

MICHAEL GERVERS

Középkori angol oklevelek datálása szóelőfordulások alapján

Az 1066-os normann hódítástól körülbelül 1300-ig tartó időszakból fennmaradt angol birtokátruházási dokumentumokon ritkán szerepel a keletkezés pontos dátuma. Ahhoz azonban, hogy egy történelmi forrást megfelelően tudjunk értelmezni, szükségünk van arra, hogy mind időben, mind topográfiailag olyan pontosan tudjuk elhelyezni, amennyire az csak lehetséges. Az oklevelekkel foglalkozó tudósok állandó gondja volt a szükséges pontosság elérése. Sir Frank Stenton 1922-ben jelentősen hozzájárult a probléma megoldásához.¹ Amikor Stenton azzal a helyzettel találta magát szemben, hogy az általa sajtó alá rendezett okleveleknek mindössze öt százalékán szerepelt vagy volt megállapítható nevek és események alapján a keletkezés dátuma, más módszert keresett, hogy az okleveleket az értelmezésükhöz elengedhetetlen dátumhoz kösse. Forrásainak vizsgálata során meggyőződött arról, hogy az oklevelek kronológiája elválaszthatatlan a formulák fejlődésétől. A rendelkezésére álló anyagok hiánya azonban csupán annyit tett számára lehetővé, hogy a 12. és a 13. századi oklevelek között a főbb formulabeli különbségeket megállapítsa. A Stenton utáni kutatók hasznosították útmutatását, de kevesen végeztek átfogó vizsgálataikhoz hasonlókat, inkább csak az ő nyomdokain folytak jelentős kutatások. Sok tudós pedig mind a mai napig egyszerűen teljesen figyelmen kívül hagyja ezt a kérdést.

W. Davies a Llandaff oklevelekről szóló tanulmányában jól fogalmazza meg a formulák szerepének jelentőségét a középkori angol birtokadományozó oklevelek datálásánál: „Szinte már közhely, hogy az ugyanazon helyről és időszakból származó hiteles oklevelek hasonló, vagy éppen ugyanazt a formulát használják a hasonló aktusok és szándékok kinyilvánítására.”² A kérdés azonban változatlan formában mindig felmerül, hogy vajon milyen mértékben érvényes más dokumentumokra az, amit a kutatók különbözőképpen 'változtatás' (*alteration*), 'szerkesztői kiigazítás' (*editorial contamination*), 'utánzat' (*fabrication*), 'javított változat' (*rebandling*), 'átdolgozás' (*reworking*) vagy egyszerűen 'hamisítvány' (*forgery*) szakkifejezésekkel illetnek. Kétség sem férhet ahhoz a tényhez, hogy ahogyan az oklevelek formulája időről időre és kiadási helytől függően változik a társadalmi körülményekkel és a jogi gyakorlattal szoros összefüggésben, úgy a korábbi oklevelek tartalmát a későbbi másolások során kiigazították az éppen aktuális viszonyoknak megfelelően.³

¹ F. M. Stenton (ed.): *Transcript of Charters relating to the Gilbertine Houses of Sixle, Ormsby, Catley, Bullington, and Alvingham*. Publication of the Lincoln Record Society for 1920. 18. (1922) Introduction (a továbbiakban: Stenton)

² Wendy Davies (ed.): *The Llandaff Charters*, Aberystwyth: The National Library of Wales. 1979. 7. (a továbbiakban: Llandaff)

³ B. R. Kemp (ed.): *British Library Manuscripts: Egerton 3031, Harley 1708 and Cotton Vespasian E XXV*. vol. 1: *General Documents and those relating to English Counties other than Berkshire*, Royal Historical Society, Camden Fourth Series, 31. (1986) 20. (a továbbiakban:

