítették (11. ábra).⁴¹ A helyben megtekinthető digitális másolattal az erdélyi nagyközönség számára is elérhetővé vált a kordokumentum.



10. ábra Az eHeritage repozitóriumi felülete



11. ábra eHeritage. A Constitutio Criminalis Theresiana 3D VR változatáról szóló hirdetés

A projekt a 3D-re és virtuális valóságra épülő rekonstrukciókon kívül foglalkozik a *kiterjesztett valóságot (Augmented Reality)*, az *érintést is bevonó virtuális valóságot (haptics)* és a *hologramkészítést (holography)* célzó fejlesztési tevékenységekkel is.

Europeana Newspapers⁴²

Az adatbázis több mint 20 milliónyi archív, ezáltal történeti forrásértékkel bíró *újságoldalt* gyűjt össze. Eredetileg a *Europeana* és a *The European Library (TEL)* oldalán is elérhettük a fájlokat, azonban utóbbin 2016. december 31-én lezárták a tartalomfeltöltést, az adatok Europeana-szerverre történő migrációja folyamatban van.⁴³ Rövid magyarázatként annyit tennék hozzá, hogy a TEL eredetileg Európa nemzeti könyvtárai és katalógusai számára kívánt közös portált biztosítani, a Europeana pedig az összes európai közgyűjtemény digitális tára.⁴⁴

A Europeana 10 milliónyi folyóiratoldala (a teljes állomány valamivel több mint fele) *karakterfelismerő programmal* (Optical Character Recognition – OCR) feldolgozott, így a szolgáltató könyvtáraknál tárolt teljes szövegben kulcsszavasán, akár személyek vagy helyszínek után is kutathatunk. A *keresésnél* a gyűjtemény típusának (Newspaper) kiválasztását elvégezve is a Europeana ismert filterrendszerével találkozunk:

- teljes szöveg vagy metaadat,
- média (szöveg és link),

- dátum,
- felhasználási jogok,
- szolgáltató ország,
- nyelv,
- aggregátor,
- intézmény.

A *megjelenítés* során (12. ábra) a fájl mellett a többi gyűjteménynél megszokott metaadatcsoportok jelennek meg az egyes újságokról:

- cím,
- osztályozás – anyagtípus (itt folyóirat) és kulcsszavak,
- tulajdonságok – méret – (itt oldalszám) és nyelv,
- idő (itt megjelenés ideje),
- eredet,
- hivatkozások és kapcsolatok,
- megnevezett entitások,
- hasonló tételek.

A NER-technológia használatával a leírásokban lévő *entitások*: micsoda, mikor, hol (13. ábra). A megnevezést tárgyszavak biztosítják, az idő és hely meghatározását a katalógusban rögzített adatok. A tartalmat a szótárakra (pl. tárgyszavaké) mutató *URI-k* (Uniform Resource Identifier) azonosítják. Számos link vagy URI is a TEL oldalára irányít, a leírás szerint a meglévő tartalmat a migrálás befejezéséig megőrzik. Magyar nyelvű folyóiratot egyelőre nem találunk a kollekcióban.

The screenshot shows a web page from Europeana Newspapers. At the top, there are navigation links: 'Gyűjtemények', 'Böngészés', 'Kiállítások', 'Blog', and a 'WEBOLDALAINK' button. Below the navigation is a search bar with the text 'Kulcsszó szerinti keresés itt'. The main content area displays a record for 'The South Wales daily post - 1910-07-30'. On the left is a thumbnail of the newspaper. To the right of the thumbnail, the title 'The South Wales daily post - 1910-07-30' is shown. Below the title, there is a description in Welsh: 'Began with issue for: [No. 1] (13 Feb. 1893). Ended with issue for: No. 17729 (15 Mar. 1930). Description based on issue for: [No. 1] (13 Feb. 1893). Cyhoeddwyd yn Abertawe gan William Llewellyn Williams. Published in Swansea by William Llewellyn Williams. Continues as: South Wales daily post and Cambria daily leader. Papur newydd dyddiol a gyhoeddwyd yn Saesneg yn bennaf. 'Roedd i'r papur dueddiadau ceidwadol. Am gyfnod, fe hawliau'r cylchrediad mwyaf o unrhyw bapur cyhoeddwyd yn Abertawe a De-orllewin Cymru. Rhoddai sylw arbennig i newyddion Abertawe a De Cymru. O tua 1906 ymlaen bu'n eiddo i'r South Wales Post Newspapers Co. Yn 1930 fe'i unwyd â'r 'Cambria Daily Leader'. Teitlau cysylltiol: Cambria Daily Leader (1861-1930); South Wales Daily Post and Cambria Daily Leader (1930-1932). A daily English language newspaper that was supportive of conservative politics. At one time, the newspaper had the greatest circulation than any other paper published in Swansea and South West Wales. The newspaper's main content included local news, with a special focus on news re-'. On the right side of the page, there are sections for 'NEXT ITEM' and 'PREVIOUS ITEM', each with a thumbnail and a title: 'The South Wales daily post - 1910-03-17' and 'The South Wales daily post - 1910-11-01'.

