

WIKITEXT AZ ELTE DATA ADATBÁZISBAN – BETHLEN MIKLÓS KANCELLÁR LEVELEZÉSÉNEK SZEMANTIKUS FELDOLGOZÁSA

Vida Bence

ELTE BTK Digitális Bölcsészeti Tanszék

vida.bence@btk.elte.hu

Sebestyén Ádám

ELTE BTK Digitális Bölcsészeti Tanszék

sebestyen.adam@btk.elte.hu

Absztrakt

Tanulmányunk Bethlen Miklós kancellár (1642–1716) Jankovics József által közreadott hivatali és magánlevelezésének digitalizálási és szemantikus elemzési folyamatát mutatja be. Adatbázisunkat a Digitális Örökség Nemzeti Laboratóriumnak és az ELTE Digitális Bölcsészeti Tanszékének ELTEdata nevű szemantikus adattárában tettük közzé. A szemantikus adatbázisban minden levél egyedi adatlapot kapott, a levelek szövegét pedig wikitext formátumban adjuk közre. A levél szövegében kódolt entitások közé tartoznak a személynevek, a geolokáció, a bibliai locusok és azonosítható politikai események. A szemantikus adatszerkezet a Wikidata adatstruktúráját követi, a szemantikus állítások és entitások szintjén is össze van linkelve a Wikidata megfelelő állításaival egy külső azonosítón keresztül. A tételek ilyen módon tetszőleges névterekhez és repozitóriumokhoz kapcsolhatók. Bár a korpusz még kezdeti stádiumban van, a 100 levél és a bennük azonosított több száz entitás alkalmassá teszi arra, hogy SPARQL-lekérdezések segítségével keressük és rendszerezzük a feldolgozott adatokat – tanulmányunkban a levelek tematikus elemzését, geolokációk szerinti elemzését, valamint kapcsolati háló szerinti elemzését végeztük el.

Kulcsszavak: digitális bölcsészet, szemantikus web, történeti adatbázis, digitalizáció, SPARQL, ELTEdata

Abstract

Wikitext in the ELTEdata Database – Semantic Analysis of the Correspondence of Chancellor Miklós Bethlen

Our study presents the digitization and semantic analysis process of the private correspondence of Chancellor Miklós Bethlen (1642–1716) published by József Jankovics. We published our database in the semantic repository called ELTEdata, maintained by the National Laboratory for Digital Heritage and the Department of Digital Humanities at Eötvös Loránd University.

vös Loránd University. In the semantic database, each letter has been assigned a unique data sheet, and the text of the letters is published in wikitext format. Entities encoded in the text of the letters include personal names, geolocations, biblical loci, and identifiable political events. The semantic data structure follows the Wikidata data structure and is linked at the level of semantic statements and entities to the corresponding statements in Wikidata via an external identifier. In this way, items can be linked to arbitrary namespaces and repositories. Although the corpus is still in its early stages, the 100 letters and the hundreds of entities identified in them make it possible to search and organize the processed data using SPARQL queries. In our study, we performed a thematic analysis of the letters, an analysis based on geolocations, and an analysis based on the network of relationships.

Keywords: digital humanities, semantic web, historical database, digitisation, SPARQL, ELTEdata

Tanulmányunk tárgya Bethlen Miklós kancellár (1642–1716) hivatali és magánlevelezésének elemzése szemantikus adatbázis építésén és SPARQL-lekérdezések alkalmazásán keresztül. A vizsgált korpusz jelenleg 100 db. levelet tartalmaz. Az első levél 1680. április 16-án kelt, az utolsó 1693. július 23-án. Valamennyi levél az 1680 és 1690 közti évtizedben keletkezett, a pontos keltezés nélküli levelek esetében a szövegkiadásban megadott intervallumra hagyatkoztunk, mivel azok éves pontossággal szerepelnek a forráskiadásban.¹ A korpusz célja a Bethlen-levelezés tartalomelemzésének elvégzése, ezt megelőzően a levelek wikitext verzióinak elkészítése, valamint szemantikus adatgazdagítás a korpuszon entitásazonosítások formájában. A korpusz történeti kontextusának bemutatását követően rátérünk a levelek feldolgozásának módszertani kérdéseire, végül bemutatjuk elemzésünk első eredményeit, kitérve a kutatás további irányaira. A korpusz létrehozásának tágabb értelemben vett célja a teljes levelezés tartalomelemzés, amelynek elvégzésével válaszokat keresünk arra, miként működtette Bethlen Miklós egész Európára kiterjedő hírhálózatát, amely nemzetközi levelezőpartnereket foglal magában, és Bethlen pályakezdetétől haláláig értékes információ forrásként, valamint hírközlő csatornaként szolgált számára, értesüléseit felhasználta politikai stratégiájának megalkotásához, politikai üzeneteit pedig közvetítette partnerei számára a nemzetközi kéziratos és nyomtatott nyilvánosságban.²

