

Digitalna humanistika in literatura

Ivana Zajc, Peter Purg



Založba Univerze v Novi Gorici

2023

Naslov dela: Digitalna humanistika in literatura
Avtorja: Ivana Zajc in Peter Purg
Recenzenta: Andrejka Žejn in Miran Hladnik
Lektorirala: Ivana Zajc
Založnik: Založba Univerze v Novi Gorici, Vipavska
13, Rožna Dolina, Nova Gorica
Leto izdaje: 2023
ISBN 978-961-7025-33-0 (PDF)

Format: PDF

<http://www.ung.si/sl/zalozba/>

<http://www.ung.si/en/publisher/>

18. 12. 2023

Publikacija v odprtem dostopu.

Izdajo učbenika je sofinanciralo Ministrstvo za visoko šolstvo, znanost in inovacije Republike Slovenije.

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID 176365827

ISBN 978-961-7025-33-0 (PDF)



To delo je objavljeno pod licenco [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0). Avtor/ji besedila dovoli/jo reproduciranje, distribuiranje, dajanje v najem in priobčitev javnosti samega izvirnega avtorskega dela pod pogoji: priznanje avtorstva, nekomercialno ter deljenje pod enakimi pogoji. Celotno besedilo licence je na voljo na spletu, na naslovu <http://creativecommons.si/licence>

CONTENTS

Kolofon	v
Nekaj uvodnih besed	vi
Kaj je digitalna humanistika in kako se povezuje z literarno vedo	1

PART I. MAIN BODY

1. Digitalno literarno/E-literatura	27
2. Oddaljeno branje književnosti	42
3. Obravnava književnosti z računalniško stilometrijo	69
4. Procesiranje naravnih jezikov, modeliranje tem in literatura	84
5. Analize omrežij v književnosti in literarni zgodovini	94
6. Literarne pripovedi z vidika digitalnih raziskav	110
7. Analiza čustev v književnosti	129
8. Računalniške analize literature in spol	137

9. Spletne podatkovne zbirke in literarno(vedno) gradivo na spletu	150
10. Digitalne kritične izdaje književne klasike	171
11. Kritična digitalna humanistika	178
Viri in literatura	187

književnega gradiva.¹ Že od svojih začetkov se področje digitalne humanistike, ki se ukvarja s književnostjo, ne ukvarja le z literarnimi vprašanji, ampak v veliki meri tudi z metodološkimi vprašanji o lastnih digitalnih orodjih.²

Področje računalniških raziskav literature se je od tedaj hitro razvijalo in je danes izjemno obsežno ter raznoliko, zato smo za predstavitev v pričujočem učbeniku izbrali nekaj osrednjih zanimanj raziskovalk in raziskovalcev. Skozi izbor zanimivih raziskav bralcu predstavljamo izbrane vsebine, ki bodo informativen uvod v nadaljnji študij področja. Ob tem se mestoma dotaknemo tudi zgodovinskih začetkov nekaterih raziskovalnih metod, ki so bile zasnovane že dolgo pred uveljavitvijo digitalne tehnologije. Primer je zasnova stilometrije z ročnim štetjem besed v posameznih besedilih, že dolgo preden so lahko to opravljali računalniški algoritmi. Del učbenika je namenjen tudi novim ustvarjalnim praksam, ki vznikajo ob tehnoloških inovacijah in s seboj prinašajo nove možnosti za umetniško izražanje.

S hitrim napredkom tehnologije in digitalizacijo informacij se tudi študij literature razvija ter odpira novim izzivom in priložnostim. Drugačne tehnologije in orodja pomenijo nov pogled na ustvarjanje in analizo literarne ter druge umetnosti. Ob uveljavljanju netradicionalnih pristopov je digitalna humanistika pogosto naletela na odpor, ena od posledic tega procesa pa je tudi delitev metod na uveljavljene, ki jih zaznamuje na primer natančno, bližnje branje besedil, in sodobne digitalne

1. Thomas 2000

2. Thomas 2000

metode. Izhodišče tega učbenika je, da v sodobnem času metodološkega pluralizma, ko področja literarnih raziskav ne obvladujejo več posamezne literarnovedne šole, taka delitev metod ni smiselna. Za načrtovanje raziskav in hipotez – pa čeprav v okviru digitalnih metod – je ključno temeljito poznavanje literarne vede, ki je poleg tega pomembno pri interpretaciji rezultatov, zato razkorak med klasičnimi in digitalnimi metodami ni utemeljen. Ne le to: digitalne raziskave književnosti oblikujejo nove tipe interdisciplinarnih raziskovalnih skupin, ki združujejo informatičarke in informatike, literarne zgodovinarke ter zgodovinarje in druge zanimive profile. S tem digitalna humanistika literarnovedne raziskave bogati ter širi razumevanje človeške kulture, umetnosti, jezika in zgodovine na načine, ki pred tem niso bili mogoči.

V tej knjigi boste spoznali, kako je mogoče s pristopi digitalne humanistike analizirati književna besedila, arhivirati literarno dediščino in še veliko več. Med temami, ki jih obravnavamo, so pristopi za obdelavo velikih podatkov, uporaba računalniškega modeliranja, vprašanja etike, dostopa do znanja preko odprtih virov in digitalnih knjižnic ter ohranjanje literarne dediščine. V prvem poglavju predstavljamo polje digitalne humanistike, pokažemo, kaj počne digitalni humanist oziroma digitalna humanistka in pokažemo, katere nove pristope prinašajo digitalne metode in orodja k raziskovanju književnosti. V drugem poglavju, ki ga je sestavil Peter Purg, se ukvarjamo s predstavitvijo e-literature s poudarkom na slovenskih umetniških projektih, v njem pa boste spoznali primere in načela ustvarjanja v digitalnem svetu, ki umetnikom ponuja

številne nove možnosti izražanja. Avtorica metodološkega dela učbenika, ki sledi, je Ivana Zajc, in sicer je najprej na vrsti nekaj poglavij, ki so posvečena digitalnih metodam za raziskovanje književnosti, tako obširnih literarnih korpusov kakor tudi posameznih besedil. Najprej v splošnem predstavljamo možnosti oddaljenega branja literature z računalniki in ga med drugim vzporejamo z bližnjim branjem, ki prav tako ostaja ključna metoda sodobnega ukvarjanja z literarnimi vprašanji. Različni tipi oddaljenega braja temeljijo na statističnih izračunih in nudijo drugačen vpogled v digitalizirano literarno gradivo, pogled "od daleč". Zatem se v posebnem poglavju posvetimo računalniški stilometriji, pojasnimo, kakšno razumevanje literarnega sloga predpostavlja, in pokažemo nekaj njenih zanimivih aplikacij. Sledi prikaz metode računalniškega procesiranja naravnega jezika in njenih aplikacij na literarnovedne raziskave. Posebno poglavje je posvečeno vse bolj aktualni temi analiz omrežij v književnosti in literarni zgodovini, ki nam omogoča, da jasneje razvidimo kompleksna omrežja, povezana z literarnim sistemom ali z motivno-tematskimi vidiki književnih besedil.

V nadaljevanju učbenik predstavlja možnosti digitalnih raziskav posameznih področij, ki zaznamujejo tudi tradicionalne literarne raziskave: gre za poglavja na temo digitalne humanistike in literarne zgodovine, analiz literarnih pripovedi z digitalnimi metodami, analiz čustev v književnosti ter digitalnih raziskav spola v literaturi. Sledi poglavje, ki predstavlja ključna prizadevanja in dosežke digitalne humanistike v navezavi na literarno dediščino, kjer predstavljamo spletne podatkovne zbirke

in digitalne knjižnice, digitalne kritične izdaje književnih besedil ter literarnovedno znanje na spletu.

Učbenik poleg vsebinskih razlag vsebuje primere izbranih raziskav, zanimive citate, hiperpovezave, slike in videe, predlaga pa tudi zanimive izzive na spletu in onkraj njega, ki jih lahko preizkusite. Ob koncu posameznih poglavij se dotaknemo tudi kritičnih pogledov na izbrano tematiko, saj je področje, s katerim se ukvarjamo v tem učbeniku, novo in dinamično, zato včasih še išče pravilne možnosti in generira mnoge strokovne razprave. Kritični digitalni humanistiki pa je posvečeno tudi zadnje poglavje učbenika, ki vsebuje izbrane teme s področja etične refleksije digitalnih raziskav in pretresa nove načine njihovega udejstvovanja v prihodnosti.

Učbenik je namenjen vsem, ki jih zanima prihodnost literarne vede in književnosti v digitalnem svetu ter vpliv tehnologije na naše razumevanje teh pomembnih civilizaijskih dosežkov. Ker pa se te spremembe z nenehnim razvojem tehnologije nikoli ne končajo, je knjiga odprta za nove in nove vsebine, ki jih bomo dodajali ob razvoju nadaljnjih inovacij in zanimivih digitalnih pojavov na področju literarnosti. Upamo, da bo delo navdih, da se nove generacije raziskovalk in raziskovalcev pridružijo skupnosti, ki si prizadeva razumeti in ohraniti neprecenljivo literarno dediščino s pomočjo digitalnih orodij, pa tudi, da spodbudi nove možnosti ustvarjalnosti v okviru e-literature.

Zahvala

Velika zahvala za nastanek tega učbenika gre spodbudi vodje Raziskovalnega centra za humanistiko Univerze v Novi Gorici red. prof. dr. Katji Mihurko, ki s svojo vizijo v raziskovalni skupini spodbuja digitalne raziskave književnosti, literarnega sistema in širše kulture.

KAJ JE DIGITALNA HUMANISTIKA IN KAKO SE POVEZUJE Z LITERARNO VEDO

Živimo v času informatizacije družbe, pa tudi informatizacije zasebnega življenja ljudi, ki si svojega vsakdana brez sodobnih informacijsko-komunikacijskih tehnologij ne znajo več predstavljati. V času t. i. velikih podatkov (ang. *big data*), katerih količina se iz leta v leto povečuje, se pojavlja vse več možnosti za njihovo obdelavo, kar je aktualno tudi za področje humanistike. S tovrstnimi interdisciplinarnimi raziskavami na preseku med humanistiko in informacijsko tehnologijo se je vzpostavilo polje digitalne humanistike.

Sodobno družbo v veliki meri zaznamuje digitalna tehnologija, ki je postala nepogrešljiv del vsakdanjega življenja: od trgovine, infrastrukture in izobraževanja do delovanja institucij in storitev, od službenih obveznosti

do prostega časa. V tem dinamičem okolju se nenehno pojavljajo nove tehnologije in inovacije, hitrejši načini komunikacije pa vplivajo na pospešeno globalizacijo sveta. Sodobna družba je družba, ki jo zaznamuje merjenje najrazličnejših vidikov, kar se je zgodilo prvič v zgodovini.¹ Humanistika pomeni študij človeškega znanja in kulture, polje digitalne humanistike pa po drugi strani obsega presek digitalnega sveta ter humanistike. Digitalna humanistika k informacijski tehnologiji pristopa z vidika humanistike in skuša po drugi strani na tradicionalna humanistična vprašanja odgovoriti z novimi tehnološkimi rešitvami.

1 Humanistika v digitalnem svetu

Ena ključnih dobrin v informacijski družbi so postali podatki, ki jih na primer velike svetovne korporacije zbirajo in analizirajo. V tem okolju je tema varstva osebnih podatkov in posameznikove zasebnosti vse bolj pereča, na tem področju pa države ter organizacije ustvarjajo nove politike in zakonodajo. Poleg tega se v družbi pojavljajo grožnje v povezavi s kibernetiko varnostjo, saj so informacijski sistemi ranljivi. V svetu, ki se pospešeno spreminja, ima pomembno vlogo humanistika, znanstvena disciplina, ki raziskuje človeka, pa tudi literaturo, kulturo, umetnost in zgodovino ter etiko in podobna področja. V času novih tehnologij in pospešenih sprememb v družbi, ki močno vplivajo na vsakdanje življenje, je humanistična kritična refleksija

1. Anderson 2008

ključna. Kot odgovor na spremembe v digitalni družbi se je oblikovala digitalna humanistika kot interdisciplinarna disciplina v svetu digitalnih tehnologij.

Digitalna humanistika “je povezava področij, znotraj katerih raziskave uporabljajo računalniške tehnologije, da odgovarjajo na vrsto vprašanj, ki so tradicionalno domena humanistike /.../ ali ki postavljajo vprašanja o računalniških tehnologijah, ki so bila tradicionalno humanistična.”²

Raznolikih odgovorov na vprašanje, kaj obsega digitalna humanistika kot izrazito meddisciplinarno področje, je s sodobnim hitrim razvojem tehnologije vse več, kar kaže tudi vse daljši seznam člankov na spletni strani organizacije European Association for Digital Humanities, ki odgovarjajo na to vprašanje.³ Med humanističnimi temami v digitalnem svetu so na primer etičnost, prosti dostop podatkov za javnost in študije spolov. Med področji digitalne humanistike je tudi kulturomika kot kvantitativna analiza kulture in trendov na podlagi digitaliziranih besedil,

2. Fitzpatrick 2010

3. What is digital humanities?

na primer tvitov in publicističnih člankov. Gre za rudarjenje podatkov za raziskovanje kulturnih fenomenov v družbi, na primer emocij ljudi v zvezi z določeno tematiko. Digitalna humanistika se po eni strani osredotoča na ustvarjanje digitalnih virov in dostop do njih, po drugi strani pa izvaja analize digitalnih materialov z naprednimi računalniškimi tehnologijami.⁴

Kaj počne digitalni humanist oziroma digitalna humanistka? Na iniciativi The European Union's Digital Research Infrastructure for the Arts



and Humanities so ustvarili Taksonomijo digitalnih raziskovalnih dejavnosti v humanistiki (TaDiRAH), ki vključuje tri široke kategorije: raziskovalne dejavnosti, raziskovalne predmete in metode. Digitalni humanisti oziroma humanistke se ukvarjajo s širokim naborom tematik, vključno z množičnim iskanjem podatkov, kodiranjem ipd., in raziskovalnih metod. Delo digitalnega humanista ali digitalne humanistke v okviru informacijskih tehnologij med drugim obsega razvoj in implementacijo algoritmov, kodiranje, pošiljanje sporočil prek različnih aplikacij, skrbništvo, upravljanje in analizo informacij, ustvarjanje in analiziranje statističnih podatkov, socialno mreženje in virtualna srečanja, zapisovanje in objavljanje informacij, uporabniško testiranje, razvoj programske opreme, spleta in

4. Biemann idr. 2014: 87–91.

vizualizacij, interpretacijo podatkov idr. Digitalna orodja ne morejo opravljati kritičnega in analitičnega dela, ki ga opravljajo ljudje, ti pa po drugi strani ne morejo opravljati dela, ki ga opravljajo oni.⁵ V določenih primerih bi lahko do rezultatov, do katerih pridejo raziskovalci z digitalnimi metodami, prišli tudi brez uporabe teh metod. Če bi hoteli na primer pokazati pogostost rabe pridevnikov v slogu določenega avtorja, bi prebrali vsa njegova dela, si označili pridevnike in napravili preštevno analizo, vendar bi nam to vzelo veliko časa. Če bi to opravili z računalniškim programom, bi torej izvedli isti tip raziskave, ki pa bi lahko bila zaradi drugačne metode širša, kompleksnejša in hitrejša.⁶ Poleg tega se z digitalnim vidikom spreminja način izvajanja raziskav v humanistiki, in sicer v smeri sodelovanja raziskovalcev različnih profilov in strokovjakov na področju informacijskih tehnologij.

Digitalna humanistika ni le aplikacija računalništva na področje kulture, opozarja Andrew Piper.⁷ Kulturna analiza nas namreč vabi k premisleku obeh kategorij – tako kulture kot tudi računalništva –, ob tem pa odpira vprašanja glede implicitnih predpostavk o objektivnosti in univerzalnosti na področju informacijskih tehnologij. Opozarja, da je znanje vselej postavljeno v določene okoliščine, v nek kraj in čas, da vedno obstaja določen subjekt, ki je nosilec znanja. Računalništvo pa prinaša spodbudo k premisleku izhodišč v humanističnih raziskavah: kaj nam zadošča kot dokaz, kakšen je odnos med teorijo in prakso ipd.⁸

5. Adwetewa-Badu 2020

6. Eve 2022: 8

7. Piper 2016

8. Piper 2016

Kultura bo polju računalništva omogočila bolj kritičen vpogled v to, kako algoritmi strukturirajo pomene.⁹

Tom Scheinfeldt¹⁰ ocenjuje, da digitalna humanistika obsega dve osrednji veji: študij književnosti in študij zgodovine, ki sta se samostojno razvijala v 90. letih 20. stoletja, s pojavom svetovnega spleta v poznih 90. letih in po letu 2000 pa sta se pričela razvijati v medsebojni povezanosti. V pričujočem učbeniku se ukvarjamo z digitalno humanistiko in literaturo, tj. s stičišči med možnostmi, ki jih ponuja digitalni svet, in raziskovanjem književnosti. Digitalne tehnologije vplivajo na področje literarnega ustvarjanja in novih umetniških praks, na polje recepcije bralcev, ki imajo možnost književna besedila brati drugače, pa tudi na področje literarnega ustvarjanja.

9. Piper 2016

10. Scheinfeldt 2008



2 Nove literarno-umetniške prakse

Pisna komunikacija na daljavo sega daleč nazaj v čase poštinih slov, in šele pred dobrimi tristo leti se je tehnologija posredovanja pisem (s poštanim sistemom, pa tudi bolj dostopnim papirjem) razvila do te mere, da so se – takrat zgolj premožnejši sloji – v medsebojnih dopisovanjih pričeli poigravati z besedami in besedilom. Temu bi lahko rekli tudi predhodnik internetnih literarnih praks, ki so se ob širši dostopnosti računalniških omrežij (vsaj na premožnejšem severozahodu) zares razmahnile šele v 90. letih 20. stoletja. Že v 19. stoletju so sicer odkrili, da lahko računalniški stroji poleg številčk uspešno (in zabavno) premetavajo tudi črke (kakor so s tem igranje eksperimentirali že v obdobju baroka, a takrat še na papirju), v 20. stoletju pa se je postopoma uveljavilo tudi računalniško procesiranje slike in zvoka.

Tako so se literarno-umetniške prakse postopoma

posluževale novih orodej in platform, ki so odpirale dotlej neslutene možnosti dostopa in dometa, povezovanja ustvarjalk in ustvarjalcev na daljavo, pa tudi poigravanja z besedilom, ter njegovim kombiniranjem z drugimi medijskimi oziroma izraznimi oblikami. Besedilo v računalniku nastopa kot podatkovni vnos, kot material – iz *črk in besed*, kar za informatiko pomeni iz *ničel in enic*, vanje pa so nenazadnje zakodirane tudi slike in zvoki. Eno od prvih del, ki je hkrati izkoristilo številne avdio-vizualne ter programerske zmožnosti računalnika, je bilo delo *Legible City* (*Berljivo mesto*) Jeffreya Shawa iz leta 1989, kjer uporabnik na prirejenem kolesu navigira skozi besedilno pokrajino. Deset let pozneje pa je interaktivna instalacija avtoric Camille Utterback in Romy Achituv z naslovom *Text Rain* (*Besedilni dež*) navidezno materialnost črk približala vse do meje (projekcije) človekovega telesa. Eksperimenti *e-mail art* (e-poštne umetnosti) so se dogajali že v 70. letih 20. stoletja in s tem nadaljevali prej omenjene tradicije večstoletnega soustvarjanja besedila na daljavo.

Takoj po razmahu svetovnega spleta sredi 90. let se je na multimedijško razširjeni platformi medsebojno povezanih računalnikov razbohotila tudi spletna literatura v vseh svojih oblikah – od t. i. hiper-romana, v katerem listanje po knjigi zamenja sledenje spletnim povezavam ter raziskovanje avdio-vizualnih elementov na ta način obogatene pripovedi, pa vse do spletne poezije, ki je poleg hiperpovezav in multimedijskih razsežnosti lahko služila tudi vse bolj raznolikim možnostim animiranja oziroma programiranja (bolj ali manj nestanovitnih) besed, črk, pa tudi poseganja v podatkovne zbirke (baze), sodelovanja

na daljavo ipd. Tudi dramsko besedilo ni ostalo imuno na nove tehnologije – eksperimentiralo se je (in se še) s t. i. klepetalnim gledališčem (chat theatre), pa tudi z različnimi internetnimi in spletnimi oblikami (kot-da-)dramskega besedila, ki vdira tudi v resnični prostor gledaliških uprizoritev, in obratno, tako v gledališčem kot v galerijskem in tudi muzejskem prostoru se srečuje s tehnologijami navidezne oziroma raziširjene resničnosti (VR, AR).

V novem tisočletju se je z razmahom družabnih omrežij in pretočnih platform literarno ustvarjanje premaknilo seveda tudi tja, e-literatura je postala pomemben prostor izražanja drugačnih, alternativnih in manjšinskih identitet, pravic manjšin, feminističnih, dekolonialnih, ekokritičnih ipd. agend. Vse te mnogotere e-literarne oblike sedaj sobivajo, tako rekoč vsak dan se razvijajo nova dela in vznikajo raznoliki projekti. Tudi na področju e-literature ter spletne umetnosti je slovenski prostor prispeval nekaj svetovno prepoznavnih imen in vrsto inovativnih projektov ter pomembnih umetniških del, ki so uvrščena v svetovne antologije ter so bila predstavljena v muzejih, galerijah ter na festivalih širom zemeljske oble.

