

Digitális Bölcsészet
2021., negyedik szám

<DIGITÁLIS BÖLCSÉSZET>



4 (2021)

Felelős szerkesztő:

Maróthy Szilvia

Szerkesztőség:

Kokas Károly, Parádi Andrea

Rovatvezetők:

Tanulmányok: Kiss Margit

Műhely: Péter Róbert

Kritika: Almási Zsolt

Labor: Maróthy Szilvia

Tanácsadó testület:

Bartók István, Fazekas István, Golden Dániel, Horváth Iván, Palkó Gábor, Pap Balázs, Sass Bálint, Seláf Levente

Korábbi munkatársaink:

Bartók Zsófia Ágnes (szerkesztő, rovatvezető), Fodor János (szerkesztő),

†Labádi Gergely (szerkesztő, rovatvezető), †Orlovsky Géza (tanácsadó testület)

ISSN 2630-9696

DOI: 10.31400/dh-hun.2021.4

Kiadja a Bakonyi Géza Alapítvány és az ELTE BTK Régi Magyar Irodalom Tanszéke (1088 Budapest, Múzeum krt. 4/A).

Felelős kiadó az ELTE BTK Régi Magyar Irodalom Tanszék vezetője.

Megjelenik az Open Journal Systems (OJS) v. 3. platformon, melynek működtetését az ELTE Egyetemi Könyvtár- és Levéltár biztosítja.



Ez a mű a Creative Commons *Nevezd meg! – Ne add el! – Így add tovább!* 2.5 Magyarország *Licenc* (<http://creativecommons.org/licenses/by-nc-sa/2.5/hu/>) feltételeinek megfelelően felhasználható.

Honlap: <http://ojs.elte.hu/digitalisbolcseszett>

Email cím: dbfolyoirat@gmail.com

Olvasószerkesztő: Bucsecs Katalin

Tördelés: Hegedüs Béla

Grafika: Hegyi Gábor

<TANULMÁNYOK>

Ted Underwood

University of Illinois Urbana-Champaign,

College of Liberal Arts & Sciences,

Department of English

tunder@illinois.edu

A műfajok életciklusai*

A kutatók még a „műfaj” jellegét illetően sem értenek egyet. Kérdéses például, hogy szöveges mivoltuk vagy társadalmi recepciójuk határozza meg inkább a műfajokat? Állandó kategóriáknak, évszázadokon átívelő történelmi konstrukcióknak vagy éppen rövid távú generációs jelenségeknek számítanak? A tanulmány véget nem érő vitákhoz kapcsolódik újszerű kvantitatív módszerekkel. Kimutatja, hogy a műfajok társadalmi és szövegalapú definíciói jól összehangolhatók: az olvasó ítélete alapján meghatározott műfajok a szöveg bizonyos jegyei alapján is felismerhetők. Erre az eredményre építve a tanulmány a műfajok szokásos élettartamával kapcsolatos ellentmondásokat vizsgálja. Kiderül, hogy a különféle műfaji modellek gyakran még akkor is összeegyeztethetők egymással, ha definíciójuk különböző korok olvasói ítélete alapján született. Így például azok a modellek, amelyek felismerik a tizenkilencedik századi „detektívtörténeteket” és „tudományos románcokat”, a huszadik századi „krimi” és „sci-fit” is felismerik. A bizonyítékok tehát arra utalnak, hogy – bár a műfajok bizonytalan történelmi konstrukciók – egy részük jóval maradandóbb, mint ahogy azt az irodalomtörténészek feltételezik. Ám ez nem minden esetben igaz. A gótikus irodalom például éppen olyan laza és változékony, mint amilyenek a kritikusok eddig is tekintették.

Kulcsszavak:

gépi tanulás, műfaj, irodalom, irodalomtörténet



A műfaj fogalma szinte egyidős az irodalomelmélettel, de az évszázadok óta húzódo viták sem tudtak túl sok egyetértést szülni a témában.¹ Ennek részben az az oka, hogy a műfaj egy adott szöveg életének különböző pontjain más és más dolognak látszik.

* William E. Underwood, „The Life Cycles of Genres,” *Journal of Cultural Analytics* 1, 1. sz. (2017), <https://www.doi.org/10.22148/16.005>.

¹ A szerző köszönetét fejezi ki a kanadai Social Sciences and Humanities Research Council által finanszírozott NovelTM projekt támogatásáért. A projektet szintén támogatta a University of Chicago Knowledge Lab és a University of Illinois, Urbana-Champaign Center for Advanced Study. A HathiTrust Digital Library feltárásában a HathiTrust Research Center nyújtott segítséget, és hasznos tanácsokat kaptam a Novel™ csapat minden tagjától és Jim Englishtől.

A retorika tudósai általában a kommunikatív cselekvés olyan mintázataira összpontosítanak, amelyekből a feljegyzések vagy a tragédiák születnek.² A szociológusokat gyakran inkább a befogadást szervező intézmények érdeklik.³ Az irodalomtudósok pedig hagyományosan maguknak a szövegeknek a mintaalkotási jellemzőivel vannak elfoglalva. A műfaj mindezen szempontjai természetesen kapcsolódnak egymáshoz. De a kapcsolatokat nem könnyű leírni.

A távoli olvasás látszólag tükrözi az irodalomtudósok mániáját, mely szerint a műfaj a szöveges műalkotások egyik alapvető jellemzője. De a kvantitatív módszerek valódi értéke abban rejlik, hogy segítségükkel a tudósok összehangolhatják a műfaj szöveges és társadalmi megközelítéseit. A jelen dolgozat egy ilyen típusú kísérleti kapcsolat felvázolására vállalkozik. Az alábbiakban a műfajt először a befogadás történetével kapcsolatos kérdésként közelítjük meg – összegyűjtve az egyes olvasók vagy intézmények által bizonyos történelmi pillanatokban csoportosított címek listáját. E címek mögé tekintve azonban magukat a szövegeket is megvizsgáljuk. A statisztikai modellezés mai gyakorlatának köszönhetően párbeszédet kezdeményezhetünk különböző szövegcsoportok között, így felmérhetjük például, hogy a gótikus* irodalom egymással versengő meghatározásai (amelyek különböző időpontokban születtek, és merőben eltérő könyvlistákat eredményeztek) valóban annyira kompatibilisek voltak-e egymással, mint azt egyes kritikusok állítják.

A történelmi összehasonlítás problémája azért is különösen sürgető, mert az irodalomtudósok mindeddig nem tudtak konszenzusra jutni a regényszerű műfajok életciklusával kapcsolatban. A gótikus irodalom kategóriája például egyes vélemények szerint 25, más vélemények szerint 250 évet ölel fel. *Graphs, Maps, Trees* című könyvében Franco Moretti átfogó képet nyújtott a műfajt érintő akadémiai stúdiumokról, és arra a következtetésre jutott, hogy a műfajoknál „igen gyakori az őrsváltás [...] ilyenkor fél tucat műfaj hirtelen eltűnik a színről, sok új műfaj bejön, majd ezek nagyjából huszonöt évig a helyükön maradnak.”⁴ A késő tizenharmadik századból származó gótikus regény 1825 körül átadja a stafétát (például) a Newgate-regénynek, ezt váltja az 1850-es évek végén a szenzációs regény, végül a tizenkilencedik század végére megérkezik a „birodalmi gótika” (pl. *Drakula*), amelyet Moretti diagramja a régebbi szerzeteses-bandítás gótikától már minden tekintetben különböző jelenségként kezel. Moretti összefüggést feltételez a sorozat huszonöt éves ritmusa és a generációk váltakozása között.

Ám bizonyos jól megalapozott recepciós hagyományok alapján a műfajok jóval hosszabb időtávon is képesek fenntartani egységes identitásukat, mint ahogy azt a

² Amy Devitt, *Writing Genres* (Carbondale: SIU Press, 2004); Carolyn Miller, „Genre as Social Action,” *Quarterly Journal of Speech* 70 (1984): 151–167, <https://doi.org/10.1080/00335638409383686>.

³ Pierre Bourdieu, *Distinction: A Social Critique of the Judgment of Taste* (Cambridge: Harvard University Press, 1984); Paul DiMaggio, „Classification in Art,” *American Sociological Review* 52 (1987): 440–455, <https://doi.org/10.2307/2095290>.

* Horace Walpole *Castle of Otranto [Otrantoi kastély]* című regénye óta a „gótikus” olyan regényműfajt jelöl, amelyben a rémület és a fantasztikus játszik nagy szerepet, és amely regények cselekménye szellemjárta, romos és félelmetes helyeken játszódik. Vö. David Mikics. *A New Handbook of Literary Terms* (New Haven–London: Yale University Press, 2007), 137. (A szerk.)

⁴ Franco Moretti, *Graphs, Maps, Trees* (London: Verso, 2005), 18.

generációnkénti ritmus megengedné. A kortárs krimi kedvelői gyakran egyforma örömmel olvassák Agatha Christie-t és Wilkie Collinst.⁵ És legalábbis a kritikusok szívesen láttatják úgy Stephen Kinget, mint aki a H. P. Lovecraft és Bram Stoker művein keresztül egészen *Az otrantói kastélyig* (*The Castle of Otranto*) visszanyúló folytonos gótikus hagyomány örököse.⁶

Tisztában vagyunk vele, hogy a fentiek mind a recepciót illető érvényes megállapítások. Morettinek igaza van, hogy a műfaj akadémiai vizsgálatai gyakran egy generációnyi időszakra terjednek ki. De az is igaz, hogy a „detektívtörténethez” hasonló kategóriák több, mint egy évszázada fontosak az olvasóknak. A szöveges elemzés egyik állítást sem fogja cáfolni, de arra talán rávilágíthat, hogy hogyan lehetnek összeegyeztethetők. Az áttekintések ellentmondásosságát feloldhatjuk például a kézenfekvő magyarázattal, hogy Morettinek a műfajok ritmusával kapcsolatban tett megállapításai az általa vizsgált századra helytállóak – de a huszadik századra már nem, mert a műfajok eddigre tartós marketinges intézményekké szilárdulnak. Moretti azonban utalt arra, hogy⁷ még egy olyan hosszú életű huszadik századi műfaj, mint a detektívtörténet is alapjában véve hasonló generációnyi szakaszok sorozatából állhat (a holmesi „ügy”, a gyanúsítottak zárt köre a vidéki házban, a bűnügyi thriller), melyeket csak e viszonylag gyenge leszármazási szál fűz össze.⁸ Előfordulhat, hogy egy ilyen sorozat eleje egyáltalán nem hasonlít a végére. Maurizio Ascari például amellett érvel, hogy „a *detektívtörténet* [detective fiction] kifejezést egyre inkább felváltja a *krimi* [crime fiction]”, és szkeptikusan kezeli azt a régi narratívát, mely „Poe-t alapító atyának tekintette, és az ő »okoskodó meséiben« látta a műfaj modelljének a kibontakozását.”⁹

Ha dönteni szeretnénk e megközelítések között, akkor lényegében arra a kérdésre keressük a választ, hogy a rövidebb életű „generációs” műfajok vajon koherensebbek-e, mint a hosszú életű műfajok. Ezen a ponton lehetnek segítségünkre a szöveges bizonyítékok. Matthew Jockers kimutatta, hogy a huszonöt-harminc éves skálán mozgó műfajok a tizenkilencedik században nyelvileg koherens jelenségek. Az úgynevezett ezüstvillás [silver-fork]^{*} vagy a szenzációs regény példáin keresztül statisztikai modell elfogadható pontossággal tud azonosítani további példákat ugyanarra a műfajra.¹⁰

⁵ P. D. James, *Talking About Detective Fiction* (New York: Vintage, 2011).

⁶ John Sears, *Stephen King's Gothic* (Chicago: University of Chicago Press, 2011), 51, 174.

⁷ Moretti, *Graphs, Maps, Trees*, 31.

⁸ A science fiction hasonló elméletéről lásd Adam Roberts, „A Brief Note on Moretti and Science Fiction,” in Jonathan Goodwin and John Holbo, eds., *Reading Graphs, Maps, Trees: Critical Responses to Franco Moretti*, 49–55 (Anderson: Parlor Press, 2011).

⁹ Maurizio Ascari, „The Dangers of Distant Reading: Reassessing Moretti's Approach to Literary Genres,” *Genre* 47, 1. sz. (2014): 1–19, 15, <https://doi.org/10.1215/00166928-2392348>.

^{*} Az „ezüst villás” [silver-fork] regények 1820–1830 között voltak népszerűek, a társadalom felső rétegének életéről szóltak (luxus, elegáns bálók). A regények az úgynevezett régens korban (1811–1820) játszódtak, azaz amikor Örült III. György helyett a fia, IV. György mint régensherceg uralkodott. Vö. David Mikics, *A New Handbook of Literary Terms* (New Haven–London: Yale University Press, 2007), 278. (A szerk.)

¹⁰ Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013), 67–81, <https://doi.org/10.5406/illinois/9780252037528.001.0001>.