1 A projekt részletes leírása megtalálható az ELTEdata honlapján: https://eltedata.elte-dh.hu/wiki/Bethlen_Mikl%C3%B3s_kancell%C3%A1r_levelez%C3%Agse Ugyanítt érhető el a feldolgozott levelek listája is. Utolsó elérés: 2025.07.01.

2 A kéziratos nyilvánosság jelenségével bővebben foglalkozik Tóth Zsombor: A kora újkori könyv antropológiája. Kéziratos irodalmi nyilvánosság Cserei Mihály (1667–1756) írás- és szöveghasználatában. Irodalomtörténeti Füzetek 178. szám. Szerk. Főrizs Gergely. Magyar Tudományos Akadémia Bölcsészettudományi Kutatóközpont Irodalomtudományi Intézet, Reciti, Budapest, 2017.

Bethlen Miklós apjának, Bethlen János kancellárnak (1613–1678) a hatására 1661 és 1664 között peregrináció keretében járta be Európát, egyetemi tanulmányokat folytatott, és diplomáciai küldetést is vállalt.³ Hazatérve az erdélyi–francia kapcsolatokat egyik közvetítőjévé vált. Politikai pályája az 1670-es évek elejéig ígéretesen ívelt felfelé, ám apja halála, valamint a családja számára bázist jelentő református székely elit csoport súlyának csökkenése nehéz helyzetbe hozta.⁴ 1676 és 1677 között Fogarasban raboskodott, miután a fejedelem elleni szervezkedéssel, Béldi Pál generális (1621–1679) mozgalmában való részvétellel vádolták meg. Szabadulása után fokozatosan tért vissza az erdélyi politikai életbe, ennek során igyekezett kihasználni a fejedelemség geopolitikai környezetében bekövetkező változásokat.

Az Erdélyi Fejedelemség 1685 után fokozatosan felszabadult az oszmán vazallusi státuszból, V. Károly lotaringiai herceg (1643–1690) vezetésével Habsburg csapatok 1687 őszére elfoglalták Erdély jelentős részét. A bevonuló csapatok vezetői, Antonio Caraffa (1642–1693), Friedrich Siegmund Scherffenberg (1644–1688), valamint később Jean-Louis Rabutin de Bussy (1642/3–1717) képviselték azt a kemény álláspontot Bécs felé, amelynek értelmében Erdélyt katonai igazgatás alá helyezték volna, tartósan kialakítva az Oszmán Birodalommal határos ütközőzónát. Elsődleges céljuk a Magyar Királyság magterületének biztosítása volt, valamint Erdély gazdasági rendszerének és nyersanyagainak végső kiaknázása, felhasználása a Habsburg utánpótlás biztosítására.

Apafi Mihály fejedelem (1632–1690) és az erdélyi nemesség több lehetséges utat is mérlegelt a konfliktuszóna kialakításának elkerülésére. A bizonytalan közigazgatási helyzet rendezésének egyik eszközeként a fejedelem tanácsa köré vizsgálóbizottságot, *deputatio*-t hoztak létre, amely a bevonuló császári csapatok kihágásait, az okozott károkat volt hivatott számba venni. Bethlen Miklós 1686-tól a *deputatio* tagjaként tevékeny részt vállalt az Erdély katonai megszállását előkészítő adóösszeírásban, amelynek a keretei közt bécsi hivatalnokok felbecsülték a fejedelemség gazdasági kapacitását, hogy kidolgozzák a bevonulás részletes feltételeit, valamint a tartós megszállás menetét.⁵ A korpuszunk részét képező levelek ennek a munkának fontos és hiteles forrásai, tartalmazzák az eljárá-

3 Az erdélyi peregrinációról bővebben ld. Tonk Sándor – Szabó Miklós: Erdélyiek egyetemjárása a középkor és a kora újkor folyamán. In: Régi és új peregrináció I–III. A III. Nemzetközi Hungarológiai Kongresszuson elhangzott előadások - Szeged, 1991. augusztus 12–16. Szerk.: Békési Imre, Jankovics József, Kósa László, Nyerges Judit, Péter László. Budapest-Szeged, Nemzetközi Magyar Filológiai Társaság – Scriptum Kft., 1993. 491–501. Sipos Gábor: „tanulásoknak okáért...” Bethlen Gábor világi elitképző programjáról. Korunk 24. évf. 3. szám (2013), 21–27.