12. ábra Europeana Newspapers. Egy folyóirat adatlapja

Megnevezett Entitások

MICSODA ▾

Fogalom terminus

<http://vocab.getty.edu/aat/300026656>

Fogalom címke

["Zeitung"] (de); ["newspapers"] (en); ["kranten"] (nl); ["diarios (noticias)"] (es);

Bővebb fogalom

<http://vocab.getty.edu/aat/300026642>

MIKOR ▶

HOL ▾

Helyszín terminus

<http://sws.geonames.org/2636432/>

Helyszín címke

["Swansea", "Swansea"] (def);

Helyszín szélesség

13. ábra Példák megnevezett entitásokra a Europeana Newspapers egy tételénél

A Europeana Newspapers hatékony módszertani segítséget nyújt a folyóirat-digitalizáló munkákhoz:

- 1. Digitalizálási projektek minőségének javítása**
Az archív folyóiratok digitális konverziója gyakran fejlesztésre szorul, így a hatékonyság értékelésére érdemes különböző mutatókat alkalmazni, úm. a digitalizálás mélysége, időtartama, költsége. A különböző módszerekből ezáltal az optimális választható ki. A *Public Materials – Deliverables* alatt találunk jó gyakorlat-leírásokat az értékelések készítéséhez.⁴⁵
- 2. Felmérés az európai nemzeti könyvtárak és nyilvános könyvtárak által digitalizált folyóirat-gyűjteményekről**
A 2012-ben elvégzett, majd 2013-ben kiegészített adatgyűjtésből kiderült, hogy a 20. századi tartalmak elérhetőségét nagyban gátolja az OCR-technológiák alkalmazásának gyakori elmaradása.
- 3. Ajánlás egy átfogó metaadat modellre**
A *Metadata Encoding & Transmission Standard (METS)*⁴⁶ és az *Analyzed Layout and Text Object (ALTO)*⁴⁷ szabványokon alapuló folyó-

irat-feltárási metaadat rendszer európai bevezetése.

- 4. Nyílt forráskódú szoftverek disszeminációja**
A *Public Materials – Technical Tools* oldalról⁴⁸ kiindulva a különböző alkalmazások közt találunk olyat, amellyel ellenőrizhetjük a digitalizált oldal felépítését, a szegmentációt, az OCR eredményeit (PAGE Viewer), továbbá a névelem-felismerésére szolgáló alkalmazást (NER) is, utóbbit a Holland Nemzeti Könyvtár⁴⁹ jövőtárból. A többi alkalmazásról táviratilag:
 - A *PAGE Metadata Scanner* megvizsgálja az oldalak tartalmát és egy statisztikát állít elő azok tulajdonságairól.
 - Az *Entity Disambiguation Tool* abban segít, hogy meghatározzuk a szövegen belül a különböző névelemeket.
 - A *Structify* egy grafikus eszköz a METS-fájlok szerkesztésére.
 - A *File Rename Tool (FRT)* a fájlok átnevezésében és a Europeanak megfelelő fájlnévszerkezet kialakításához nyújt segítséget.

- A *Binarization and Conversion Tool (BCT)* JPEG2000 vagy JPEG fájlokat hoz létre a mesterképekből, mivel ezeket a formátumokat részesíti előnyben a Europeana Newspapers projekt.
- A *File Analyzer Tool (FAT)* egy ellenőrző eszköz, megelőzi, hogy a felismerés során hosszú várakozási idő merüljön fel, és hatékonyra teszi a fájlok feldolgozását.

5. Marketingkommunikáció

Az elektronikus formában elérhető gyűjtemények népszerűsítése információs napokkal és workshopokkal.

Látható, hogy a Europeana Newspapers nemcsak az archív folyóiratokat tartalmazó digitalizált európai történeti gyűjtemények egy felületről történő elérését biztosítja, hanem a *könyvtári (hírlap) digitalizációs tevékenység* meghatározó uniós kutató- és fejlesztőközpontja is.

Forced Labor 1939-1945⁵⁰

A *náci Németország* több mint 20 millió embert sújtott *kényszermunkával*. Közülük 26 ország 590 túlélőjével készült felvételt tárol az adatbázis 583 interjúban (7 páros riport) német, angol, orosz és cseh nyelvű felületen.