Érdekes lenne kideríteni, hogy az olyan hosszabb kiterjedésű jelenségek, mint a detektívtörténet vagy a science fiction kisebb vagy nagyobb nyelvi koherenciát mutatnak-e, mint ezek a harmincéves kategóriák. Ösztönösen azt várnánk, hogy a hosszabb életű kategóriák határai elmosódnak; nehéz elhinni, hogy *A Holdgyémánt* (*The Moonstone*, 1868) és *A hosszú álom* (*The Big Sleep*, 1939) között szövegszinten túl sok hasonlóság lenne. Ha kiderül, hogy a „szenzációs regényeket” könnyebb felismerni, mint a „detektívtörténeteket”, akkor lesz némi bizonyítékunk arra, hogy Morettinek igaza volt a műfajok mögöttes generációnkénti logikájával kapcsolatban. Ilyesfajta bizonyítékokkal nem zárunk ki a hosszabb távú folytonosság lehetőségét: végül is nem tudjuk, hogy a könyveknek kell-e szövegesen is hasonlítaniuk egymásra ahhoz, hogy ugyanahhoz a műfajhoz tartozzanak. De ha azt látjuk, hogy a szöveges koherencia rövidebb időtávokon volt a legerősebb, abból annyi következtetést legalábbis levonhatunk, hogy az egy generációnyi műfajokban megtalálható egy bizonyos *fajtájú* koherencia, amely a hosszabb életű műfajokból hiányzik.

Másrészt legalább ugyanilyen ésszerű mértékben számíthatunk ezektől eltérő történelmi pályákra is. Például kifejezetten sok energiát szántak arra a kutatók, hogy leírják a műfaji konvenciók fokozatos szabványosulását a huszadik század elején, és a műfaj megszilárdulásának sarkalatos pontjaként emelték ki a műfajspecifikus ponyva megjelenését vagy az olyan kritikai kinyilatkoztatásokat, mint Ronald Knox a detektívtörténetek szabályait bemutató *Decalogue*-ja (*Tízparancsolat*, 1929).¹¹ Bizonyos műfajok esetében szerintük a folyamat még ennél is tovább tarthatott. Gary K. Wolfe azt állítja, hogy „a sci-fi regény” az 1940-es évek előtt „nem tudott egy műfajjára összeállni”.¹² Ha ez az irodalomtörténeti rekonstrukció helyes, akkor nem egymástól eltérő, egymást követő generációs fázisokra számíthatunk, hanem a határok folyamatos keményedésére, amelyek a huszadik század közepén már sokkal világosabban elkülönülnek, mint a század elején. Erre a történetre számítottam, amikor belevágtam a jelen projektbe.

E kérdések megválaszolásához a recepció tizennyolc különböző helyszínén gyűjtöttem össze a műfajhoz rendelt címek listáját. A listák egy része a közelmúlt kritikai álláspontjára épít, más listákat írók és szerkesztők jelöltek ki a huszadik század elején, míg egyes listák nagy számú könyvtári katalóguskészítő gyakorlatát tükrözik (lásd az *A függelék*). Bár minden egyes lista némileg eltérően határozza meg a tárgyát, nagyjából három kategóriába sorolhatók: detektívtörténet [detective fiction] (vagy „rejtélyes történet” [mystery] vagy „krimi” [crime fiction]), (a szintén többféleképpen definiált) sci-fi [science fiction] és a gótikus [Gothic]. (Vitatható, hogy a „gótikus” műfajnak számít-e egyáltalán – de éppen emiatt érdekes eset.) A Chicago Text Lab és a HathiTrust Digital Library gyűjteményére támaszkodva én magam is kigyűjtöttem az ezekben a kategóriákba tartozó szövegeket.¹³ Ha összehasonlítjuk a különböző recepció helyszínekkel és az idővonal különböző szegmenseivel társított szövegcsoporto-

¹¹ Ronald Knox, „Introduction’ to *The Best Detective Stories of 1928–29* (1929)”, reprinted in Howard Haycraft, ed., *Murder for Pleasure: The Life and Times of the Detective Story*, Revised Edition (New York: Biblio and Tannen, 1976).

¹² Gary K. Wolfe, *Evaporating Genres: Essays on Fantastic Literature* (Middletown: Wesleyan University Press, 2011), 21.

¹³ Ha óvatlanok vagyunk, a szövegek normalizálási folyamata elrejtheti előlünk a buktatókat. Megpróbáltam például eltávolítani a címnegyedeket, az előszókat és az élőfejeket, amelyek egyébként tartalmazhatnak olyan explicit műfaji címkéket, amelyekre a modell kíváncsi lenne. Az is aggo-

kat, meg tudjuk állapítani, hogy a különböző kategóriák pontosan mennyire voltak stabilak.

Az ebből a kísérletből körvonalazódó történet az imént vázolt alternatív elképzelések – a generációk váltakozása vagy a műfaji konvenciók fokozatos megszilárdulása – egyikéhez sem tudott szépen igazodni. Kevés bizonyítékot látok a Moretti elmélete által előre jelzett generációnkénti hullámokra. Sőt, még csak arról sincs szó, hogy az időrendileg egy koncentrált műfajba (mint például a „szenzációs regény, 1860–1880”) tartozó könyvek szükségszerűen jobban hasonlítanak egymásra, mint a hosszú idővonalat lefedő könyvek. A detektívtörténet és a sci-fi legalább akkora szöveges koherenciát mutat, mint a Moretti-féle rövidebb életű műfajok, és ezt a koherenciát nagyon hosszú ideig (160 vagy akár 200 évig is) fenn tudják tartani. Tehát, azt hiszem, félretehjük azt a (termékeny) feltételezést, mely szerint a huszonöt éves generációs ciklusok különösebb jelentőséggel bírnának a műfajok tanulmányozása szempontjából.

De várakozásaim ellenére nem találtam sok bizonyítékot a fokozatos megszilárdulást feltételező elméletre sem. Bár az kétségtelenül igaz, hogy a műfajt irányító kiadói intézmények fokozatosan fejlődtek, úgy tűnik, tévedtem, amikor arra számítottam, hogy a műfajok közötti szöveges különbségek is fokozatosan alakulnak ki. A detektívtörténetek esetében például azok a szöveges eltérések, melyek a huszadik századi leleplezős történeteket más műfajoktól megkülönböztetik, nagyon világosan visszavezethetők *A Morgue utcai kettős gyilkosságig* – de sokkal tovább nem. A detektívtörténet abban az értelemben fokozatosan terjedt, hogy Poe és Vidocq kezdetben még elszigetelt, másodrangú, utánpótlás nélküli alakok voltak, műfajspecifikus magazinokról és könyvklubokról pedig végképp nem lehetett még beszélni. De a szöveges mintázatoknak nem kell olyan fokozatosan fejlődniük, mint az intézményeknek. Poe történeteiben már számos olyan jellemző megtalálható, amelyek a huszadik századi krimi elkülönítik a többi műfajtól.

Prediktív modellezés

A számítógép tulajdonképpen azért lép színre ebben a dolgozatban, hogy a segítségével próbáljuk meg kibogozni a kortárs műfajelméletek szülte kusza problémákat. Ha azonosítani tudnánk egyetlen olyan formai alapelvet, amely alapján egyszer s mindenkorra meg lehetne határozni a műfajok mibenlétét, azzal jelentősen megkönnyíthetnénk a kritikusi feladatunkat. Elég lenne annyit mondanunk, hogy sci-fi az, amit Darko Suvin „a kognitív elidegenítés irodalmaként”¹⁴ definiál. Az olvasók között azonban ritka az egyetértés egy adott műfaj meghatározó jegyeit illetően, míg a különböző közösségek ugyanazon művekkel kapcsolatban gyakran más-más dolgot értékelnek. A teoretikusok gyanúja szerint egyre valószínűbb, hogy a műfaj valójában „családi hasonlóságon” alapul, amelyet egy sor egymással átfedésben álló

dalomra adhat okot, hogy a szövegek két különböző könyvtárból (HathiTrust és Chicago Text Lab) származnak.

¹⁴ Darko Suvin, „On the Poetics of the Science Fiction Genre,” *College English* 34, 3. sz. (1972): 372–382, 372, <https://doi.org/10.2307/375141>.

funkció teremt meg.¹⁵ Ráadásul a műfajok történelmi alkotások: fontos jellemzőik változhatnak.¹⁶

Összefoglalva: egyre inkább úgy tűnik, hogy a műfaj nem egyetlen, megfigyelhető és leírható tárgy. Talán inkább különböző módon kapcsolódó művek közötti összefüggések változékony rendszereként kell elképzelnünk, amelyek különböző mértékben hasonlítanak egymásra. Az ehhez hasonló problémák olyan módszertant igényelnek, amely óvatos az ontológiai feltételezésekkel, a részleteket illetően pedig türelmes. A prediktív modellezés éppen ilyen módszer. Leo Breiman szerint a prediktív modellek leginkább abban térnek el a szokványos statisztikai módszerektől (és, hozzátenném, a hagyományos kritikai eljárásoktól), hogy előzetes elvárások nélkül próbálják azonosítani a vizsgált jelenséget ténylegesen okozó és magyarázó mögöttes tényezőket.¹⁷ A műfaj esetében ez azt jelenti, hogy a műfaj meghatározása helyett egy olyan modell azonosítására törekszünk, amely képes reprodukálni az egyes valós megfigyelők ítéleteit. A méretre vonatkozó jelzők (például „hatalmas” [huge], „gigantikus” [gigantic], de az „apró” [tiny] is) például a legmegbízhatóbb olyan szöveges kulcsok közé tartoznak, amelyek jelzik, hogy egy könyvet sci-fi-nek fognak nevezni. Kevesen definiálnák a scifit a méreteken való merengésként, de kiderült, hogy a (bizonyos forrásokban) sci-fi-ként minősített művek valójában feltűnően sokat tárgyalják ezt a témát. Adjunk hozzá még néhány száz szót, és talán kapunk egy olyan statisztikai modellt, amely azonosítani tudja az ugyanezekben a forrásokban „science fiction” névvel illetett műveket, akkor is, ha a műfaj alapját képező definíció továbbra is nehezen fogalmazható meg (vagy talán soha nem is létezett).

Újabban Hoyt Long és Richard Jean So hasonlóképpen prediktív modellekkel igyekeztek felderíteni a haiku stílus „látens, nem explicit nyomait” az angol költészetben.¹⁸ Az ilyen projektekben a gépi tanulás célja nem is annyira az, hogy megnöveljük az általunk figyelembe vett könyvek számát, hanem hogy észlelhessük és összehasonlíthassuk az elmosódott családi hasonlóságokat, amelyek redukció nélkül nehezen lennének verbálisan definiálhatók. Sarkosabban fogalmazva: a számítógépes módszerek hasznossá teszik a kortárs műfajelméletet. Megszabadulhatunk a berögzült meghatározásoktól, és a műfaj tanulmányozását ehelyett csak bizonyos történelmi szereplők változó gyakorlatára alapozhatjuk – de még mindig olyan műfaji modelleket kapunk, amelyek értelmes összehasonlításokat és szembeállításokat tesznek lehetővé.

Mivel a prediktív modellben az egyes változókhoz nincs ok-okozati erő rendelve, a tulajdonságok kiválasztása nem elsődleges fontosságú. Az alábbi kifejtésben ugyan használok szavakat kulcsfogalomként, de nem szeretném azt sugallni, hogy a műfaj nyelvi jelenség. Mert nem az. A műfaj egy tág értelemben vett társadalmi jelenség, és a szavak történetesen kényelmes prediktív kulcsként működnek; segítségükkel nyo-

¹⁵ Paul Kincaid, „On the Origins of Genre,” *Extrapolation* 44, 4 sz. (2003): 409–419, 413–414, <https://doi.org/10.3828/extr.2003.44.4.04>.

¹⁶ John Rieder, „On Defining SF, or Not: Genre Theory, SF, and History,” *Science Fiction Studies* 37, 2. sz. (July 2010): 190–209, 193.

¹⁷ Leo Breiman, „Statistical Modeling: The Two Cultures,” *Statistical Science* 16, 3. sz. (2001): 199–231, <https://doi.org/10.1214/ss/1009213726>.

¹⁸ Hoyt Long and Richard Jean So, „Literary Pattern Recognition: Modernism between Close Reading and Machine Learning,” *Critical Inquiry* 42, 2. sz. (2016): 235–267, 266, <https://doi.org/10.1086/684353>.

mon követhetjük a különféle kiválasztási gyakorlatok közötti implicit hasonlóságokat és különbségeket. Tetszés szerint akár a szöveg más jellemzőit is használhatnánk. Egyes kutatók írásjeleket vagy karakterhálózatokat használtak a műfaj előrejelzésére; amikor egy 850000 kötetből álló gyűjteményben próbáltam műfajokat azonosítani, az oldalformátumhoz kapcsolódó jellemzőkben gondolkodtam.¹⁹ De annak a projektnek az volt a célja, hogy maximalizáljuk a modell teljes prediktív pontosságát, mivel ismeretlen vizekre vetettük ki a hálónkat, és egyszerűen csak a lehető legnagyobb fogást akartuk elérni.