4 Az időszak erdélyi székely elit csoportjainak felemelkedéséről, működéséről, valamint Bethlen változó helyzetéről ld. Balogh Judit: Bethlen Miklós kapcsolatrendszere a székely elitben. In: Horn Ildikó – Laczházi Gyula szerk.: Reformer vagy lázadó? Bethlen Miklós és kora. L'Harmattan Kiadó, Budapest, 2020. 315–347.

5 Bővebben ld. Fazekas István: *A Magyar (Udvari) Kancellária és hivatalnokai 1527–1690 között*. Akadémiai doktori értekezés. Budapest, 2018.

sok menetét, a vallomások részleteit, a jelen lévő hivatalnokok és katonai előjárók véleményét, Bethlen saját helyzetértékelése mellett. Fontos forrásai nemcsak a deputatio tevékenységének, hanem Bethlen politikai szemléletének is – a leírtak politikai diskurzusban elhangzó érvrendszerek alapjait, programadó beszédek kereteit tartalmazzák.

Az itt tárgyalt korpusz alapjául Bethlen Miklós kancellár levelezésének Jankovics József-féle szövegkiadása szolgált.⁶ A levelek döntő többsége magyar nyelven íródott, a mintában tizenkét latin nyelvű levél található, valamint egy levél utóírata német nyelvű. Formalizált feldolgozásukat nehezíti, hogy írásmódjuk nem következetes, nem alkalmaznak interpunkciót, a kisbetű és nagybetű használat is esetleges. Ezeket az egyenetlenségeket igyekeztünk egyenesíteni a wikitext verzió elkészítésénél. A korpusz építésének első fázisában a leveleket a Jankovics-féle forráskiadásból véve digitalizáltuk, az OCR elvégzése után javítottuk az esetleges karakterfelismerési hibákat, és elkészítettük a levelek digitális másolatát. Ezt követően a levelek szövegén és a levelek szerkezeti felépítésén is módosításokat végeztünk a forráskiadáshoz képest, hogy a wikitext változat elkészítésére alkalmas normalizált szöveghez jussunk.⁷

A digitalizált forrásokat az ELTEdata felületén adtuk közre. Ily módon a szövegek nem XML-, hanem a Wikimédia saját, wikitext nevű formátumában lettek létrehozva. A levél szövegében kódolt entitások közé tartoznak a személynevek, a geolokáció, a bibliai locustok és az azonosítható politikai események. A kódolt entitások köre a későbbiekben szabadon bővíthető. Mivel a célunk a tesztanyaggal nem egy digitális szövegkiadás előkészítése volt, továbbá számos, az eredeti forrásszövegben sokszor csak közvetett módon, utalás formájában jelen lévő adatot akartunk leképezni (például a levelekben említett eseményeket, diplomáciai egyezményeket, politikai változásokat), használhatóbbnak bizonyult a wikitext formátum.⁸ Ennek segítségével lehetőség nyílik az annotációk megjelenítésére, entitásazonosítás révén a levelek politikatörténeti dimenzióira vonatkozó információk hozzáadására, és az adatbázisban rögzített adatok lekérdezésére, vizualizálására is. Jóllehet, mindezekre a TEI-XML is alkalmas, a kutatás jelen állásánál rugalmasabbnak,

6 A forráskiadás pontos adatai: Bethlen Miklós levelei (1657-1698) I., II. Összegyűjtötte, sajtó alá rendezte, a bevezető tanulmányt és a tárgyi jegyzeteket írta Jankovics József. A magyar nyelvi jegyzeteket írta Nényei Gáborné, fordította Kulcsár Péter. (Régi magyar prózai emlékek). Régi Magyar Prózai Emlékek sorozat, Akadémiai Kiadó, Budapest, 1987. A Régi Magyar Prózai Emlékek sorozatot az Akadémiai Kiadó alapította, 1968-tól 1993-ig jelentek meg a sorozat kötetei. Alapító szerkesztője Tolnai Gábor volt.