3 Raziskovanje književnosti z digitalnimi orodji in kvantitativni podatki v literarni vedi

Z digitalnim obratom družbe so se pričele pojavljati tudi različne nove metode raziskovanja v literarni vedi, kjer so

raziskovalci tradicionalno temeljili predvsem na bližnjem branju. Pri slednjem se posameznik osredotoči na podrobnosti in globoko razumevanje besedila, s poudarkom na analizi zapletenih konceptov ter odnosov med besedami in povedmi. Digitalna orodja uvajajo t. i. računalniško branje oziroma strojno branje, ki vsebine danih besedil analizira z uporabo jezikovnih modelov in umetne inteligence. Računalniško branje je proces, s katerim računalniki iz določenih virov podatke pridobivajo in jih shranjujejo ter jih obdelujejo s kvantitativnimi metodami. Tako izluščijo določene informacije, prepoznajo posamezne vzorce, tematike ipd. Digitalna orodja raziskovalcem omogočajo, da hitro in natančno analizirajo obsežne besedilne zbirke, pa tudi posamezna literarna besedila. Gre za kvantitativne analize, ki dopolnjujejo tradicionalne metode literarne vede, raziskovalec ali raziskovalka pa rezultate kritično interpretira, preveri na primer morebitno pristranost digitalnih orodij v interpretaciji literarnih del.

3.1 Literarna veda in možnosti digitalnega

Raziskovanje književnosti s sodobnimi orodji informacijske tehnologije je v sodobnem času postalo podpodročje literarne vede. Literarna veda sicer tradicionalno obsega tudi druga področja: literarno zgodovino, literarno teorijo in literarno interpretacijo. Poleg tega posebno polje znotraj literarne vede zavzemajo didaktika književnosti, primerjalna književnost in tekstologija, medtem ko

metodologija literarne vede obsega samorefleksijo metod v literarni vedi, vključno z digitalnimi metodami.

Na vseh teh področjih je mogoče opazovati, kako se literatura oplaja z možnostmi digitalnih tehnologij. Raziskovanje književnosti s sodobnimi orodji informacijske tehnologije ima različna poimenovanja, znano je na primer kot kulturna analitika, literarno podatkovno rudarjenje, kvantitativni formalizem, literarno besedilno rudarjenje, računalniško besedilna analiza, računalniška kritika, algoritemska literarna veda, socialno računalništvo v literarnih vedah in računalniška literarna veda.¹¹

“Moramo vsi postati digitalni humanisti? Seveda ne. Vsekakor bo nove tipe dokazov, ki jih digitalne metode lahko prinesejo na področje literarne vede, vse težje ignorirati.”¹²

»Bo literarna veda še naprej predvsem interpretativna? Ali obstajajo novi modeli

11. Nan 2019: 601

12. Eve 2022: 153

interpretacije? Bo morala literarna veda in s tem literarne študije dobiti nove oblike? Ali bo imel premik od tiskane k elektronski kulturi posledice za koncept literature in s tem za literarno vedo?”¹³

Na področju digitalnih metod v literarni vedi je zelo razširjeno t. i. oddaljeno branje oziroma makroanaliza književnosti. Gre za računalniške analize obsežnih korpusov literarnih besedil z uporabo računalniških orodij. Predmet raziskovanja literarnovednih vprašanj z digitalnimi orodji so v širšem smislu teksti, kar pa ne pomeni, da se ta tip raziskovanja omejuje le na oddaljeno branje obsežnejše zbirke literarnih besedil. Predmet oddaljenega branja je lahko korpus besedil ali posamezno besedilo oz. njegovimi deli (npr. v romanu analizo dialogov v primerjavi z besedilom pripovedovalca), zanimajo nas lahko katalogi metapodatkov o literarnih delih (npr. spol avtoric oz. avtorjev) ali izvajamo oddaljeno branje posameznih tipov literarnih likov ipd. Poleg literarnih besedil pa se lahko posvečamo tudi literarnozgodovinskim virom, na primer digitaliziranim dnevnikom književnih ustvarjalcev in ustvarjalk, kjer reflektirajo svoje življenje ter ustvarjanje.

Izziv: razmislek o rabi digitalnih metod v litrarni vedi

Zakaj bi sploh uporabljali digitalne metode za raziskovanje kulture, zakaj je zdaj pravi čas za to, zakaj to ni le modna muha, ampak nujnost, se sprašuje raziskovalec Andrew Piper.¹⁴ S pomočjo spleta skušajte najti po tri argumente za in po tri argumente proti uporabi digitalnih metod za raziskovanje področja književnosti znotraj kulture. Posamezne argumente zapišite in pretehtajte, ali so po vaši presoji utemeljeni.

3.2 Kvantitativno raziskovanje književnosti

Raziskovanje književnosti z digitalnimi metodami je kvantitativno, saj temelji na statističnih izračunih. Vendar so se kvantitativne metode literarnih raziskav pojavile že pred uporabo računalniških tehnologij za analizo književnosti, tudi v slovenskem prostoru.

Primer: slovenska povest preštevno

Miran Hladnik je preveril, katere lastnosti slovensko povest razlikujejo od romana. Obravnaval je korpus sedmih pripovednoproznih del z zgodovinsko temo pisatelja Josipa Jurčiča in preštevno ugotovil naslednje določnice povesti v Jurčičevem opusu:¹⁵



- v povprečju so romani sicer daljši od povesti, približno polovica romanov pa je enako dolga ali pa celo krajša od povesti;
- dogajalni čas je v romanih krajši kot pri povesti (3,1 leta : 4,4 leta), število poglavij pa je večje;
- pripovedni tempo – tj. gostota dogajanja, ki jo označujemo z razmerjem med pripovednim časom, ki se meri s količino besedila, in dogajalnim časom – je za polovico poglavij pri povesti in romanu enak, in sicer zmerni tempo, tj. pripovedna scena s pripovednim poročilom, medtem ko je v drugi polovici poglavij pri romanu večja uravnoteženost med pospešenim, hitrim in zelo hitrim tempom, pri povesti pa gre večinoma za pospešeni tempo;
- v romanu je oseb in prizorišč več kot v povesti.

Izziv: mala preštevna analiza slovenskih romanov in povesti

Oblikujte preprosto preštevno analizo slovenskih romanov tako, da sledite navodilom:

- Izberite nekaj ključnih slovenskih avtorjev oziroma avtoric iz 19. in z začetka 20. stoletja stoletja, ko se je pričela vzpostavljati daljša slovenska proza z romani in povestmi;
- na spletu ali v tiskanih virih preverite, koliko naslovov povesti in romanov je posamezen avtor oziroma avtorica izdal/-a tekom svojega življenja;
- v besedilnem urejevalniku oblikujte tabelo, kamor vpišete ime in priimek avtorja ter število izdanih del po naslednjem ključu:

Ime in priimek	Število izdanih romanov in povesti
Josip Jurčič	38
...	...

- poigrajte se v besedilnem urejevalniku in napravite različne vizualizacije podatkov v tabeli. Dodajte še druge podatke, v tabeli ločite na primer na primer romane in povesti ter dodajte informacije o drugih literarnih vrstah in zvrsteh.

Sodobni literarni raziskovalec oziroma raziskovalka, ki se pri svojem delu poslužuje digitalnih metod, razume kompleksna vprašanja z različnih področij literarne vede, pa tudi možnosti, ki jih nudijo digitalna orodja. V tem

smislu gre za kombiniranje kvantitativnih in kvalitativnih raziskovalnih metod.

“Sprememba, ki jo doživljamo, je ravno v tem, da je kvantitativne in kvalitativne dokaze vse lažje kombinirati, kar zabriše meje med disciplinami. Zdaj delamo na metodološkem kontinuumu, ki sega od zgodovine in literature prek jezikoslovja ter sociologije. Raziskovalci se seveda še vedno svobodno specializirajo za dele kontinuumu in specializacija je še vedno dragocena. Vendar nas nič ne ovira pri tem, da bi dosegali širši razpon. Ker je tudi področje humanistike kontinuum, bi morali svobodno uporabiti kakršno koli kombinacijo metod, ki nam omogoča, da določen problem rešimo.”¹⁶

Kvantitativni podatki so se v literarni vedi uporabljali že prej, vendar je z vzponom digitalnih informacijskih tehnologij drugače, saj imamo na voljo digitalne zbirke podatkov in samodejno iskanje informacij. Kar se tiče jezika in sloga, je mogoče izvajati raziskave, o katerih so pretekle generacije lahko le sanjale.¹⁷ Računalniške analize se dobro

16. Underwood 2016

17. Moretti 2013: 212

odrežejo pri primerjavah, saj nam lahko statistične metode pomagajo interpretirati besedila, posebej kadar preverjamo hipoteze in interpretacije, ki so v preteklosti temeljile na človeškem branju in nekoliko naključnem opazovanju zanimivih lastnosti besedil.¹⁸ V določenih primerih bi sicer lahko do rezultatov, do katerih pridejo raziskovalci z digitalnimi metodami, prišli tudi brez uporabe teh metod. Če bi hoteli na primer pokazati pogostost rabe pridevnikov v slogu določenega avtorja, bi prebrali vsa njegova dela, si označili pridevnike in napravili preštevno analizo, vendar bi nam to vzelo veliko časa. Če bi to opravili z računalniškim programom, bi torej izvedli isti tip raziskave, ki pa bi lahko bila zaradi drugačne metode širša, kompleksnejša in hitrejša.¹⁹ Poleg tega se z digitalnim vidikom spreminja način izvajanja raziskav v humanistiki, in sicer v smeri sodelovanja raziskovalcev različnih profilov in strokovjakov na področju informacijskih tehnologij. Računalniško branje v literarno vedo, ki jo je tradicionalno zaznamovalo kvalitativno raziskovanje, prinaša kvantitativne pristope. Primerjavo med tema tipoma pristopov prikazuje naslednja preglednica:

18. Hockey 2000: 66

19. Eve 2022: 8

Tabela: Razlike med kvantitativnim in kvalitativnim raziskovanjem v literarni vedi

	kvantitativno raziskovanje	kvalitativno raziskovanje
odnos med teorijo in raziskovanjem	strukturirane stopnje, ki si logično sledijo, deduktivno raziskovanje (teorija je pred opazovanjem)	odprto, interaktivno
koncepti	operativni	orientirani, odprti, razvojni
psihološka interakcija med raziskovalcem in opazovanim fenomenom	znanstveno raziskovanje, izločeno, nevtravno	empatična identifikacija in perspektiva do objekta študija
reprezentativnost	statistično reprezentativen vzorec	posamični primeri, ki niso statistično pomembni
matematične in statistične tehnike	intenzivno uporabljene	niso uporabljene
implikacije rezultatov	posplošene	specifične

Vir: Salgaro 2018: 51

3.3 Vizualizacije podatkov

V digitalnih raziskavah literature se pogosto pojavljajo novi tipi vizualizacij podatkov, ki so jih v okviru literarne vede tradicionalno manj uporabljali. Čeprav se spoznanja na področju literarne vede širijo predvsem z besedili –

članki in knjigami raziskovalcev in raziskovalk -, pa se je v preteklosti občasno posluževala tudi ilustracij, ki so podkrepile njene vsebine, vendar so sodobne vizualizacije literarnih raziskav v okviru digitalne humanistike drugačne, saj načeloma temeljijo na kvantificiranih podatkih, kar na primer velja za najbolj tipične prikaze: mape, zemljevide in drevesa.²⁰ Gre za dvo- ali tridimenzionalne prostorske oblike, ki poglobljajo raziskavo, besedilo pa ilustrirajo ali ga s svojo informativnostjo celo nadomeščajo.²¹ Ti prikazi bodisi predstavljajo podatke in njihovo interpretacijo bodisi posredujejo novo znanje. V spletnem okolju so vizualni prikazi lahko tudi interaktivni in nas vabijo k raziskovanju tematik, ki nas še posebej zanimajo.

Med grafičnimi predstavitvami podatkov, ki se uporabljajo v literarni vedi, Miran Hladnik navaja naslednje tipe:²²

- grafikon: histogram, črtni, stolpčni, palični, tortni graf raztrosa (imenovan tudi korelacijski diagram ali raztreseni oziroma razsevni grafikon), statistična distribucija (verjetnostna razporeditev), graf funkcije;
- diagram: časovni ali kronološki, drevesni (hierarhični diagram ali dendrogram, ki ga generira klastrska analiza ali grupiranje), omrežni, organizacijski, diagram poteka, genealoški in Vennov diagram;

20. Eve 2022: 67–68.

21. Stanley 2003: 9

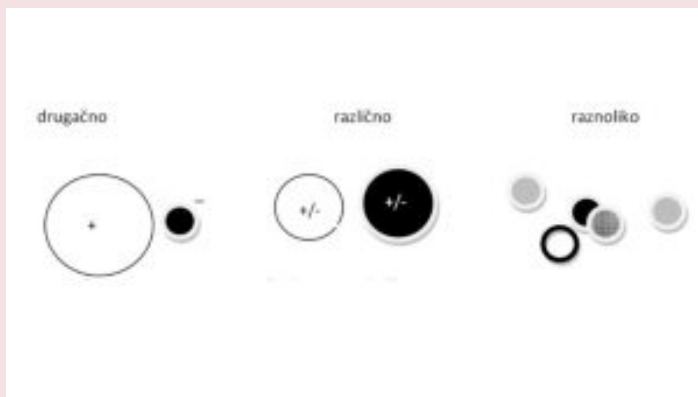
22. Hladnik 2016

- kartogram, shema, piktogram.

Primer: piktogram drugačnosti

Od abstraktnih pojmov do zgodovinskih dogajanj – na področju humanistike se pogosto pojavljajo vsebine, ki jih lahko prikažemo s poenostavljeno sliko, tako pa kompleksne vsebine predstavimo na razumljiv način.

“Literarni razvoj se je dogajal z izumljanjem vedno novih, od tradicije drugačnih načinov sporočanja in ob stalnem prehajanju drugačnega v normalnost *mainstreama* in v kanon,” piše Miran Hladnik, ki s sliko ponazori razlike med pojmi drugačno, različno in raznoliko v slovenščini.²³



(vir: Hladnik 2016)

Izziv: Izdelajte piktogram

S piktogramom, ki ga izdelate v enostavnem program Slikar, ponazorite razdelitev literarnih vrst, zvrsti in žanrov v literarni teoriji. Lahko si pomagate z viri, na primer monografijo Janka Kosa: *Literarna teorija*.

Primer: vizualizacije poezije na grafitih v Rimu

Književnost se dotika tudi mnogih fiktivnih in resničnih geografskih lokacij, kjer potekajo literarne zgodbe. Literarna veda se je že tradicionalno ukvarjala s prostorskimi razsežnostmi tako v okviru fiktivnih literarnih svetov kakor tudi v okviru avtorskih biografij. Zemljevidi, povezani s književnimi vsebinami, so bili del literarnih študij, še preden se je v zgodnjih 90. letih 20. stoletja pojavil t. i. prostorski obrat v literarni vedi, ki je odprl nove možnosti za stike med literarno vedo in geografijo.²⁴ Na preseku med literarno vedo, geografijo in digitalno humanistiko nastajajo projekti digitalnega mapiranja podatkov s področja literature. Pristopi digitalnega mapiranja omogočajo globlje razumevanje kronologije in medsebojnih povezav med literarnimi liki.²⁵ Hkrati se na ta način izpostavijo vrzeli med fiktivnimi svetovi, ki prostor

24. Urbanc in Juvan 2012: 297–316

25. Eve 2022: 108

uporabljajo kot sredstvo za strukturiranje pripovedi, in zemljevidi, ki skušajo predstavljati zunajbesedilni geografski svet.²⁶

Mapiranje pa se ne povezuje nujno s fiktivnimi zgodbami literarnih del, ampak tudi s procesi ustvarjanja, ki so nujno povezani z določenim konkretnim okoljem. Primer je spodnji zemljevid Rima, kjer so s pikami označeni primeri poezije na grafitih na določenih lokacijah mesta, z različnimi barvami pa so označene tematike besedil (zelena je eksistencialna tema, vijolična je socialna tema, modra pomeni tematiko umetnosti, rumena ljubezensko temo, siva pa temo spolnosti).²⁷

26. Eve 2022: 128

27. Vitali 2018: 105–136



Slika: Zemljevid Rima z označenimi lokacijami pesmi-grafitov; vir: Vitali 2018: 124

Izziv: oblikujte zgodbo na platformah StoryMaps in Tiki-toki Online Timeline Maker

- Med svojimi najljubšimi literarnimi avtorji ali avtoricami izberite primer književnih besedil, ki omenjajo resnične zemljepisne lokacije. Vpišite

jih v spletni zemljevid na platformi [StoryMaps](#) in jih povežite z različnimi medijskimi vsebinami, na primer s fotografijami krajev, ilustracijami likov fiktivnih zgodb različnih ustvarjalcev, zvokovnimi posnetki branja odlomkov, ki govorijo o teh krajih, video posnetki okolice itd.

- Na spletni strani [Tiki-toki Online Timeline Maker](#) oblikujte časovnico življenjske zgodbe izbrane avtorice ali avtorja. Uporabite čim več virov in vključite slike različnih oseb, krajev, književnih izdaj its. Poigrate se lahko z možnostjo oblikovanja tridimenzionalne časovnice, ki gledalca popelje skozi zgodovinski razvoj v obliki prostorskega premika, pa tudi z interaktivno multimedijško časovnico, ki obiskovalcu ponudi informacije v različnih oblikah.

4 Kritični pogledi na digitalno humanistiko v literarni vedi

Digitalna humanistika je dokaj novo področje raziskav, ki se na tehnološkem polju hitro razvija in ki ga zaznamujejo eksperimenti, zato se srečuje tudi s težavami – od družbenih krivic, ki jih odražajo algoritmi, do vprašljivih rezultatov analiz z digitalnimi orodji.

Nekatere izmed odmevnih kritik digitalne humanistike razgrnemo v nadaljevanju.



Za digitalne raziskave književnosti je nujno, da so literarna dela digitalizirana in shranjena v ustreznih formatih, ki pa najpogosteje niso prosto dostopni za javnost, oziroma je pretvorba dokumentov, ki so zaščiteni z avtorskimi pravicami, v te formate v nekaterih državah kriminalizirana.²⁸ To je ovira pri digitalno humanističnih raziskavah književnosti in morda tudi eden od razlogov, da se mnoge raziskave posvečajo starejšim tekstom, ki so jim avtorske pravice že potekle.

Posamezni raziskovalci in raziskovalke opozarjajo, da digitalna humanistika ne upošteva v zadostni meri kategorij rase, etničnosti, spola, seksualnosti in razreda. Po njihovem mnenju se zaradi ustroja računalniških algoritmov pojavljajo krivične prakse pri delovanju umetne inteligence in analizi velikih podatkov, zato sta v boju proti neenakostim ključni odprta dokumentacija in razložljivost podatkov za javnost.²⁹ Čeprav eden od kritičnih pogledov digitalnim metodam analize književnosti očita apolitičnost, pa praksa – na primer feministične študije na podlagi digitalnih metod – kaže,

28. Paul Eve: 2022: 11

29. Prescott 2023

da se raziskovalci in raziskovalke z računalniškimi algoritmi lotevajo tudi globoko političnih tematik.³⁰

Še ena kritika metod digitalne humanistike je, da ustvarja vzorce in modele tam, kjer niso pomenljivi: “/š/tevilke lahko vržemo v največje računalniške oblake, kar jih je svet kdaj videl, in pustimo, da statistični algoritmi najdejo vzorce tam, kjer jih znanost ne more.”³¹ Eden od odgovorov na ta kritični vidik je, da programi z velikimi podatki ravnajo na temelju statističnih in matematičnih teorij, ne pa naključno.³²

Digitalna orodja se sicer pogosto naslanjajo na različne teorije, na polju literarnih raziskav pa tudi na vrsto metod literarne vede.³³ Kljub temu kritike pravijo, da digitalne raziskave književnosti ne temeljijo na ustaljenih literarnovednih izhodiščih in zgolj prikazujejo določene rezultate analiz, namesto da bi s temi metodami preizkušale hipoteze in z uporabo statističnih orodij poskušale prikazati vzročno-posledične povezave med pojavi.³⁴

30. Eve 2019

31. Anderson 2008

32. Mayer-Schönberger in Cukier 2013

33. Warwick 2016: 538–52

34. Nan 2019: 605

1.

DIGITALNO LITERARNO/ E-LITERATURA

Digitalna kultura tradicionalnim oblikam jezikovnega izraza zapoveduje in hkrati prepoveduje preskok na internet, kjer se literaturo prodaja, prevaja, komentira in, morda najpomembneje, stimulira njena srečevanja z drugimi mediji in umetniškimi oblikami. Kakšne izrazne in komunikacijske možnosti se ponujajo literarnemu izrazu v omrežjih računalnikov in mobilnikov? Kateri so morebitni zadržki pred novomedijsko evforijo na literarnem področju? Kaj nudijo specifične oblike digitalne in internetne literature tisti ali tistemu, ki ustvarja prozo, poezijo ali dramo?