A 32 projektcsapat által 2005-ben és 2009-ben főleg Közép- és Kelet-Európában (Ukrajna, Lengyelország, Oroszország) készített felvételeknek szinte kivétel nélkül elérhető a szöveges változata, 170-ből német fordítás is készült.

Az oldalon található *keresési és böngészési* lehetőségek a következők:

- név;
- csoportok (pl. koncentrációs táborok túlélői, politikailag üldözöttek);
- foglalkoztatási szektor (pl. mezőgazdaság, építőipar);
- internálás fajtája (pl. koncentrációs tábor, gettó, börtön);
- interjú nyelve;
- internáláskori lakhely (ország).

A feldolgozás befejezése után *interjúrészletekre* is lehet majd keresni témák szerint: tárgyszavakkal látják el a felvételeket, majd hozzákapcsolják ah-

hoz az időponthoz, ahol szó volt az adott témakör-ről, és ezekre a részletekre lehet ugrani a videóban.

Az interjúkat regisztráció után érhetjük el,⁵¹ a *hátéranyagokat* (cikkek, rövidfilmek, szakértőkkel való beszélgetések, oktatási programok) belépés nélkül is megtekinthetjük. A *14. ábra* egy magyar interjúalannal készült felvétel *adatlapjába*, míg a *15. ábra* pedig az adatbázis *interaktív térképébe* enged betekintést.

PELAGIOS⁵²

A betűszavas feloldáson (Pelagios: Enable Linked Ancient Geodata In Open Systems) kívül metaforaként is értelmezhető a PELAGIOS, mely jelentheti: „*a tengerről*.” A tenger, amely az ipari korszak előtt a közlekedés fő útvonalaival biztosította, hasonlóan, ahogy egy digitális forrás összekötheti az *antik helyszínekre vonatkozó hivatkozásokat*. A *Pelagios Commons* a *Linked Open Geographic Data*⁵³ bölcsészeti alkalmazására szerveződött szakmai közösség és infrastruktúra. Három Open Data megoldásra épülő, ezáltal szabadon elérhető alkalmazást tart fenn.

Az első, a *Recogito* (Link Data menü) egy online felület dokumentumok kollaborációs *annotálására*. Hasonló a már említett CENDARI Working Spaces szolgáltatásához. Ebbe is tölthetünk fel saját fájlokat, dokumentumokat. A NER-szolgáltatás automatikusan bejelöli a helyneveket és a személyneveket, az eseményeket (még) nem. Képet is csatolhatunk, létrehozhatunk annotációkat, különböző tag-eket (helyek, személyek és események megjelölésére). Az automatikus helyazonosításnál a rendszer felkínálja a javaslat jóváhagyását vagy változtatását. Az annotációkat CSV és RDF formátumban, a helyeket GeoJSON és KML típusban *exportálhatjuk*, az annotált szövegeket pedig TEI/XML formátumban (csak a helynév-annotációkat) tudjuk menteni, mivel azokat tárolja ehhez szükséges formátumban. A létrehozott dokumentumot meg is oszthatjuk, beállíthatjuk a meghívott közreműködők jogosultságait. A szolgáltatást a *Stanford CoreNLP* alkalmazás⁵⁴ biztosítja, de egy plugin segítségével sajátot is használhatunk erre a célra.

Eva B.: Ungarische Jüdin, nach der KZ-Haft in die USA emigriert

Die ungarische Jüdin wurde als Kind nach Auschwitz und in ein Außenlager von Flossenbürg transportiert. Nach 1945 emigrierte sie in die USA und arbeitete in der Krankenpflege.

- geboren 1931 in Chust an der ungarisch-tschechoslowakischen-ukrainischen Grenze
- im März 1944 als 13-Jährige nach Auschwitz deportiert, wo die ganze Familie ermordet wird
- überlebt knapp die Selektionen
- im Herbst 1944 Transport in ein Außenlager des KZ Flossenbürg in Hainichen/Sachsen, dort Zwangsarbeit in einer Munitionsfabrik
- 1945 Todesmarsch nach Theresienstadt und Befreiung durch die Rote Armee
- zunächst Rückkehr nach Ungarn, 1947 Emigration in die USA
- Arbeit als Krankenpflege-Lehrerin

Interview-Daten:

- [Interview za562 »](#)
- englischsprachiges Audio-Interview
- am 24./26. Juli 2006 durch Sara Ghitis
- Interview-Dauer: 5 Stunden
- [Teilsammlung "USA – Breman Atlanta"](#)