Másrészeiről azonban, szemben a nagyobb számban előforduló magánoklevelekkel, ezeknek a változtatásoknak a megjelenése kevésbé valószínű a pápai, királyi, illetve a püspöki oklevelekben, mivel ez utóbbiak kiváltságokat adományoznak. Továbbá statisztikai számítások segítségével kimutatták, hogy ezek a változtatások túlságosan ritkán fordulnak elő ahhoz, hogy az eredeti megfogalmazásokat felismerhetetlenné tegyék.⁴ Így aztán az okleveleket egészként kezelve a különböző módosítások azonosíthatók az oklevélben szereplő bizonyos szavak használata és a formulák alapján. A későbbi változtatások általában a helységnevek helyesírásának korhoz igazításából, az adományozások terminológiájának rövidítéséből, valamint a tanúnévsorok lerövidítéséből vagy éppen teljes elhagyásából tevődnek össze.⁵ Bár a rövidítések minden bizonnyal csökkentik a vizsgálható formulák számát, ám a változtatások egyike sem módosítja számottevően magukat a formulákat.

A dokumentumokban rögzített pontos dátumok, datálható események hiányában a történészek számos más, belső vagy külsőleges nyomra akadhatnak, amelyek segítséget nyújthatnak az oklevelek körülbelüli vagy hozzávetőleges (*circa*) keletkezési dátumának meghatározásában. A belső utalások alapján való datálás legelterjedtebb módszere az, amikor az egy dátummal nem rendelkező dokumentumban előforduló neveket más, dátummal rendelkező dokumentumok neveivel vetik össze.⁶ Nyilvánvaló, hogy minél több név, illetve lehetséges keletkezési időintervallum adódik a vizsgálat alapján, annál nagyobb a pontos datálás lehetősége. Előfordul azonban, hogy a korábban keletkezett források tanúnévsorát későbbi dokumentumok listáihoz, vagy más, olyan listákhoz csatolták,⁷ amelyekben a nevek jöllehet ugyanazok, de korábbi vagy későbbi generációkból származnak.⁸ A legnagyobb nehézséget azonban az okozza, hogy a tanúnévsorok hiányosak, vagy teljesen elhagyták őket sok *cartularium*ban, különösen a később összeállítottakban.⁹ Ilyen esetekben még a dokumentumok körülbelüli keletkezési dátumának meghatározása is meglehetősen nehéz.

Végül az eredeti dokumentumok és bizonyos mértékig a *cartulariumok* esetében¹⁰ alkalmazni lehet a kézírások összehasonlítását, de ennek a módszernek is megvan a maga korlátai. Stenton egyetértett azzal a véleménynel, hogy a kézírás vizsgálata értékes adalékokkal szolgálhat a datáláshoz, de hozzátette, hogy „ha a

Reading I.)

⁴ Llandaff 17., 91.

⁵ Una Rees (ed.): *The Cartulary of Haughmond Abbey*. Cardiff, Shropshire Archaeological Society and University of Wales Press, 1985. 4.; Reading I. 4., 7.

⁶ J. L. Fischer (ed.): *Cartularium Prioratus de Colne*, Essex Archaeological Society, Occasional Publications, 1, Colchester, 1946. IV.; Llandaff, 31., 33-34.; Reading I. 4.

⁷ U. Rees (ed.): *Cartulary of Shrewsbury Abbey*. 2 vols. Aberystwyth, 1975. XVII.

⁸ M. Gervers (ed.): *The Cartulary of the Knights of St. John of Jerusalem in England*, [pt. 2]: *Prima Camera*, Essex, *Records of Social and Economic History*, n. s. 23. London: Oxford University Press for the British Academy, 1996. jegyzetek nos. 79., 84., 94. (a továbbiakban: Gervers 1996.)

⁹ A. N. Webb (ed.): *An Edition of the Cartulary of Burscough Priory*, Chetham Society, 3rd series, 18 (1970) 14. (a továbbiakban: Borscough); Reading I. 4.