1 Posebnosti digitalne literarne pisave¹

Kljub pomislekom, da bodo tehnološke novosti, kot so

1. Pričujoče poglavje se opira na prispevek Purg in Pranjčič 2023

internet in mobilniki, omejile ali celo prekinile tako pesniško kot tudi prozno in dramsko literarno izražanje, se je izkazalo, da je (še posebej umreženo) računalništvo že v fazi interneta, ter nato po letu 1993 v obliki svetovnega spleta pravzaprav obogatilo literarno prakso in njene raznolike izraze.²

Dileme o tem, ali bo papirna knjiga izumrla in kaj bo prinesla elektronska knjiga, niso več relevantne, ključno pa je vprašanje, kako razumeti nove oblike pisav ter proizvodnjo ali vznik, pa tudi kroženje znanja, značilnega za današnjo digitalno kulturo. Pri tem je potrebno na tehnološko informiran način kritično obravnavati nova razmerja družbene moči in kulturnih potencialov ter nujno tudi razvijati strategije upora znotraj e-literarnih praks, ki včasih resda kritično napadajo, spet drugič pa tudi zgolj brezzobo zabavajo, se larpurlartistično naslanjajo nad ekonomskimi oziroma političnimi dominantami digitalnih platform ali konec koncev neogljenjo tekmujejo z njimi.

Izziv: zbirke Organizacije za elektronsko literaturo

Preglejte vse dosedanje [zbirke Organizacije za elektronsko literaturo](#) (ELO) od leta 2006 dalje. Iz vsake od zbirk izberite po eno delo, ki vas pritegne. Izbrana e-literarna dela primerjajte med seboj po tem, kakšno spletno

ali drugo tehnologijo uporabljajo in kako, katerih tem se lotevajo, kako uporabljajo jezik in kako sliko ter morda zvok, pa tudi kako se ti dopolnjujejo, kako zastavijo različne oblike interakcij znotraj dela oziroma z bralko ali bralcem, kako se povezujejo z zunanjim svetom, morda z drugimi novomedijskimi okolji, platformami, kako bi bilo mogoče opisati njihovo estetiko ipd.

2 Razvoj literarno-umetniških praks

Bliskovit razvoj jezika in z njim povezanih izraznih sredstev se v kontekstu novih medijev odraža že v vsakodnevni komunikaciji. Prav tako so se vrste znakov in pisav glede na tiste, ki smo jih uporabljali v 20. stoletju, z digitalnimi orodji in platformami pomembno razširile – in sicer v smislu, da niso več zgolj vizualne ali avdiovizualne, temveč tudi trirazsežnostne, interaktivne, s tem pa (lahko) občinstvo vključijo v povratno zanko z besedilom še na dodaten, fizičen način, hkrati pa pri tem pogosto reflektirajo tudi medij kot tak. To pomeni, da navigacijske možnosti novih medijev bralki ali bralcu omogočajo izvajanje dejanj – procesov odločanja na podlagi avtonomne izbire – v nekakšni vsaj navidez odprti obliki pripovedi, kjer se je bralka ali bralec samostojno odloča o poti branja, torej izbiri hipertekstualnih linkov ter načinu interakcije z avdiovizualnim gradivom. Pri tem ne gre le za klikanje po spletiščih ali drsanje po aplikacijah, temveč tudi za celotelesne interakcije v instalcijskih postavitvah,

bodisi s tehnologijami navidezne oziroma razširjene resničnosti (različni prilagojeni vmesniki, npr. igralne palice ali konzole) bodisi s prostim telesom (zaznava oziroma zajem gibanja, npr. sistem Kinect).

Z razmahom mobilne telefonije in mobilnikom kot zmogljivim multimedijским vmesnikom, torej računalnikom v malem, zadnji dve desetletji v ospredje močno stopa gibljiva slika, pa tudi možnosti telesnih potopitev, kot so navidezna resničnost (VR) ali mešana (MR) oz. razširjena (AR) resničnost. Literatura se tako v okviru vsesplošnega razcveta medijske oziroma digitalne kulture zliva in prepleta z nebesed(il)nimi kulturnimi kodi, kot so glasba, moda, večrazsežnostni (video)film, igričarstvo.³ Med e-litearne prakse tako lahko štejemo dela, ki izhajajo iz ustvarjalnega ukvarjanja z digitalnimi tehnologijami in imajo za posledico takšne primarno jezikovne izdelke (ali tudi procese), ki vključujejo računalništvo, digitalno sliko, animacijo, večpredstavnost, interaktivnost in/ali računalniška oziroma mobilna omrežja ter odstopajo od estetskih omejitev in kulturega okvirja natisnjene ali računalniške strani premočrtnega besedila.⁴

Elektronsko literaturo lahko razumemo kot na pisanje osredotočeno umetnost, ki vključuje izrazni potencial elektronskih in digitalnih

3. Strehovec 2016

4. Rettberg, 2019

medijev.

Naraščajoča raznolikost praks elektronske literature – to razumemo kot »na pisanje osredotočeno umetnost, ki vključuje izrazni potencial elektronskih in digitalnih medijev«⁵ danes zajema vse več neliterarnih besedil (npr. podkasti in digitalno novinarstvo, transmedijsko ter interaktivno pripovedništvo). Kljub temu ostaja bistvena komponenta e-literature pisava oz. “umetniška angažiranost pisnega jezika v digitalnem mediju.”⁶ Tretja generacija e-literature, ki se je začela leta 2005 z razmahom t. i. spleta 2.0 (premik k uporabniško generiranim vsebinam) ter postopno prevlado družbenih omrežij in traja še danes,⁷ se razvija v kontekstu novejših tehnologij, kot so zaslони na dotik, dinamično generirane spletne vsebine in družabne aplikacije, tudi memi, pred kratkim pa še tehnologije veriženja blokov in nenazadnje veliki jezikovni modeli, za katere več ne potrebujemo posebnih medijsko-produkcijskih veščin ali podrobnejših znanj o njihovem delovanju oziroma ta tako na izkušnjo branja kot na izkušnjo pisanja ne morejo (več) pomembno vplivati.

5. Flores 2021: 27

6. Flores 2021: 28

7. Flores 2021: 28

digital(izira)na literatura

Tako človeški kot strojni jezik je mogoče zapisati s pomočjo znakov.

Koda črk in koda števil sta medsebojno izmenljivi oziroma prevedljivi.

Digitalna medijska tehnika deluje kot bodisi specifični bodisi alternativni produkcijski, transportni, pomnilni in predvajalni medij literarnega.

* disketa, CD-ROM, DVD, trdi disk, "elektronska knjiga" itd. kot lokalni materialni pomnilniki (pretežno prenosljivi predmeti na konkretnih krajih);

* internet in Vsesvetovni splet (WWW) kot globalna platforma medsebojno povezanih računalnikov (iz "slehernega" kraja in ob "slehernem" času dostopni pretežno neprenosljivi predmeti – procesorji in pomnilniki);

* novi mediji postanejo nosilci literarnega: mobilna telefonija (mobilniki, dlančniki), klasična (elektronska) in spletna (digitalna) televizija ter radio, računalniško vodene prostorske instalacije, performansi, računalniške ter spletne igre itd.

literatura na internetu

("tradicionalni" medij literature v "novem" umreženega računalnika)

Povečini delujeta internet in Vsesvetovni kot globalno (na večih mestih hkrati) dostopna knjigarna, torej v distribucijski funkciji onstran produktivnih in receptivnih posebnosti. Linearni "enomedijski" literarni tekst ter metatekst sicer izključno novih transakcijskih, transportnih možnosti digitalnega zapisa.

Vir: Purg, 2005

3 Dva primera iz slovenskega e-literarnega prostora

Tovrstno znanje sta nedvomno potrebovala začetnika in še danes najvidnejša predstavnika slovenske spletne umetnosti in e-literature – Vuk Ćosić in Jaka Železnikar. Oba avtorja svoje poglobljeno poznavanje interneta in računalništva z bolj ali manj ironično igrivostjo uporabljata tako za eksperimentiranje z besedno formo kot za družbeno-kritični komentar. Ćosić in Železnikar sta začela ustvarjati v t. i. drugi generaciji elektronske literature, katere začetek lahko postavimo v obdobje sredi 90. let 20. stoletja, ki označuje širši razmah rabe interneta na ravni svetovnega spleta; to (razmeroma) sodobno elektronsko literaturo označujeta multimedijškost in interaktivnost.

Izmed njunih del na presečišču e-literature in spletne umetnosti bomo izpostavili predvsem dva projekta, ki sta zanimiva zaradi vsebinskih in tehničnih posodobitev, ki sta jih avtorja izvedla v razponu več kot dveh desetletij. Prav kontinuiteta njunega dela skozi različne generacije e-literature lahko prikaže značilnosti in relevantnost poetike ter specifične (diksutiranja, kritiziranja) pisave v določenem elektronskem mediju, pa tudi refleksije družbenih vprašanj in realnosti časa nastanka določenega dela.

3.1 Jaka Železnikar: *Interaktivacija* (1997–2007–2017)

Jaka Železnikar je v skoraj treh desetletjih ustvaril prek petdeset del e-literature in spletne umetnosti, med katerimi prednjačijo tako formalni kot semantični eksperimenti, ki se odzivajo na ključni lastnosti spletnega medija, tj. interaktivnost in večmedijskost. Železnikarjevo prvo vidnejše delo e-poezije je *Interaktivalija* (1997), ki je nastalo še v kontekstu zgodnjega obdobja spletne umetnosti oz. t. i. net-arta, ki so ga slovenski avtorji (med njimi še posebej že omenjeni Čosić) pomembno zaznamovali na globalni ravni. Ta “prva zbirka spletne, interaktivne poezije v slovenskem kulturnem prostoru”, kot jo poimenuje sam avtor, na igriv, a nazoren, možnostim novega medija naklonjen način predstavi raznolike različice interakcije s pesniškim besedilom. Vanj uporabnik lahko posega in ga sooblikuje s svojimi izbirami prek računalnika, opremljenega z zaslonom, tipkovnico in miško. V tako interakcijo je uporabnik pravzaprav primoran, saj brez nje literarnega dela ni: udeležba aktivnega branja-kot-klikanja oziroma vtipkovanja je bistvena in obvezna sestavina tega literarnega dela. Vendar pa povezava s spletom pri tem ni bila nujno potrebna – delo je v celoti delovalo tudi z diskete. Nekaj let kasneje je nastalo delo *Pesem za Echelon* (2001), kjer je Železnikar uporabniku na igriv način ponudil algoritmično (vsakokratno) prilagojeno pesem za internetni prisluškovalni program Echelon, ki so ga v tistem času razkrinkali in je prvič širše razprl vprašanja o zasebnosti v internetni komunikaciji.

Podobno kot Čosić se je tudi Železnikar nemudoma odzival na pojavitve novih internetnih tehnologij, npr. Twitterja (*And Your Hair/Tweet*, 2008), pozneje pa tudi

na vprašanja podatkovne varnosti in politične ali gospodarske (zlo)rabe osebnih, predvsem prostorskih podatkov, ki jih omogočajo npr. geolokacijske tehnologije prenosnega telefona (*Fragments of Distances*, 2010; *Private Property – Access Forbidden*, 2021). Njegov projekt *Interaktivnija* se je pričel leta 1997, leta 2007 pa je doživel drugo izvedbo z e-knjigo *Interaktivnij*. Gre za spletno pesniško zbirko generativne vizualne poezije, katere “drobci”, izdani v e-knjigi, predstavljajo avtorjev izbor posnetkov fragmentov iteracij vsake od osmih e-pesmi, sestavljenih iz tipografskih znakov s pomočjo generativnega algoritma, ki po besedah avtorja “ob vsakem ogledu pesem, ki ohranja osnovno strukturo in nekaj semantičnih elementov, nekoliko vizualno spremeni.”⁸ Desetletje pozneje je Železnikar izvedel trajno intermedijsko postavitev *Spomini Interaktivnija in interaktivna spletna poezija* (2017) v ljubljanski Vodnikovi domačiji, kjer lutka-robot po imenu Interaktivnij (z uporabo geolociranja) na določenih točkah ob zgradbi pripoveduje o zgodnjem internetu, z njim povezano interaktivno spletno poezijo ter o Valentinu Vodniku.



Slika: Zaslonski posnetek spletne strani Jaka Železnikar – Interaktivacija – 1997; vir: zajem zaslona, uporaba z dovoljenjem avtorja



Slika: Lutka-robot "Interaktivnij" – 2017; vir: fotografija iz arhiva avtorja, uporaba z dovoljenjem

3.2 Vuk Ćosić: *Nacija – Kultura* (2000–2022)

Ćosića štejemo med pionirje spletne umetnosti, s svojimi deli se predstavlja v številnih mednarodnih

zbirkah in pregledih intermedijske ter spletne umetnosti, pogosto pa svoje poglobljeno poznavanje interneta uporablja tudi v aktivistične namene. Njegovo delo *Nacija – Kultura* je v svoji zgodnji različici nastalo leta 2000, predstavljeno je bilo ob dvestoti obletnici rojstva Franceta Prešerna v okviru razstave Pevec, ne bogat ali sloveč v Narodnem muzeju Slovenije. Leta 2022 je avtor na pobudo Moderne galerije zamisel znova obudil tako s tehnološkimi kot konceptualnimi posodobitvami, ki odražajo moč in relevantnost elektronske literature. Namesto izvorno uporabljenega slovenskega brskalnika Matkurja je avtor kot izvor dejanskih uporabniških vnosov tokrat uporabil družabno platformo Twitter ter kratka besedila obdelal z orodji za ravnanje z velepodatki in tehnologijami strojnega učenja. Pri tem so pod vplivom algoritma nastale sonetne besedilne forme, ki so vsebino zajemale tako rekoč neposredno iz javnega političnega diskurza, s tem pa so delovale politično kritične – četudi izmeščene iz polja avtorjevega mnenja skozi uporabo tehnologij strojnega učenja – kot tiste pred dvema desetletjema. Te so zajemale iskalne nize na slovensko govorečem spletu zgolj po časovnem načelu (opoldan in opolnoči). Pri tem je zanimivo, da sta obe iteraciji dela doživeli tako večjo galerijsko postavitev kot tudi tiskano knjižno izdajo. Gre za premik od celotnega avtorjevega besednega zaklada ali pač miselnega oziroma notranjega sveta k paradigmi besedilnosti, ki temelji na računalniški zbirki podatkov, pri čemer pride do poudarka na zbiranju, primerjanju in pozicioniranju informacij ter celo njihovemu samodejnemu generiranju.

S svojim evidentno kulturno- in tehnološko-kritičnim

delom *Nacija – Kultura* (pomišljaj v naslovu je bil zapisan tudi z besedo »minus« na način Nacija Minus Kultura) avtor skozi preplet konceptualno-produkcijske ironije ter idejne (hiper)doslednosti razkriva mnogotere razpoke v sodobnem medijskem (“duh ljudstva”) pa tudi političnem diskurzu (“strojni dialog”), ki odražajo vse močnejše kontraste, a hkrati tudi naraščajočo zmešnjavo oziroma vsaj redefiniranje vrednot in pomenskih vrednosti v slovenskem jeziku. Pri tem Čosić svojo poetiko dodatno plasti z rabo posebej za projekt razvite tipografije, ki temelji na bohoričici, in (strojno podprto) apropiacijo oblike soneta. V tem kontekstu sonet ne predstavlja le ultimativnega formalistično osmišljenega slovenskega jezika, ampak tudi konstituira mit o njegovi emancipacijski vzpostavitvi.

Zanimivo je, da se tako Železnikar kot Čosić v svojih delih pomembno navezujeta na dva od ključnih avtorjev slovenskega razsvetljenskega oziroma romantičnega pesništva, ki sta prelomno zaznamovala ustvarjalnost slovenskega jezika tako v njenem družbenem oziroma kulturnem kot tudi umetniškem pomenu. Kljub temu da se ob naglem razvoju sodobnih tehnologij in njihove umetniške apropiacije dve desetletji spletne ustvarjalnosti zdita paradoksalno daljši od dveh stoletij, je daljnosežnost tukaj opisanih sprememb in prelomov več kot očitna. Ob tem velja še omeniti, da se Železnikar ob raznolikih in vse vidnejših kuratorskih vlogah redno udejstvuje tudi kot gostujoči pedagog mladih digitalnih umetnikov, Čosić pa se vedno znova izpostavi kot kulturni komentator in politični aktivist, kar nenazadnje tudi

njunemu e-pesniškemu delu odmerja relevantno mesto v družbi.

Izziv: e-literarno delo

Zamislite si e-literarno delo, ki uporablja novomedijska okolja in orodja na način, ki vam je najbolj blizu. Katere literarne zvrsti, oblike in vsebine prepoznate? Kakšen jezik uporablja, kakšnih tem se loteva, kako predstavi svoja vprašanja, nastopajoče ...? Kakšna znanja in veščine ter opremo bi potrebovali za izdelavo takšnega e-literarnega dela?

4 Kritični pogled na področje e-literature

Ne moremo mimo dejstva, da na področju e-literature (ter njenega popisovanja, odkrivanja, zgodovinenja) prevladujejo avtorji (redkeje avtorice) iz tehnološko razvitih držav in družbenih okolij. Na tem področju kulturne in družbene manjšine ter ne-bele ustvarjalce in ustvarjalke najdemo še razmeroma redko, četudi se trend v zadnjem desetletju obrača – a spet praviloma le v bolj razvitih družbenih okoljih! Nenazadnje gre za posledico kritike in vsaj začetnega uravnovešanja zgodovinsko privilegirane dostopa do digitalnih tehnologij (t. i. digitalni razkorak), neenakega izobraževanja na področju

programiranja, digitalne pismenosti in naprednih digitalnih veščin ter medijske kulture.

2.

ODDALJENO BRANJE KNJIŽEVNOSTI

Oddaljeno branje, imenovano tudi makroanaliza oziroma algoritemska kritika je analiza velikih zbirk literarnih besedil, lahko tudi na tisoče ali več, z uporabo računalniških orodij. Za razliko od bližnjega branja, ki ga izvajamo sami, se oddaljeno branje kot kvantitativna metoda raziskovanja književnosti naslanja predvsem na sodobne informacijske tehnologije.

Termin oddaljeno branje je okrog leta 2000 vpeljal Franco Moretti, uveljavila pa sta se tudi izraza makroanaliza, ki ga je uvedel raziskovalec Matthew L. Jockers, in algoritemska kritika raziskovalca Stephena Ramsayja. Oddaljeno branje ima sicer dolgo zgodovino, ki sega onkraj uporabe spleta in računalniških tehnologij.¹ V sodobnem smislu gre za uporabo kvantitativnih metod

1. Underwood 2017

pri raziskovanju posameznih književnih besedil, večjih literarnih korpusov ali drugih podatkov, povezanih z literarnim sistemom. Metode temeljijo na digitalnih orodjih in statističnih izračunih. S tem tipom računalniške analize je mogoče razpoznati vzorce v literarnih delih, na primer teme, sloge pisanja, čustva in gradnjo književnih likov, s čimer lahko bolj razumemo literarno ustvarjanje skozi čas. Pri sodobnem oddaljenem branju, ki ga raziskovalci in raziskovalke izvajajo z računalniškimi tehnologijami, ne gre le za preprosto štetje besed ali obsegov, ampak za razumevanje pisanje kot polje odnosov, ki jih je mogoče modelirati z uporabo enačb, kar pa ni zgolj premik v okviru literarne vede, ampak je del intelektualne zgodovine naše sodobnosti.²

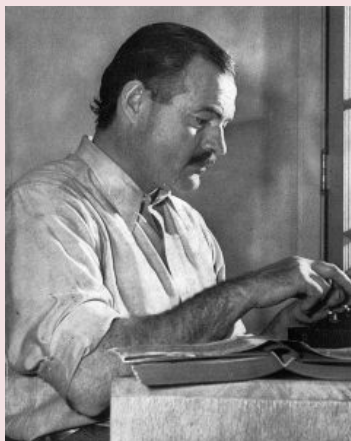
1 Metode oddaljenega branja

Z razvojem tehnologije se širijo tudi možnosti oddaljenega branja književnosti. Te metode uporabljajo strojno učenje kot področje umetne inteligence, ki se osredotoča na razvoj računalniških sistemov, ki se lahko "učijo" in izboljšujejo iz izkušenj ter podatkov, namesto da bi jih programerji ročno programirali. Tako samodejno odkrivajo vzorce in pravila iz podatkov brez programiranja za vsako specifično nalogo. Poleg tega lahko metode oddaljenega branja temeljijo tudi na statistični analizi besed, tj. metodi, ki se uporablja za kvantitativno analizo besed in besednih zvez v književnih delih. Štetje pogostosti besed ali besednih zvez v literarnem korpusu pokaže, v kolikšni meri so določeni izrazi razširjeni v različnih literarnih obdobjih ali žanrih. Tudi na ta način

je mogoče odkriti vzorce in informacije, ki v besedilih “s prostim očesom” niso vidni. Nadalje metode oddaljenega branja za analizo literarnih del uporabljajo tudi računalniško obdelavo naravnega jezika (tj. NLP), ki omogoča globlje razumevanje vsebine danega korpusa književnih besedil. Na ta način je mogoče na primer ugotavljati čustvenost literarnih besedil ali njihovih delov, pa tudi čustvene vsebine, ki jih s svojimi glasovi v delih izražajo literarni liki ali pripovedovalci.

Primer: Specifičen literarni slog Ernesta Hemingwayja

Nobelov nagrajenec Ernest Hemingway je znan po tem, da je pri pisanju uporabljal ožji nabor besed, kar pomeni, da besedišče v njegovih literarnih delih ni raznoliko. To potrjuje tudi primerjalna analiza, ki je preverjala delež edinstvenih besed v književnih odlomkih, kar imenujemo bogastvo besedišča. Večje bogastvo besedišča pomeni manj ponavljanja istih besed v književnem delu, manjše leksikalno bogastvo pa pomeni več ponavljanja.