Eva B., 2006 in den USA



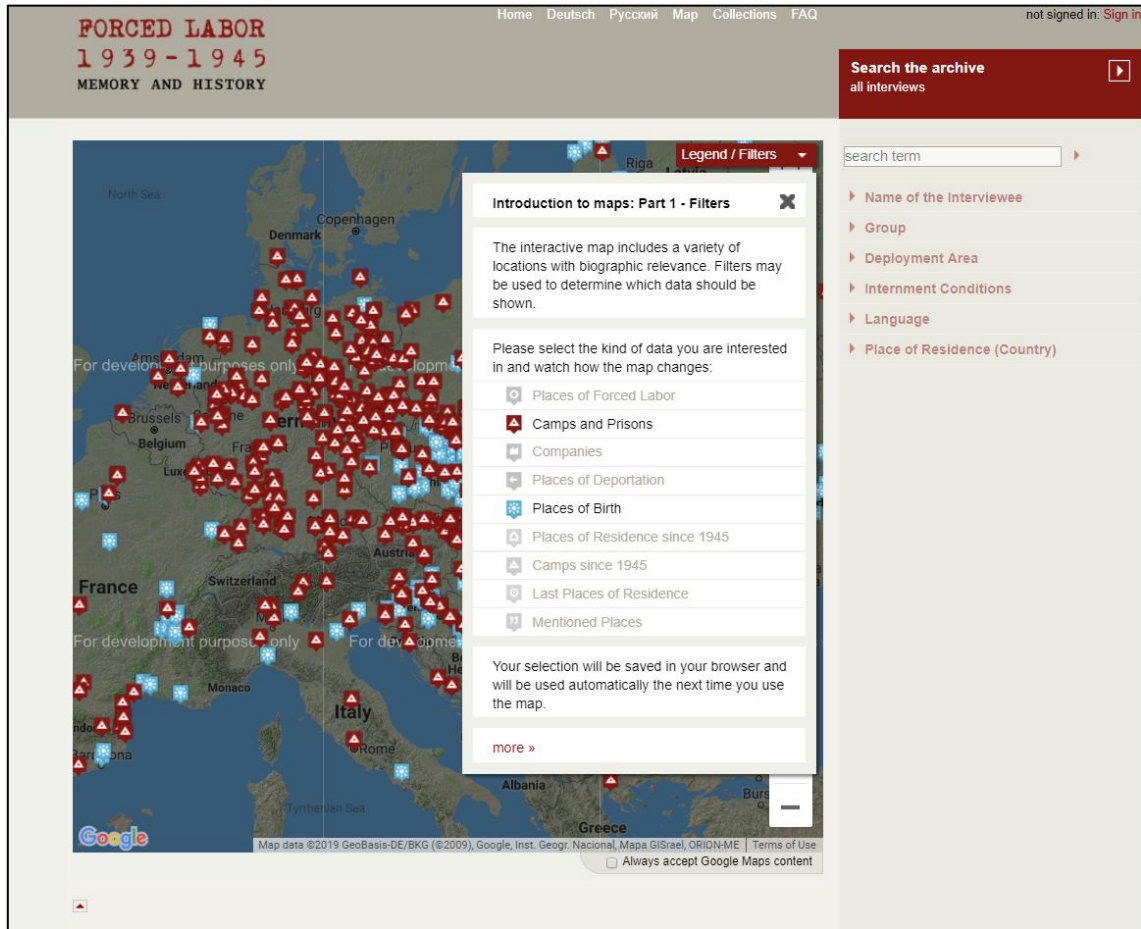
Eva B. (links) mit ihrer Familie, 1941

14. ábra Forced Labor 1939-1945. Egy magyar interjúalannal készült felvétel ismertetése