¹⁰ C. D. Ross (ed.): *The Cartulary of Cirencester Abbey*, Gloucestershire, i.: Oxford University Press, 1964. 221-22.; III., ed. M. Devine 1977. 812.

kézírás és a formulák vizsgálata egymásnak ellentmondó eredményeket produkál, akkor általában a formulák vizsgálatából származó eredményeknek kell hitelt adni.”¹¹

Amikor Stenton 1922-ben közzétette eredményeit, a magánoklevelek datálása a bennük szereplő nevek azonosítása révén alig volt célravezetőbb a formulák segítségével végzett kormeghatározásnál, mivel az oklevelekben szereplő tanú- és méltóságnévsorok viszonylag rövidek voltak. Azóta azonban történelmi dokumentumok tízezrei láttak napvilágot, ami a nevek segítségével való datálást egyre pontosabb módszerré teszi. Ennek eredményeképpen a formulák alapján történő kormeghatározást csak végső esetben használják, ha a nevek hiányoznak, vagy nem azonosíthatók. Sok esetben, amikor a *cartularium*ból hiányzik a tanúnévsor, a szerkesztők megelégedtek azzal, hogy csak a 'korai', 'a század közepe tájékáról', vagy a 'kései' jelzőt biggyesszék oda a 12. és 13. századi oklevelek mellé, vagy olyan hozzávetőleges időintervallumot adjanak meg a kiállítás lehetséges idejeként, amely felölel több mint egy fél évszázadot.¹² A kronológiailag ilyen pontatlanul elhelyezett történelmi források értéke azonban kutatási szempontból erősen megkérdőjelezhető.

A kutatók mindezidáig idegenkedtek a Stenton által kifejlesztett oklevélformula-vizsgálat datálási módszerként való alkalmazásától, és ez csak részben magyarázható azzal a ténnyel, hogy egyre nagyobb számban állnak rendelkezésünkre azonosítható nevek. Ezt az idegenkedést azzal a feltételezéssel is magyarázhatjuk, és ezt a nézetet maga Stenton is alátámasztja, hogy ha egyszer a formulák kialakultak, akkor változatlanul éltek tovább. Hatvan évvel Stenton után M. Clanchy 'rögzült frazeológiáról' (*set phraseology*) beszél, amely világosan kimutatható a birtokadományozó oklevelekben.¹³ Ezzel a közkeletű elképzeléssel szemben azonban a középkori oklevelek formulái és kifejezései soha nem voltak teljesen 'rögzültek'. Bár ezeknek a dokumentumoknak a tartalma egyre inkább formálissá vált, de a formulák magukban és egymáshoz képest is meglehetősen változatos képet mutatnak, és ez a kép változik, ahogyan haladunk előre az időben.¹⁴ És nemcsak a formulák változtak, hanem a nyelvezet is, amelyen íródtak. Ezek a változások gyakran tükrözték azokat a társadalmi, politikai, gazdasági körülményeket, amelyek közepette az adott dokumentum létrejött. Bizonyítható az is, hogy stílusbeli preferenciák kapcsolhatók egyes korokhoz, sőt bizonyos esetekben egyes írókhoz is. A középkori oklevelek nyelvezete, formája és tartalma állandó változásban volt, olyannyira, hogy közelebről megvizsgálva és hasonló dokumentumokkal összehasonlítva egy adott oklevél pontosan utalhat kiállításának időpontjára. Az oklevelek számítógépes feldolgozásával kimutatható, hogy minden kontextusba helyezhető szó és szókapcsolat rendelkezik egyedi, a keletkezésére utaló jellemzővel, egyfajta 'ujjlenyomattal'. Ezeknek az 'ujjlenyomatoknak' az összessége egy dokumentumon belül meghatározhatja a dokumentum keletkezési dátumát. Egy kísérleti adatbázisból (1. táblázat) származó szóelőfordulási ujjlenyomatok például mutatják a 2–4. táblázatok.

¹¹ Stenton XXXII.

¹² Burscough; Reading I.

¹³ M. T. Clanchy, *From Memory to Written Record, England, 1066–1307*. 2nd Edition, Blackwell, Oxford (UK) & Cambridge (Mass.), 1993. 327.