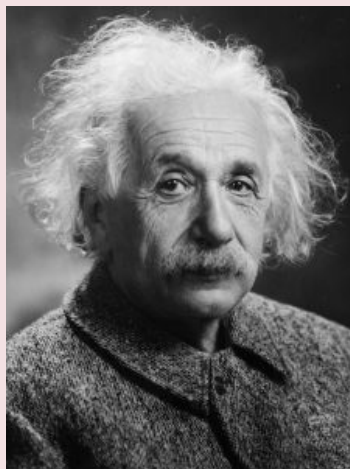
Hemingwayjeva dela tako v primerjavi z deli podobnih avtorjev kažejo najmanjše bogastvo besedišča.³

Slika desno: Ernest Hemingway; vir: javna domena, Wikimedia Commons

Oddaljeno branje lahko izvedemo tudi brez pomoči računalniške tehnologije, in sicer s preštevniimi metodami, ki vključujejo sistematično zbiranje določenega tipa informacij. Na ta način dobimo nov vpogled v predmet našega zanimanja.

Primer: Lastnosti fiktivnih likov, ki se ukvarjajo s fiziko

Daniel Dotson v preštevni analizi preverja, kako so fiziki in fizičarke prikazani v romanih, pa tudi v filmih in televizijskih programih. Analiza je zaobsegla osemdeset likov in pri njih preverjala prisotnost določenih lastnosti. Raziskava je pokazala, da je najpogostejša skupna točka teh likov pogum, kar je značilnost okrog 72 % obravnavanih likov. Naslednje lastnosti, ki se ponavljajo, so trma, odsotnost stika z realnostjo, obsesivnost in arogantnost. Fizičarke imajo sicer v obravnavanih reprezentacijah pozitivnejše lastnosti kot fiziki.⁴



Slika desno: Albert Einstein; vir: javna domena, Wikimedia Commons

2 Oddaljeno branje drugih kulturnih pojavov

Oddaljenega branja pa ne uporabljajo le raziskovalke in raziskovalci književnosti, ampak tudi drugih področij: tako je mogoče na primer “prebrati” na primer objave na družbenih omrežjih, ki kažejo razpoloženje in odnos ljudi do določene tematike. Spletna stran [Pudding](#) na primer z

4. Dotson 2009

oddaljenim branjem spremlja različne tematike sodobne družbe – od raziskav [naslovnice sodobnih popularnih ljubezenskih romanov](#) do [oddaljenega branja ameriških učnih načrtov](#) na področju literature na univerzitetnih programih z namenom identifikacije kanonskih besedil. Oddaljeno branje raziskave aplicirajo tudi na nebesedilne vire, na primer na filme in slike, ki jih lahko računalniški algoritmi že dobro analizirajo. Poleg tega se oddaljeno branje uporablja tudi na področju zgodovine, kjer sicer prav tako prevladuje bližnje branje in natančno ukvarjanje s posameznimi dokumenti, vendar lahko z oddaljenim branjem raziskovalci in raziskovalke razvidijo vzorce v zgodovinskih virih, kar jim daje drugačen vpogled v dogajanje v preteklosti. Zgodovinarji oziroma zgodovinarke lahko na primer v besedilih najdejo ključne besede, kar jim pomaga pri analizi vse bolj obsežnih korpusov, ki nastajajo z digitalizacijo gradiva.

3 Bližnje in oddaljeno branje književnosti: primerjava

Franco Moretti opozarja, da lahko z bližnjim branjem zajamemo le omejeno količino literarnih del, zaradi omejitev možnosti človeškega branja pa ostaja v literarni zgodovini veliko besedil, ki so pozabljena in jih raziskovalci ne obravnavajo, t. i. “veliko neprebrano”, zato je literarna zgodovina za raziskovalca “klavnica literature”.⁵ S tem se je Moretti usmeril v nov predmet preučevanja v literarni vedi z opiranjem na analizo svetovnih sistemov in ga konceptualiziral z uporabo

5. Moretti 2000

modelov grafa, zemljevida in drevesa, ob tem pa razmišljal o uporabi naravoslovnih in družboslovnih ved.⁶ Ker je vsa literarna dela iz preteklosti, pa tudi vse naslove, ki se izdajajo v sodobnosti, nemogoče prebrati, so besedila, ki so v kanonu izpostavljena, omejena, njihova izbira pa je pristranska.⁷ Kot opozarja Andrew Piper, radicionalno literarnozgodovinske raziskave delujejo kot metonimije, pri čemer del nadomesti celoto, kadar posamezno delo ponazarja lastnosti širšega literarnozgodovinskega obdobja, kar pa je le omejen dokaz, da so lastnosti literarnozgodovinskega obdobja zares utemeljene, saj so po mnenju avtorja dokazi pomanjkljivi.⁸ Kljub temu ne zagovarja tega, da bi se posplošitvam odrekli, ampak raziskovalce spodbuja, da reflektirajo konstruiranost znanja, ki ga oblikujejo s tovrstnimi posplošitvami, ki temeljijo zgolj na majhnih delih celote. S tega vidika je znanje vselej oblikovano na podlagi omejenega dela, ki nikoli ne more biti popolnoma povezan s celoto⁹

Če se bližnje branje posveča izbranim literarnim besedilom, ki so navadno kakovostna in/ali kanonizirana, pa lahko z oddaljenim branjem v raziskave vključimo tudi manj kakovostna dela ali književna besedila, ki so bila v zgodovini zaradi določenih razlogov spregledana. Ne gre nujno za to, da bi podobo uveljavljenega kanona spremenili in vanj vključili druga dela, kakor za to, da lahko na ta način na celotno literarno zgodovino in književna dela – kanonizirana in nekanonizirana –

6. Habjan 2012: 92

7. Eve 2022: 9

8. Piper 2016: 3

9. Piper 2016: 4

gledamo drugače.¹⁰ Poleg tega lahko analiziramo hibridna sodobna besedila, na primer pestrost besedišča v tujih reperskih komadih.¹¹

“Bližnje branje (angl. *close reading*) je metonimija za natančno interpretacijo posamične besedne umetnine (najpogosteje izbrane iz kanona), oddaljeno branje pa je metafora za analizo arhiva, tj. množice tekstov, v kateri nas ne zanimata ne pomen ne struktura enega samega besedila, temveč distribucija oziroma pogostnost določene lastnosti, strukture ali elementa v izbrani besedilni množici.”¹²

Bližnje branje temelji na subjektivni analizi posameznega literarnega dela, medtem ko lahko z oddaljenim branjem po drugi strani zajamemo obsežne zbirke besedil in v književnosti iščemo vzorce ter zakonitosti. Oddaljeno branje raziskovalcem omogoča, da se ukvarjajo z večjim številom književnih del, ob tem pa literaturo obravnavajo kot družbeni fenomen, ki se odraža v velikih besedilnih korpusih. Z oddaljenim branjem tako v obsežnih zbirkah literarnih del prepoznavamo vzorce, ki jih z bližnjim branjem posameznega dela ni mogoče

10. Moretti 2000b: 207–208

11. Daniels: *The largest*

12. Juvan, Šorli in Žejn 2021: 57

zaznati. Gre za metodološki prelom, ki spreminja podobo literarne vede, vendar raziskovalci oz. raziskovalke opozarjajo, da je ključno, da računalniške metode oddaljenega branja kombiniramo z drugimi, ustaljenimi metodami raziskovanja književnosti, kar dokazujejo mnoge raziskave, ki kombinirajo oba tipa branja. Oddaljenemu branju pogosto sledi kvalitativna interpretacija rezultatov na podlagi spoznanj literarne vede, spet drugač so literarnovedna izhodišča temelj za hipoteze, ki jih z oddaljenim branjem preverjamo.

Pionirski projekt v 40. letih 20. stoletja

Projekt Index Thomisticus, ki ga je vodil raziskovalec Roberto Busa, se je začel leta 1946 in v 34 letih so analizirali podobnosti v besedilih, ki so bila pripisana Tomažu Akvinskemu.

Analizo so izvedli na podlagi pojavljanja 10.631.980 besed, ki so jih sprva zapisovali na luknjane kartice, računalniški pomnilniški medij, ki so ga v začetnem obdobju razvoja informacijskih tehnologij uporabljali za vnos in shranjevanje podatkov.

Na ta način so besedila prenesli v obliko, ki je bila berljiva za tedanjo računalniško tehnologijo.

Ugotovili so, da je od analiziranih skupno 179 besedil 118 del tudi zares napisal Akvinski, 61 besedil pa ni bilo njegovih in so mu jih napačno pripisali ali pa je šlo za dela, ki jih je začel pisati Akvinski, dokončal pa nekdo drug.

Slika desno: Tomaž Akvinski, avtor Carlo Crivelli (1476), javna domena



Če poteka bližnje branje književnosti linearno, pa lahko z računalniškim, oddaljenim branjem oblikujemo drugačne, nelinearne reprezentacije literarnih del, ki pokažejo nove, s “prostim očesom” nevidne razsežnosti gradiva.¹³ Primer je mapiranje izrazov znotraj književnega

dela, ki ustvarja pregledno strukturno-tematsko sliko besedila, kar je neke vrste novi računalniški formalizem.¹⁴

Ko se v okviru digitalnih metod branja nanašajo na formalizem, imajo raziskovalci in raziskovalke v mislih analitični pristop, ki se osredotoča na notranje dele literarnega besedila, na primer strukturo, jezik in obliko, medtem ko se ne dotika zunanjih dejavnikov, na primer avtorjevega namena ali zgodovinskega konteksta.

“Če besedilo razdelimo na njegove glavne funkcionalne enote, posamezne besede, in nato te besede preštejemo, so rezultati običajno predvidljivi: na začetku seznama glede na število pojavitev v besedilu vedno najdemo veznike in lastna imena. Če ti dve kategoriji vzamemo iz enačbe in poskušamo najti povezave med drugimi, bomo morda videli, kaj je dejanski fokus besedila”¹⁵.

Čeprav se sodobne študije književnosti z digitalnimi orodji osredotočajo na oddaljeno branje velikih korpusov literarnih besedil, pa je mogoče s temi orodji natančno “brati” tudi posamezno delo. Pri tem lahko na primer opazujemo spremembe besedišča v literarnem delu, ki ustrezajo žanrskim premikom, stilne spremembe tekom književnega besedila, uredniške posege v različne izdaje

14. Eve 2019

15. Pojoga idr. 2020

besedila in druge vidike.¹⁶ Oddaljeno branje posameznega literarnega dela pomeni kvantitativno analizo tega besedila z uporabo računalniških metod. Za razliko od bližnjega branja tega romana, ki predvideva preučevanje podrobnosti besedila, si njegovo oddaljeno branje prizadeva odkriti osrednje teme, jezikovne značilnosti ali strukturne elemente, ki s tradicionalnimi tehnikami branja morda niso očitni.

Primer: oddaljeno branje posameznega romana

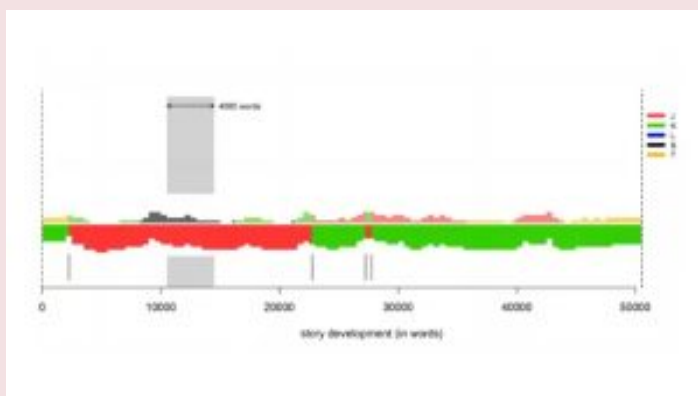
- **Oddaljeno branje romana Josipa Jurčiča in Janka Kersnika *Rokovnjači***

Primer tovrstne slike književnega besedila “od daleč” je stilometrična analiza romana *Rokovnjači*, ki ga je začel pisati Josip Jurčič, dokončal pa Janko Kersnik. Zaradi medsebojnih povezav avtorjev slovenskega realizma so literarni zgodovinarji drugi del romana razlagali kot Kersnikov poklon Jurčiču, saj naj bi Kersnik pri nadaljevanju romana upošteval Jurčičev slog in njegove zapiske. Po mnenju slovenski literarnega zgodovinarja Antona Slodnjaka je Kersnik Jurčičev slog oponašal do te mere, da »površen bralec niti ne opazi, kdaj preneha prvi in kdaj začne pripovedovati drugi pripovednik«, kljub temu pa raziskovalec opaža, da je Kersnik na trenutke ponovno.¹⁷ Tezo o prevzemanju Jurčičevega sloga pri Kersniku smo skušali preveriti z metodo računalniške

16. Eve 2019

17. Slodnjak 1961: 127

stilometrije, imenovano »rolling stylometry«. S strojnim učenjem je algoritem na podlagi korpusa daljših prozних pripovedi v stilometrični analizi usvojil stilne lastnosti različnih avtorjev slovenskega realizma (Josipa Jurčiča, Janka Kersnika, Frana Levstika, Ivana Tavčarja) in v romanu identificiral avtorstvo. Spodnji prikaz se nanaša na razvoj romana glede na število besed in kaže, da je stilometrična analiza prvi del Rokovnjačev pripisala Jurčiču (z rdečo barvo), drugi del pa Kersniku (z zeleno barvo), zato lahko sklepamo, da je Kersnikov slog v romanu razviden in se kaže celo v delu na začetku besedila.¹⁸



Slika: Stilometrična analiza romana Rokovnjači; vir: Mandić in Zajc 2020

- **Oddaljeno branje romana Leva Nikolajeviča Tolstoja *Vojna in mir***

David McClure je z oddaljenim branjem romana Leva

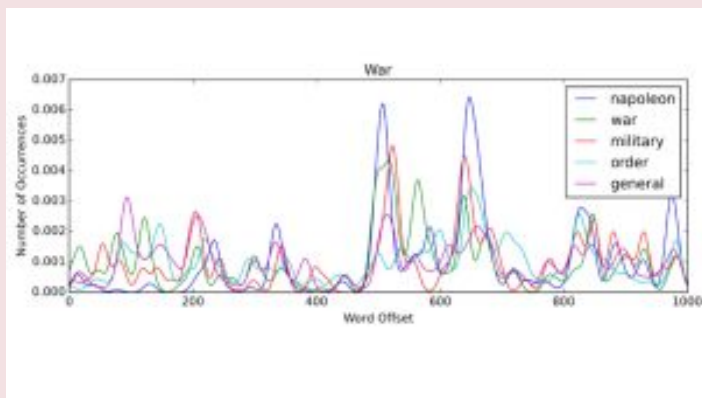
Nikolajeviča Tolstoja *Vojna in mir* pokazal, da je tematsko razdeljen na vsaj dva dela, ki ju ponazarja že sam naslov: na vojno in mir. Vizualizacijo tematske členitve romana na spodnji sliki je mogoče podrobneje pogledati v [prikazu na povezavi](#), ki ga je raziskovalec oblikoval s programom Textplot. Besedilo romana je program torej pretvoril v mrežo izrazov, da bi izluščil njegove osrednje tematike. Če se besede, kot so mama, družba in pogovor, pojavljajo v okviru teme miru, ki se v prikazu razporedi na desni strani (smer ni važna, ključno je združevanje posameznih izrazov v gruče), pa besede, kot so zmaga, sovražnik in bitka, najdemo v okviru teme vojne na levi strani. Na zgornjem delu prikaza se razporedijo besede, ki so povezane s tretjo tematiko, tj. zgodovino, ki v romanu pojasnjuje historične dogodke.



Slika: Tematska razdelitev romana Leva Nikolajeviča Tolstoja *Vojna in mir*; vir: McClure 2014; izdano pod licenco Creative Commons

Nadalje je McClure roman *Vojna in mir* analiziral s Kernelovo metodo, ki meri pogostost pojavljanja izbranih izrazov. Spodnji prikaz na primer kaže rabo besed Napoleon, vojna, vojska, red in general v kronološkem poteku celotnega romana. Izrazi so označeni s črtami določene barve, ki prikazujejo pogostost pojavljanja posamezne besede v kronološkem razvoju romana. Pokaže se razloček med deli romana, kjer so te besede pogostejše

in so namenjeni vojni temi, in deli, kjer so redkejšje in so namenjeni temi miru.¹⁹



Slika: Število besed (vertikalna os) v poteku romana *Vojna in mir* (horizontalna os); vir: McClure 2014; izdano pod licenco Creative Commons

Izziv: Oddaljeno branje zgodbe o Martinu Krpanu

Na spletu poiščite spletno stran [Voyant Tools](#), ki omogoča izvedbo enostavnega oddaljenega branja. Na spletišču Wikivir poiščite zgodbo Frana Levstika Martin Krpan z Vrha in besedilo skopirajte v okence orodja Voyant tools. Če zgodbe ne posante dobro, jo (ponovno) proberitte.

19. McClure 2014

- Ko se bo na strani prikazala analiza besedila, se vam bo na levi strani izrisal besedni oblak, ki prikazuje najpogosteje uporabljene besede v zgodbi. Pogostost rabe besede v pripovedi o Martinu Krpanu se ujema z velikostjo zapisa besede v besednem oblaku. Katere besede so najpogosteje uporabljene? Poigrajte se z večjim ali manjšim številom prikazanih besed v besednem oblaku. V sredinskem okencu, kjer je besedilo zgodbe, kliknite na besedo “Krpan” in na desni strani se bo izrisal graf, ki prikazuje pogostost uporabe te besede tekom poetka pripovedi. Kako pa je z besedo cesar?
- Levo spodaj preverite: Koliko je vseh besed v zgodbi? Koliko je unikatnih, tj. zgolj enkrat uporabljenih besed? Koliko besed ima v povprečju ena poved te pripovedi?
- Desno spodaj lahko pojavljanje besed tekom zgodbe prikažete z zanimivih grafom, ki potek pripovedi označi z vodoravno črto, omembe besede pa glede na pogostost s krogi ustreznih velikosti (glej funkcijo: Bubblelines). S funkcijo, ki išče povezave med besedami (glej: Correlations), pa lahko preverite, katere besede v zgodbi se povezujejo z določenimi drugimi besedami. Preverite različne besede, na primer izraza “junaka” oziroma “junakov”. Kako si

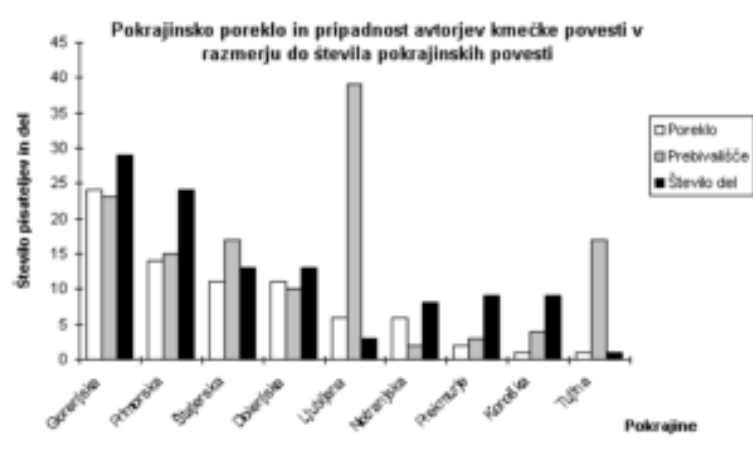
razlagate te povezave v zgodbi?

4 Raziskovanje metapodatkov

Ob ukvarjanju z literaturo lahko zbiramo metapodatke o literarnih besedilih, ki jih strukturirano razvrščamo v katalogih. Metapodatki so – kot pove že predpona te besede – podatki o podatkih. V okviru književnosti so to na primer dejstva o literarnem besedilu, njegovem nastanku in drugih vidikih, npr. naslovi in podnaslovi del, čas njihovega izida, nadaljnje izdaje dela, jezik, žanrske oznake, tematike književnega dela, geografske določnice nastanka literarnih besedil ali literarnih zgodb, pa tudi bibliografije avtorjev oziroma avtoric. Spodnja vizualizacija rezultatov raziskave Mirana Hladnika je na primer nastala na podlagi metapodatkov o avtorjih slovenske kmečke povesti, in sicer vključuje metapodatke o njihovem poreklu in prebivališču, pa tudi o številu literarnih del, ki se dogajajo v posamezni slovenski pokrajini.²⁰ Hladnik med drugim ugotavlja, da so bili avtorji slovenske kmečke povesti rojeni večinoma na podeželju, in sicer največ na Gorenjskem (30 %), na Štajerskem (22 %) in Primorskem (20 %), medtem ko se največ kmečkih povesti odvija na Gorenjskem.²¹

20. Hladnik 1998

21. Hladnik 1998

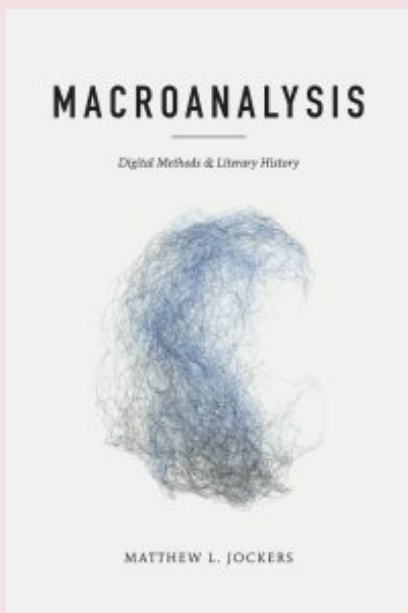


Slika: Pokrajinsko poreklo in pripadnost avtorjev v razmerju do števila pokrajinskih povesti; vir: Hladnik 1998

Za tovrstne raziskave je ključno, da metapodatki izbrane kategorije opisujejo jasno, strnjeno in organizirano. Nekatere od teh informacij so pomembne za organizacijo zbirk literarnih besedil v digitalnih knjižnicah, lahko pa jih uporabimo tudi v raziskovalne namene. Na podlagi izbranih informacij je namreč mogoče z rudarjenjem podatkov oblikovati nove sintetične ugotovitve o literarnem sistemu.²²

22. prim. Jockers 2013: 35

Primer: od grafikonov do zemljevidov



Jockers je izvedel odmevno študijo, utemeljeno na seznamu 758 literarnih del, vezanih na ameriške Irce, ki so bila izdana v časovnem razponu 250 let. Seznam je dopolnil s podatki o spolu, starosti, kraju bivanja in rojstnem kraju avtorja oz. avtorice, geografskimi koordinatami kraja, iz katerega je izhajal avtor oziroma avtorica, in kraja, kjer se posamezno literarno delo odvija.