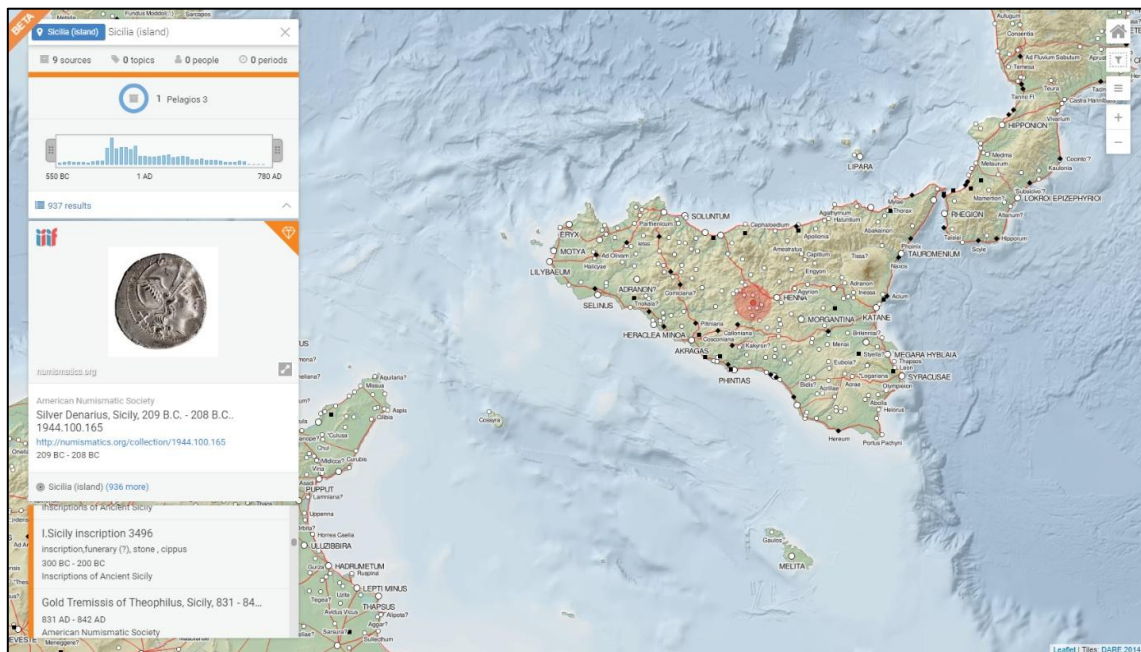
A második, a *Peripleo*⁵⁵ (Expore Data – Search Peripleo menü) *interaktív térképen* tartalmazza az összes ókori helynevet és modern elnevezését. A térképen szerepelnek a kapcsolódó források (16. ábra), például egy csatlakozott gyűjtemény linkje vagy a *Fasti Online* általi ásatási helyszínek. Az *I. Sicily* epigráfiai projekt⁵⁶ is adatforrás, így a szicíliai településeknél a feliratokra vonatkozó információkhoz is eljuthatunk.

A harmadik, a *Pelagios Map Tiles* (Expore Data – Map Tiles menü) egy *dinamikus térkép*, melynek választható funkciói:

1. *Digital Map of the Roman Empire*: a Római Birodalom digitális atlasza antik helynevekkel, ahol a különböző történelmi korszakok szerint jeleníthetjük meg a településeket;
2. *Modern places – Open Street Map (OSM)*⁵⁷ alkalmazásra épülő térkép a mai helynevekkel;
3. *Aerial*: műholdkép-nézet a helyszínek adatait webre és mobil eszközökre egyaránt optimalizált felületen kínáló térítéses *Mapbox*⁵⁸ szolgáltatással.



15. ábra Forced Labor 1939-1945. Térképes nézet



16. ábra Periplo. Szicília ókori térképe

Tom Elliott és Sean Gillies, a Pleiades közösség New York University-beli munkatársai szerint „a számítástechnika, a klasszikusok és a geográfia történetét több tudományterület és azok gyakorlói közötti gazdag és eredményes dialógusnak lehet tekinteni.”⁵⁹

Fasti Online⁶⁰

A *Mediterraneum* ásatásaira koncentrázó régészeti adatbázis három egységből áll, amelyeket a főoldalról érhetünk el:

1. *Excavation*: ásatási helyek 2000-től nyilvántartva;
2. *Archaeological Conservation*: az ásatások során előkerült leletek megőrzésével kapcsolatos adatforrások;
3. *Survey*: régészeti terepfelméréseket tartalmazó adatbázis.

Mindnek az alapja egy *interaktív térkép*, amelyen a településre való kattintással fedezhetjük fel az adott hely régészeti feltárásának adatait kronológiai sorrendben (17. ábra):

- régészeti leletek,
- időszakok,
- ásatási alkalmak.

A három platform keresőjét összekapcsolták, az adatok jól strukturáltak, a feltöltést útmutatók szabályozzák annak érdekében, hogy az adatokat egységes szerkezetben adják meg az ásatások koordinálói. A térképes mellett a *haladó keresés* (Advanced Search) mező:

- ország,
- ásatási helynév,
- a régészeti lelet típusa,
- dátumintervallum.