7 Az átíráshoz az alábbi forráskiadás átírási elveit vettük alapul: Enyedi György prédikációi I. (67–99. beszéd) 2. Kolozsvári Kódex, Marosvásárhelyi Kódex, Sárospataki Kódex, Conciones vetustissimae. Sajtó alá rendezte, az előszót és a jegyzeteket írta Lovas Borbála. MTA-ELTE HECE, Magyar Unitárius Egyház, Budapest, 2016. 23–29.

8 XML és wikitext különbségeiről, hátrányairól és előnyeiről lásd Fellegi Zsófia összefoglalóját az ELTEdata Tudásáramlás című gyűjteményének technikai háttéréről: Fellegi Zsófia, *Wikidata – ELTEdata – Wikimédia: A szövegkiadások technikai háttérre*, <https://eltdedata.elte-dh.hu/wiki/WIKIDATA-ELTE-ELTE-ELTE-WIKIMEDIA>

az adatszerkezet kialakításának a szempontjából könnyebben kezelhetőnek mutatkozott a wikitext formátum.

Rátérve a levelek szemantikus adatszerkezetének kódolására, minden levélhez készült egy adatlap és egy a levél szövegét tartalmazó aloldal, így több irányból is közelíthetünk a szemantikus adatstruktúrához. A wikitext verzió és a levél adatlapját is link kapcsolja össze, így a korpuszt használva az eredeti átirási módban közreadott misszilis anyag is elérhetővé, és a korpuszban szereplő levelek szövegével összevethetővé, összeolvashatóvá válik. Jelenleg a következő entitások vannak kódolva és kérdezhetőek le a szemantikus adatstruktúrában: geolokáció, dátumok, személynevek (címezett, feladó, a levél szövegében szereplő nevek), azonosítható politikatörténeti események, méltóságok, logisztika, kereskedelem. Mivel a korpuszt az ELTEdata felületén tároljuk, a szemantikus adatszerkezet a Wikidata adatstruktúráját követi, a szemantikus állítások és entitások szintjén is össze van linkelve a Wikidata megfelelő állításaival egy külső azonosítón keresztül. A tételek ilyen módon természetesen névterekhez és repozitóriumokhoz kapcsolhatók. Az adatstruktúra kialakítása az RDF-nyelvet követve, tulajdonság (property) – értékpárok (value) létrehozásával valósul meg. A teszt-korpusz adatbevitelét kézzel végeztük, de tervben van félautomatikus adatbevitel alkalmazása, miután az adatbevitel folyamatát sikeresen normalizáltuk.⁹

A digitalizált forrásszöveg közzététele mellett a levelek tartalmát szemantikusan kivonatolt adatlapokon képeztük le.¹⁰ Ez az adatlap az alapvető információkon túl (osztály, gyűjtemény, keltezéssel kapcsolatos adatok) feltünteti a levél nyelvét, az eredeti, nyomtatott forráskiadást, az onnan átvett sorszámot, a levél témáját, az említett személyeket, helyszíneket (értve ezalatt településeket és történelmi közigazgatási egységeket) továbbá a szóba kerülő fontosabb eseményeket is. A forrásszöveg entitásként való létrehozása lehetővé teszi, hogy a felhasználó a levelekből kivonatolt adatokon különféle lekérdezéseket futtasson, rendszerezze vagy vizualizálja a szemantikusan leképezett információt.

Az adatlapon létrehozott entitások nagy része a forrásszövegben is jelölve van, hasonlóan egy Wikipédia-oldalhoz, azzal a különbséggel, hogy itt nem lapok, hanem entitások vannak hozzáadva a szöveghez. Az entitásazonosítás sok esetben a lábjegyzeteket is kiváltja. Példának okáért, a korpusz leggyakrabban említett személyei, I. Lipót császár és Apafi Mihály erdélyi fejedelem, szinte soha nincsenek néven nevezve, többnyire csupán a rangjukra vonatkozó szavak jelölik őket, úgy mint *őfelsége, fejedelem, urunk őnagysága*

⁹ A félautomata adatbevitel lehetőségeiről ld. Sebestyén 2022, 180–181.