Metapodatki so bili temelj analize z digitalnimi metodami, rezultati pa so bili vizualizirani z grafikoni, ki prikazujejo

literarno produkcijo ameriških Ircev v času z različnih vidikov, od spola do regije, pa tudi z zemljevidi, ki prikazujejo geografske razsežnosti njihovega ustvarjanja.²³

5 Oddaljeno branje na strani Google Books

Spletna stran Google Books, ki vsebuje na milijone literarnih del, med drugim nudi tudi funkcijo [Ngram Viewer](#), ki omogoča analizo digitaliziranih literarnih besedil.²⁴ Tako lahko uporabnik v korpusu besedil v določenem časovnem obdobju preveri rabo izbranih besed ali besednih zvez. V iskalnik vpišemo posamezen izraz ali besedno zvezo in s tem preverimo, v kolikšni meri se pojavlja v korpusu gradiva. Raziskovalci so sicer opazili mnoge omejitve in izrazili kritike glede delovanja platforme, saj spletna stran uporabniku ne omogoča, da bi dobil v vpogled dela, ki jih analizira, korpusa besedil pa ne more prenesti na svoj računalnik. Kritike izpostavljajo, da je korpus, na katerem temelji iskanje, neuravnotežen, saj vsebuje preveč znanstvene literature, da so mnoga dela napačno kategorizirana, pa tudi da ne vsebuje metapodatkov o vključenih besedilih.²⁵

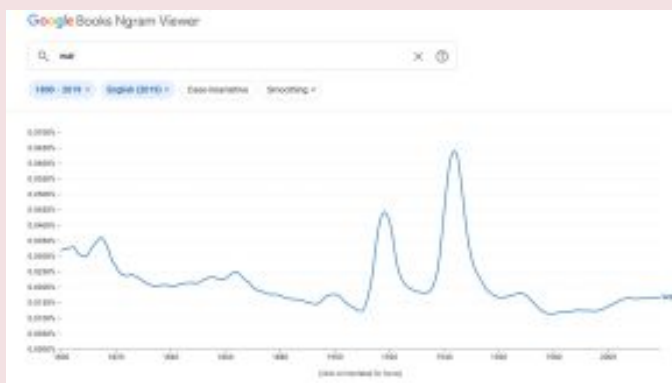
23. prim. Jockers 2013

24. Lin 2012: 169–174; Michel idr. 2010: 176–182

25. Kopleinig 2017: 169–188

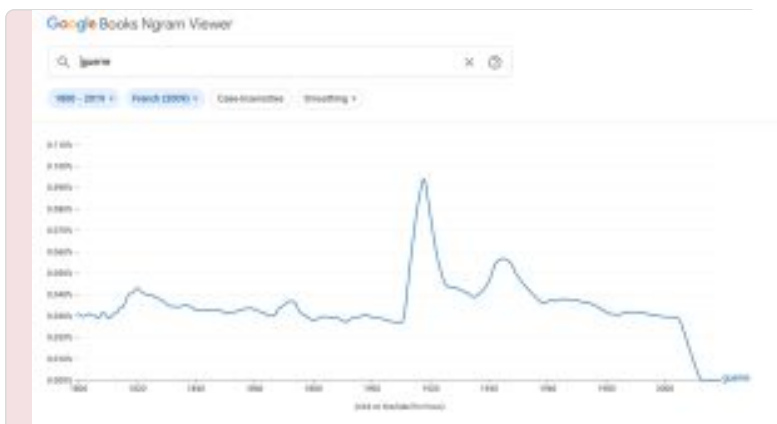
Primer: Pojavljanje besed v Google Books

Na spletni strani, ki trenutno še ne deluje za dela v slovenskem jeziku, smo preizkusili uporabo besede “vojna” med letoma 1800 in 2019. Spodnji graf na horizontalni osi prikazuje letnice izida knjižnih publikacij, na vertikalni osi pa kaže, kolikšen procent vseh besed v danes korpusu besedil, napisanih v angleščini med tema letoma, predstavlja beseda “vojna”. Iz grafa je jasno, da se beseda v več jezikih mnogo pogosteje uporablja v časovnih obdobjih blizu prve in druge svetovne vojne, kot je razvidno iz spodnje slike.



Slika: Pojavljanje izraza “war” (slo.: vojna) v Google Books med letoma 1800 in 2019

Preverili smo tudi pojavljanje besede vojna (*guerre*) v gradivu v francoskem jeziku in dobili podoben rezultat.



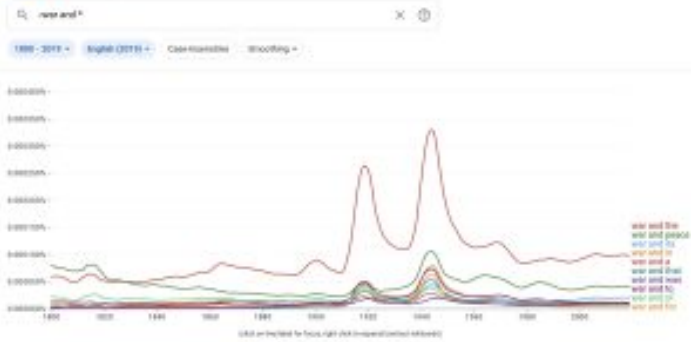
Slika: Pojavljanje izraza “guerre” (slo.: vojna) v Google Books med letoma 1800 in 2019

Dodatne možnosti:

Spletna stran ponuja tudi možnosti iskanja različic določene besede (ang. inflection search), različice zapisa z velikimi oziroma malimi tiskanimi črkami (ang. case insensitive search), pri primerih dveh različnih besed, ki ju zapisujemo enako, je mogoče določiti, ali iščemo samostalnik, glagol ali katero drugo besedno vrsto.

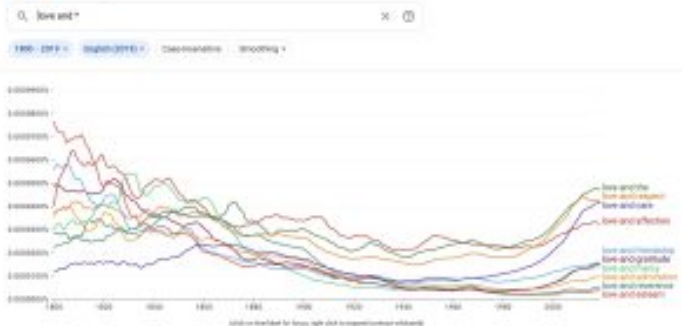
Ko namesto besede vstavite zvezdico (*), bo Ngram Viewer prikazal prvih deset zamenjav za besedo, ki ste jo zamenjali z zvezdico. Naslednji primer kaže, katere besede so najpogostejše v drugem delu besedne zveze “vojna in”. Kot vidimo, se vojna najpogostejše pojavlja skupaj z nepolnopravnimi besedami, npr. z vezniki, pa tudi z besedo “mir” (ang. peace).

Google Books Ngram Viewer



Z mnogo več polnopomenskimi besedami se povezuje izraz ljubezen (ang. love), kot je razvidno iz naslednjega grafa. Gre za izraze: ljubezen, ugled, skrb, naklonjenost, prijateljstvo, hvaležnost, usmiljenje, občudovanje, oboževanje in spoštovanje.

Google Books Ngram Viewer



Izziv: Izrazi, značilni za (literarno)zgodovinsko obdobje – dva primera

Katera beseda ali besedna zveza bi bila po Vašem mnenju značilna za obdobje romantike v literaturi? Preizkusite njeno rabo na spletni strani Google Books.

Kako se na primeru rabe določenih izrazov od leta 1800 do sodobnosti na spletni strani Google Books kažejo spremembe v tehnološkem napredku družbe? Preverite na primer rabo besede konj (ang. horse), živali, ki je ljudem služila kot prevozno sredstvo, dokler je niso nadomestili avtomobili – se to odraža v pogostosti rabe tega izraza v angleških pisnih virih?

6 Nekateri kritični pogledi na oddaljeno branje

Branje literature s pomočjo računalniških tehnik je kontroverzno,²⁶ prav zato je pomembno, da pri ukvarjanju s književnostjo poznamo



tako možnosti kot tudi omejitve posamezne metode in tehnike digitalne humanistike. Ena najbolj očitnih omejitev oddaljenega branja je dejstvo, da

26. [footnote]Eve 2019: 2

temelji izključno na literarnih besedilih, ki so digitalizirana in s tem na voljo za računalniške analize, medtem ko so obsežne zbirke besedil, ki (še) niso bile digitalizirane, iz raziskav izvzete. Poleg tega kritike oddaljenega branja izpostavljajo različne težave, med drugim zapostavljanje globljega razumevanja književnosti in posameznih literarnih besedil, ki ga omogoča bližnje branje. Bralec z bližnjim branjem besedilo doživlja in občuti, medtem ko je čustvena komponenta pri oddaljenem branju izvzeta. Pojavlja se tudi vprašanje interpretacije ironičnih pomenov in metaforičnega jezika. Nenazadnje makroanalize pogosto spregledajo nekatere kontekste, ki so bistveni za razumevanje literature, od kulturnih do zgodovinskih in družbenih. Stanley Fish²⁷ je do oddaljenega branja kritičen, saj po njegovem mnenju raziskovalci najprej oblikujejo računalniške analize, zatem pa preverijo, ali se prilegajo njihovi hipotezi. Gayatri Spivak ima zadržke do oddaljenega branja z vidika postkolonialnih teorij, saj meni, da ne predstavlja nujno perspektive tistih del iz literarne zgodovine, ki so pozabljena in ki jih nihče ne bere več, ob tem pa opozarja, da se s tem področjem digitalne humanistike ustvarjajo avtoritativni

totalizacijski vzorci, ki so odvisni od nepreverjenih trditev maloštevilčnih skupin ljudi.²⁸

3.

OBRAVNAVA KNJIŽEVNOSTI Z RAČUNALNIŠKO STILOMETRIJO

Računalniška stilometrija oz. računalniška stilistika je raziskovalna metoda, ki tekste obravnava na podlagi digitalnih orodij s kvantitativnimi pristopi. Metoda temelji na strojnem učenju na podlagi korpusa digitaliziranih literarnih besedil, na primer romanov. Rezultat je statističen prikaz podobnosti in razlik med vključenimi književnimi besedili.

Literarni raziskovalci lahko poleg “običajnega” branja pri svojih študijah uporabljajo tudi računalniško branje, kar jim omogoča, da v raziskavo zajamejo večje število književnih del, pa tudi, da raziskujejo vidike besedil, ki jih z bližnjim branjem ne bi opazili. Rezultat računalniškega

branja je posplošitev o določeni skupini besedil, ki jo prikažemo z različnimi tipi vizualizacij, kar je lahko v pomoč pri raziskavah literarnih korpusov.

1 Metoda računalniške stilometrije

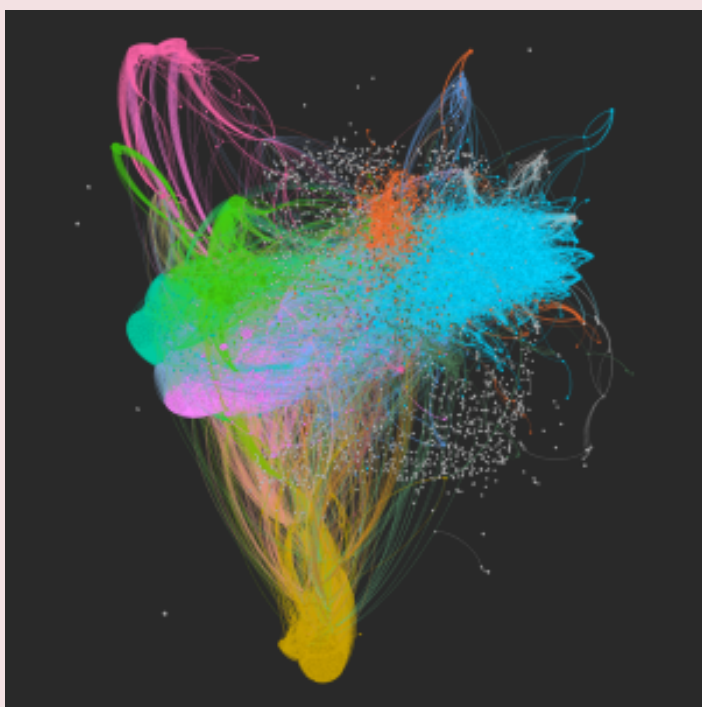
Računalniška stilometrija je metoda analize besedil, ki temelji na njihovi statistični računalniški obdelavi. Ta metoda je najpogosteje izvedena v programskih jezikih »R«, Java in Python, z njo pa je mogoče prepoznati vzorce v večjih korpusih književnih besedil in prikazati odnose med besedili, njihovimi avtorji, skupinami avtorjev ali raznimi drugimi kategorijami, ki jih je mogoče pripisati vključenim literarnim delom. Gre za analize **relativne frekventnosti najpogostejših besed v korpusu literarnih besedil**, pa tudi za analizo **povprečne dolžine povedi ali besed ter rabe in razporeditve locil¹**, pa tudi stavčne skladnje ter povezav med pomeni besed. Rezultati tovrstnih analiz so vizualizacije podatkov, ki glede na bližino oziroma oddaljenost posameznih literarnih besedil kažejo, kateri tipi tekstov so si medsebojno bolj ali manj podobni. Vendar z računalniško stilometrijo ne ugotavljamo le razlik med različnimi avtorskimi stili, pač pa tudi na primer slogovni razvoj posameznega avtorja ali avtorice. Poleg tega se besedila v vizualnih prikazih lahko grupirajo glede na različne žanre ali celo spol piscev in pisk. Če

1. Eder idr. 2017: 1–2

preverjamo literarni razvoj posamezne avtorice ali avtorja oziroma skupine v času, pa govorimo o stilokronometriji.

Primer: raziskava 5917 srednjevisokonemških rokopisov

Obsežna [računalniška stilometrična analiza](#) Gustava Fernandez de Rive, ki je bila izvedena s paketom Stylo v programskem jeziku R, je zajela kar 5917 digitaliziranih srednjevisokonemških rokopisov. Srednje visoka nemščina spada v zgodovinsko obdobje razvoja nemškega jezika in se razlikuje od stare visoke nemščine, ki je predhodila temu obdobju, in od nove visoke nemščine, ki se je razvila po srednjevisokonemškem obdobju. V tem obdobju so nastala mnoga pomembna dela nemškega srednjega veka, na primer *Pesem o Nibelungih* in različne lirske pesmi. Omenjena analiza je zajela najrazličnejša besedila tega obdobja, od literarnih do neliterarnih. Rezultat raziskave je mogoče videti na spodnji vizualizaciji: besedila, označena s pikicami, se razlikujejo po zvrsteh, ki jih zaznamujejo različne barve. Analiza je pokazala, da se besedila tega obsežnega korpusa delijo glede na različne žanre, tj. lirika, ljubezenska literatura, verska literatura, profana literatura in poučna literatura.



Slika: vizualizacija 5917 digitaliziranih srednjevisokonemških rokopisov; vir: Gustavo Fernandez Riva 2018

2 Računalniško branje in literarni slog

“Kar se tiče
jezika in sloga,
lahko delamo

Slog literarnega besedila vključuje razporeditev njegovih posameznih elementov, od najmanjših (zlogi, rime, ritem, besede) do največjih (verzi, pripovedna zgradba, odstavki, poglavja in širša

zgradba besedila).³ Slogovni naturalizem temelji na izhodišču, da imajo avtorji v svojem pisanju niz profilnih značilnosti, nad katerimi nikoli ne morejo imeti popolnega nadzora.⁴ Za posameznega literarnega avtorja oz. avtorico je značilen individualen način pisanja, ki ga je mogoče

stvari, o katerih so pretekle generacije lahko le sanjale.” Franco Moretti²

statistično določiti. Ta teza sicer izhaja že iz 19. stoletja, ko so v duhu empirične literarne vede preštevno določali literarni stil. Tudi sodobne raziskave z metodo računalniške stilometrije kažejo, da daje posamezen pisec pri ustvarjanju prednost nekaterim besednim vzorcem pred drugimi, pri tem pa oblikuje specifičen slog.

Primer: stilometrična analiza 90 slovenskih romanov

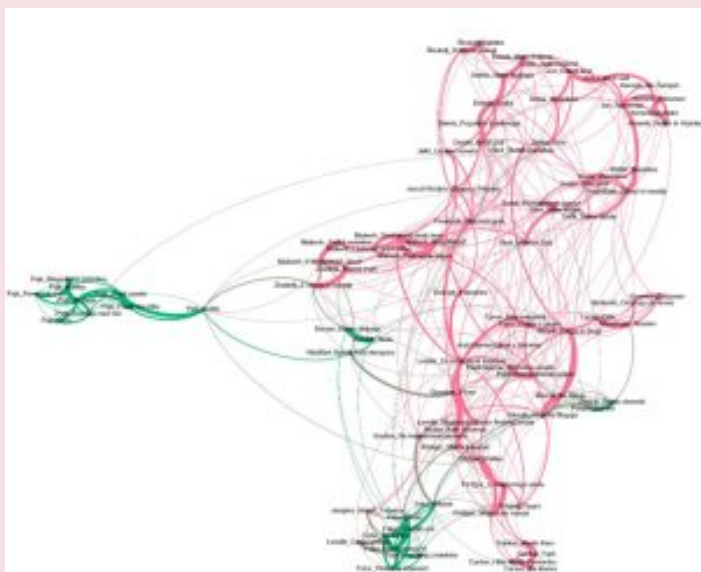
Oblikovali smo korpus 90 slovenskih romanov iz 19. in začetka 20. stoletja, kar pomeni, da smo romane shranili v dokumente, urejene na način, ki je dostopen programu. Za analizo besedil smo uporabili paket Stylo v programskem jeziku R.

2. Moretti 2013: 212

3. Allison idr. 2013

4. Eve 2022: 32

Spodnja slika prikazuje rezultate raziskave: posamezna točka označuje določeno literarno besedilo, ki vsebuje pripis avtorja in naslova, s črtami so označene povezave med njimi glede na splošno slogovno podobnost. Z zeleno barvo so označena dela avtoric, z rdečo pa dela avtorjev, vendar vidimo, da spol pisca ni ključen signal, ki dela razvršča v skupine, marveč je najizrazitejši signal njihovo avtorstvo: primer so besedila Ivana Cankarja ali Lee Fatur, ki so se razvrstila v skupini v spodnjem delu prikaza. Skrajno levo so se razvrstila dela pisateljice Pavline Pajk, iz česar lahko sklepamo, da so izrazito drugačna od ostalih besedil v našem korpusu.



Slika: stilometrična analiza korpusa 90 slovenskih romanov iz 19. in začetka 20. stoletja; vir: lasten

Vaja: realizem in modernizem

Korpus, na katerem je bila izvedena zgornja stilometrična analiza, vsebuje 90 romanov slovenskih avtoric in avtorjev iz dveh literarnozgodovinskih obdobj: modernizma ter realizma. Preverite, če se avtorice oziroma avtorji iz teh dveh obdobjih v zgornji vizualizaciji delijo v dve skupini glede na pripadnost literarnozgodovinskemu obdobju ali to ni ključni signal, ki določa bližino in oddaljenost besedil v tem primeru. Razpravljajte, kako bi interpretirali ta rezultat?