A jelen projekt célja más. Címkézett példákkal dolgozom, nem pedig címkézetlenekeket próbálok elkapni. Ha az itt leírt modellek mindegyikét javítani lehetne 1%-kal, az semmit sem változtatna az érvelésen. Ami igazán számít, az a különböző szövegcsoporthoz tartozó határok *relatív* erőssége. Ezért fektettem kevés energiát a pontosság optimalizálására; inkább a minél jobb olvashatóságra és a következetességre törekedtem. Az itt leírt modellek mind ugyanazt a tulajdonsághalmazt használják, amelyet úgy kaptunk meg, hogy egyszerűen vettük a teljes gyűjtemény 10000 leggyakoribb szavát (dokumentumgyakoriság szerint), minden műfajban és a meghatározott műfaj nélküli művekben is. (Vehettük volna a leggyakoribb 5000 vagy 3000 szót is, a pontosság így körülbelül 1%-kal változna.) A tanulási algoritmushoz L2-szabályos logisztikus regressziót használok, egy jól ismert algoritmust, mely viszonylag egyszerű becsléseket ad a jellemzők fontosságáról.²⁰

A pusztán pontosság iránt tanúsított gavalléros hozzáállásunknak persze megvannak a maga határai. Ha a statisztikai modellek egyáltalán nem tudnák megjósolni a műfajt, nyilvánvalóan nem szolgálnának hasznos bizonyítékokkal. Ebbe a problémába azonban itt nem fogunk beleütközni. Az itt tárgyalt modellek 70–93%-os pontossággal tudnak előrejelzéseket tenni, inkább e tartomány felső vége felé csoportosulva. És bár most egy olyan műfajt fogunk jellemezni, amelyet mindössze 76%-os pontossággal, „viszonylag laza” csoportosulásként jósolnak meg a modellek, és ezt állítjuk szembe egy 91%-os biztonsággal felismerhető műfajjal, az igazság az, hogy ezek a számok mind nagyon jelentős hasonlóságokra utalnak egy szövegcsoporthoz tartozó műveken belül: jóval a szokásos társadalmi-tudományos hatásnagysági küszöbértékek felett vagyunk.²¹

Bár minden műfajnál ugyanazt a módszert alkalmaztam, nem tudom előre garantálni, hogy a módszer minden műfajnál egyformán működőképes. Ha egy műfaj különösen nehéz a szókincs alapján „megfogni”, akkor nehézkes lehet a szózsák [bag-of-words] modell segítségével osztályozni. A sci-fi például különleges problémákat vet fel, mivel a tengeralattjárók egy idő után már nem számítanak futurisztikus elképzeléseknek, és a holnap újabb csodái váltják fel őket. A gyakorlatban ritkán találkozunk ilyen problémákkal. A lexikális modelleknek nem okoz nehézséget megtalálni a tematikusan eltérő műveket összekapcsoló közös formális elemeket. Általában a

¹⁹ [üres lábjegyzet az eredetiben]

²⁰ Fabian Pedregosa et al., „Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research* 12 (2011): 2825–2830; Az előállított képek alapja: Hadley Wickham, *ggplot2: Elegant Graphics for Data Analysis* (New York: Springer, 2009), <https://doi.org/10.1007/978-0-387-98141-3>.

²¹ Ha a prediktív pontosságot Cohen-féle „d” értékre alakítjuk, 65% felett már minden a „nagy hatás” kategóriába tartozik.

szövegek közötti olyan hasonlóságokat mutatnak ki, amelyek szorosán követik a kritikai intuíciónkat. De el is térhetnek a kritikusok elvárásaitól (amelyek viszont nem egységesek). A pozitív eltérések könnyen értelmezhetők: a lexikális modell által fedezett folytonosságok mindjárt kizárják azt a tételt, hogy két recepció helyszínnek nincsenek közös vonásai. A negatív eltérésekkel kapcsolatban lassabban fog kialakulni bizonyosság, mivel nem lehet *eleve* kizárni a modellt elkerülő hasonlóságok létezését. Abban, hogy a lazábbnak tekintett csoportosulások szöveges szinten valóban kevésbé koherensek, csak akkor kezdhetünk el bízni, ha módszereink már számos különböző műfaj koherenciáját felismerték.

Ha egy tanulási algoritmusnak elegendő változót adunk, akkor az lényegében „memorizálni” tud egy adathalmazt, és így valószerűtlenül pontos előrejelzéseket készít a korábban már látott példákra. Tehát egy számos változóval operáló modellt valójában csak visszatartott példákkal lehet tesztelni. A jelen dolgozatban használt modelleket folyamatosan visszatartott szerzők keresztellenőrzésével értékeltük ki.²² Vagyis megmutatjuk a modellnek az egy halmazban szereplő összes szerzőt (egy kivételével), majd a modellt a nem látott szerző művein teszteljük. A folyamatot addig ismétljük, amíg az összes szerzőt lefedtük, és ki tudjuk számítani az egész halmazra vonatkozó pontosságot.

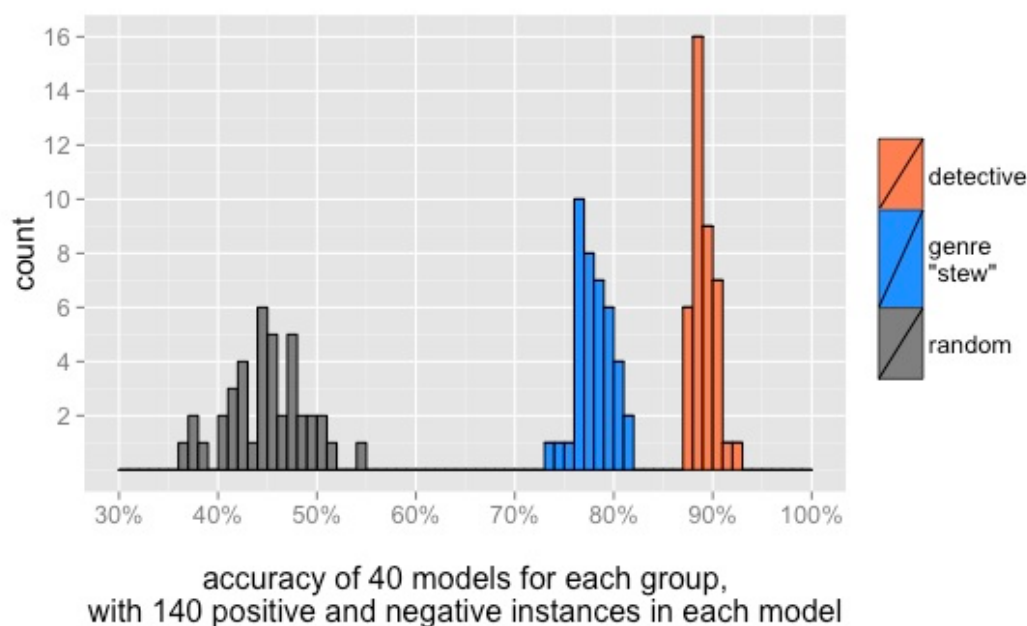
A műfajok háromnál több dimenziós térben találhatók

A „műfaj” szó alapvetően egy térképszerű mentális képet idéz fel, ahol a táj kisebb parcellákra van felosztva, és minden irodalmi mű szigorúan csak egy parcellában található meg. Ám a fiktív szövegek kategorizálásának a gyakorlata egyáltalán nem ilyen térképet eredményez, és én sem törekszem egy ilyen térkép előállítására. A *fehér ruhás nőhöz* (*The Woman in White*, 1859) hasonló regényeket egyes megfigyelők a gótikus irodalomhoz, mások a szenzációs regényhez sorolják. Az én metaadataim között mindkét besoroláshoz kapcsolódó címkék szerepelnek. Más regények egyetlen határozott műfajhoz sem kapcsolódnak: valójában a tizenkilencedik századi művek többsége soha nem kapott konkrét besorolást. A fikciós művek tehát tartozhatnak sok műfajhoz vagy egyetlen egyhez sem. Ahelyett, hogy megpróbálnék felfedezni egyetlen olyan sémát, amely szerint ez az egész tér szerveződik, külön összehasonlítások sorát futtatom, és mindig olyan műveket csoportosítok, amelyek egy bizonyos műfaji besorolást tükröző címkét kaptak, és ezt a halmazt hasonlítom össze egy azonos méretű kontraszthalmazzal. A kontraszthalmaz általában véletlenszerűen lett kiválasztva valamely digitális könyvtárból (azt leszámítva, hogy milyen mértékben zár ki a pozitív halmazban szereplő címkéket), és olyan időbeli eloszlás szerint, mely a lehető legszorosabban egyezik a pozitív halmaz eloszlásával.

Ha nincs érdemi különbség a két összehasonlított kategória között, akkor a modell várakozásaink szerint a véletlenszerű találgatásnál (50%-os pontosság) nem sokkal pontosabb előrejelzéseket készít. A hitelesség ellenőrzése végett célszerűnek látszik e teszt lefuttatásával kezdeni. Véletlenszerűen két „csapatba” soroltam az ebben a cikk-

²² Ennek kifejtését lásd: D. Sculley and Bradely M. Pasanek, „Meaning and Mining: The Impact of Implicit Assumptions in Data-Mining for the Humanities,” *Literary and Linguistic Computing* 23, 4. sz. (2008): 409–424, <https://doi.org/10.1093/llc/fqn019>.

ben felhasznált összes szerzőt, mindkét „csapatból” véletlenszerűen kiválasztottam 140 kötetet, majd megpróbáltam betanítani a modellt a két csoport megkülönböztetésére. Az 1. ábra negyven kísérlet eredményeit ábrázolja (szürkével). Ahogy látható, az átlagos pontosság csak egy kicsivel alacsonyabb (45%), mintha véletlenszerűen tippeltünk volna. Amikor egy osztályozási algoritmus próbálja megtalálni a különbségeket két véletlenszerűen kiválasztott csoport között, képes ugyan még felfedezni egy halvány mintát, de az esetleges minta nem fog értelmes kapcsolatban állni a visszatartott példákkal, és egy ilyen haszontalan szabály könnyen olyan elfogultságot adhat az előrejelzéseknek, amely a találgatásnál is rosszabb. Így tehát elenyésző a veszélye annak, hogy az algoritmus látszólagos különbséget azonosít két olyan csoport között, amelyeknél ilyen különbségek nem állnak fenn.



1. ábra. 3 vélelmezett „műfaj” 40 modelljének pontosságát ábrázoló hisztogram. Minden modellhez 140 pozitív eset lett véletlenszerűen kiválasztva egy hosszabb listáról. A „detektívtörténet” lista a következő szakaszban kifejtett módszerekkel lett kialakítva.

De van egy másik fajta alapeszt is, amelyet le kell futtatnunk. Mi történik, ha az általunk tanulmányozott műfajok bármelyikével felcímkézett valamennyi műből egyetlen szörnyű ragut készítünk, és ezt a szuperhalmazt hasonlítjuk össze az összes véletlenszerűen kiválasztott, semmilyen műfaji címkével nem társított művel? Ahogy az 1. ábrán látható, a modell gyakran felismeri a „műfaji raguból” érkező köteteket, jóllehet a gótikus, a Newgate, a szenzációs, a detektív és a sci-fi ezen kombinációja valószínűleg egyáltalán nem képez egy szokványosan koherens műfajnak nevezett csoportot. A modell előrejelzései átlagosan 78%-ban helyesek.²³ Ha rápillantunk né-

²³ Mivel a jelen dolgozatban mindig két egyenlő méretű kategóriát fogunk összehasonlítani, nem fogom megkülönböztetni a precizitást a felidézéstől.

hány olyan szóra, melyek az ebben a műfaj-szuperhalmazban megfigyelhető tagságra utalnak, könnyen megérthető, mi is történik. Az ide tartozó egyes műfajok nem teljesen különböznek egymástól. Közös bennük a szenzációs tárgy és néhány olyan kellék vagy cselekményszöveési eszköz, amelyek nagyobb valószínűséggel fordulnak elő bennük, mint egy véletlenszerűen kiválasztott műben. A prediktív szavak listájának tetején áll például a „gyilkosság” [murder], a „szörnyű” [ghastly], a „zár” [lock], a „kulcs [key]”, az „elmélet” [theory] és a „laboratórium” [laboratory]. Nehezebben azonosíthatók a véletlenszerű kontraszthalmaz jellegzetes szavai, de (összehasonlítás-képpen legalábbis) általában a hétköznapi családi élet különböző vonatkozásait idézik fel („házas” [married], „hibáztat” [blame], „reggelek” [mornings], „büszke” [proud], „barátok” [friends], „délutánok” [afternoons]). A kép talán megváltozna, ha olyan műfajokat tanulmányoznánk, mint a fejlődésregény. De az ehhez a cikkhez összeállított adatkészletben nagy különbségek figyelhetők meg a műfaji fikció, mint olyan, illetve a véletlenszerűen kiválasztott, viszonylag köznapi háttér között.

Ha jobban meg akarnánk érteni ezt a különbséget, nem lenne elég egy tizezernél több jellemzőt tartalmazó lista alját és tetejét megnézni. Kétségtelenül ki lehetne alakítani egy olyan szemantikai keretet, amely támogatná és pontosítaná a „szenzációs” és a „köznapi” témát érintő, rendszertelen következtetéseimet. Ám ez helyet és időt venne igénybe, és ennek a dolgozatnak nem az az elsődleges célja, hogy a tárgyalt műfajok mindegyikének a tartalmáról új leírást kínáljon. Így tehát, bár a jellemzők teljes listája elérhető egy online kód- és adatmellékletben,²⁴ a leírásaim rövidek lesznek, és mindegyik modellnél csak néhány olyan prediktív szóra térek ki, amelyek általános benyomást nyújtanak a tárgyalt szembeállításról.