¹⁰ A jelenlegi projekt egyik újdonsága, hogy a szemantikus adatlapok létrehozása és a forráskiadás egyszerre valósul meg. Az ELTEdata eddigi gyűjteményeiről lásd az alábbi konferenciaelőadásokat: Sebestyén Ádám, „Az ELTEdata szemantikus adatbázis legújabb fejlesztései”, in *Workshop 2022*, szerk. Tick József, Kokas Károly, Holl András, 179–185 (Budapest: HUNGARNET Egyesület, 2022). Sebestyén Ádám, Sárközi-Lindner Zsófia, „Történeti források szemantikus feldolgozása – Az ELTEdata adatbázis új gyűjteményei”, in *Workshop 2024*, szerk. Tick József, Kokas Károly, Holl András, 105–112 (Budapest: HUNGARNET Egyesület, 2024).

és hasonlók. Ennél is homályosabbak az olyan esetek, mikor Bethlen fontos katonai vezetőkre hivatkozik kizárólag a titulusukkal (*generális, commissarius*). A Jankovics-féle nyomtatott kiadásban a névmutató nyújt támpontot a személyek azonosításához, és tünteti fel, adott levél melyik oldalon említi a szóban forgó személyt, a wikipediaként létrehozott levél azonban közvetlenül a forrásszövegben azonosítja és teszi egyértelművé, kiről van szó. Így módon, amennyiben egy levél például két különböző tábornokra is hivatkozik a titulusuk megnevezésével, az entitásazonosítás jegyzetek nélkül is egyértelművé teszi, kiről van szó. Hasonlóképp lehet feloldani a szerződéseket, egyezményeket takaró latin szavakat, például a *tracta* esetében, mely túlnyomórészt vagy az úgynevezett Haller-féle diplomára, vagy ennek előzményeire, a fejedelemség és a bécsi udvar közti diplomáciai találkozókra, tárgyalásokra utal.

Az entitások azonosításán túl lehetőség van egy fogalmi jelölésrendszer kialakítására is, mindez azonban egyelőre a távlati célkitűzések közé tartozik. Az ELTEdata Tudásáramlás című digitális forráskiadását alapul véve jelölnénk a levelek fogalmilag azonosítható szakaszait.¹¹ Példának okáért, ha egy adott szakasz valamiről tudósít, vagy épp a diplomáciai és politikai nyelvezet jegyeit hordozza, akkor ennek megfelelően kapna egy jelölést, csatolva hozzá adott fogalom ELTEdata-adatlapját. Ennek szellemében egyszerre valósulna meg a szövegkiadás és a szövegfeldolgozás, feltárva a Bethlen-levelek mélyebb szemantikai rétegeit.

Bár a korpusz még kezdeti stádiumban van, a 100 levél és a bennük azonosított több száz entitás alkalmassá teszi arra, hogy SPARQL-lekérdezések segítségével keressük és rendszerezzük a feldolgozott adatokat. A látványosabb adatvizualizációk egy úgynevezett aggregátor funkciót alkalmazó lekérdezéssel készültek, melynek segítségével a gyakoriságra vonatkozó adatokat lehet vizsgálni. Példának okáért a felhasználó lekérdezheti a levelekben említett helyszíneket, az aggregátor funkcióval pedig számszerűsítheti, melyek ezek közül a leggyakrabban említett települések.¹² Természetesen figyelembe kell venni, hogy amennyiben az osztálynál a települést adjuk meg (tehát szűkítjük a lekérdezés paramétereit), az olyan történelmi régiók, mint Erdély, illetve az egyes vármegyék, nem fognak szerepelni