¹⁴ R. B. Patterson, (ed.): *Earldom in Gloucester Charters. The Charters and Scribes of the Earls and Countesses of Gloucester to A. D. 1217*, Oxford, Clarendon Press, 1973. 21., 30.

Kutatásaim és publikációim magját az elmúlt huszonöt évben a 12. és 13. századi angol oklevelek sajtó alá rendezése alkotta. A kiállítási évszámmal nem rendelkező oklevelek datálása a szerkesztés fontos momentumja. Az eddig datálásra használt módszerek a legidőigényesebb, és egyben a legpontatlanabb lépései a szerkesztői munkának. Amennyiben az általam kifejlesztett, az oklevelek datálását a változó szóelőfordulások számítógépes vizsgálatára alapozó projekt sikeresnek bizonyul, akkor lehetővé válik, hogy bármilyen oklevél kiállítási időpontját megbízható pontossággal határozzunk meg. A középkori Angliát tanulmányozó kutatók mindegyike hasznát látja majd a nagyobb pontosságnak, hiszen alapvetően okleveles forrásokkal dolgoznak. A módszer természetesen azok számára is hasznos lehet, akik saját kutatásaik során szeretné alkalmazni a nagyobb pontosság elérése érdekében.

A kutatás alapjául szolgáló elmélet abból az egyszerű tényből ered, hogy a nyelv időben változik, és hogy a nyelvben beálló változások, amelyek megjelennek az oklevelek szövegeiben, tetten érhetők, és az az időtartam pedig, amíg ezek jelen vannak a nyelvben, meghatározható. Az állandó változások természetes következménye, hogy azok elég gyakran szerepelnek a dokumentumok szövegében ahhoz, hogy mennyiségileg is mérhetőek legyenek, ami aztán levetíthető egy meghatározott kronológiára. Ha már az egy adatbázisban előforduló szókapcsolatok összességének datálása sikerült, akkor ezeknek az előfordulásoknak a datálatlan oklevelekkel történő összehasonlításával, a közös pontok, illetve egyezések alapján megállapítható ezek kiállításának valószínű dátuma vagy az az időintervallum, amelyben a datálatlan dokumentum nagy valószínűséggel keletkezett. Ha két dokumentum szóelőfordulásai teljesen megegyeznek, akkor a két dokumentumot ugyanazon helyen és azonos időben állították ki, vagyis elméletileg ugyanarról a dokumentumról van szó.

A módszer

A dátummal nem rendelkező oklevelek szóelőfordulások alapján történő datálása során három, jól elkülöníthető fázisról beszélhetünk. Az első ezek közül egy *ORACLE 7* adatbázis létrehozása pontosan datált, publikált 12. és 13. századi okleveles forrásokból származó szóelőfordulásokból. 4500 ilyen dokumentum tartalmának számítógépes azonosítását és feldolgozását tervezzük, amelyeket a stílusbeli azonosság kedvéért délkelet-angliai forrásokból választottunk ki.

A szóelőfordulások vagy szókapcsolatok olyan két-, vagy többszavas csoportokból állnak, amelyek többször előfordulnak az adatbázisban szereplő oklevelek szövegében. Minél hosszabb egy szókapcsolat, annál kevesebb előfordulásra és ennek megfelelően rövidebb feldolgozási időre lehet számítani, a szókapcsolatban szereplő szavak számával fordítottan arányosan. Hogy elkerüljük azt a hibát, hogy egy szókapcsolat előfordulását többször is megszámloljuk, azokat a szóelőfordulásokat, amelyek minden esetben azonos tartalommal és sorrendben beágyazhatók egy hosszabb szókapcsolatba, az adatbázis elkészítésekor nem vesszük külön figyelembe. Az egyszavas, önálló szintaktikai és kronológiai 'ujjlenyomattal' rendelkező előfordulások kezdetben nem kerülnek az adatbázisba. Szintén kizárjuk az olyan előfor-