A műfajok újradefiniálása helyett ebben az esszében alapvetően a műfajok változó élettartamával és szöveges koherenciájuk fokozataival kapcsolatban hozok fel érveket. E célból nem az a fontos, hogyan jellemezzük a nyelvezetet, inkább az, hogy hogyan értelmezzük a szövegcsoportok közötti hasonlóság fokát. Különösen fontos megérteni, hogy a műfajok hasonlóságának a tere háromnál több dimenzióra terjedhet ki. Az egyik tengelyen egymáshoz közel álló elemek más szempontból nagyon messze lehetnek egymástól. Láttuk például, hogy ha összekeverjük a detektívtörténetet, a sci-fit és más műfajokat, akkor az esetek 78%-ában megkülönböztethetők a véletlenszerűen kiválasztott művektől. Csábító lenne ebből arra következtetni, hogy a detektívtörténet és a sci-fi „alapvetően folyamatos” vagy „78%-ban hasonló”; akár el is kezdhethetnénk felépíteni egy narratívát, amely Sherlock Holmes kémiai iskolázottságát a két műfaj közötti döntő, eddig figyelmen kívül hagyott kapcsolatnak tekintené. Valójában azonban a különbségek e műfajok között még erősebbek, mint a véletlenszerűen kiválasztott háttérrel szemben fennálló közös különbözőségük. Ha betanítunk egy modellt, hogy megkülönböztesse a detektívtörténetet a sci-fitől vagy a gótikustól, akkor az esetek több mint 93%-ában igaza lesz (mindkét esetben).

A prediktív modellek kicsit olyanok, mint az emberek: mindig találnak olyan módszereket, amelyek szerint a két műcsoport hasonló, de olyanokat is, amelyek szerint különböznek. Ha tudni akarjuk, hogy a detektívtörténet és a sci-fi értelmesen egy kalap alá vehető-e, az egyszerű kétoldali összehasonlítás soha nem fog válaszolni a

²⁴ Underwood, „Code and Data to Support The Life Cycles of Genres,” 12 May 2016, <https://github.com/tedunderwood/fiction>, <https://doi.org/10.22148/16.005>.

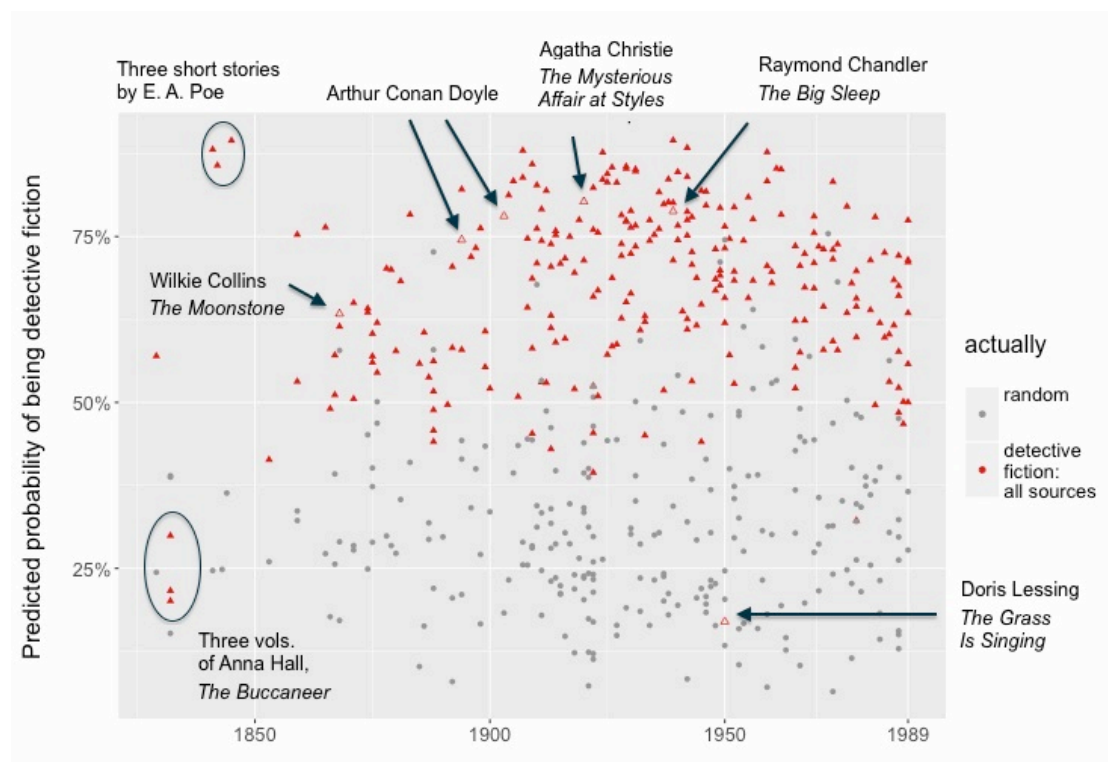
kérdésre. Érdemesebb lehet több összehasonlítás relatív erejére figyelni. Az 1. ábrán például azt látjuk, hogy a detektívtörténet lényegesen könnyebb megkülönböztetni a véletlenszerű kontraszthalmaztól, mint a „műfaji ragunktól”, ami már arra utalhat, hogy ez egy szorosabb kötésű kategória. Alternatív megoldásként feltehetnénk egy háromoldalú kérdést, amelynek segítségével két műfajt ugyanabban a koordináta-rendszerben vizsgálhatunk. Például *ugyanúgy* különböznek-e a detektívtörténetek a többi fikciós műtől, mint a sci-fi? A prediktív modellek jól tudják egy bizonyítékhalmoz alapján a másikat kikövetkeztetni. Így betaníthatjuk a modellünket a detektívtörténet és egy véletlenszerűen kiválasztott háttér közötti megkülönböztetésre, majd kérhetjük ugyanazt a modellt a sci-fi műfajú művek és ugyanazon háttér megkülönböztetésére. Ahogy az várható, a modell teljesen megbukik: az esetek kevesebb, mint a felében lesz igaza. Bár a két műfajban néhány dolog közös (például az elméletek és a laboratóriumok), a *legtöbb* olyan jellemzőjük, amely megkülönbözteti őket a háttértől, eltérő. Amikor el kell döntenem, hogy két műfaji modell hasonló-e, az alábbiakban ezt a tesztet fogom a legmegbízhatóbbnak tekinteni. Annyit biztosan meg tud mondani, hogy a szörnyű „műfaji ragu” elkülöníthető detektívtörténetre és sci-fire. A következő kérdés az, hogy maga a „detektívtörténet” feloszlik-e további alműfajokra vagy bizonyos recepciók helyszínek köré csoportosuló művekre, amelyek jobban különböznek egymástól, mint amennyire hasonlítanak.

A detektívtörténet

A huszadik század második feléből származó műfaji listák összeállítása során először a Kongresszusi Könyvtár (Library of Congress, Washington) műfaj/forma címsoraira támaszkodtam. A köteteket különböző könyvtárosok látták el ezekkel a címkékkel, és sok különböző ember műfajokkal kapcsolatos hallgatólagoz feltevéseit tükrözik. A detektívtörténet esetében több különböző címsort egy kategóriába soroltam, többek között a „Rejtélyes történet” [Mystery fiction], a „Detektívtörténet és rejtélyes történet” [Detective and mystery fiction] és a „Nyomozók” [Detectives] tárgyi címsort (ha többnyire fikciós művekre alkalmazták őket). Ahogy megyünk vissza 1940 előttre, a címkék egyre szórványosabbá válnak; olyan művekkel találkozunk, amelyeket eredetileg akkor katalogizáltak, amikor a Kongresszusi Könyvtár még nem a jelenlegi formájában használta a rendszerét. E művek közül csak néhányat katalogizáltak újra a modern, bőbeszédű módszerrel. Így tehát a korábbi időszak tekintetében leginkább bibliográfiákra és kritikai tanulmányokra kell támaszkodnunk.

Több óriási detektívtörténeti bibliográfia is létezik, igazi kihívás olyat találni, amely elég kicsi az átíráshoz. A háború előtti detektívtörténettel kapcsolatban leginkább az Indiana University Library által 1973-ban szervezett „A krimi első száz éve 1841–1941” kiállítás katalógusára támaszkodtam. A kiállítás regények mellett novellásköteteket (és néhány önálló novellát) is katalogizál. A lefordított művekre is kiterjed. Én is hasonlóan nyitott leszek a jelen esszé során. A Jules Verne kaliberű írók Franciaországon kívül is elemi mértékben alakították a műfajt, így a fordítások kizárásával sokat veszítenénk. Sőt, kétlem, hogy az itt vizsgált mintázatok bármilyen lefordíthatatlan problémát támasztanak: mint látni fogjuk, Verne még fordításban is a sci-fi kivételesen tipikus képviselője marad.

Előfordulhat persze, hogy a detektívtörténet egyetlen (1941 előtti kötetekre korlátozó) kiállítási katalógusa olyan képet alakít ki a műfajról, amely lényegileg különbözik a sokféle ember által katalogizált, 1829–1989 közötti kötetek által tükrözöttektől. De a statisztikai modellezés éppen az ilyen típusú kérdéseket teszi vizsgálhatóvá. Ha csak az indianai kiállítás 88 kötetét (amelyeket meg tudtam szerezni digitálisan) modellezzük, akkor igen magas, 90,9%-os pontosságot érhetünk el. A Kongresszusi Könyvtár műfaji címkéivel ellátott 177 kötet már vegyesebb képet mutat, és csak az esetek 87,3%-ában ismerhetők fel. Ha egyesítjük a két halmazt, 249 kötetet kapunk (mivel 16 mindkét halmazban szerepelt), amelyeket 91,0%-os biztonsággal tudunk felismerni. Így a különböző módokon kiválasztott csoportok keverése nem csökkenti a pontosságot; ez egy „szintlépést” biztosító kompromisszum.



2. ábra. A „detektívtörténet” halmazból való származás várható valószínűsége; 91,0%-os átfogó pontosság.

De, mint már említettem, az algoritmikus modellek hatékonyan találják meg a közös elemeket a művek egy-egy csoportjában. A két kategória (A és B) közötti hasonlóság valódi próbája az, ha megkérünk egy A és C közötti szembeállításra betanított modellt arra, hogy B-t is különböztesse meg C-től. Ha például azt kérjük a Kongresszusi Könyvtár detektívtörténet kategóriáján betanított modelltől, hogy különböztesse meg az indianai kiállítás anyagát egy hasonló véletlenszerű háttértől, azt még mindig 89,3%-os pontossággal tudná megtenni. Ez valóban igazolja, hogy a detektívtörténet esetében nagyrészt egybevágó definíciókkal van dolgunk.

Az általunk használt modell valószínűségi jellege miatt láthatóvá válik, hogy a detektívtörténet képviselői közül kik a különösen jellegzetesek vagy a különösen

nehezen osztályozhatók. Kivetíthetjük a köteteket az y tengelyre, amely azt jelzi, hogy a modell milyen bizonyossággal tudja a „detektívtörténet” halmazba sorolni őket. A 2. ábrán ezt készítettem el mind a 249, vagy egyes könyvtárosok által felcímkézett, vagy az indianiai kiállításon szerepeltetett kötettel.

Az egyik nagyon szembetűnő dolog Edgar Allan Poe három, az 1840-es évek elejéről származó történetének a pozíciója. E művek nemcsak a saját idejükben, hanem a teljes 1829–1989 közötti idővonalat meghatározó normák szerint is a műfaj példaértékű modelljeinek látszanak. Bár az, hogy Poe tartósan sablon tud maradni, összhangban áll a detektívtörténetek egyik elterjedt származástörténetével,²⁵ a dologról a kritikusok között nincs egyetértés. Moretti például azt állítja, hogy a detektívtörténet csak 1890 körül érte el ma ismert formáját,²⁶ Ascari pedig nyíltan tagadja, hogy Poe még mindig fontos lenne a krimi szempontjából.²⁷ Ilyesfajta folytonosságra egy statisztikai modellben egyáltalán nem számítottam. Sőt, inkább azt vártam, hogy a detektívtörténet határai egyre jobban elmosódnak, ahogy visszamegyünk Conan Doyle elé, abba az időszakba, amikor a leleplező történetek még gyakran összeolvadtak más műfajokkal, például a szenzációs regénnyel. Talán valamivel több elmosódás figyelhető meg az 1860–1870-es években, mint a huszadik század közepén. (És az 1830-as években néhány kirívó hiba is megfigyelhető: azok a katalóguskészítők, akik megpróbálták odáig nyújtani a „detektívtörténet” kategóriát, hogy egy 1832-es regény is beférjen, végső soron tönkreteszik a koncepciót.) De a kritikusok által általában prototípusként azonosított korai művek (*A Morgue utcai kettős gyilkosság*, *A Holdgyémánt*) továbbra is példaértékűek ebben a modellben. Ez a bizonyíték nem feltétlenül jelöl ki egy „eredetet”, és azt sem bizonyítja, hogy bizonyos írók határozták volna meg a műfajt. Csupán annyit bizonyít, hogy a huszadik század végén a detektívtörténet prototípusait azonosító kritikusok helyesen jártak el, amikor a saját koruk gyakorlatát visszavetítették a múltba.

De a bizonyíték azt is mutathatja, hogy a detektívtörténet folytonossága több, mint egy sor genealógiai kapcsolat a különböző formák között. Tegyük fel például, hogy összevonjuk az indianai kiállítást és a Kongresszusi Könyvtár címkéit egyetlen, 249 szövegből álló csoportba, de az 1930-as évnél időrend szerint kettéosztjuk a csoportot. Milyen mértékben változik a detektívtörténet definíciója a két időszak között? Ha az 1930 utáni 130 kötetet modellezzük, 88,5%-os pontosságot kapunk. De ha a modellt az 1930-ig megjelent 119 kötet alapján tanítjuk be, majd e modell alapján készítünk az 1930 utáni művekre vonatkozó előrejelzéseket, akkor a modell pontossága még mindig 86,9%-os. Az 1930-ig a detektívtörténetet jelölő verbális különbségek nagyrészt a későbbiekben is jellemzőek maradnak. Ez nem jelenti azt, hogy a műfaj nem változott. A kemény [hardboiled] detektív fő divatja például 1930-ban még sehol sincs, ez kétségtelenül fontos változás. De a *detektívtörténet* és az irodalmi mező többi része közötti különbség természete már nem változott drámai mértékben. A műfaj határai stabilak. (Érdemes megjegyezni, hogy ez annak ellenére igaz, hogy az 1930 előtti

²⁵ Stephen Rachman, „Poe and the Origins of Detective Fiction,” in Catherine Ross Nickerson, ed., *The Cambridge Companion to American Crime Fiction* (New York: Cambridge University Press, 2010), 17–28, <https://doi.org/10.1017/CCOL9780521199377.003>.