11 A kiadás elérhető az alábbi linken:

<https://eltedata.elte-dh.hu/wiki/Tud%C3%A1s%C3%A1raml%C3%A1s>

12 A lekérdezés, valamint annak eredménye az a lábbi linken elérhető és futtatható:

[https://query.elte-dh.hu/embed.html#SELECT%20%3Fplacesmentioned%20%3FplacesmentionedLabel%20\(COUNT\(%3Fletter\)%20AS%20%3Fcount\)%60A%60A%60AWHERE%60A%60A%67B%60A%620%620%3Fletter%20wdt%3AP2%20wd%3AQ380%3B%60A%620%620%620%620%620%620%620%620%620%620wdt%3AP54%20wd%3AQ39885%3B%60A%620%620%620%620%620%620%620%620%620%620p%3AP197%20%5Bps%3AP197%20%3Fplacesmentioned%20%5D.%60A%620%620%3Fplacesmentioned%20wdt%3AP2%20wd%3AQ349.%60A%620%620%60A%620%620SERVICE%20wikibase%3Alabel%20%67B%20obd%3AServiceParam%20wikibase%3Alanguage%20%622%5BAUTO_LANGUAGE%5D%62Cen%622.%20%67D%60A%620%620%60A%620%620%67D%60A%60AGROUP%20BY%20%3Fplacesmentioned%20%3FplacesmentionedLabel%60A%60AORDER%20BY%20DESC\(%3Fcount\)](https://query.elte-dh.hu/embed.html#SELECT%20%3Fplacesmentioned%20%3FplacesmentionedLabel%20(COUNT(%3Fletter)%20AS%20%3Fcount)%60A%60A%60AWHERE%60A%60A%67B%60A%620%620%3Fletter%20wdt%3AP2%20wd%3AQ380%3B%60A%620%620%620%620%620%620%620%620%620%620wdt%3AP54%20wd%3AQ39885%3B%60A%620%620%620%620%620%620%620%620%620%620p%3AP197%20%5Bps%3AP197%20%3Fplacesmentioned%20%5D.%60A%620%620%3Fplacesmentioned%20wdt%3AP2%20wd%3AQ349.%60A%620%620%60A%620%620SERVICE%20wikibase%3Alabel%20%67B%20obd%3AServiceParam%20wikibase%3Alanguage%20%622%5BAUTO_LANGUAGE%5D%62Cen%622.%20%67D%60A%620%620%60A%620%620%67D%60A%60AGROUP%20BY%20%3Fplacesmentioned%20%3FplacesmentionedLabel%60A%60AORDER%20BY%20DESC(%3Fcount))

a találatok között. Ha ezekkel a beállításokkal futtatjuk a lekérdezést, kiderül, hogy a tesztanyag leveleiben a legtöbbet említett település Nagyszében, amely 1686-ot követően a levelek domináns keltezési helyeként mutatja Bethlen hivatali tevékenységének súlypontját.

Az efféle gyakoriságot vizsgáló lekérdezések arra is alkalmasak, hogy segítségükkel a levelekben említett személyeket csoportosítsuk. Amennyiben volt rá adat, feltüntettük a foglalkozáson túl a személyek felekezeti hovatartozását és társadalmi státuszukat is, adott tehát, hogy ezen szűrők alapján is rendszerezzük Bethlen politikai és magánéleti kapcsolatait.

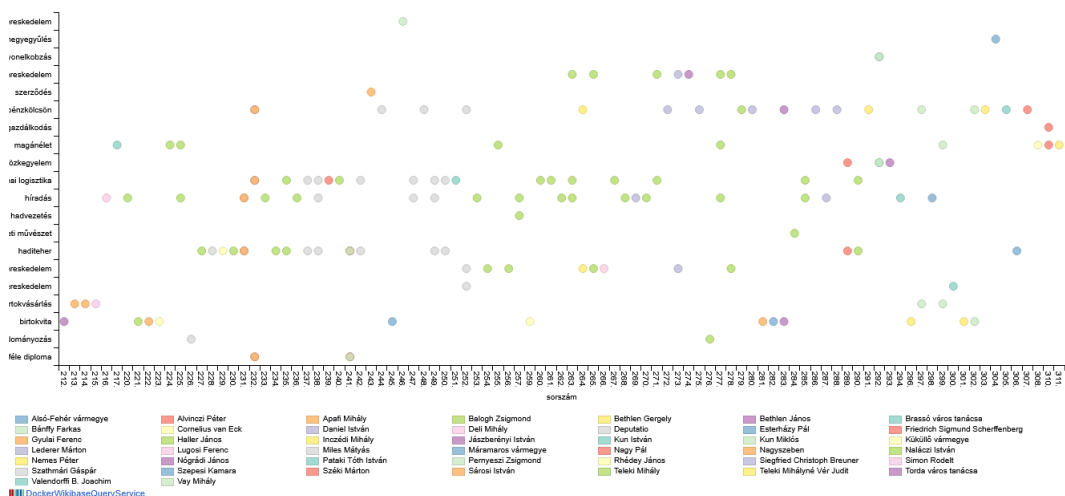
Ugyancsak aggregátor funkció segítségével szépen vizualizálható a levelek központi témáinak eloszlása. Ebből látszik, hogy a tesztanyagban feldolgozott Bethlen-levelek nagy része katonai–diplomáciai, valamint pénzügyi kérdésekkel foglalkozik. Mivel a korpusz jelentős része a császári hadak által elkövetett visszaélések kivizsgálására összpontosít, nem meglepő módon gyakori témák az egymással szoros összefüggésben lévő katonai logisztika (hadseregélelmezés és egyéb gazdasági folyamatok összessége), illetve a haditeher. Ehhez lazán kötődik a pénzkölcsön, mely egyaránt szerepelhet országos politikai és egyéni pénzügyi kontextusban. A leggyakrabban előforduló levéltémát, a híradást, olyan szövegeknél tüntettük fel, ahol Bethlen egyértelműen utal a tudósítás aktusára („Tegnap bőven tudósítottam, uram kegyelmeteket”), vagy a kontextusból világosan kiderül, hogy híradásról van szó (példa erre az összes olyan eset, amikor kortárs európai katonai és politikai fejleményekről tájékoztat). Előfordul, hogy Bethlen a konkrét médiumot, a hírközlés korabeli formáit is megnevezi (a novellára, mint újságtípusra néhány levelében hivatkozik). A korpusz bővülésével érdemes lesz majd ezeket a distinkciókat is figyelembe venni, és adott esetben különbséget tenni a szóbeli tájékoztatás, valamint annak írott formája között.