Juola navaja predpostavko večine sodobnih raziskovalcev, da ima vsak posameznik značilen vzorec uporabe jezika v svojih zapisih, vendar so lahko ti »avtorski prstni odtisi« na drugi strani bolj zapleteni kot nespremenljive statistike, kot je povprečna dolžina besede ali velikost besedišča.⁵ Na prvi pogled morda manj pomembne besede – členki, predlogi in vezniki – so se izkazale za posebej ključne pri določanju osebnega sloga posameznikov in posameznic, saj se literarna besedila po njihovi rabi zelo razlikujejo, ne glede na tematiko pisanja.⁶

Signal je računalniško izmerjen podatek, ki je v

5. Juola 2008: 239

6. Stamatatos 2018: 540

zgornjem primeru prikazan v obliki vizualizacije, na primer očitne razdalje med določenimi skupinami analiziranih besedil, ki kaže na njihovo slogovno različnost. Koncept je po drugi strani fenomen, s katerim določen signal razlagamo, na primer dejstvo, da spadajo besedila, med katerimi so očitne razdalje, v različne literarne žanre. Argumente v digitalni humanistiki oblikujemo na podlagi konceptov, medtem ko lahko z računalniki izmerimo le signale, pojasnjujeta Ryan Heuser in Long Le-Khac.⁷

Primer: *Sreča v nesreči*

Prva izvirna slovenska povest je delo Janeza Ciglerja *Sreča v nesreči* (1836), na avtorja pa naj bi po mnenju slovenske literarne zgodovine močno vplivalo ustvarjanje bavarskega pisatelja Christopha Schmidta. To so argumentirali z vsebinskimi vzporednicami med deli obeh avtorjev, manj pa z analiziranjem stilnih podobnosti.⁸

Stilometrična analiza je vključila poslovenjena dela obeh avtorjev in prevode še nekaterih drugih piscev, ki so imeli po cenah literarne zgodovine pomemben opliv na tedanji razvoj pripovedne proze. Andrejka Žejn med drugim opozarja, da se pri poslovenjenih besedilih kaže vpliv prevajalskega sloga. Na vizualizaciji rezultatov se dela obeh avtorjev razdelita v ločeni skupini. Raziskava je torej pokazala, da je imel Janez Cigler svojstven slog, ki se je

7. 2011: 81

8. Žejn 2020

razlikoval od Schmidovega sloga, ki naj bi ga domnevno posnemal. Jasen je torej Ciglerjev primat pri vzpostavljanju slovenske pripovedne proze.⁹

3 Določanje avtorstva literarnih del

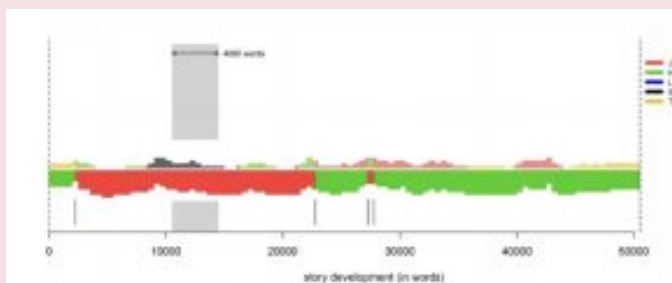
Sodobne raziskovalke in raziskovalci računalniško stilometrijo najpogosteje uporabljajo za preverjanje avtorstva literarnih del in ne preseneča, da so metodo v preteklosti uporabljali tudi v kriminalistiki za prepoznavanje avtorstva anonimnih besedil, ki so se znašla v središču zločinov.

Računalniško branje Jurčič-Kersnikovega romana *Rokovnjači*

Zgodovinski roman *Rokovnjači*, ki je pričel izhajati v *Ljubljanskem zvonu* leta 1881 in opisuje življenje razbojnikov na Kranjskem v začetku 19. stoletja, je začel pisati Josip Jurčič, ki je ustvaril prvih 11 poglavij. Po njegovi smrti je roman dokončal pisatelj Janko Kersnik.

Literarnozgodovinske analize utemeljujejo, da je Kersnik pri svojem delu izhajal iz Jurčičevih zapiskov in oponašal njegov slog, kar smo preverili s stilometrično analizo. Uporabili smo metodo t. i. *rolling stylometry*: na podlagi

strojnega učenja je program usvojil slog različnih slovenskih avtorjev obdobja realizma, zatem pa preveril, kateri slog prevlada v romanu *Rokovnjači*. Izkazalo se je, da v prvem delu romana prevlada Jurčičev slog, medtem ko v drugem delu prevlada Kersnikov slog, kar je na spodnji sliki razvidno na spodnjem delu črte, kjer vidimo, kateri avtorski slog je v določenem delu besedila prevladujoč, na zgornjem delu črte pa so vidni drugi slogovni vplivi.¹⁰



Slika: Računalniško branje romana *Rokovnjači*; razlaga barvnih oznak na sliki: rdeča – Jurčič, zelena – Kersnik, modra – Levstik, črna – Stritar, oranžna – Tavčar. (vir: Zajc in Mandić 2020: 11)

5 Raziskovanje literarnih ugank

5.1 Je William Shakespeare res napisal vsa dramska dela, ki mu jih pripisujejo?

William Shakespeare (1564-1616), največji angleški dramatik in član gledališke skupine Lord Chamberlain's Men oz. King's Men, naj bi napisal množico dramskih besedil. Že več kot stoletje se pojavlja t. i.

shakespearejansko vprašanje, ali je res avtor obsežnega opusa, ki ga pripisujemo Shakespearu, le en, oz. ali je sloviti angleški dramatik sploh obstajal? Je šlo morda le za psevdonim, za katerim se je skrivala skupina piscev? Večina shakespeareologov te dvome zavrača na podlagi (literarno-)zgodovinskih podatkov, avtorstvo pa so preverili tudi z računalniško stilometrijo. Še pred izumom računalnikov, proti koncu 19. stoletja je z ročnim štetjem besed Shakespearjev stil preveril poljski raziskovalec Wincenty Lutoslawski. Statistične metode so pokazale tipične vzorce v Shakespearovih delih, ki jih v dramskih besedilih drugih avtorjev niso našli, čeprav so poznejše stilometrične analize pokazale, da je Shakespeare pri ustvarjanju določenih dramskih del najverjetneje sodeloval z drugimi dramatikami, saj so se pokazale sledi njihovega značilnega sloga. Primer je Christofer Marlowe, dramatik in Shakespearov sodobnik. Statistične računalniške raziskave kažejo na stilno podobnost med



njunimi deli, zato sklepajo, da sta vplivala drug na drugega ali celo mestoma sodelovala pri pisanju, kar se kaže predvsem v prvem delu Shakespearove tragedije *Henrik VI*. Vsekakor pa Marlowe in Shakespeare glede na rezultate stilometričnih analiz nista bila ista oseba, ampak sta imela kot dramatika vsak svoj prepoznaven slog.¹¹

5.2 Kdo je skrivnostni Robert Galbraith?

Leta 2013 je znana avtorica serije o Harryju Potterju, ki obsega osem knjig, Joanne K. Rowling (roj. 1965) izdala nov detektivski roman pod psevdonimom Robert Galbraith. Z metodo računalniške stilometrije so raziskovalci¹² ugotovili, da gre v resnici za slovito avtorico. Pozneje je priznala, da si je želela s psevdonimom “začeti od začetka”, tudi zatem, ko je bila razkrinkana, pa je nadaljevala z uporabo nadimka, češ da ji je všeč, da ima *alter ego*.

5.3 Kdo je v resnici Elena Ferrante?

Avtorstvo znane romaneskne trilogije *Genialna prijateljica* je neznanka, saj je ime, ki je zapisano na platnicah, psevdonimno. Da bi odkrila, kdo je Elena Ferrante, je skupina raziskovalcev uporabila metodo računalniške stilometrije na korpusu sodobnih avtorjev in avtoric. V skupini 150 sodobnih romanov so se romani Elene Ferrante razporedili ob romane pisatelja Domenica Starnoneja, neapeljskega romanopisca, za katerega so v

11. Fox, Ehmoda in Charniak 2012

12. Juola 2015

preteklosti že sumili, da je napisal trilogijo.¹³ Kljub temu rezultat ni dokončen, saj je mogoče, da resnični pisec ali piska ni bil vključen/-a v analizo.

Izziv

Domislite se še kake literarne uganke, ki bi jo lahko rešili s pomočjo računalniške stilometrije.

6 Nekaj kritičnih pogledov na računalniške analize besedilnih korpusov

Stilometrična analiza upošteva in primerja pogostost pojavljanja besed v vseh delih, vključenih v korpus, vendar iz tabel ali vizualnega prikaza stilistične bližine med vključenimi besedili ni razvidno, kateri formalni elementi (npr. metafore, besedne zveze, pogosta raba določenih besed), katere strukture (npr. mnogovezje, ponavljanja) in načini tvorbe literarnega besedila so značilni za posamezno avtorico ali avtorja. Zaradi tega razloga je v tovrstnih raziskavah zelo pomembna kvalitativna interpretacija kvantitativnih podatkov. Pomembno je upoštevati, da lahko z računalniškim branjem pridobimo le določen tip podatkov: digitalna orodja ne morejo opravljati kritičnega in analitičnega dela, ki ga

13. Eder et al. 2018

opravljamo mi, po drugi strani pa tudi mi ne moremo opravljati dela, ki ga opravljajo ta orodja.¹⁴ Računalniško stilometrijo kot kvantitativno metodo literarne vede kombiniramo s kvalitativno interpretacijo rezultatov.

Digitalizacija literarnih besedil, na katerih temeljijo stilometrična analiza in druge računalniške metode v humanistiki, se srečuje s problemom neenakomerne delitve med jeziki. Računalniška stilometrija namreč predvideva obsežne korpuse digitaliziranih besedil, zato obravnava predvsem literarna dela večjih svetovnih jezikov, zapostavlja pa jezike z manj govorniki in skromnejšo infrastrukturo.

“Računalniško podprta stilometrija se je namreč uveljavila z raziskavami literature v angleškem jeziku, ki je, kar se tiče razpoložljivosti elektronskega korpusa in orodij za procesiranje naravnih jezikov, tudi najbolj razvit jezik, njena prednost je še, da praktično ne pozna slovnične pregibnosti.”¹⁵

Digitaliziranje literarnih del je precej zahtevno delo, saj je potrebno besedila v procesu digitalizacije preveriti, po drugi strani pa je gradivo omejeno, saj morajo književnim

14. Adwetewa-Badu 2020

15. Z ˇ ejn 2020

besedilom poteči avtorske pravice. Poleg tega je potrebno za nadaljnjo obravnavo s stilometričnimi metodami zbrati in pripraviti korpus, kar ni avtomatizirano in je zato podvrženo napakam. Zaradi novih možnosti raziskav s stilometričnimi orodji, ki so pogosto še rudimentarna, se velikokrat pojavljajo pomanjkljivosti ali napake v programskih kodah, ki raziskavo zaustavijo.

Metod za računalniško analizo besedil je vse več, zato obstaja tudi nevarnost, da raziskovalec izbere tisto, ki najbolj podpre njegovo tezo.¹⁶ Poleg tega lahko pridemo do nekoliko različnih rezultatov, če spreminjamo parametre raziskave, zato se pojavlja vprašanje, ali bo računalniška stilometrija razvila veljavne formalne postopke, ki bodo ponudili prepričljive rezultate, posebej za občutljivo področje določanja avtorstva besedil.

4.

PROCESIRANJE NARAVNIH JEZIKOV, MODELIRANJE TEM IN LITERATURA

Naravni jeziki so tisti jeziki, ki jih ljudje vsakodnevno uporabljamo in z njimi komuniciramo, pri tem pa se nenehno spreminjajo, medtem ko so umetni jeziki namenjeni za specifične naloge in so nespremenljivi, na primer programski jeziki. Obdelava naravnih jezikov (ang. *Natural Language Processing* ali s kratico NLP) se ukvarja z računalniško obdelavo človeških jezikov in pomeni most med naravnimi in računalniškimi jeziki.

Procesiranje naravnih jezikov uporablja modele strojnega učenja za analizo, interpretacijo ali ustvarjanje naravnih jezikov. Obdelava naravnih jezikov je uporabna za številne vsakodnevne

naloge, na primer za prepoznavanje govora in avtomatsko prevajanje besedil.

1 Iskanje tematik znotraj literarnega korpusa

Raziskave na področju procesiranja naravnih jezikov so namenjene oblikovanju algoritmov in orodij za avtomatsko obdelavo jezika, osredotočajo se na izboljšave meril znotraj svojih procesov ter razvoj optimalnih algoritmov. S preizkušanjem njihove učinkovitosti na podlagi standardnih korpusov besedil je možno na sistematičen način primerjati različne pristope. Raziskave znotraj digitalne humanistike se po drugi strani ne osredotočajo v tolikšni meri na same algoritme, ampak bolj na rezultate, na podlagi katerih raziskovalke in raziskovalci odgovarjajo na svoja raziskovalna vprašanja. Zaradi tega razloga se vsaka posamezna študija osredotoča na ožji nabor podatkov, ki je pogosto edinstven.¹ Pred izdelavo tematskih modelov je običajno, da korpus urednimo, in sicer je ključno, da izločimo t. i. prazne besede (ang. *stop words*):² nepolnopomenske besede, imena mesecev, dni v tednu, literarnih junakov ipd.³

Primer: korpus ELTeC

1. McGillivray idr. 2020

2. Seznam slovenski praznih besed je dostopen na [tej](#) povezavi.

3. Mandić 2023

Korpus *ELTeC-slv* je literarni korpus slovenskih proznih besedil (romanov, povesti in novel), ki obsega 100 naslovov del, objavljenih med letoma 1836 in 1921.⁴ Korpus je del obsežnega korpusa, ki ga sestavljajo manjši korpusi v 17 evropskih jezikih. Vsi korpusi so oblikovani po enakih parametrih.

2 Modeliranje tem v književnih besedilih

Ko procesiranje naravnih jezikov uporabimo na korpusih književnih del, lahko dobimo na primer pregled nad temi besedili s samodejnim izločanjem osrednjih tematik danega besedila in odnosov med njimi.



Pri analizi besedil gre za samodejno iskanje besed in besednih zvez, ki povzemajo vsebino dela ali korpusa književnih del, s tem pa lahko raziskovalke in raziskovalci izluščijo ključne tematske poudarke tekstov. Gre za preverjanje, kako pogosto se določene besede ali besedne zveze pojavljajo v danem književnem besedilu ali literarnem korpusu, s tem pa je mogoče ugotoviti, katere teme in motivi v delih prevladujejo.

Z modeliranjem tem tako ugotovimo, katere teme so v določenem literarnem besedilu ali skupini besedil ključne in kateri so osrednji vsebinski poudarki. Poleg tega lahko na ta način na primer ugotavljamo, katere teme so bile najbolj zanimive za literate in literatke v določenem zgodovinskem obdobju. Tematski modeli so uporabni za različne namene v literarnovednih raziskavah: za združevanje dokumentov v skupine, organiziranje velikih korpusov besedilnih podatkov in pridobivanje informacij iz besedil.⁵

Primer: vizualizacije literarne produkcije ameriških Ircev

Raziskava Matthewa Jockersa in Davida Mimne je najprej na podlagi metapodatkov (npr. leta izdaje, avtorji in pogostost določenih besed) z računalniško analizo preverjala vzorce v obsežnem naboru britanskih romanov iz 19. stoletja. Raziskovala sta na primer, kako se je uporaba določenih besed ali tematik v romanih spreminjala skozi čas ali kakšni so bili različni pristopi k pripovedovanju. Študija uporablja metapodatke in računalniške metode za odkrivanje pred tem nepoznanih vzorcev v obsežni zbirki romanov, kar omogoča vpogled v transformacije britanske književnosti. Med rezultati raziskave je dejstvo, da se tematike literarnih del spreminjajo glede na nacionalnost in spol avtoric ter avtorjev, pa tudi glede na čas izida književnega besedila. Raziskava je na primer pokazala, da je dvakrat večja možnost, da bo o tematiki ženske mode pisala literarna avtorica kakor avtor.⁶

5. Mandić 2023

6. Jockers in Mimno 2013

3 Modeliranje tem in literarnovedna vprašanja

Modeliranje tem v literarnih delih torej omogoča raziskovalkam in raziskovalcem, da razširijo svoje metode analize, s tem pa obogatijo razumevanje književnih del ter literarne zgodovine. To področje se še naprej razvija in prispeva k bolj poglobljeni analizi ter interpretaciji literature. Tematike lahko raziskovalec ali raziskovalka išče glede na določene vsebine, ki jih želi osvetliti ali medsebojno povezati. Če želimo, da so naše raziskovalne hipoteze v okviru književnih raziskav relevantne, je potrebno dobro poznavanje področja raziskav. Enako velja tudi za interpretacijo rezultatov modeliranja tem.

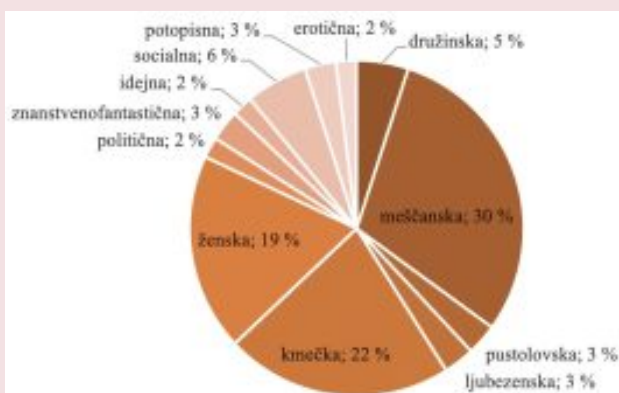
Primer: tematike, povezane z industrializacijo, meščanstvom, političnim udejstvovanjem in nacionalnimi idejami v slovenski literaturi 19. stoletja

Lucija Mandić⁷ je izvedla modeliranje tem na slovenskem korpusu *ELTeC-slv*, cilj raziskave pa je bil ugotoviti, katere teme so povezane z industrializacijo, meščanstvom, političnim udejstvovanjem in nacionalnimi idejami v daljši slovenski pripovedni prozi 19. stoletja. Študija izlušči

naslednje teme v besedilnem korpusu, ki jih lahko povežemo z omenjenimi tematikami:

- politika: volitev, komisar, glavar, kandidat, shod, stranka, narod, volilec, poslanec, večer;
- nacionalna ideja: narod, država, stoletje, jezik, pesnik, zgodovina, korist, nega, rod, duh;
- industrializacija: zdravnik, gospod, seme, nadzornik, delavec, ravnatelj, tovarna, vrata, delo, glava;
- plemstvo: baron, grofica, gospod, dan, misel, beseda, družba, čas, srce, gospa;
- podeželje: pot, deklica, glava, večer, beseda, vas, mladenič, oče, mlin, raz;
- družina: mati, oče, dan, otrok, leto, sin, pismo, hiša, denar, mož;
- sentimentalnost: oči, glava, pogled, obraz, ljubezen, gospod, duša, ženska, lice, miza;
- meščanstvo: gospod, beseda, doktor, duh, društvo, prijatelj, sluga, gospa, oči, odvetnik.

Kot je razvidno s spodnje slike, se v korpusu največkrat pojavljata meščanka in kmečka tema, sledi jima ženska tema.



Slika: Sestava Korpusa glede na tematiko; vir: Mandić 2023

Mandić med drugim ugotavlja, da so teme “politika”, “nacionalna ideja”, “meščanstvo” in “sentimentalnost” v večji meri prisotne v delih z oznako roman kot v delih z oznako povest ali novela.⁸

3 Modeliranje tem in čustvene vsebine literarnih del

Čustva igrajo ključno vlogo pri razumevanju in analizi književnih besedil, raziskujemo pa jih lahko tudi z modeliranjem tem v literarnih korpusih. Slednje se povezuje z zaznavanjem čustvenih vidikov besedil prek čustvenih izrazov, s čimer dosežemo bolj celostno

8. Mandić 2023

razumevanje vsebine, njene čustvene izrazne moči in vpliva na bralce. Gre na primer za raziskovanje, ali je besedilo pozitivno, negativno ali nevtravno naravnano in katera čustva v tekstu prevladujejo. Poleg tega je mogoče ugotavljati, kako so v besedilih čustveno obarvane določene tematike. Tovrstno modeliranje tem seveda ni rezervirano le za književnost: uporablja se za analize obsežnih objav na družabnih omrežjih, v časopisih, revijah, spletnih blogih ipd. Obsežna analiza tem v slovenskih medijih je na primer pokazala, da ima večina tem s politično tematiko negativen ali nevtralen sentiment, medtem ko je pozitivnih zelo malo.

Primer: Čustva v zbirki o Harryju Potterju

Zanimiva raziskava je merila in vizualizirala čustveno intenzivnost v šestih knjigah zbirke o Harry Potterju: *Kamen modrosti*, *Dvorana skrivnosti*, *Jetnik iz Azkabana*, *Ognjeni kelih*, *Feniksov red* in *Princ mešane krvi*. Kot je razvidno iz rezultatov raziskave, se čustveni naboj tekom zbirke stopnjuje, najbolj emocionalno intenziven pa je del z naslovom *Princ mešane krvi*. Trije najbolj čustveni trenutki v popularni zgodbi o čarovniku so sodeč po analizi tekom vseh obravnavanih nadaljevanj zbirke naslednji: kadar Harryja Potterja izbere Ognjeni kelih, kadar umre lik Cedric Diggory in kadar umre lik Dumbledore.⁹

3 Nekaj kritičnih pogledov na procesiranje naravnih jezikov in literaturo

Naravno procesiranje jezikov se srečuje z metodološkimi in tehničnimi izzivi, povezanimi med drugim z zgodovinskimi oblikami jezikov ter z večjezičnostjo, pa tudi s potrebami po bolj razvitih metodah digitalizacije besedil in polavtomatskim označevanjem tekstov.¹⁰ Ker je pri tej metodologiji, ki se intenzivno razvija, poudarek na razvoju novih računalniških sistemov ali izboljšanju obstoječih, je zelo pomembno, da se ti ovrednotijo na standardnih zbirkah podatkov z uporabo ponovljivih metod. Vendar to pomeni, da se raziskovalke in raziskovalci osredotočajo na omejene nabore podatkov zgolj v izbranih jezikih, zato se razvijajo orodja, ki so optimizirana izključno za te nabore podatkov in jezike, pri tem pa izključujejo vse druge jezike, ki so s tega vidika manj razviti.¹¹ Orodja za naravno procesiranje jezikov so za širši krog uporabnikov precej zahtevna, da jih usvojimo, potrebujemo veliko časa. Med manj zahtevnimi orodji so po drugi strani na voljo predvsem orodja, ki so hkrati precej manj natančna, kakršno je Ngram Viewer na platformi Google Books.¹²

10. McGillivray idr. 2020

11. McGillivray idr. 2020

12. prim. Greenfield 2013

5.