dulásokat, amelyeknek az adatbázison belül nincsenek megfelelőik, mivel ezek nem rendelkeznek 'ujjlenyomattal'. Bár fölöttébb valószínű, hogy ezeknek az egyedi szókapcsolatoknak megtalálnánk az adatbázison kívüli előfordulását más, datálásra váró dokumentumokban, de az ezek által szolgáltatott bizonyíték elégtelennek tűnik ahhoz, hogy csak ezért minden egyes alkalommal megnövekedjék az adatbázis feldolgozási ideje a bővítés miatt. Másrészt pedig a megnövekedett adatbázis valószínűsíti ugyanazon szóelőfordulások többszöri figyelembevételét. Ezért az adatbázis tesztelésére kizárólag már datált, az adatbázisban magában nem szereplő dokumentumokat használunk.

A feldolgozás előtt minden személy- és helységnevet 'P' betűvel, minden mennyiségre utaló számot pedig a '#' szimbólummal helyettesítünk, hogy elkerüljük azoknak a szóelőfordulás-láncoknak a megszakítását, amelyekben ezek szerepelnek. Ahol az olvasat nem köthető következetes időbeni változáshoz, ott az olvasatot az adatbázisnak megfelelően egységesítjük.

A második lépés azoknak a szóelőfordulásoknak a kiválasztása az adatbázisban még nem szereplő dokumentumból, amelyek az adatbázison belül is előfordulnak. Az adatbázisból kiválasztott minden szóelőforduláshoz csatoljuk a származási dokumentum számát, dátumát, típusát és forrását. A beágyazott szóelőfordulásokat figyelmen kívül hagyjuk.

A harmadik, végső stádiumban, a datálatlan oklevelek keletkezési dátumának meghatározásakor a – második lépésben kigyűjtött – szóelőfordulásokat numerikusan (az előfordulás gyakorisága szerint) és kronológiailag (időbeni előfordulás szerint) csoportosítjuk, a szókapcsolatokban szereplő szavak száma szerint. Mivel azoknak a hosszabb szókapcsolatoknak az egybeesése kevésbé esetleges, amelyek mind az adatbázisban, mind a datálatlan dokumentumban előfordulnak, mint a hasonló, rövidebb szavaké, ezért ezek numerikus mutatóit a mértani közép súlyozásaiként alkalmazzuk a kronológiai mutatókra. További finomítás érhető el a szókapcsolatok fontossági sorrendjének kialakításával, a nyelvben előforduló gyakoriságuk alapján. Ebben az esetben, ha egy négyzavas előfordulással található egyezés, amely ritkán előforduló szót vagy szavakat tartalmaz, azt súlyozottabban kell figyelembe venni, mint azt a négyzavas előfordulást, amely gyakoribb szavakból áll. A szavak relatív gyakorisága abból az oklevél-adatbázisból határozható meg, amelyből magukat a szóelőfordulásokat kigyűjtöttük.

A súlyozás elvégzése után a kapott adatokat összegezzük. Annak érdekében, hogy az adatbázisban szereplő dokumentumok egyenletlen időbeli eloszlásából származó pontatlanságokat kiküszöböljük, minden öt éves periódus összegzett adatai külön táblázatba kerülnek. Az adott öt éves periódusból a legmagasabb pont, amely egy gyakori szókapcsolat első és utolsó előfordulásának dátumából számítható ki, jelzi a dokumentum legvalószínűbb kiállítási idejét. Az előzetes vizsgálatok azt mutatják, hogy a számítógép segítségével megállapított dátumok általában valamivel későbbiek, mint a valódi kiállítási időpontok, nyilván azért, mert a szóelőfordulások eltűnése az okiratokból lassabb, mint állandóságuk. A valódi és a számítógép által meghatározott dátumok átlagos eltéréseinek megállapítását célzó megfelelő tesztek elvégzése után a program módosításra kerül a pontatlanságok kiküszöbölése végett. A dátumok megállapításánál használt számeredményeket statisztikai program ellenőrzi majd,

amely meghatározza az eloszlási görbéket, elvégzi az elmélet tesztelését és megbízhatósági felmérést végez. Minden egyes megszerkesztett eloszlási görbe magassága és hosszúsága megkönnyíti majd a datálási eredmények pontosságának meghatározását.