Mivel gyakori, hogy Bethlen egy levelében több témát is érint, érdemes lehet megvizsgálni ezek lehetséges kapcsolódásait. Az itt bemutatott adatvizualizációk, a TreeMap, Scatterchart és a gráf megjelenítések a wikidata által nyújtott vizualizációs eszközök, amelyek az ELTEdata adatbázisai számára is elérhetők. Az első ábrán látható TreeMap-vizualizáció a címzetteket és a központi témákat mutatja. Néhány részletet kiemelve láthatóvá válik, hogy a témák alapján a levelek címzettjei is csoportosíthatóak. A *deputatio* esetében egyértelmű a katonai logisztika, valamint az ezzel rokon témák túlsúlya. A Szathmári Gáspárhoz írt levelekben viszont, aki a nagyenyedi kollégium vicekurátoraként az intézmény fenntartásáért volt felelős, az olyan témák dominálnak, mint a pénzkölcsön, adományozás, vagy a különböző javakkal folytatott kereskedelem. Kiemelendő még Bethlen első feleségének apja, Kun István, akinek – mivel Thököly Imre híve volt – végül el kellett hagynia Erdélyt. Ezt a konfliktust mutatja, hogy a neki írott levelekben a magánélet mellett az uralkodóval való szembeszegülés lehetséges retorziói kerülnek elő, úgymint a vagyonelkobzás, illetve a közkegyelem elnyerésének lehetőségei.



1. ábra: Levéltémák és címezettek (TreeMap)

A címzettek és beszédtemák kapcsolatát a szórásdiagram is remekül illusztrálja. A 2. ábra vízszintes tengelye a beszédek sorszámát jelzi, a függőleges pedig az egyes témákat. A pontok különböző színei a címzettek label-jeinek feleltethetők meg.



2. ábra: Levéltémák és címzettek (Scatter chart)

A kapcsolati háló összetettségét jól szemléltetik a gráf-adatvizualizációk. A SPARQL-ben generált egyszerű gráfban a levelek és a személyek csomópontokként funkcionálnak, az élek pedig a kettő közti kapcsolatot, az említések számát jelzik. Ennek segítségével szépen

kirajzolódik, hogy egy periférián lévő, vagy egy sokszor említett személyről van-e szó, képet adva Bethlen kapcsolati hálójának alakulásáról is.¹³

A tesztanyag során tapasztaltak és az elért eredmények is igazolják, hogy érdemes hosszabb távon foglalkozni Bethlen Miklós kancellár levelezésének szemantikus feldolgozásával. Ami a távlati célkitűzéseket illeti, az adatgazdagítás és a korpusz bővítése a legfontosabb feladat. Mivel a Bethlen-levélgűjtemény kellően terjedelmes, érdemes lehet életpályájának egy adott szakaszához kapcsolódva válogatni a levelek között. Így nemcsak életének és politikájának egyes korszakai válnának jól elkülöníthetővé, hanem az is a figyelem középpontjába kerülne, hogyan alakult kapcsolati hálója politikai vagy magánéleti változásokat követően. Mivel a levelekből strukturált adatkészlet nem áll rendelkezésre, korpuszbővítés esetében végig kell gondolni az automatizálás lehetőségeit is.