²⁶ Moretti, *Graphs, Maps, Trees*, 31.

²⁷ Ascari, „The Dangers of Distant Reading,” 15.

véletlenszerű kontrasztanyag javarészt a HathiTrust könyvtárból, az 1930 utáni anyag viszont többnyire a Chicago Text Lab könyvtárából származik. A gyűjteményeket eltérő módon választották ki, de a különbségek nem elég nagyok ahhoz, hogy belezavarjanak a műfaji jelbe.)

A modellünk bizonyítékokat hoz arra, hogy a folytonosság elég erős ahhoz, hogy valódi problémát jelentsen a jelenleg uralkodó műfajelméleti nominalizmus számára. Azt az (érvényes) feltevést, hogy a műfajok nem feltétlen foghatók össze egy világos definíció vagy már létező esszencia mentén, gyakran egy-két lépéssel még tovább viszik, és azt sugallják, hogy a műfajokat csak egy genealógiai szál kapcsolja össze – hogy a múlt és a jövő pusztán a „különböző gyakorlatokat folytató közösségek” közötti folyamatos manőverezésként kapcsolódik össze.²⁸ Maurizio Ascari bírálja Morettit, amiért „pozitivistá” módon közelíti meg a detektívtörténetet, és emlékeztet rá, hogy „a huszadik század folyamán a detektívtörténet alapjaiban megváltozott”, sőt, „minden irodalmi műfaj szüntelenül változik”.²⁹ Mark Bould és Sherryl Vint (2009) szerint „bár gyakran úgy érzékeljük őket, a műfajok soha nem a világban már eleve meglévő dolgok, amelyeket később a műfaj szakértői tanulmányoznak, hanem különböző állítások és gyakorlatok szövetéből álló, képlékeny, testetlen szerkezetek” – vagy még merészebben fogalmazva, „nincs olyan dolog, hogy sci-fi.”³⁰ Amikor elkezdtem ezt a tanulmányt, a hasonló állítások egy részével talán magam is azonosulni tudtam; a műfaj stabilitásával kapcsolatos szkepticizmus összességében előnyösebbnek tűnt, mint a definíciókkal kapcsolatos parttalan vita. A prediktív modellek azonban lehetővé teszik a közéletet. Kezdetünk óvatos lépésekkel, meghatározott történelmi szereplők által kijelölt vélelmezett határokkal, majd empirikusan rákérdezhetünk, hogy hallgatóságos kiválasztási kritériumaik mennyire egyeznek vagy térnek el egymástól. A detektívtörténetek esetében a különböző emberek által különböző időpontokban gondozott szövegliszták igen nagy mértékben összeegyeztethetők. Sőt, a műfaj múltjának modellje kiválóan előre tudja jelezni a jövőjét.

De pontosan mi *volt* az a „detektívtörténet”-definíció, amely a háttérben működött? A válasz javarészt egyáltalán nem sokkoló. A „rendőrség” [police], a „gyilkosság” [murder], a „nyomozás” [investigation] és a „bűntény” [crime] szabja meg a műfaj tematikai alapvetését. A „gyanú” [suspicion], a „bizonyíték” [evidence], a „bizonyítani” [prove], az „elmélet” [theory], a „véletlen” [coincidence] (és, hogy egy finomabb példával is szolgáljunk, a „bárki” [whoever]) szavak alapozzák meg a cselekményt hajtó, kétségre és bizonyításra építő szerkezetet. Ha egy kicsit mélyebbre ásunk a modellben, kevésbé nyilvánvaló részletekre bukkanhatunk. Így például fontos kulcsfogalmak jönnek az építészet és a lakberendezés területéről: az „ajtó” [door], a „szoba” [room], az „ablak” [window], az „íróasztal” [desk] mind nagy prediktív erővel bíró szavak. A skála másik végén pedig a gyermekkor és az oktatást leíró szavak szerepelnek („született” [born], „nőtt” [grew], „tanított” [taught], „gyermekek” [children], „tanárok” [teacher]), melyek nagy bizonyossággal jelzik, hogy az adott kötet nem detektívtörténet. Talán

²⁸ Rieder, „On Defining SF, or Not,” 15.

²⁹ Ascari, „The Dangers of Distant Reading,” 7, 15.

³⁰ Mark Bould and Sherryl Vint, „There is No Such Thing as Science Fiction,” in James Gunn, Marleen Barr, and Matthew Candelaria, eds., *Reading Science Fiction* (New York: Palgrave, 2009), 43–51, 48, 43.

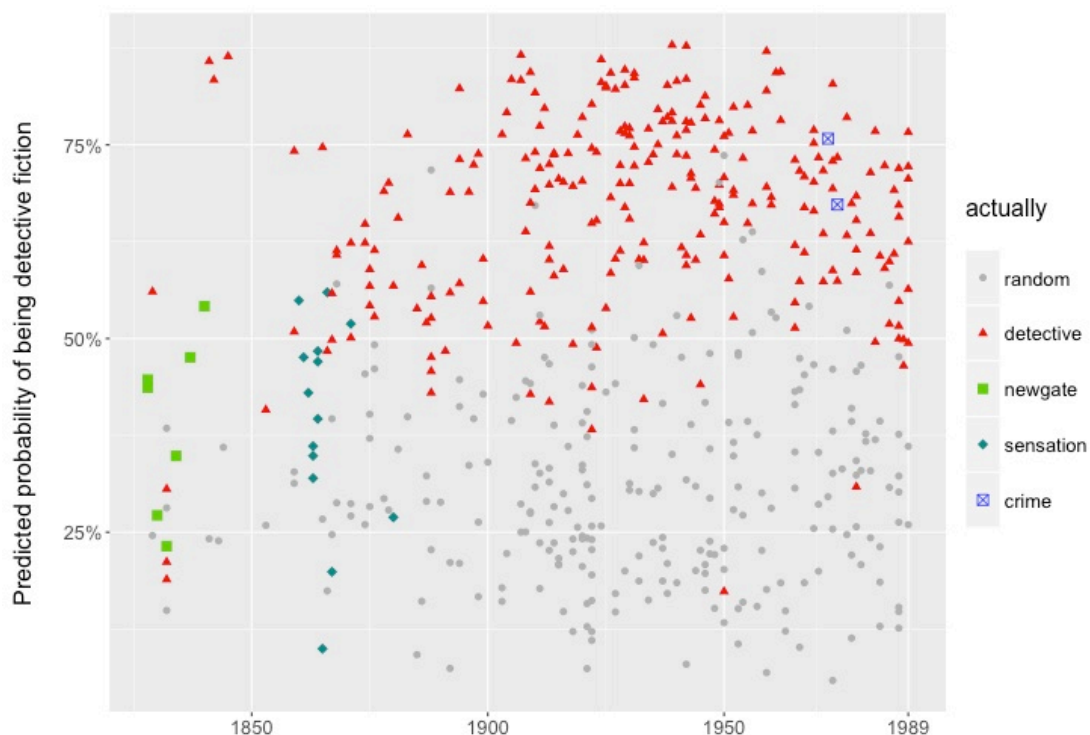
amiatt, hogy a műfaj jobbára egy bizonyos titokzatos eseményre koncentrálnak (vagy mert hajlamos novella formájában megjelenni), az életrajzi keretek viszonylag szűkösek.

A detektívtörténeetről tehát megállapítottuk, hogy 160 éven keresztül (1829–1989) szövegesen koherens maradt. De addig nem teszteltünk le alaposan egy modellt, amíg nem találjuk meg, hol omlik össze. Mi történik például, ha a nyomozó alakja helyett a bűnözői miliót állítjuk előtérbe? A „detektívtörténet” és a „krimi” közötti határ nagyon képlékeny, és Patricia Highsmith regényei, illetve az úriember tolvaj Arthur J. Raffles alakja szinte meg is kérdőjelezi e határok létjogosultságát. Mi történik, ha felveszünk néhány „krimi” címkével igen, de „detektívtörténet” címkével el nem látott huszadik századi regényt, valamint egy csoport 1820–1840 közötti Newgate-regényt? Elvégre a detektívtörténetek egyik klasszikus feldolgozása a *Twist Olivérral* kezdődik.³¹ A szenzációsregényeket is azonosították már a detektívtörténetek előfutáraként;³² *A Holdgyémánt* az indianai kiállításon is szerepelt, de megpróbálhatnánk további szenzációs regényeket hozzáadni, hogy lássuk, azok is beleférnek-e ebbe a kategóriába.

Bizonyos helyeken a modellünk könnyen magába fogad más kategóriákat, más helyeken viszont egyáltalán nem rugalmas. A 3. ábrán egy olyan modell látható, amelyet a 2. ábránál felhasznált 249 detektívtörténettel tanítottam be, de azt is engedélyeztem neki, hogy a betanításhoz használt halmazon kívüli más művekről is készítsen előrejelzéseket. A „krimi” címkével ellátott újabb regényekről (például Patricia Highsmith, *A Dog's Ransom [Silány váltságdíj]*) kiderült, hogy kifejezetten jól összeegyeztethetők a detektívtörténeti modellünkkel. A tizenkilencedik századi szenzációs regények és a Newgate-regények azonban nem férnek bele ugyanabba a szövegdobozba. Mindez nem azt bizonyítja, hogy D. A. Miller tévesen tárgyalta együtt a Newgate-regényeket és a detektívtörténeteket a *The Novel and the Police* című könyvében. Végző soron semmilyen törvény nem írja elő, hogy az irodalomtörténeti fogalmaknak a nyelvezet szintjén is felismerhetőnek kell lenniük. Ha úgy szeretnénk definiálni a krimi nevű műfajt, mint amely magába foglalja a Newgate-regényt, bátran megtehetjük, és a fogalom megvilágító erejű lehet. De így már egy kissé más típusú műfajról fogunk beszélni, és ebből a műfajból hiányzik a nyelvi homogenitásnak az a foka, amely Edgar Allan Poe-t Agatha Christie-vel és Patricia Highsmith-szel egyaránt összeköti. A vizsgálat lényege más szóval nem az, hogy eldöntse, mit lehet vagy mit nem lehet „műfajnak” nevezni, hanem hogy segítsen megkülönböztetni azokat a különböző mintákat, amelyek leírására az irodalomtörténészek használni szokták a szót.

³¹ D. A. Miller, *The Novel and the Police* (Berkeley: University of California Press, 1988).

³² Christopher Pittard, „From Sensation to the Strand,” in Charles J. Rzepka and Lee Horsley, eds., *A Companion to Crime Fiction* (Oxford: Wiley-Blackwell, 2010), 105–116, <https://doi.org/10.1002/9781444317916.ch7>.



3. ábra. Egy csak a detektívtörténetekkel és véletlenszerű példákkal betanított modell előrejelzései három másik kategóriában.

A gótikus irodalom

A gótikus fikció eredete, stílszerűen, kései leszármazottaikat már pusztán halovány emlékek formájában kísértő ősalakok rejtélyes történeteként körvonalazódik. A kritikusok látszólag igencsak biztosak benne, hogy a gótikus regény egy 1760-tól talán 1830-ig tartó, kifejezetten koherens jelenség volt. De ahogy haladunk tovább a tizenkilencedik századba, egyre kevésbé egyértelmű, hogy a gótikus továbbra is folyamatos hagyomány marad-e. Mint már említettem, Franco Moretti a tizenkilencedik századi gótikus irodalmat a század elején és végén uralkodó két műfajra osztja fel. A huszadik századi Amerikában a „déli gótikus” irodalmat gyakran különálló irodalmi jelenségként kezelik. Erős érvek szólnak ezen túl egy kifejezetten női gótikus hagyomány létezésére mellett, amely DuMaurier *Rebecca* és Brontë *Jane Eyre* című művéig vezethető vissza.³³ Egyes kritikai hagyományok viszont ragaszkodnak a felsorolt dolgok közötti folytonossághoz, és a gótikus irodalom kategóriáját addig tágítják, hogy az már a kortárs „horror” kiadói kategóriát is magába foglalja.³⁴ Kaphatók olyan, a gótikus irodalmat gyűjtő antológiák, melyek az *Otranto* és Anne Rice közötti teljes távolságot lefedik, így a recepció gyakorlati oldalát illetően kell lennie *valamilyen* folytonosság-

³³ Juliann Fleenor, ed., *The Female Gothic* (Montreal: Eden, 1983).

³⁴ Mark Edmundson, *Nightmare on Main Street: Angels, Sadomasochism, and the Culture of Gothic* (Cambridge: Harvard University Press, 1999).