Colle Oppio, Terme di Traiano 2000 2004 2009 2010 2011 2012 2013

Bibliography

Monuments

Baths Domus Cemetery

Periods

Roman Republican Roman Imperial Late Antique Medieval Modern Contemporary

Season

2000 In March 2000, the director of the British School at Rom, Professor Andrew Wallace-Hadrill, and the Soprintendenza and Comune di Roma, commissioned a geophysical survey over part of the Colle Oppio in the centre of Rome.
The trial resistivity survey w
... Read More

2004 As a result of the archaeological discoveries made between 1998 and 1999 in the gallery beneath the southwestern corner of the Baths of Trajan on the Oppian Hill (that included among other finds the building with the fresco "Painted City" and the wall mosa... Read More

2009 Following the work undertaken between 1998-99, and the discoveries made in those years, funding was obtained under the L. 396/90 for Roma Capitale, which financed campaigns in 2003-2006 and 2007-2008.
A museum is to be created in the gallery below the s
... Read More

2010 The 2010 excavations investigated the room named of the "Vendemmia" after a large patch of preserved wall mosaic, and part of the mosaic decorating its vault. The mosaic was part of the decorative scheme of one large building, probably constructed in the F... Read More

2011 In 2011, excavation continued at the centre of the gallery, also revealing various surfaces associated with the Trajanic construction site. In the room housing the Muse and Philosopher wall mosaic, the eastern sector, cut diagonally by the Trajanic walls, ... Read More

2012 In 2012, the excavation continued in the northernmost part of the room revealing most of the mosaic's lower register. This was in a better state of preservation and a series of four female figures can be distinguished.

17. ábra Fasti Online. Egy ásatás adatlapja

A FOLD&R (Fasti On Line Documents & Research) alatt olvasható három elektronikus lektorált *szakfolyóirat*:

1. FOLD&R *Italy Series* (428 közlemény),
2. FOLD&R *Archaeological Conservation Series* (5 közlemény),
3. FOLD&R *Archaeological Survey Series* (9 közlemény).

A tanulmányok bibliográfiai adatai mellett külön link mutat az adatbázis vonatkozó rekordjaira, így a Fasti kiváló példa a tudományok művelésének elméleti és gyakorlati szinergiájára.

Összegzés

A bölcsészettudományokon belül a történeti stúdiók az egyik legnagyobb szegmenst képviselik. Az informatika térhódítása nyomán *a nagy tömegű, különböző forrásokban és formátumban rendelkezésre álló múltbéli adatok* – különösen a személyek, a helyszínek, az események és az időpontok – *kompatibilis módon történő hálózatba szervezése új, eddig ismeretlen összefüggések felismerésével járhat*. Ez sok esetben kiegészítheti, finomíthatja, de akár meg is változtathatja a régi korokról alkotott tudományos világképet. A dokumentumok (a szövegek mellett képek, sőt hangfájlok stb.) és a tárgyi emlékek (pl. használati eszközök, épületek) virtuális reprezentációja általi pusztá közzététele rendszerint már kevésnek bizonyul, a *multidiszciplinaritásra épülő információtechnológiai eredmények* felhasználásával sok szempontú, automatizált feldolgozást és az *összefüggéseket korszerű vizualizációs eszközökkel* kísért megjelenítését követeli meg korunk. A jelen tanulmányunkban felvonultatott, *történelemtudományra fókuszáló tematikus portálok* is bizonyítják, hogy a programokat egyértelműen *nemzetközi kooperációval* és (pénzügyi) támogatással, a közgyűjteményi, illetve a könyvtári adatcsere szabványokra alapozva érdemes megvalósítani. Közleményünk kifejezett célja, hogy az elemzett projektek mintául szolgálhassanak a magyarországi tartalomfejlesztésekhez, s felhívjuk a figyelmet arra, hogy az adatfeltöltésen túl a hazai kutatók is bekapcsolódhatnak a már meglévő alkalmazások tökéletesítésébe. Nem elhanyagolható szempont az sem, hogy a kulturális vállalkozások célközönsége nem szűkíthető le a történész, a régész, a nyelvész stb. szakközönségre és a fejlesztők, közgyűjteményi szakértők körére sem. Olyan *többfunkciós*, egyszerre tudományos igényű, ugyanakkor egyszerűen kezelhető, gyakorta látványos, a figyelmet megragadó *forráshálózatokról* értekeztünk, amelyek

nem tekinthetők a klasszikus értelemben vett referenz-szolgáltatásoknak, hiszen a laikus érdeklődők használatára is számítanak: *a tájékoztatói eszközök új evolúciós lépcsőjének lehetünk tanúi*.