ANALIZE OMREŽIJ V KNJIŽEVNOSTI IN LITERARNI ZGODOVINI

Kakšni so odnosi oziroma povezave med različnimi elementi, povezanimi s produkcijo in recepcijo književnih besedil ter njihovo vsebino, raziskujemo z metodami analize omrežij, ki jih na podlagi statističnih izračunov izvajamo z digitalnimi orodji. Na ta način ugotavljamo, kakšni so odnosi med različnimi literarnimi avtorji in avtoricami, časopisi, knjižnimi izdajami, literarnozgodovinskimi obdobji ter podobnimi vidiki.

Kako se medsebojno mrežno povezujejo liki v literarnih zgodbah, književne avtorice in avtorji ter intelektualke in intelektualci, kakšna so omrežja znotraj literarnega sistema: mreže založb, tiskarn, uredništev, literarnih

kritičark in kritikov? V tem poglavju predstavljamo kvantitativno in računalniško podprto analizo omrežij na literarnovednem področju, ki omogočajo boljše prepoznavanje vzorcev, povezav ter tematik in so informativni temelj za kvalitativne analize.

Analize omrežij v književnosti in literarni zgodovini predvidevajo sodelovanje ter medsebojne izmenjave med različnimi disciplinami.¹ Raziskovalkam in raziskovalcem omogočajo, da drugače, “od daleč” vidijo na primer odnose med liki v književnih zgodbah. Tako posredujejo ogromno količino podatkov naenkrat, v enem samem pogledu.²

1 Digitalne raziskave literarne zgodovine z vidika omrežij

Z digitalnimi orodji lahko zajamemo velike količine (literarno)zgodovinskih podatkov in prepoznavamo vzorce, ki nam omogočajo nov googled v literarno zgodovino, na primer z raziskavami razvoja književnih oblik, žanrov ali različnih literarnih postopkov ter tematik, ki so zanimale ustvarjalke in ustvarjalce. Sodobne raziskave lahko književna besedila na primer mapirajo v geografske prostore ali jih postavijo v določene zgodovinske, kulturne, politične in socialne kontekste. Poleg tega vizualizacije, ki so informativne in lahko tudi interaktivne, raziskovalna odkritja, s tem pa tudi področje

1. Anhert idr. 2020

2. Anhert idr. 2020

literature približajo širši javnosti. Računalniške analize omrežij, s katerimi algoritmi iz literarnih del ali drugih dokumentov samodejno črpajo podatke, pokažejo, kako se obravnavane teme vpletajo v širše kulturno tkivo. Poleg tega tovrstne raziskave pokažejo na primer, kako se oblikujejo literarne smeri in tokovi ter kako odnosi med določenimi elementi vplivajo na ustvarjanje, interpretacijo in razumevanje književnosti.

Primer: vizualizacija literarne produkcije

Raziskovalna skupina iz Lausanne je oblikovala **vizualno predstavitev ustvarjalnega procesa**, ki mu je sledil pisatelj Gustave Roud. Gre za t. i. genetske mreže, ker se nanašajo na analizo geneze književnih besedil. Raziskava se je pričela v arhivih, kjer so pregledali avtorjeve dnevnike, beležke in druge zapise njegovih zamisli, ki so spremljale literarno ustvarjanje. Na tej podlagi je nastala zbirka podatkov, ki vključuje tudi odnose med posameznimi elementi. Z vidika povezav med deli gradiva so nastale nove interpretacije pisateljevega opusa.³

Prvi korak pri analizah omrežij v literaturi je, da raziskovalci ali raziskovalke podatke o literarnem sistemu in drugih vidikih, ki jih zanimajo, jasno strukturirajo, zatem pa jih

obdelajo z računalniškimi programi. Omrežje je vizualizacija teh odnosov in prikazuje, kako so določeni elementi medsebojno povezani v določenem času ter prostoru. To vključuje različne tipe podatkov: razmišljanje o literaturi z vidika omrežij je interdisciplinarno, saj presega posamezna področja raziskav.⁴ Pomemben je razvoj novih in novih tipov vizualizacij znanja: “Da bi razumeli, kako je mogoče računalniške metode vključiti v humanistične discipline, potrebujemo boljše metafore prostora.”⁵

Analiza omrežij predpostavlja tudi nekoliko drugačen pogled na književnost kot sklop medsebojnih povezav med različnimi akterji. Ti pristopi v literarno vedo prinašajo nov zorni kot in za razliko od hierarhij ter struktur, ki imajo jasno središče, pomenijo decentralizacijo.⁶



2 Literarni vplivi z digitalnega vidika

Področje medsebojnih vplivov med literarnimi tokovi in smermi, pa tudi posameznimi avtoricami oziroma avtorji v zgodovini je že tradicionalno polje literarnovednih

4. Anhert idr. 2020

5. Gavin 2016

6. Jagoda 2017

raziskav. Na podlagi digitalnih metod in orodij lahko dobimo drugačen vpogled v te dinamike. Z računalniki lahko raziskovanje literarnih vplivov poglobimo in sistematiziramo ter ga s tem bolje razumemo, in sicer se pokaže, da je proces literarnih vplivov zelo pluralen.⁷ Z digitalnimi tehnologijami najdemo več primerov navezovanja na določen motiv ali temo ali več primerov ponavljanja posameznih fraz, ki so jih piske in pisci povzemali drug od drugega. Nova odkritja postavljajo pod vprašaj tradicionalne predstave o posameznem viru, ki ima neposreden vpliv na druge avtorice in avtorje, saj lahko digitalne analize odkrijejo številne druge navezave.⁸

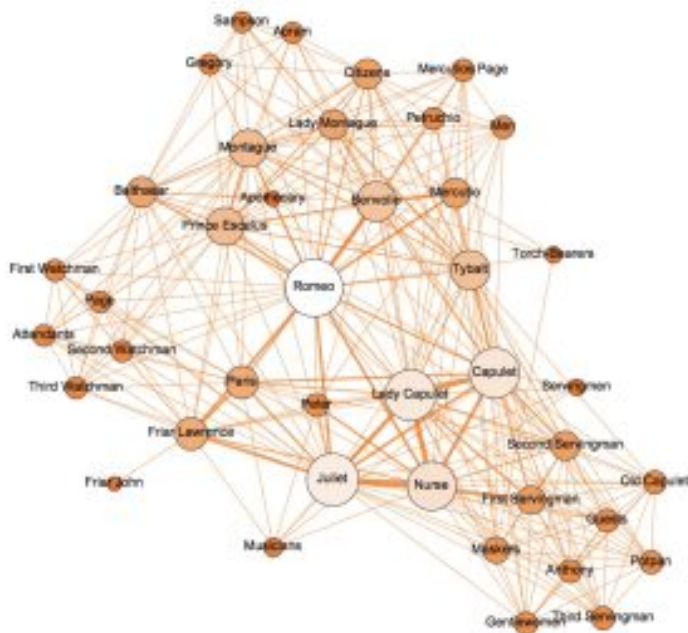
3 Odnosi med liki v fiktivnih zgodbah

Kako se liki v literarnih delih medsebojno povezujejo? Kateri liki so osrednji in v kakšne skupine se delijo? Na ta in podobna vprašanja lahko odgovorimo z analizo omrežij likov v književnih besedilih. Tovrstna raziskava tragedij Williama Shakespearea na primer jasno pokaže medsebojno ločene skupine likov, ki oblikujejo dramsko napetost: Montegi in Kapuleti v *Romeu in Juliji*, Trojanci in Grki v *Troilus in Kresidi*, triumvirske stranke in Egipčani v *Antoniju in Kleopatri*, Volščani in Rimljani v *Koriolanu* ter skupini zarotnikov v *Juliju Cezarju*.⁹

7. Shore 2016

8. Shore 2016

9. Grandjean 2015



Slika: Omrežje likov v tragediji Williama Shakespearea Romeo in Julija; vir: Grandjean 2015

Pripovedi, posebej daljše, so včasih zapletene in zato je z bližnjim branjem težko ugotoviti, kakšne so dejansko povezave med literarnimi osebami, na primer kateri liki so najbolj povezani z drugimi književnimi osebami. Statistične analize, ki vključujejo število posameznih govornih izmenjav med liki ali njihovih srečevanj, pa kompleksne

mreže relacij med liki v fiktivnem svetu prikažejo na razumljiv in intuitiven način.

Primer: mreža odnosov v seriji romanov z 2000 liki

Za bralce popularne literature so zanimive analize omrežja obsežnih književnih del, na primer analiza omrežja fantazijske zbirke romanov v petih delih *Igra prestolov* ameriškega pisatelja Georgea R. R. Martina. V analizi obsežne zbirke romanov, kjer nastopa okrog 2000 literarnih oseb, se je pokazalo, da sta lika, ki sta najbolj povezana z drugimi liki knjižne serije, Tiron in John Snow, sledijo pa ostali osrednji liki: Daenerys, Robert, Robb, Sansa in Jamie.¹⁰ Kljub velikemu številu književnih likov je analiza socialnih omrežij tekom petih romanov pokazala, da besedilo sestavljajo zorni koti zgolj 15 osrednjih likov, preostali liki so stranski. Socialne mreže posameznih likov so sestavljene iz okrog 150 posameznikov, kar je pogosto tudi v strukturi realnih družbenih vezi in s tem bralcu omogoča, da razmišlja v okvirih, ki so mu znani z vsakdanjega življenja. Zgodba torej odseva vsakdanje izkušnje bralcev in se s tem prilagaja njihovim kognitivnim sposobnostim obvladovanja odnosov v družbi.¹¹

3 Odnosi med člani kulturnega sistema

Analize omrežij v književnosti in literarni

10. Andrew Beveridge in Jie Shan 2016: 18-22.

11. Gessey-Jones 2020

zgodovini lahko pokažejo tudi na odnose med knjižnimi avtorji in avtoricami v posameznem literarno zgodovinskem času ter osvetlijo njihovo medsebojno sodelovanje, na primer medbesedilne povezave med literarnimi deli prek slicevanj in citatov. V širšem smislu lahko raziskujemo, kako se v določenem časovnem obdobju širijo razni slogovni prijemi ali literarna gibanja. Pogled na literaturo z vidika omrežij tudi literarno avtorstvo predstavlja v drugi luči, in sicer kot kolektivni ustvarjalni proces, ob tem pa literarno besedilo razlaga kot široko razvejan model najrazličnejših družbenih odnosov, v katerem že od začetka nastajanja literarnega dela najdemo več različnih udeležencev.¹² Literarnovedne študije tradicionalno temeljijo na konceptu naroda, kar pogosto razumejo v smislu mistificirane avtohtonosti – idejo, da so literatke in literati ter njihova dela povezani s svojim rojstnim krajem. Caroline Levine¹³ predlaga, da namesto kategorije naroda uporabljamo kategorijo omrežja, s čimer se pokažejo nove razsežnosti literarnega sistema na določenem območju v določenem obdobju. Ne le to – književne ustvarjalke in ustvarjalci se čezmejno povezujejo še z drugimi akterji v kulturnem

12. Thomalla idr. 2019

13. Levine 2013

sistemu: osebami, ki delajo v založništvu, v medijih, na univerzi in drugih kulturnih ustanovah. Omrežja lahko tako opazujemo z vidika različnih delovalnikov sistema književnosti, na primer omrežja okrog literarnih revij (avtorstvo, uredništvo, založništvo itd.), intelektualna omrežja sovplivanja med posamezniki in posameznicami (piskami, pisci, urednicami, uredniki, kritičarkami, kritiki itd.), omrežja tematik v literarnih delih in drugih tekstih. Pri analizi omrežja pišočih posameznic in posameznikov postanejo razvidne kompleksne skupine, ki določajo literarni sistem.

Primer: omrežja pisem

[Elektronska zbirka Pisma](#), ki nastaja na Raziskovalnem centru za humanistiko Univerze v Novi Gorici, vsebuje 1560 pisem iz 19. in 20. stoletja. Obsega pisma avtoric in avtorjev slovenske moderne, pa tudi druge korespondence, na primer pisma družine Kremenšek. Ta mreža pisem vsebuje zapise družinskih članov in njihovih prijateljev ter znancev. Spodnje omrežje je vizualizacija te korespondence, ki je nastala na podlagi metapodatkov, s katerimi so opremljena pisma. Prikazuje pisemske povezave med posameznicami in posamezniki.



Marinković, Mila, 2023: Data visualizations in the “PISMA” collection: Cultural Heritage at the intersection of informatics and literary science. Tipkopis.

Eden najbolj znanih primerov analize omrežja pisem je sicer projekt [Mapiranje republike pisem](#) z Univerze Stanford. Namen tega projekta je bil rekonstruirati in analizirati medsebojne povezave intelektualcev in pisateljic in pisateljev iz obdobja razsvetljenstva. Cilj projekta je bil ustvariti digitalno vizualizacijo obsežne korespondenčne mreže med omenjenimi posameznicami in posamezniki v 17. in 18. stoletju, ki je temeljila na zbranih metapodatkih iz tisočih pisem, ki so si jih izmenjevali. Med vključenimi v raziskavo so bile pomembne osebnosti, kot so M. de Voltaire, Jean-Jacques Rousseau in Gottfried Wilhelm Leibniz.

Metapodatki so med drugim vključevali strukturirane informacije o pošiljatelju, prejemniku, datumu in lokaciji vsakega pisma. Avtorice oziroma avtorji pisem so predstavljali posamezno vozlišče v omrežju. Velikost obsega njihovih korespondenc je na vizualizacijah pisemskih omrežij na primer označena z velikostjo kroga v vozlišču posamezne piske ali pisca, črte, ki imena medsebojno povezujejo, pa kažejo tokove pisemskih poti. Analiza je pokazala, kdo je bil v središču intelektualne izmenjave in kakšen je bil intelektualni razvoj v tem obdobju, kako so ideje in informacije potovale po Evropi prek dopisovanja ter kako je potekal vzpon in padec določenih središč intelektualne dejavnosti, pa tudi kako se je omrežje razvijalo skozi čas.¹⁴

Analize omrežij na področju umetnosti in humanistike niso le formalistični opis določenih povezav med posameznimi elementi, ampak ponujajo možnost za zaznavanje implicitne ter eksplicitne prisotnosti najrazličnejših delovalnikov, ki so povezani z literarnimi deli, s tem pa raziskovalce in raziskovalke spodbujajo, da presedlajo od ustaljenih analitičnih kategorij in začnejo razmišljati o širših vidikih, ki vplivajo na literaturo.¹⁵ Analize omrežij nudijo poidlago za raziskovanje procesov razširjanja idej in inovacij v literarnem sistemu.

Primer: Omrežje razvoja literarne kritike v Angliji

14. Edelstein idr. 2017

15. Hachmann 2023

Kritika literarnih del se je v Angliji kot publicistična zvrst razvila v poznem 17. stoletju. Tedaj so jo razumeli kot nevtralen mehanizem, ki presega nesoglasja med posameznimi literatkami in literati ter je pokazatelj literarne vrednosti. Analiza omrežja kritik, ki se je osredotočila na njihovo produkcijo, je pokazala, da so literarno kritiko pomembno podpirali knjigotržci in prodajalci knjig, ki so promovirali dela razmeroma majhne peščice plodovitih pisateljic in pisateljev znotraj veliko širšega polja pišočih.¹⁶

4 Povezave med temami v besedilih

Katere teme ali ključne besede se pojavljajo v literarnih delih in kako se medsebojno povezujejo? To je še eno vprašanje, na katero lahko odgovarjamo z računalniškimi analizami omrežij. Tovrstne tematske analize, ki sicer temeljijo na kvantitativnih metodah, raziskovalkam in raziskovalcem omogočajo kvalitativno interpretacijo gradiva, saj omogočajo natančen vpogled v pomske mreže tekstov. Opazujemo lahko razvoj tematik v literarnih delih skozi čas in ugotavljamo, katere teme so zanimale ustvarjalke in ustvarjalce. Poleg tega lahko tovrstna analiza zajame tudi zgodovinske dokumente, kakršna so pisma.

Primer: Povezave med temami v elektronski zbirki Pisma

Raziskava je temeljila na zbirki pisem avtoric moderne v okviru že omenjene spletne zbirke Pisma in je bila izvedena z orodjem ClowdFlows, ki so ga razvili na Odseku za tehnologije znanja na Institutu Jozef Stefan. Cilj raziskave je bil ugotoviti, s katerimi besedami se povezujejo izrazi na pomenskem polju ljubezni; eden od rezultatov raziskave – analiza povezav besede “srce” – je vizualiziran na spodnji sliki. Omenjeno pomensko polje smo izbrali na podlagi bližnjega branja pisem, pri katerem smo ugotovili, da je bila ljubezen v korespondencah pisateljic, urednic in intelektualk v času slovenske moderne tema, ki je bila velikokrat zatrta. Avtorice so na primer, ko je beseda nanesla na ljubezen, hitro spremenile tematiko pisanja. Spodnja slika prikazuje izraze, ki kažejo, s katerimi drugimi besedami se v pismih povezuje izraz “srce”. Pokazalo se je na primer, da se beseda “srce” povezuje z besedo “skrivati”, beseda “ljubezen” pa z izrazom “skrivnost”. To je še en pokazatelj, da je bila ljubezen kot čustvo v tem kontekstu podvržena skrivanju.



Analiza povezav besede “srce” v zbirki pism avtoric moderne v okviru spletne zbirke Pisma; vir: lasten

Zbirka pism pisateljic, urednic in intelektualk v času slovenske moderne Pisma obsega tudi metapodatke o tematikah, ki jih vsebujejo posamezna pisma. S spodnjega prikaza je razvidno, s katerimi drugimi temami se povezuje tematika ljubezni. Kot vidimo, so te teme najrazličnejše, od bolj predvidljivih, kot so zakon, spolnost n osebna razmerja, do manj predvidljivih, kot so delavnost, domoljubje in politika. Med drugim najdemo tudi temo cenzure. Ugotovitve kažejo, da se je tema ljubezni v pismih pogosto povezovala z najrazličnejšimi drugimi temami.

takšna omrežja ne morejo zagotoviti popolnoma natančnih predstavitev danega korpusa ali odgovorov na raziskovalna vprašanja. Vizualna mreža namreč predstavlja le del sicer izjemno kompleksnih relacij med posameznimi elementi literarnega sistema, kljub temu pa daje vpogled v kompleksne povezave z vidika izbranih kategorij. Pri kakršni koli uporabi kvantitativne, računalniško podprte analize omrežij je ključno, da raziskovalci in raziskovalke zbirke podatkov in vizualizacije rezultatov podprejo s teoretičnimi uvidi.¹⁷

6.

LITERARNE PRIPOVEDI Z VIDIKA DIGITALNIH RAZISKAV

Priповed je poročilo o dogajanju, ki se v posameznem literarnem delu razteza čez določen časovni razpon, ki ga zaznamujejo posamezni dogodki in literarni liki. Je ena ključnih lastnosti velikega dela besedil iz svetovne literarne tradicije od epov prek romanov do dramskih besedil in drugih književnih vrst ter žanrov. Z digitalnimi metodami lahko pripovedni vidik raziskujemo na različne načine, na primer z avtomatsko tvorbo povzetkov literarnih del, vizualizacijo pripovednih lokov in mrež odnosov med književnimi liki.

Eden ključnih vidikov književnosti je pripoved, ki jo lahko s sodobnimi računalniškimi metodami analiziramo z različnih zornih kotov – nekatere od njih predstavljamo

v tem poglavju. Digitalne analize pripovednih literarnih besedil raziskovalkam in raziskovalcem omogočajo, da z oddaljenim pogledom pridejo do novih uvidov o zgradbi zgodb, in sicer je možno raziskovati najrazličnejše vidike pripovedi, na primer nove pripovedne tehnike, geografske vidike pripovedi, čustvene vidike pripovedi, variante besedil, pa tudi njihove adaptacije v drugih medijih.

1 Pripovedni lok in emocionalni lok

Da se pripovedi razvijajo v lokih, ki jih zaznamujeta stopnjevanje in opuščanje napetosti, so opazili že mnogi bralci. Tudi v polju literarne vede velja soglasje, da pripovedi navadno sestavljajo začetek, jedro ter konec, kar naj bi jim dajalo določeno strukturo. Kot izhaja iz tega razumevanja in kot kažejo računalniške analize tega vidika pripovedi, se struktura zgodbe razlikuje od njene vsebine: dva romana z različnim dogajanjem in liki imata lahko enako strukturo.¹ Če bi loka v obeh romanih torej vizualizirali, bi dobili precej podobni sliki.