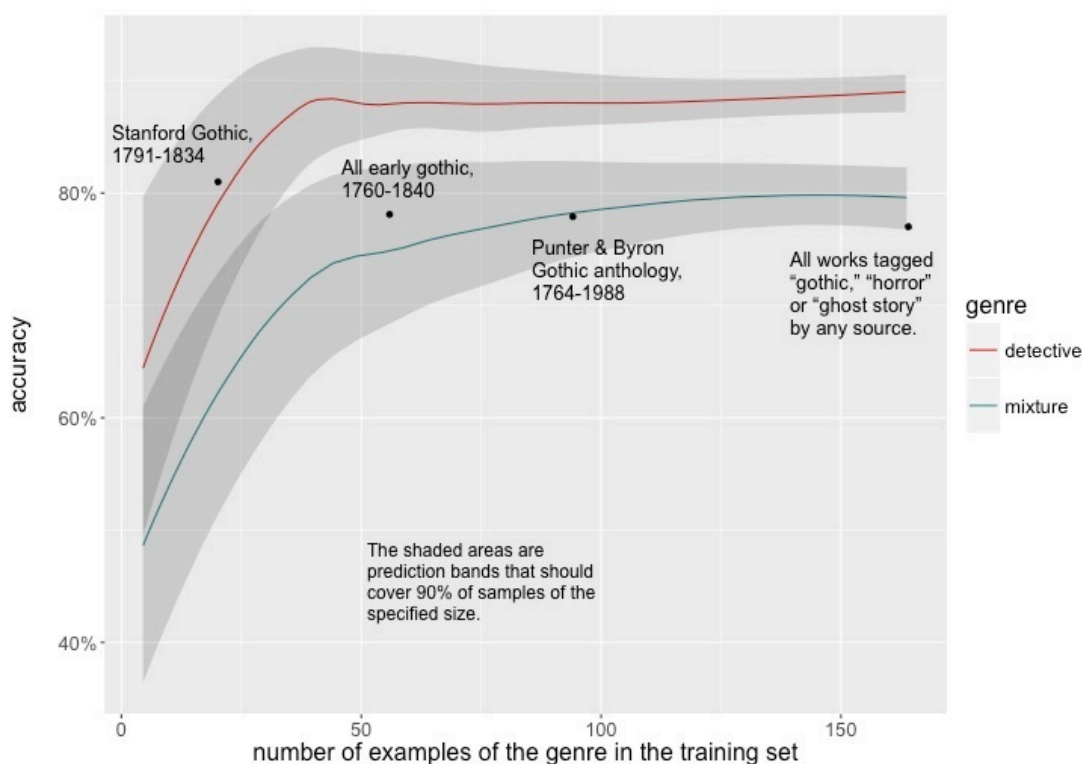
nak, nevezzük azt műfajnak, módnak, rajongói (szub)kultúrának (fandomnak) vagy témák laza sorozatának.

Mivel megalapozottan kételkedhetünk abban, hogy a „gótikus” irodalom valóban egyetlen, erősen egységet mutató hagyománynak tekinthető-e, ebben az esetben különösen fontos volt, hogy össze tudjuk hasonlítani a különböző forrásokból érkező bizonyítékokat. Az 1840 előtti időszak tekintetében nagy mértékben támaszkodtam a Stanford Literary Lab gótikus irodalomról készített listájára;³⁵ az 1940 utáni időszak kapcsán pedig növekvő mértékben tudtam a Kongresszusi Könyvtár „horrorhoz” vagy „kísértethistóriához” kapcsolódó műfaji címkéire építeni. Magam is számos művet kigyűjtöttem a *The Gothic* című Blackwell-kézikönyvben (szerkesztette David Punter és Glennis Byron, 2004) szereplő említések alapján, amely Walpole-tól egészen Brett Easton Ellisig próbálja meg feltérképezni a gótikus hagyományt, és őket egy sor közvetítő, köztük Charlotte Brontë, Henry James és H. P. Lovecraft alakján keresztül köti össze.

Egyik lista sem mutat a detektívtörténet esetében észlelt mértékű koherenciát. A legkisebb méretű mintát sikerült a legpontosabban előre jelezni: a Stanford Literary Lab által gótikusként azonosított 21 művet (1791 és 1834 között) 81,0%-os pontossággal sikerült felismerni. Az összes listát magába foglaló szuperhalmazt volt a legnehezebb modellezni: a 165 kötetet csak 77,0%-os pontossággal tudtuk felismertetni. A 81% és 77% közötti számszerű eltérés itt kicsit drámaibb, mint első hallásra gondolnánk, mivel itt nem almát hasonlítunk össze almával. A pontosság várakozásaink szerint a modellezett készlet méretével párhuzamosan nő. E növekedést érzékeltetendő különböző görbéken ábrázolom a más műfajok esetében különböző mintaméreteknél megfigyelt átlagos pontosságot.

A gótikus regények stanfordi részhalmaza szövegesen éppolyan koherens, mint a detektívtörténetek hasonló méretű mintája. A gótikus fikció évszázadokon átívelő mintái ellenben méretükhöz képest kirívóan gyengén teljesítenek; nem könnyebb modellezni őket, mint a projektben vizsgált minden műfaj közös halmazát. Ez arra utal, hogy a romantika korának gótikus regényei nyelvileg nem sok közös vonást mutatnak a huszadik századi horror és természetfeletti fikció hagyományával. Ezt úgy erősíthetjük meg, ha az egyik ilyen időszak alapján tanítunk be egy modellt, majd a másik időszakra alkalmazzuk azt; ez alig teljesít jobban, mint a véletlenszerű halmaz.

³⁵ A „stanfordi gótika” esetében használt metaadatokat a Stanford Literary Lab fejlesztette ki; és alighanem sok ember, köztük Ryan Heuser és Matthew L. Jockers munkájáról van szó.



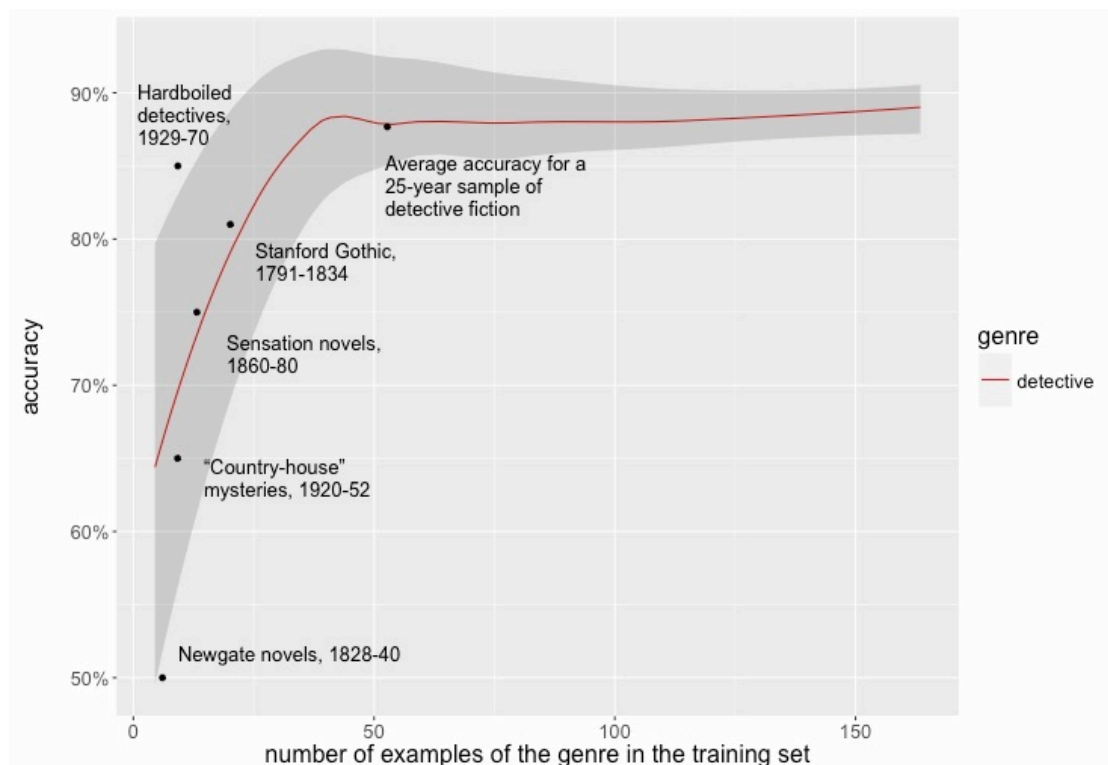
4. ábra. A gótikus fikció különböző mintái más műfajok különböző mintaméretek esetében mutatott pontossági tartományához viszonyítva.

Nem hinném, hogy ez túlzottan meglepi az olvasókat. Nagyon kevés kritikus állítja azt, hogy a gótikus olyan szorosan kötött műfaj, mint a szenzációsregény vagy a detektívtörténet. Punter és Byron (2004) még egy két évszázados antológia építése során is elismeri annak a lehetőségét, hogy „nagyon kevés tényleges irodalmi szöveg van, amelyek »gótikusak«; vagyis a gótikus talán inkább bizonyos elszórtan észlelhető pillanatokon, trópusokon, ismétlődő motívumokon keresztül jelentkezik a nyugati irodalmi hagyományon belül”.³⁶ A gótikus fikció két évszázados skálán történő modellezésének a nehézségei részben megerősítik ezt a gyanút, de talán módszerünk hitelességi ellenőrzésének is tekinthetők. (Ha egy módszer nem ismeri fel a nehéz eseteket, megbízhatósága más esetekben is megkérdőjelezhető.) Ez arra is emlékeztethet bennünket, hogy bizonyos folytonosságokat egy nyelvezeten alapuló modell nem tud észlelni. A kritikusok között egyetértés figyelhető meg a gótikus által jelölt dolgok spektrumát illetően, és azt, ha nem is éppen műfajnak, de legalábbis egy módnak vagy laza tematikus hasonlóságnak tekintik. Modellünk azonban ezt a spektrumot nem látja egységesebbnek, mint a népszerű műfajú művek véletlenszerűen kiválasztott halmazát.

Az egyik nyilvánvaló magyarázat erre az lehetne, hogy a gótikus fogalma az idővonal túl nagy részére terjed ki, és a nyelv egyszerűen túl sokat változik két évszázad alatt ahhoz, hogy a műfajt ilyen távlatban nyelvileg modellezni lehessen. De van olyan példánk is, ahol ez mintha nem jelentene problémát – a 4. ábra például arra emlékeztet

³⁶ David Punter and Glennis Byron, *The Gothic* (Malden: Blackwell, 2004), xviii.

bennünket, hogy a detektívtörténet 1829-től 1989-ig egészen jól egyben marad. És valóban, ahogy azt az 5. ábra mutatja, nem igazán van bizonyítékunk arra, hogy a kronológiailag koncentrált műfajok nyelvi értelemben általában koherensebbek, mint az a mintánk, amely a detektívtörténetek 160 évét fedi le.



5. ábra. Több, körülbelül generációnyi méretű műfaj helye, a detektívtörténetek 1829 és 1989 közötti véletlenszerű mintáján belül megfigyelt pontossági tartományt ábrázoló görbéhez képest. A sötétebb sáv egy, a detektívtörténetek véletlenszerű mintája alapján készített modell 90%-át lefedő előrejelzési sáv.

Ha a detektívtörténetek mintáját egy, a gótikussal, a Newgate-regénnyel vagy a szenzációstörténetekkel megegyező méretű betanítási halmazra csökkentjük, akkor a detektívtörténet éppolyan koherensnek tűnik, mint e kronológiailag koncentrált műfajok, annak ellenére, hogy a köteteket több, mint 150 éves időszakból válogattuk ki. Ez a döntő bizonyíték Franco Moretti azon feltételezése ellen, hogy a műfajok élettartama egy generációra terjedne ki. Úgy tűnik, hogy az egy generációnál jóval több ideig élő műfajokat olyan szöveges hasonlóságok tartják egybe, melyek éppen olyan erősek, mint a rövidebb életű műfajokat egybetartó hasonlóságok. Mivel nehéz bizonyítani, hogy ezek a modellek a szövegek közötti összes lehetséges hasonlóságot rögzítik, fogalmazzunk óvatosan: ha van is valamilyen generációs ritmus a műfajok történetében, ez a módszer nem észlelte azt – pedig úgy tűnik, az összes olyan mintát fel tudja ismerni, melyet a tudósok műfajnak szoktak nevezni.

Vajon még koherensebbé tehetnénk a detektívtörténetet, ha egy szűkebb generációs időtávra koncentrálnánk? Nem igazán. Ha a detektívtörténet hosszú ívét huszonöt éves szakaszokra bontjuk, majd azokat külön-külön modellezzük, olyan átlagos pon-

tosságot kapunk, amely nagyon közel áll az egész idővonalról vett hasonló méretű minta pontosságához. Legalább egy, a kritikusok által is elismert alműfaj esetében valamivel magasabb pontosság érhető el: ez nem más, mint a „kemény krimi”. Még egy viszonylag kisebb, tíz darab 1929–1970 közötti kemény krimiből álló mintát is 85%-os pontossággal lehet kiemelni a halmazból. E regények stilisztikai homogeneitása talán társadalmi homogeneitásukkal áll összefüggésben. Dorothy Hughes az egyetlen nő a kemény krimiket tartalmazó századközepi mintánkban. A kemény krimi és a vidéki ház-hagyomány* közötti különbség azonban egyáltalán nem mutat generációs megosztottságot; inkább a nemek és az Atlanti-óceán áll a háttérben. Ám ez sem egy leküzdhetetlen szakadék. A vidéki ház-hagyományon betanított modell éppolyan pontosan (85%) azonosítja a kemény nyomozókat, mint a magukkal a kemény krimikkel betanított modell. Kevés okunk van tehát arra következtetni, hogy ezek az alműfajok függetlenedtek volna a „detektívtörténet” átfogó fogalmától.