Hivatkozások

- ¹ Projects. EADH: <https://eadh.org/projects>
- ² 1914-1918-online. International Encyclopedia of the First World War: <http://eadh.org/projects/1914-1918-online>
- ³ ALCIDE (Analysis of Language and Content In a Digital Environment): <http://eadh.org/projects/alcide>
- ⁴ CENDARI (Collaborative European Digital Archive Infrastructure): <http://eadh.org/projects/cendari>
- ⁵ OURAGE (Cultural Opposition – Understanding the Cultural Heritage of Dissent in the Former Socialist Countries): <http://eadh.org/projects/courage>
- ⁶ eHeritage (Expanding the Research and Innovation Capacity in Cultural Heritage Virtual Reality Applications): <http://eadh.org/projects/eheritage>
- ⁷ Europeana Newspapers: <http://eadh.org/projects/europeana-newspapers>
- ⁸ Forced Labour 1939-1945. Memory and History: <https://eadh.org/projects/forced-labour-1939-1945-memory-and-history>
- ⁹ PELAGIOS (Pelagios: Enable Linked Ancient Geodata In Open Systems): <http://eadh.org/projects/pelagios>
- ¹⁰ Fasti Online: <http://eadh.org/projects/fasti-online>
- ¹¹ BALÁZS Ágnes – SEBŐK Miklós: Névelemfelismerés. In: Kvantitatív szövegelemzés és szövegbányászat a politikatudományban. Szerk. Sebők Miklós. Budapest, L'Harmattan Kiadó. 2016. 51-61. p. ISBN 978-963-414-229-4 és HOOLAND, Seth van – De WILDE, Max – VERBORGH, Ruben – STEINER, Thomas – Van de WALLE, Rik: Exploring entity recognition and disambiguation for cultural heritage collections. = Literary and Linguistic Computing. 30. évf. 2015. 2. sz. 262-279. p. <https://doi.org/10.1093/lc/ftq067>
- ¹² Named Entities Recognition Annotator Tool for Europeana Newspapers. GitHub: <https://github.com/KBNLresearch/europeanap-ner>
- ¹³ GitHub: <https://github.com>
- ¹⁴ Access to Memory (AtoM): <https://www.accesstomemory.org/en>
- ¹⁵ Freie Universität Berlin: <https://www.fu-berlin.de>
- ¹⁶ 1914-1918-online. International Encyclopedia of the First World War: <https://encyclopedia.1914-1918-online.net>

- 17 Open Encyclopedia System (OES):
<http://www.open-encyclopedia-system.org>
- 18 Center for Digital Systems:
<https://www.fu-berlin.de/en/einrichtungen/verwaltung/cedis/index.html>
- 19 Center for Modern Greece (CeMoG):
<http://www.cemog.fu-berlin.de/en/index.html>
- 20 Bayerische Staatsbibliothek:
<https://www.bsb-muenchen.de/en>
- 21 Semantic MediaWiki (SMW):
https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki
- 22 1914-1918 Online connects with CENDARI:
<http://www.cendari.eu/about/news/1914-1918-online-connects-cendari>
- 23 Digital Object Identifier (DOI) System:
<https://www.doi.org>
- 24 Analysis of Language and Content In a Digital Environment (ALCIDÉ):
<http://dh.fbk.eu/projects/alcide-analysis-language-and-content-digital-environment>
- 25 Istituto Storico Italo-Germanico (ISIG):
<https://isig.fbk.eu>
- 26 „Alcide de Gasperi 1945 és 1953 között miniszterelnökként és külügyminiszterként vezető szerepet játszott Olaszország bel- és külpolitikájának formálásában.” Forrás: Alcide de Gasperi: az európai demokrácia és szabadság elhivatott szószólója. Európai Bizottság. Az EU alapító atyái:
https://europa.eu/european-union/sites/europaeu/files/docs/body/alcide_de_gasperi_hu.pdf
- 27 TextPro – Text processing Tools: <http://textpro.fbk.eu>
- 28 DUDÁS László: Alkalmazott mesterséges intelligencia. Közread. a Miskolci Egyetem. Budapest, Kempen Farkas Hallgatói Információs Központ Digitális Tankönyvtár. 2011. 65-67. p.
https://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_alkalmazott_mesterseges_intelligencia/0046_alkalmazott_mesterseges_intelligencia.pdf
- 29 PÉTER Róbert: A Big Data kihívás és lehetőség a bölcsészettudományokban: digitális szövegek és metaadatok távoli olvasása. = Magyar Tudomány. 177. évf. 2016. 11. sz. 1323-1330. p.
<http://www.matud.iif.hu/2016/11/08.htm>
- 30 Highcharts: <https://www.highcharts.com>
- 31 Data-Driven Documents (D3.js): <https://d3js.org>
- 32 Leaflet: <https://leafletjs.com>
- 33 CENDARI (Collaborative European Digital Archive Infrastructure): <http://www.cendari.eu>
- 34 Encoded Archival Description (EAD). The Library of Congress: <https://www.loc.gov/ead>
- 35 COURAGE – Gyűjtemények hálózata:
<http://hu.cultural-opposition.eu>
- 36 APOR Balázs – APOR Péter – HORVÁTH Sándor szerk.: The Handbook of COURAGE: Cultural Opposition and its Heritage in Eastern Europe. Budapest, Institute of History, Research Centre for the Humanities, Hungarian Academy of Sciences. 2018. 636 p. <http://doi.org/10.24389/handbook>
- 37 PI. APOR Péter – BÓDI Lóránt – HORVÁTH Sándor – MICSIK András – SCHEIBNER Tamás: COURAGE: az ellenzéki kultúra adatbázisa. In: Valóságos könyvtár – könyvtári valóság. Könyvtár- és információtudományi tanulmányok 2018. Szerk. Kiszl Péter – Csík Tibor. Budapest, ELTE BTK Könyvtár- és Információtudományi Intézet. 2018. 345-350. p. <http://doi.org/10.21862/vkkv2018.345>
- 38 eHeritage: <http://www.eheritage.org>
- 39 Sketchfab: <https://sketchfab.com>
- 40 Muzeul Județean de Istorie Brașov:
<http://www.brasovistorie.ro>
- 41 A korabeli büntető törvénykönyvből az Országgyűlési Könyvtárban is található példány: Constitutio Criminalis Theresiana (Theresiana). Országgyűlési Könyvtár:
<https://www.ogyk.hu/hu/blog/posts/constitutio-criminalis-theresiana-theresiana>
- 42 Europeana Newspapers:
<http://www.europeana-newspapers.eu>
- 43 The European Library:
<http://www.theeuropeanlibrary.org/tel4/newspapers>
- 44 Bővebben ld.: KOVÁCSNÉ KORENY Ágnes: Hozzáfértés és újrahaznosítás: a TEL és a Europeana a tudományos kutatás szolgálatában. In: Valóságos könyvtár – könyvtári valóság. Könyvtár- és információtudományi tanulmányok 2016. Szerk. Kiszl Péter – Boda Gáborné Köntös Nelli. Budapest, ELTE BTK Könyvtár- és Információtudományi Intézet. 2017. 225-232. p. <http://hdl.handle.net/10831/34590>
- 45 Deliverables. Europeana Newspapers:
<http://www.europeana-newspapers.eu/public-materials/deliverables>
- 46 Metadata Encoding & Transmission Standard (METS). The Library of Congress:
<http://www.loc.gov/standards/mets>
- 47 Technical Metadata for Layout and Text Objects. The Library of Congress:
<https://www.loc.gov/standards/alto>
- 48 Tools. Europeana Newspapers:
<http://www.europeana-newspapers.eu/public-materials/tools>