A fentebb leírt módszer kipróbálása már a kísérleti stádiumnál tart, és az eredmények biztatóak. A programot kipróbáltuk a Jeruzsálemi Szent János Lovagrend 1175, Essex-ből származó oklevelén (BL Cotton ms Nero E VI).¹⁵ Mivel ezeknek az okleveleknek nagy része datálatlan volt, illetve csak hozzávetőleges (*circa*) kiállítási dátummal rendelkeztek, ezért a szóállományukból létrehozott szóelőfordulási kronológia nem kielégítően pontos datálatlan, 12. és 13. századi okiratok datálására. Mindazonáltal, a londoni Szent Pál Székesegyházból (STP)¹⁶ származó, korai, 333-as jelzetű oklevél keletkezési dátumát, amelyet egyébként 1239-re datálnak, a program 1245-re valószínűsítette (lásd az 5. táblázatot). Az eredeti keletkezési dátumtól való öt éven belüli eltérés elfogadhatónak tekinthető a projekt szempontjából. Mihelyt egy datált oklevelekből származó szóelőfordulás-adatbázis elkészül, legalább száz az adatbázison kívüli, datált dokumentummal teszteljük le, hogy egységes stílusú módszert fejlesszünk ki, amely eredményesen alkalmazható lesz minden oklevél esetében.

Nemrégiben egy ezzel összefüggő, ám meglehetősen eltérő módszert tett közzé F. D. Neiman és N. W. Alcock,¹⁷ középkori angol oklevelek számítógépes datálására. Módszerük, amely személyek történelmi forrásokban ismétlődő felbukkanásán alapszik, nem más, mint az összefüggések elemzésén keresztül elvégzett archeológiai sorrendbe állítás. Bár teljes mértékben tisztában vagyunk módszerük lehetőségeivel, mégis úgy véljük, hogy az sokkal bonyolultabb, és a vizsgálatba bevonható oklevelek száma is korlátozottabb, mint a DEEDS projekt által kifejlesztett módszer esetén. Ennek oka az, hogy a fennmaradt oklevelek nagy részének hiányzik a tanúnévsora, ahol az oklevelekben szereplő nevek többsége található, és ezért elsősorban a szóelőfordulások alapján történő datálást kell alkalmaznunk. Neiman és Alcock módszere azokban az esetekben alkalmazható, amikor dokumentumok nagy csoportjában gyakran fordulnak elő ugyanazok a nevek, míg az itt leírt datálási eljárás olyan esetekben is alkalmazható, amikor egyetlen, ismeretlen provenienciájú dokumentumról van szó.

Közvetlen célunk az, hogy egy olyan számítógépes módszert fejlesszünk ki, amely lehetővé teszi a University of Toronto DEEDS projektjében résztvevő kutatók számára, hogy azokat a 12. és 13. századi, angol birtokadományozó oklevelek és dokumentumok nagy részét, amelyek nem tartalmaznak utalásokat kiadási dátumukra, viszonylagos pontossággal datálhassák. Amennyiben más történészek is használni kívánják a szóelőfordulásból összeállított adatbázist hasonló dokumentumok datálására, mint ahogyan reményeink szerint lesznek ilyenek, úgy ezt

¹⁵ M. Gervers (ed.): *The Cartulary of the Knights of St. John of Jerusalem in England*, [pt. 1]: *Secunda Camera, Essex, Records of Social and Economic History*, n. s. 6, London: Oxford University Press for the British Academy, 1982.; Gervers 1996.

¹⁶ Marion Gibbs (ed.): *Early charters of the cathedral church of St. Paul*, London, Royal Historical Society, Camden Third Series, 1939. no. 333., 270-72.

¹⁷ F. D. Neiman & N. W. Alcock „Archeological Seriation by Correspondence Analysis: An Application to Historical Documents”, *History & Computing*, 7/1 (1995) 1-21.