Sci-fi

Eddig amellet érveltem, hogy az egy évszázadnál tovább fennálló műfaji fogalmak nyelvileg éppolyan koherensek lehetnek, mint a néhány évtizedig fennálló műfajok. Eddig azonban a detektívtörténet az egyetlen példám, és okkal gyaníthatjuk, hogy a detektívtörténet/rejtélyes történet/krimi műfaj szokatlanul stabil premisszákra épül. Mindig adott a bűntény, mindig adott a nyomozó, mindig adott a nyomozás. A sci-fi viszont mintha nagyobb kihívást jelentene, hiszen e műfaj premisszái eredendően változékonyak. A műfaj Verne-féle prototípusai gyakran írnak le léggömböket, tengeraltjárókat és további olyan szállítóeszközöket, amelyek ma már nem számítanak tudományos-fantasztikusnak. A műfaj legújabb képviselői már nagyon másfajta technológiákra építenek. Elsőre nem tűnik nyilvánvalónak, hogy Gibson *Neurománcában*, Wells *Az időgépeben* és Mary Shelley *Frankensteinjében* túl sok lenne a közös szókincs. Számos szkeptikus műfaji elmélet éppen a sci-fi változékonyságából indult ki – ahogy arra *A sci-fi nem létezik [There Is No Such Thing As Science Fiction]* cím emlékeztet minket.³⁷ Összefoglalva: ösztönösen nem világos, hogy a sci-fi esetében azt várhatjuk-e, hogy hosszú időtávon át is egységes marad, mint a detektívtörténet, vagy éppen hogy szétesik, mint a gótikus irodalom.

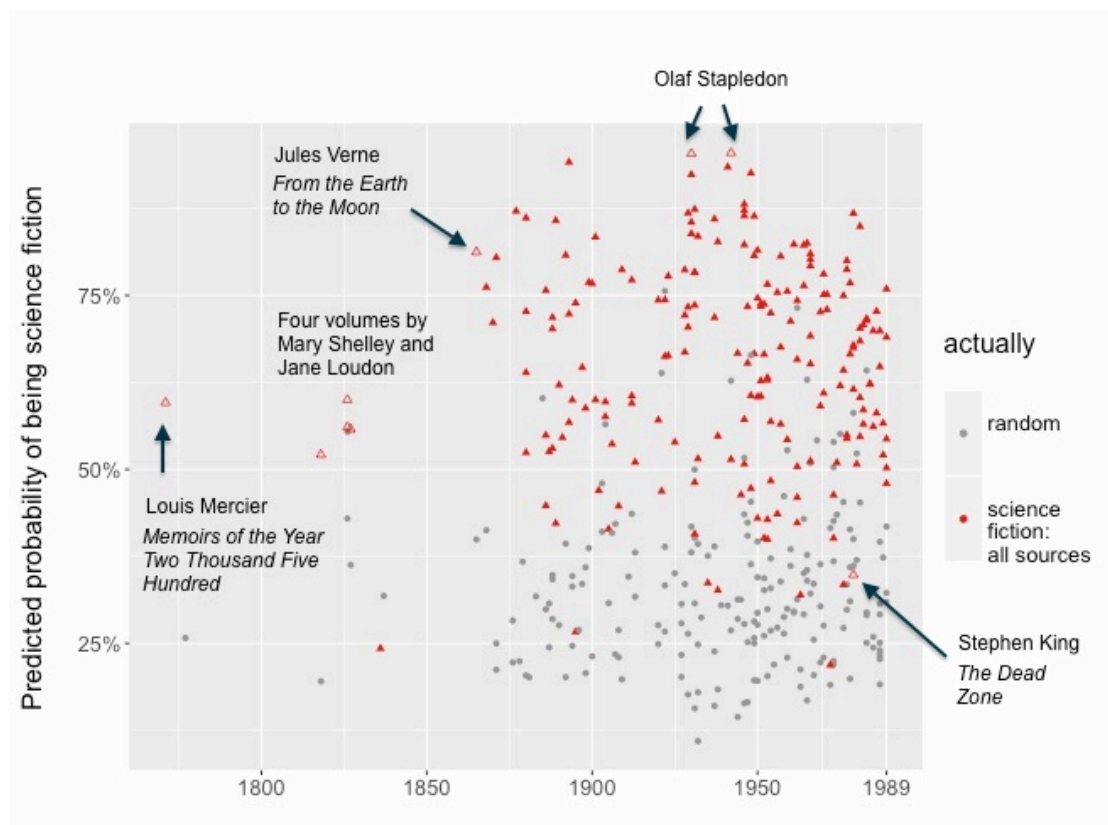
Mivel a sci-fi korai történetéről többféle narratíva létezik, ezt az időszakot különböző források alapján dolgoztam fel. A *The Anatomy of Wonder* című közismert bibliográfia a sci-fi korai történetéről szóló fejezeteket is tartalmaz, melyek szerzője a sci-fi író Brian Stableford. Stableford műfaj története erősen hangsúlyozza H. G. Wells és a jövőbeli háború hagyományának szerepét, ugyanakkor kissé óvatosabb az olyan elődökkel kapcsolatban, mint Mary Shelley és Jane Loudon. (A sci-fi számos tör-

* A vidéki ház-hagyomány a detektívtörténetek egyik alműfaja. A detektívtörténetek egy része olyan helyen játszódik, hogy a lehetséges gyanúsítottak egy zárt halmazból kerülhetnek ki, nincs lehetőség külsős elkövetőre, mert a gyilkosság egy vonaton, hajón, repülőn stb. játszódik. Ilyen a vidéki ház hagyomány is, ahol a ház lakói a gyanúsítottak. Általában vidéki arisztokrata, dzsentri házakról van szó. Agatha Christi is írt ilyen regényt. (A szerk.)

³⁷ Bould és Vint: „No Such Thing as Science Fiction”; lásd még: Kincaid, „On the Origins of Genre”; Rieder, „On Defining SF, or Not.”

ténéséhez hasonlóan Stableford is hajlamos a tudományos tartalom révén meghatározni a műfajt, és ő is szkeptikus azokkal a művekkel kapcsolatban, melyekből az ilyen tartalom hiányzik.) Mivel a nők műfajon belüli megjelenéséről is teljesebb képet szerettem volna kapni, egy, a nők korai sci-fiben betöltött szerepéről szóló bibliográfiára is építettem, amelyet Mary Mark Ockerbloom (a University of Pennsylvania könyvtárának munkatársa) állított össze. Eltérő fogalmi hangsúlyaik ellenére e források olyan szöveglisztákat kínálnak, amelyek nyelvi értelemben nagyon hasonló módon modellezhetők. Ha az egyik lista modelljét összekapcsoljuk a könyvtárosok által felcímkézett huszadik századi szövegekkel, akkor a modell 90%-os vagy nagyobb pontossággal tudja előre jelezni a másik listát.

Ha valamennyi bibliográfiai forrásunkat egyesítjük, akkor egy 196 kötetből álló, az 1771 és 1989 közötti időszakot lefedő listát kapunk, amely 88,3%-os pontossággal modellezhető. A műfaj határai egy kicsit kevésbé világosak, mint a detektívtörténet esetében, de koherenciája még így is egyértelműen közelebb áll ehhez a műfajhoz, mint a gótikushoz. (A pontosság még akkor is körülbelül háromszor közelebb áll a detektívtörténethez, ha az összes műfaji mintát azonos számú kötetre csökkentjük.)



6. ábra. Tudományos-fantasztikus művek 1771–1989 között, 88,3%-os pontossággal osztályozva.

Hogyan lehet egy ilyen hosszú és ezerarcú történettel bíró műfajt pusztán a szavak megszámlálása alapján modellezni? A *Frankenstein* a romantikus korszak szövege, amely még magára a „tudományra” sem hivatkozik túl gyakran. Kiderült azonban,

hogyan vannak olyan szóbeli nyomok, amelyek 170 éven keresztül is fennmaradtak. A méretre való utalások („hatalmas”, „messze”, „nagyobb”) nagyon jellemzőek a sci-fi-re, csakúgy, mint a nagy számok („ezrek”). A „földre” és az „emberi” dolgokra mutató öntudatos hivatkozásokat általában kiegészíti a „teremtmények” társasága, melyektől az emberiség különbözik, és gyakori a semleges birtokos névmás [its], mivel gyakran olyan szereplőkkel találkozunk, akik nem rendelkeznek könnyen felismerhető emberi nemmel. Ezt egyáltalán nem szánom a műfaj kimerítő leírásának – csak ízelítő a modellből.

A 6. ábra egyik vizuálisan kiugró jellemzője, hogy a piros háromszögek 1950 után csekély mértékben csökkenni kezdenek. Korántsem biztos, hogy ez statisztikailag szignifikáns tendencia, de érdekes módon hasonló tendencia figyelhető meg a detektívtörténet esetében is (3. ábra). Mindkét esetben úgy tűnik, hogy a háború utáni kötetek fokozatosan elveszítik az 1930-as és 1940-es évekre jellemző markáns műfaji elkülönülésüket. Van okunk óvatosságra: ez egy finom trend, amit akár a kiválasztás viszontagságai is okozhatnak, hiszen a század közepén megjelenő kötetek, ha rendelkeznek műfajspecifikus rajongói közönséggel, esetenként nagyobb valószínűséggel állnak rendelkezésre digitálisan. Az is elképzelhető, hogy egy idővonal közepén szereplő művek általában jobban illeszkednek a modellhez, mint az idővonal két végén lévő művek (bár ez a minta nem volt nyilvánvaló a korábbi, ezzel a módszerrel lebonyolított kutatásoknál). Ám ha ez egy valós tendencia, akkor elképzelhető, hogy Gary Wolfe hipotézise köszön itt vissza, mely szerint a fantasztikus műfajok határai az utóbbi időben instabillá váltak. „A fantasy elpárolog [...] egyre csak sugároz, átjárja a körülötte lévő levegőt, és furcsa szagot áraszt az irodalmi légkörben.”³⁸ Vannak azonban kisebb különbségek Wolfe téziséhez képest, és a „mágikus realizmus” kapcsán megfogalmazott, idevágó megállapításokhoz képest is. Úgy tűnik, az ezekben a modellekben megfigyelhető váltás közelebb vitte a műfajspecifikus fikciót az irodalmi mező további részeihez. Az ezzel ellentétes tendencia – hogy a műfaji trópusokat a fősodorba tartozó irodalmi szerzők játékosan felhasználják – még nem különösebben látható. Talán 1990 után válna láthatóvá.

A másik, ami esetleges várakozásaink ellenére sem jelenik meg ebben a modellben, az a műfaji konvenciók fokozatos megszilárdulása, pedig a sci-fi szakértői nem kevés időt töltenek e folyamat vizsgálatával. E műfaj történetének kutatói ritkán hajlandóak akkora jelentőséget tulajdonítani Verne és Shelley műveinek, mint amennyit a detektívtörténetek kutatói tulajdonítanak Poe-nak. A történeti áttekintések jelentős részének az a narratív alapvetése, hogy a sci-fi eredetileg csökevényes jelenség volt (amely elszórtan jelent meg az utópia, a planetáris románc stb. műfajokban), míg bizonyos 1925 és 1950 közötti ponyvamagazinok és antológiák új köntösbe nem öltöztették és új irányokat adtak neki. Hugo Gernsback *Amazing Stories* (Csodálatos történetek, 1926) című műve gyakran központi szerepet játszik. Wolfe például azt mondja, hogy a „sci-fi, annak ellenére, hogy az egész tizenkilencedik században élő erőteljes örökségre épít, 1926 után lényegében megtervezett műfaj lett”.³⁹ De még e pont után is „a sci-fi regény kitartóan elkerülte a rejtélyes történetekhez és a westernekhez hasonló műfaji koherenciát”, míg 1943-ban meg nem jelent a *Pocket Book of Science*

³⁸ Gary K. Wolfe, *Evaporating Genres*, viii.

³⁹ Uo., 34.

Fiction (Sci-fi zsebkönyv).⁴⁰ Egyik ilyen kritikus megszilárdulási pillanat sem látható a modellben. A nyelv tekintetében a Verne és Gernsback közötti fél évszázad (1875–1925) ugyanolyan koherensnek és a fikció más formáitól ugyanannyira eltérőnek tűnik, mint az 1926 utáni időszak. Természetesen lehetséges, hogy a modell rossz. Talán a sci-fi pusztán nyelvi alapú megkülönböztethetősége nem olyan fontos, mint a megszilárdulás egyéb formái. De az is lehetségesnek tűnik, hogy a tudományos-fantasztikus irodalomtörténet-írást indokolatlanul lenyűgözte a Gernsback által bevezetett „tudományos-fantasztikus” kifejezés vagy a ponyvakorszak romantikája vagy bizonyos technológiák (például az űrrepülés) szimbolikus jelentősége, és ez összességében háttérbe szorította azt a tényt, hogy a műfaj alapvetően összhangban áll a tudományos románc korábbi hagyományaival.

Mi a tanulság?