Irodalomjegyzék

Balogh, Judit: Bethlen Miklós kapcsolatrendszere a székely elitben. In: Horn Ildikó – Laczházi Gyula szerk.: *Reformer vagy lázadó? Bethlen Miklós és kora*. L'Harmattan Kiadó, Budapest, 2020. 315–347.

Bethlen Miklós levelei (1657-1698) I., II. Összegyűjtötte, sajtó alá rendezte, a bevezető tanulmányt és a tárgyi jegyzeteket írta Jankovics József. A magyar nyelvi jegyzeteket írta Nényei Gáborné, fordította Kulcsár Péter. (Régi magyar prózai emlékek). Régi Magyar Prózai Emlékek sorozat, Akadémiai Kiadó, Budapest, 1987.

Fazekas, István: *A Magyar (Udvari) Kancellária és hivatalnokai 1527–1690 között*. Akadémiai doktori értekezés. Budapest, 2018.

Fellegi, Zsófia: Wikidata – ELTEdata – Wikimédia: A szövegkiadások technikai háttere, <https://eltedata.elte-dh.hu/wiki/WIKIDATA-ELTEDATA-WIKIMEDIA>

Sebestyén, Ádám: „Az ELTEdata szemantikus adatbázis legújabb fejlesztései”, in *Networkshop 2022*, szerk. Tick József, Kokas Károly, Holl András, 179–185 (Budapest: HUNGARNET Egyesület, 2022).

Sebestyén, Ádám – Sárközi-Lindner, Zsófia: „Történeti források szemantikus feldolgozása – Az ELTEdata adatbázis új gyűjteményei”, in *Networkshop 2024*, szerk. Tick József, Kokas Károly, Holl András, 105–112 (Budapest: HUNGARNET Egyesület, 2024).

Sipos Gábor: „tanulásoknak okáért...” Bethlen Gábor világi elitképző programjáról. *Korunk* 24. évf. 3. szám (2013), 21–27.

13 A lekérdezés, valamint annak eredménye az alábbi linken elérhető és futtatható:
https://query.elte-dh.hu/embed.html#%23defaultView%3AGraph%0A%0ASELECT%20%3Fp%20%3FpLabel%20%3Forsors%3%Aim%20%3Fw%20%3FwLabel%0AWHERE%20%0A%67B%0A%09%3Fp%20wdt%3AP2%20wd%3AQ380%3B%0A%20%20%20%20%20%20%20%20wdt%3AP54%20wd%3AQ39885%3B%0A%20%20%20%20%20%20%20%20wdt%3AP196%20%3Forsors%3%Aim.%0A%09%3Fp%20wdt%3AP194%20%3Fw%20.%0A%20%20%0A%20%20SERVICE%20wikibase%3Alabel%20%67B%20bd%3AserviceParam%20wikibase%3Alanguage%20%22%5BAUTO_LANGUAGE%5D%2Cen%22.%20%67D%0A%20%20%67D

Tonk, Sándor – Szabó, Miklós: Erdélyiek egyetemjárása a középkor és a kora újkor folyamán. In: Régi és új peregrináció I-III. A III. Nemzetközi Hungarológiai Kongresszuson elhangzott előadások - Szeged, 1991. augusztus 12-16. Szerk.: Békési Imre, Jankovics József, Kósa László, Nyerges Judit, Péter László. Budapest-Szeged, Nemzetközi Magyar Filológiai Társaság – Scriptum Kft., 1993. 491–501.

Tóth Zsombor: A kora újkori könyv antropológiája. Kéziratok irodalmi nyilvánosság Cserei Mihály (1667–1756) írás- és szöveghasználatában. Irodalomtörténet Füzetek 178. szám. Szerk. Főrész Gergely. Magyar Tudományos Akadémia Bölcsészettudományi Kutatóközpont Irodalomtudományi Intézet, Reciti, Budapest, 2017.