⁴⁹ Named entity recognition for digitised historical newspapers. KB Research. Research at the National Library of the Netherlands. March 3, 2014: <http://blog.kbresearch.nl/2014/03/03/ner-newspapers>

⁵⁰ Zwangsarbeit 1939-1945: <https://www.zwangsarbeit-archiv.de>

⁵¹ Vö.: VHA (Visual History Archive) Online. USC Shoah Foundation: <http://vhaonline.usc.edu/login>

⁵² Pelagios Commons: <http://commons.pelagios.org>

⁵³ Linked Open Geodata: <http://linkedgeodata.org/About>

⁵⁴ tanford CoreNLP – Natural language software: <https://stanfordnlp.github.io/CoreNLP>

⁵⁵ „[peri'pleo]. Ancient Greek for «to sail (or swim) around» in the sense of exploration or discovery.”
Forrás: <https://peripleo.pelagios.org>

⁵⁶ I.Sicity: <https://isicity.wordpress.com>

⁵⁷ OpenStreetMap: <https://www.openstreetmap.org>

⁵⁸ Mapbox: <https://www.mapbox.com>

⁵⁹ „We view the history of computing, classics and geography as a rich and profitable dialogue between many disciplines and practitioners.” Forrás: ELLIOTT, Tom – GILLIES, Sean: Digital Geography

and Classics. = Digital Humanities Quaterly. 3. évf. 2009. 1. sz.

<http://www.digitalhumanities.org/dhq/vol/3/1/000031/000031.html>

⁶⁰ Fasti Online: <http://www.fastionline.org>

Beérkezett: 2018. X. 20-án.



Kiszl Péter

habilitált egyetemi docens,
az ELTE BTK Könyvtár- és
Információtudományi Intézetének
igazgatója, Információtudományi
Tanszékének és Könyvtártudományi
doktori programjának vezetője.

E-mail: kiszl.peter@btk.elte.hu



Mátyás Melinda

informatikus könyvtáros,
MTMT adminisztrátor,
ELTE Egyetemi Könyvtár és Levéltár,
Egyetemi Könyvtár,
Gyűjteményszervezési Osztály.

E-mail: matyas.melinda@lib.elte.hu