Az ebben a cikkben összegyűjtött bizonyítékok három meglévő műfajelméletet vonnak kétségbe. E célpontok közül talán Franco Moretti feltevése a legkevésbé fontos, mely szerint a műfaj egy generációs ciklus: Moretti ezt eleve szabadkozva, afféle spekulációként kínálta fel, és csak néhány más kutató fogadta el.⁴¹ Az a feltevés, hogy a műfaji határok fokozatosan „megszilárdulnak” a huszadik század elején, már komolyabb kérdés. Az, hogy a ponyva adott formát a korábban „koherenssé válni képtelen” sokszínű hagyománynak, nagyon bevett a sci-fi kutatásban.⁴² Nem gondolom, hogy ezt az állítást már megcáfoltam volna, de a nyelvi modellek mindenesetre a bizonyítékok feltűnő hiányát mutatják: a sci-fi jellegzetes nyelvezete, úgy tűnik, már a megszilárduló intézmények előtt kialakult. A harmadik műfajelmélet, amelyet megkérdőjeleztem, az az újabban népszerű elképzelés, mely szerint a műfajok története nem több, mint egy egymástól eltérő kulturális formákat összekötő genealógiai szál. A prediktív modellek nyíltan megkérdőjelezték ezt az állítást. Ha az 1930 előtti detektívtörténeteken betanított modell később is fel tudja ismerni a detektívtörténeteket (és a krimiket), akkor a műfajt az irodalmi mező többi részétől elválasztó különbségek viszonylag stabilak.

De újra hangsúlyoznom kell, hogy a stabil műfaji *határ* nem ugyanaz, mint az adott határon belül található tartalmak stabil *definíciója*. Bár a műfajokhoz kapcsolódó konkrét szavak gyakran lenyűgözőek, az erre vonatkozó részleteket szándékosan mellékesen kezeltem itt, hogy elkerüljem a definíciós állításokat. Sőt, ez a tanulmány éppen azért támaszkodik prediktív modellekre, mert így a definíció problematikáját elkerülve inkább a műfajjal kapcsolatos kortárs szkepticizmus alapját képező, óvatos, nominalista alapvetésekből indulhattam ki. A prediktív modellek például könnyen magukba építik azt az elképzelést, hogy a műfajok egyetlen meghatározó jellemző helyett számos tulajdonságból álló „családi hasonlóságokat” foglalnak magukba.⁴³ A prediktív modellek továbbá azzal a feltételezéssel is kifejezetten jól összeegyeztethetők, mely szerint a műfajok nem feltárássra váró formai mélystruktúrákból, hanem egymással

⁴⁰ Uo., 21.

⁴¹ Roberts, „A Brief Note on Moretti.”

⁴² Wolfe, *Evaporating Genres*, 21.

⁴³ Kincaid, „On the Origins of Genre.”

versengő, szubjektív „címkézési” cselekményekből állnak.⁴⁴ A jelen cikk azért épít a számítógépes módszerekre, mert a segítségükkel tudunk ilyen típusú pluralisztikus és távlati alapokra támaszkodni. Máskülönben (számomra legalábbis) nehéz lett volna jellemezni és összehasonlítani a sok különböző megfigyelő által azonosított összetett családi hasonlóságok erejét.

De még egy távlati alapvetésekből kiinduló vizsgálat is képes lehet végül azt felfedni, hogy az egymással versengő címkézési cselekmények valójában bizonyos esetekben hallgatólagosan összeegyeztethetők. És az is előfordulhat, hogy még az időbeli változékonyság alaptételéből kiinduló vizsgálatok is azt mutatják ki végül, hogy egyes műfajok határai akár másfél évszázadon keresztül is stabilak maradtak. Véleményem szerint ezt láttuk a detektívtörténetek és a sci-fi esetében (bár talán mindkét műfaj esetében igaz, hogy a huszadik század végére kezdenek „elpárologni”).

Nem várhatunk el minden esetben ugyanolyan stabilitást. A gótika nevű jelenségcsoport például kevésbé hajlamos az összeolvadásra. Számos tizenkilencedik századi műfaj (például a Newgate-regény és a szenzációstörténetek) pedig, úgy tűnik, éppen olyan rövid életű, mint ahogy azt Moretti állította. Még a sci-fi is valamivel változékonnyabb, mint a detektívtörténet. Ha lehet valamilyen fő következtetést levonni ebből a tanulmányból, az az lenne, hogy a „műfajnak” nevezett entitásoknak különböző típusai vannak, élekciklusaik különbözőek lehetnek, és különböző fokú szöveges koherenciát mutathatnak. Az irodalomtudósok intuitíve már elismerték ezt például a „műfajok” és a „módok” megkülönböztetése révén. E bináris megkülönböztetés azonban nem tudja kielégítő módon leírni az itt feltárt történelmi folytonosságot. Bár ez a cikk elutasítja Franco Moretti azon feltételezését, hogy a műfajok generációkhoz hasonló ritmust követnek, azt hiszem, egyúttal igazolja azt az átfogóbb állítását, mely szerint a kvantitatív módszerek rugalmasabb és az általunk vizsgált anyag összetettségére jobban reagáló leíró forrásokat adhatnak az irodalomtudósok kezébe.

Fordította: Maczelka Csaba

⁴⁴ Rieder, „On Defining SF, or Not,” 192–193.

The Life Cycles of Genres

Critics disagree even about the kind of thing a “genre” is. For instance, are genres defined mostly by textual form or by social reception? Are they permanent categories, century-sized historical constructs, or short-lived generational phenomena? This essay tries to address some of these persistent debates with recent quantitative methods. It shows, first, that social and textual definitions of genre can align fairly well: genres defined by the judgments of specific readers can also be recognized by textual clues. Building on that success, it investigates controversies about the typical lifespan of a genre. It turns out that different models of a genre are often compatible, even if defined by the judgments of readers in different periods. For instance, the models that can recognize nineteenth-century “detective stories” and “scientific romance” also recognize twentieth century “crime fiction” and “science fiction.” This evidence suggests that—while genres are contingent historical constructs—many of them are more durable than literary historians have been willing to claim. But this is not universally true. The Gothic, for instance, is just as loose and mutable as critics have typically believed.

Keywords:

machine learning, genre, literature, literary history

A) függelék: metaadatok

A modellezéshez használt metaadatok a GitHub repozitórium `finalmeta.csv` fájljában található.

Minden egyes mű korlátlan számú „műfaji címkét” viselhet, amelyek különböző, az adott művel összekapcsolt csoportokat jelölnek.

Az alábbi táblázat a címkék jelentését mutatja be: összesen 962 szövegre vonatkoznak. A szövegek nagy része kötetnyi méretű, de előfordul néhány novella is.

| címke | #szövegek | dátum | leírás vagy forrás |
|------------|-----------|-----------|--|
| det100 | 89 | 1829–1941 | <i>The First Hundred Years of Detective Fiction, 1841–1941</i> . 1973. Lilly Library, Bloomington, IN. http://www.indiana.edu/~liblilly/etexts/detective/ |
| chimyst | 146 | 1923–1989 | A könyvtárosok által a „detektív” [detective] vagy a „rejtélyes történet” [mystery fiction] kategóriába sorolt művek a Chicago Text Lab gyűjteményében. |
| locdetmyst | 45 | 1832–1922 | A könyvtárosok által „detektív- és rejtélyes történet” [detective and mystery fiction] kategóriába sorolt művek a HathiTrust gyűjteményében. |

| | | | |
|--------------|----|-----------|---|
| locdetective | 16 | 1865–1912 | A könyvtárosok által a „detektívek” [Detectives] kategóriába sorolt művek. Gyakran esetnapló-jellegű történetek. |
| crime | 2 | 1972–1974 | A könyvtárosok által a „krimi” [crime fiction] kategóriába felvett, de a „detektívtörténetek” [detective fiction] kategóriából kihagyott művek. |
| cozy | 10 | 1920–1952 | Olyan művek, amelyek szerzői a <i>The Mystery Readers’ Advisory: The Librarian’s Clues to Murder and Mayhem</i> című kötetben (John Charles, Joanna Morrison és Candace Clark, Chicago: ALA, 2002) vidéki házban játszódó rejtélyes történetek íróiként szerepelnek. |
| hardboiled | 10 | 1929–1970 | Geoffrey O’Brien, <i>Hardboiled America</i> (New York: Van Nostrand Reinhold, 1981) című kötetének melléklete. |
| newgate | 7 | 1828–1840 | Keith Hollingsworth, <i>The Newgate Novel</i> , Detroit 1963. |
| sensation | 14 | 1860–1880 | „The Sensation Novel,” Winifred Hughes, in Patrick Brantlinger and William B. Thesing, eds., <i>A Companion to the Victorian Novel</i> (Blackwell, 2002). |
| lockandkey | 10 | 1800–1903 | A <i>The Lock and Key Library: Classic Mystery and Detective Stories</i> című, Julian Hawthorne által szerkesztett (New York, 1909) antológiába beválogatott szövegek. Olyan írók is szerepelnek itt, mint Dosztojevszkij, akit valószínűleg még 1909-ben sem tekintettek krimiírónak, a cikkben használt „detektívtörténet” modellbe nem vettük fel. |
| pbgothic | 96 | 1764–1988 | David Punter and Glennis Byron, <i>The Gothic</i> (Malden: Blackwell, 2004), „Kronológia”. |
| stangothic | 21 | 1791–1834 | A Stanford Literary Lab metaadataiban a „gótikus irodalom” [Gothic] címkével ellátott művek kisebb halmaza. |
| lochorror | 4 | 1818 | A könyvtárosok által a „horror” [horror] kategóriába sorolt művek a HathiTrust gyűjteményében. |
| chiorror | 23 | 1933–1989 | A könyvtárosok által a „horror” [horror] kategóriába sorolt művek a Chicago Text Lab gyűjteményében. |
| locghost | 28 | 1826–1922 | A könyvtárosok által a „kísértethistóriák” [ghost stories] kategóriába sorolt művek. |

| | | | |
|------------|-----|-----------|---|
| locscifi | 21 | 1836–1909 | A könyvtárosok által a „sci-fi” [science fiction] kategóriába sorolt művek a HathiTrust gyűjteményében. |
| chiscifi | 144 | 1901–1989 | A könyvtárosok által a „sci-fi” [science fiction] kategóriába sorolt művek a Chicago Text Lab gyűjteményében. |
| femscifi | 9 | 1818–1922 | Ockerbloom, Mary Mark. 2015. „Pre-1950 Utopias and Science Fiction by Women.” |
| anatscifi | 36 | 1771–1922 | Stableford, Brian. 2004. „The Emergence of Science Fiction” és „Science Fiction Between the Wars.” In <i>Anatomy of Wonder</i> , ed. Neil Barron, 5th edition, 3–44. |
| chiutopia | 13 | 1920–1976 | A könyvtárosok által az „utópiák” [Utopias] kategóriába sorolt művek a Chicago Text Lab gyűjteményében, a jelen tanulmányban nincs beépítve a „sci-fi” kategóriába. |
| chifantasy | 53 | 1901–1989 | A könyvtárosok által a „fantasy” [fantasy] vagy „fantasztikus irodalom” [fantasy fiction] kategóriába sorolt művek, a jelen tanulmány célkitűzéseinek megfelelően nincs beépítve a „sci-fi” kategóriába. |
| – | – | – | – |
| juvenile | 23 | 1904–1922 | Ifjúsági közönségnek szánt művek; kigyűjtve, de a jelen tanulmány nem használja. |
| drop | 33 | 1838–1922 | Olyan művek, amelyeket végül nem használtam fel, de az átláthatóság érdekében meghagytam a metaadatokat. Leggyakrabban az ifjúsági jelleg miatt maradtak ki. |
| random | 169 | 1769–1922 | A HathiTrust Digital Library anyagából véletlenszerűen, a NEH [National Endowment for the Humanities] által finanszírozott „Understanding Genre in a Collection of a Million Volumes” című projekt keretében kidolgozott, fikcióval kapcsolatos metaadatok alapján kiválasztott művek. A „véletlenszerű kiválasztás” ebben az esetben azt jelenti, hogy a kötetek véletlenszerűen lettek kiválasztva, ám a szerző jóváhagyta vagy elutasította a kiválasztott szövegeket, hogy elkerülje a fikciós szövegeket, klasszikus költészetet, ifjúsági műveket tartalmazó kóbor köteteket. |

| | | | |
|-----------|-----|-----------|---|
| chirandom | 202 | 1920–1989 | A Chicago Text Lab anyagából véletlenszerűen kiválasztott művek. A kiválasztás ebben az esetben közelebb állt a valódi véletlenszerű kiválasztáshoz. ⁴⁵ |
| teamred | 484 | 1760–1989 | A hitelesség ellenőrzésére véletlenszerűen kiválasztott szerzők. |
| teamblock | 500 | 1764–1989 | A hitelesség ellenőrzésére véletlenszerűen kiválasztott szerzők. |
| stew | 224 | 1764–1989 | Véletlenszerűen kiválasztott, a gótikus irodalom [Gothic], a sci-fi [science fiction] és a krimi/detektívtörténet [crime/detective] hagyományát kiegyensúlyozottan reprezentáló kötetek halmaza, célja egy szörnyű műfaji ragu elkészítése. |

B függelék

A B függelék az a GitHub repozitórium, amely a projekthez tartozó kódokat, adatokat és metaadatokat tartalmazza, és az alábbi címen érhető el:

<https://github.com/tedunderwood/fiction>

⁴⁵ Megjegyzendő, hogy a két „véletlenszerű” [random] címke más műfaji címkékkal együtt is előfordulhat. A véletlenszerűen kiválasztott kötetek például a „chimyst” kategóriának is tagjai lehetnek – ilyen esetekben csak akkor hagytam ki az adott művet a negatív (kontroll) halmazból, ha a „chimyst” szerepelt a pozitív halmazban.