Primer: pripovedni lok

S pripovednim lokom kot eno ključnih lastnosti pripovedi se je ukvarjala obsežna raziskava, ki je z računalniško podprtimi metodami jezikovne analize merila več strukturnih in psiholoških kategorij jezika. Obravnavani korpus je obsegal 40.000 t. i. tradicionalnih fiktivnih pripovedi, kakršne najdemo na primer v romanih in

1. Boyd idr. 2020

filmskih scenarijih. Poleg tega je raziskava zajela besedila, ki niso literarna, a imajo določene elemente pripovedi, in sicer je korpus obsegal 20.000 t. i. netradicionalnih pripovedi. Gre na primer za znanstvena poročila v časopisnih člankih, govore TED in mnenja vrhovnega sodišča. Raziskava je pokazala, da se v tradicionalnih pripovedih dosledno ponavlja temeljna struktura zgodbe, ki je razkrila tri primarne procese: uprizoritev, napredovanje zapleta in kognitivno napetost. Dokazov, da bi bila prisotnost normativnih struktur zgodbe kakor koli povezana z njeno priljubljenostjo, raziskava ni našla. Analiza t. i. netradicionalnih pripovedi, ki temeljijo na dejstvih, je razkrila strukture, ki so se razlikovale od t. i. tradicionalnih pripovedi, tj. pripovedi, ki temeljijo na fiktivnih zgodbah.²

Predpostavko, da se pripovedi odvijajo v obliki pripovednega loka, so raziskovalci z računalskimi metodami preizkusili tudi na drugačen način, in sicer ker je dogajanje v zgodbah predvsem čustvene narave, so preverili, kako se tekom zgodbe razvijajo čustvene dimenzije dogajanja v obliki emocionalnega loka.



Gre za graf, ki prikazuje pojavnost pozitivnih oziroma

negativnih dogodkov skozi razvoj pripovedi. Emocionalni lok je torej eden od vidikov zgodbe poleg kronološkega dogajanja, ki se odvija, in pomenov, ki jih prepoznavajo bralci.³ To ni niti zgodba niti struktura, saj se lahko podobna emocionalna loka izrišeta pri pripovedih z zelo različnimi zgodbami ali strukturami; gre za nihanja čustvene intenzivnosti in čustvenega naboja (negativnega ali pozitivnega) v zgodbah.⁴

Primer: Raziskava čustev v zgodbah

Analizo čustvenih lokov v različnih literarnih pripovedih, katere rezultati so dostopni na spletni strani [Hedonometer](#), je raziskovalna skupina izvedla z istoimenskim orodjem, ki preverja čustvene izraze v posameznih delih literarnega besedila in izriše graf emocionalnega nihanja v besedilu. Raziskava je zajela 1327 najbolj popularnih knjig spletne zbirke Project Gutenberg, ob tem pa pokazala, da se dela glede na potek emocionalnega loka delijo na šest tipov:⁵

- “od revščine k bogastvu” (vzpon),
- tragedija ali “od bogastva v revščino” (padec),
- “človek v luknji” (padec-vzpon),
- “Ikar” (vzpon-padec),
- “Pepelka” (vzpon-padec-vzpon),

3. Reagan idr. 2016

4. Reagan idr. 2016: 2

5. Reagan idr. 2016

- “Ojdip” (padec-vzpon-padec).

Izziv:

Na spletni strani Hedonometer preverite tovrstni lok za delo Lewisa Carola [Alica v čudežni deželi](#). Odgovorite na vprašanja: V kateri tip zgodbe bi uvrstili delo glede na potek emocionalnega loka? Ali je beseda zajec (ang. rabbit) povezana z negativnimi ali s pozitivnimi čustvi v zgodbi?

Primerjajte graf emocionalnega loka v delu *Alica v čudežni deželi*, ki predstavlja izrazito stopnjevanje od negativnih občutkov do pozitivnih občutkov in nazaj k negativnim, z grafom emocionalnega loka v romanu [Ana Karenina](#) Leva Nikolajeviča Tolstoja. Komentirajte razlike z vidika vsebine vsake od zgodb.

2 (Ne)srečni konci zgodb

Računalniške metode raziskovanja književnosti dajejo vpogled v zgodbene strukture “od daleč” na način, ki z bližnjim



branjem ni dostopen. Tako lahko na primer identificiramo in primerjamo tipe koncev v velikem številu romanov (so srečni ali ne?), pa tudi dinamiko celotnih pripovedi.

Pomemben del literarnih zgodb so njihovi zaključki, ki lahko za osrednje like pomenijo nekaj pozitivnega ali negativnega, kar pomeni, da je konec zgodbe srečen ali tragičen. Računalniški algoritmi lahko prepoznajo ta dva tipa zaključkov zgodb v romanesknih pripovedih. Ena od raziskav je preverjala, ali računalniški programi konce 212 nemških romanov klasificirajo enako kot usposobljeni bralci.⁶ Strategija, ki so jo raziskovalci ubrali, je preverjanje pozitivnega ali negativnega čustvenega naboja zgodb. Ugotovili so, da je program pri prepoznavanju koncev uspešen.

3 Povzetki literarnih del

Avtomatsko oblikovanje povzetkov različnih tipov besedil z računalniškimi programi, na primer pravnih ali publicističnih besedil, je razširjena praksa, znotraj katere pa ima povzemanje literarnih del svoje specifike. V književnih besedilih najdemo namreč pogosto rabo metaforičnega in figurativnega jezika, zamolke določenih delov vsebine, ki jih s svojim sklepanjem glede na druge dele besedila zapolnjujejo bralci, rabo dialogov ipd., zaradi teh lastnosti književnosti pa je oblikovanje avtomatiziranih povzetkov z računalniškimi orodji dosti bolj zapleteno kot za besedila, ki so v večji meri strukturirana.⁷

6. Zehe idr. 2016

7. Kazantseva in Szpakowicz 2010: 72

V središču literarnih pripovedi je pojem časa in odvijanja dogodkov. Ko razumemo zgodbo, lahko razumemo časovni vrstni red dogajanja dogodkov. Dano besedilo ima lahko več zgodb; ko takšno besedilo razumemo, lahko razločimo te različne zgodbe. Tako razumevanje zgodbe iz besedila vključuje izgradnjo globalnega modela zaporedja dogodkov v besedilu, pa tudi strukture ugnezdenih zgodb. Končno lahko razumemo čustvena stanja, povezana z agenti, in se lahko vživimo v določene rezultate. Te ravni človeškega razumevanja lahko predlagajo metode za modeliranje strukture pripovedi za uporabo pri povzemanju.

4 Povezave med fiktivnimi liki

Pripovedi so umetno ustvarjene skupnosti, fiktivno oblikovane družbene oblike, ki jih v njih najdemo, pa je mogoče kvantificirati in analizirati, pri čemer se nam izrišejo mreže povezav med literarnimi liki.⁸ Nove raziskave pripovedi z računalniškimi tehnologijami so tako poleg ustaljenih kategorij, kakršna je na primer dihonomija med zgodbo in pripovedjo,⁹ uvedle tudi koncept mreže vezi med književnimi osebami. V taki socialni mreži posamezne literarne like predstavljajo vozlišča, povezave med njimi pa so označene s črtami, ki so v primeru močnejših vezi odebeljene, v primeru šibkejših vezi pa tanke. Izhodišče tega tipa vizualizacij je, da je mogoče povezave med fiktivnimi liki v literarnih

8. Sack 2012

9. prim. Zupan Sosič 2013

besedilih zaznati in kvantificirati na podlagi njihovih interakcij, na primer z dialogi med njimi.

V primeru raziskovanja skupnih dialogov med liki računalniški algoritem zaznava, kdaj se dva lika pogovarjata, in sicer identificira imena fiktivnih oseb ter jim pripisuje posamezne dialoške dele besedila.¹⁰ To pomeni, da algoritem identificira, katera od literarnih oseb govori določen del besedila. Zaznavanje govora je ključno, vendar ga je težko avtomatizirati, saj v pripovednih literarnih delih večino izjav v dialogih liki izgovarjajo, ne da bi jih pripoved eksplicitno poimenovala. Včasih so imenovani s parafrazo ali pa je vir govora razviden iz konteksta. Zaradi tega razloga je avtomatizirano zaznavanje subjektov v literarnih dialogih zahteven proces.

Romane raziskovalke in raziskovalci pogosto razlagajo z vidika družbenih razmerij in skupnosti, v katerih literarni liki živijo. Raziskave omrežij med književnimi osebami so na primer pokazale, da dolgi romani vsebujejo večja ali bolj zapletena družbena omrežja kot kratki.¹¹

Primer: Je več dialogov na vasi ali v mestu? Družbene vezi v angleških romanih iz 19. stoletja

Ker se študije romanov v 19. stoletju pogosto ukvarjajo s socialnim krogom protagonistov, angleška študija¹² preverja nekatere od teh teorij. Med njimi so

10. David Elson, Nicholas Dames in Kathleen McKeown 2010

11. Alexander 2019

12. Elson idr. 2010

predvidevanja, da se dialogi z vse večjim številom likov v romanih krčijo, saj naj bi liki zaradi množičnosti dobili vse manj priložnosti za medsebojne dialoge. Spet druga teroija pravi, da je v romanih, ki se dogajajo v mestnem okolju, manj dialogov, v tistih, ki se dogajajo v vaškem okolju pa naj bi jih bilo več.¹³ Program, ki samodejno zaznava dialoge v literarnih delih in identificira njihove govorce, je preveril mreže govornih interakcij v 60 angleških romanih, med katerimi so klasiki, kot sta Charles Dickens in Jane Austen. Razislava je preverjala gostoto mreže povezav med liki in morebitne družbeno izključene like, ki imajo omejeno komunikacijo z drugimi.

Raziskava je tezo, da posamezni liki v romanih, ki imajo več likov, tudi manj govorijo, ovrгла, saj na tem področju z večjim številom literarnih oseb ne prihaja do bistvenih sprememb. Prav tako se je izkazalo, da v romanih, ki se dogajajo v mestnem okolju, ni manj dialogov v primerjavi s tistimi, ki se dogajajo v vaškem okolju. V obeh tipih romanov so namreč našli podobno število literarnih oseb in njihovih dialoških stikov. Raziskava je pokazala tudi, da vsebujejo romani, ki imajo prvoosebna pripovedovalca, mnogo manj dialoških delov, kjer dobijo glas drugi romaneskni liki, v primerjavi z romani s tretjeosebno pripovedovalcem, v katerih različni liki v večji meri sodelujejo v dialoških izmenjavah.¹⁴

Dialogi pa niso edini pokazatelj stikov med književnimi osebami – družbena omrežja v literarnih delih lahko

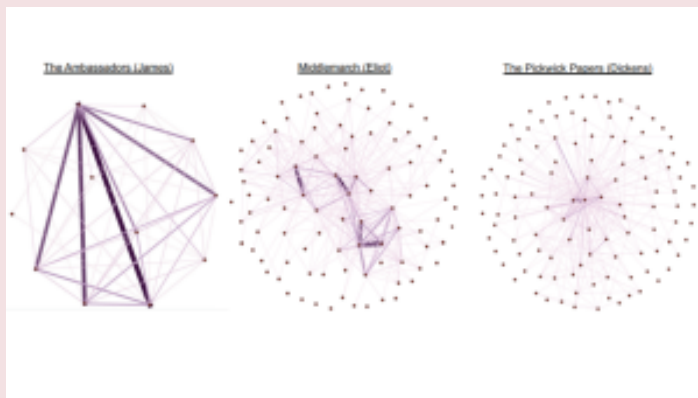
13. Elson idr. 2010: 138

14. Elson idr. 2010: 145

analiziramo tudi na podlagi sopojavljanja književnih likov v posameznih odstavkih literarnih besedil.¹⁵

Primer: Kateri liki se v romanih srečujejo?

V okviru Sackove raziskave na primer “močne vezi” med likoma v romanih pomenijo sopojavljanje teh likov v 20 ali več odstavkih besedila. Odebeljene črte v spodnjih vizualizacijah, ki prikazujejo vezi v romanih Henryja Jamesa, Georgea Eliota in Charlesa Dickensa, označujejo močnejše vezi med danimi liki, tanjše črte pa označujejo šibkejše vezi.



Slika: Dela Henryja Jamesa *The Ambassadors*, Georgea Eliota in *The Pickwick Papers* Charlesa Dickensa¹⁶

Zgornje slike prikazujejo tri tipične strukturne tipe vezi med literarnimi liki v romanih, ki jih je zaznala raziskava. Delo *The Pickwick Papers* Charlesa Dickensa (na zgornji

15. Sack 2012: 38

16. Sack 2012: 38

sliki desno) je pikareska, ki je strukturirana v obliki izletov po okolici, ki so medsebojno nepovezani. Omrežje je temu primerno oblikovano v obliki zvezde s socialnimi srečanji, kjer prevladuje eno samo središče znotraj tesno povezane ožje skupine likov (odebeljene črte), obdano z razpršeno mrežo drugih, šibkejših vozlišč (tanke črte), ki pa so medsebojno nepovezane. Roman *Middlemarch* Georgea Eliota (na zgornji sliki v sredini) prav tako prikazuje veliko omrežje vezi med liki, ki pa so v večji meri medsebojno prepleteni, kar posebej velja za osrednje like, ki so izrazito povezani, medtem ko se s stranskimi liki povezujejo v manjši meri. Vezi med osrednjimi liki so močne (odebeljene črte) in kažejo obliko triade: dva lika, ki sta prijatelja s tretjim likom, sta pogosto tudi prijatelja ali znanca. Tudi vezi v romanu Henryja Jamesa *The Ambassadors* (na zgornji sliki levo) so razporejene v obliki triade, ob tem pa opazimo, da si liki, ki so medsebojno izrazito povezani, delijo tudi skupne prijatelje oziroma znance. Vsa vozlišča so povezana drug z drugim in vse literarne osebe se medsebojno poznajo, kar je kontrast tipu romana, kakršen je *The Pickwick Papers* Charlesa Dickensa, kjer se književni liki medsebojno ne poznajo, povezuje pa jih skupina osrednjih likov, ki prihaja z vsakim v stik.¹⁷

5 Ko se modernizem sreča z digitalno humanistiko

Računalniško modeliranje lahko uporabimo za zasledovanje sprememb v svetovni književnosti in zaznavanje obsežnih vzorcev v procesu različnih literarnih premen. Primer so t. i. nove študije modernizma, mnoge raziskave obdobja literarnega modernizma, ki je v

17. Sack 2012: 39

literarno polje prinesel specifične slogovne inovacije. Tako študije na primer raziskujejo pripovedne tehnike, značilne za modernizem: nenadne zasuke zgodbe nazaj ali naprej v času, spremembe ravni pripovedi, tok zavesti in prosto posrednega govora kot specifičnega tipa tretjeosebne pripovedovanja, ki vključuje nekatere prvine neposrednega govora v prvi osebi. Te tehnike lahko raziskave zasledujejo v različnih časovnih obdobjih in literarnih žanrih.¹⁸

Primeri: modernistične pripovedi – od toka zavesti do prikaza mest

- **Širjenje modernističnega toka zavesti v književnosti**

Tok zavesti je pomemben postopek literarnozgodovinskega obdobja modernizma, ki ga je mogoče raziskovati tudi z digitalnimi metodami. Da je to dosegla, je raziskovalna skupina tok zavesti najprej opredelila prek nabora jezikovnih značilnosti in retoričnih elementov, ki ga zaznamujejo. Na tej podlagi so oblikovali model, ki je sledil gibanju tega modernističnega postopka prek geografskih meja (od Evrope do Amerike in Azije) različnih žanrov (od modernističnega romana do žanrov popularne kulture) in širom različnih jezikov (od angleščine do japonščine in mnogih drugih). Raziskava je pokazala, da se je tok zavesti po svetu širil po vzorcu valov, vendar ne gre za en sam

vzorec širjenja in vpliva, ampak za razdrobljen model pojavljani toka zavesti v različnih variacijah.¹⁹

- **Tridimenzionalni modeli prikaza mest v modernističnih pripovedih**

[Projekt Univerze v Viktoriji](#) raziskuje, kako modernistična pripoved fragmentira in izkrivlja zgodovinski čas ter prostor. Z oblikovanjem tridimenzionalnim modelom mest prikazuje subjektivno doživljanje prostora in odgovarja na vprašanje, kako so znana upodobljena mesta (npr. Pariz, London, Dublin in New York) na različne načine prikazana v delih različnih modernističnih avtoric ter avtorjev. Podatkovni model je izhajal iz števila besed v posameznem književnem besedilu, ki se nanašajo na posamezno lokacijo. S to hkrati analogno preštevno in hkrati algoritemsko metodo je raziskovalna skupina na primer v izdaji *Uliksesa* Jamesa Joycea preštela besede, ki se nanašajo na določene prepoznavne predele Dublina v času nastanka besedila. To število so zatem delili s skupnim številom vseh besed v knjigi in dobili razmerja za vsako področje, delo na primer 30.262 besed (ali približno 13 % besedila romana) dogajanju na naslovu Ecclesovi ulica 7. Na podlagi tabele teh podatkov so na zemljevidu iz leta 1925 ustrezno poudarili posamezna območja, ki so v romaneskni zgodbi pomembna.²⁰

6 Kako napisati uspešnico

19. Long in Jean 2016

20. Tanigawa 2013

Več digitalnih raziskav književnosti se ukvarja z vprašanjem, kaj je značilno za knjižne uspešnice in zakaj v množici knjig izstopajo. Gre za kompleksna vprašanja, ki se jih digitalne raziskave književnosti lotevajo z vidika posameznih elementov besedil.

Primer: Ključ do uspeha detektivskih romanov

Že Franco Moretti je ugotavljal, katere značilnosti detektivskih romanov jim zagotavljajo uspeh. Jernej Habjan v članku predstavlja njegovo raziskavo uspešnic o Sherlocku Holmesu pisatelja Conana Doylea. Sprva angleški romanopisec ni bil najbolj uspešen avtor, vendar je njegovo delo o Holmesu leta 1892 izšlo v 10.000 izvodih, leta 1893 pa je eden njegovih detektivskih romanov dosegel milijon bralcev samo v Veliki Britaniji. Kot ugotavlja Moretti, je postal avtor uspešen, ko je začel v svojih detektivskih pripovedih učinkovito uporabljati namige, ki so jih bralci razvozlavali.²¹



Slika: Conan Doyle, vir: java domena, Wikipedia

21. Moretti 2005; Habjan 2012

Moretti²² je izdelal drevo evolucije zgodnje detektivske zgodbe, ki vključuje tudi namige v pripovedih. O tem je opazil, da postajajo z vključevanjem namigov te zgodbe vse bolj popularne. Ob tem Moretti opazi tudi, da ostajajo Doylovi namigi v zgodbah pogosto nerazrešeni, kar razlaga z dejstvom, da pisatelj tega sredstva še ni znal popolnoma učinkovito uporabljati.²³

Z vprašanjem, katere so lastnosti, ki literarne uspešnice ločujejo od drugih književnih del, se je ukvarjala tudi raziskava Bena Blatta.²⁴ Preverjala je sestavine knjižnih uspešnic glede na nasvete za pisanje, ki so jih predlagali avtorji in avtorice uspešnic. Raziskovalec z uporabo digitalnih metod ugotavlja, ali pisci in piske tudi dejansko upoštevajo lastne nasvete za oblikovanje uspešnic in ali načela pisanja, ki jih predlagajo, najdemo tudi v drugih knjigah, ki so v javnosti dosegle največjo branost.²⁵

Primer: Koliko klicajev naj vsebuje prodajna uspešnica?

Uspešni pisatelj **Elmore Leonard**, ameriški romanopisec, pisec kratkih zgodb in scenarist, je tekom

22. 2005: 70–78

23. Moretti 2005; Habjan 2012

24. Ben Blatt, Nabokov's favorite word is Mauve. 84–97.

25. Blatt 2017: 83.

svoje kariere napisal več kot 40 romanov, skoraj polovica njegovih del pa je bila adaptirana v filmski scenarij. Uspešni literarni avtor je nadebudnim piscem in piscem dal zelo konkreten nasvet. Za uspešno oblikovanje pripovedi Leonard predlaga, da piska ali pisec na vsakih 100.000 besed uporabi le dva ali tri klicaje, ne več. Koliko klicajev pa vsebujejo njegovi romani? V skupno 45 naslovih najdemo 3,4 milijona besed in če bi sledil lastnemu nasvetu, bi v celotni karieri uporabil le 102 klicaja, v resnici pa jih je v vseh svojih delih uporabil 1651, kar je šestnajstkrat več od njegovega predloga.²⁶ Sklepamo lahko torej, da se recept za uspeh romanov ne skriva v številu klicajev, uporabljenih v besedilu. Kljub temu je raziskava pogostosti uporabe klicajev v delih 50 komercialno uspešnih ameriških avtorjev in avtoric pokazala, da je Elmore Leonard uporabil najmanj klicajev med vsemi pisci in pisce. Razlike med povprečji klicajev na 100.000 besed so bile pri raznih avtoricah in avtorjih zelo različne: če je Elmore Leonard v povprečji uporabil 49 klicajev, jih je James Joyce uporabil 1105. Pa Leonard res ni upošteval lastnega nasveta za pisanje? Kot pokaže raziskava pogostosti uporabe klicajev v njegovih delih, se je njihovo število v času njegove kariere močno znižalo: na začetku jih je uporabljal mnogo, prosti koncu in posebej po javni objavi omenjenega nasveta pa občutno manj.²⁷

Vaja: število klicajev

26. Ben Blatt, Nabokov's favorite word is Mauve. 83-87.

27. Blatt 2017: 86-87

Na platformi Wikivir poiščite romana Josipa Stritarja *Zorin*. S funkcijo iskanja preverite, koliko klicajev roman vsebuje. Besedilo romana prenesite v program za urejanje besedil in preverite njegovo dolžino. Izračunajte, koliko klicajev delo vsebuje na 100.000 besed. To vrednost primerjajte s pogostostjo rabe klicajev pri Elmoru Leonardu in Jamesu Joycu.



Preberite naslednji odlomek romana Josipa Stritarja *Zorin* in pri tem uporabite metodo bližnjega branja ter odgovorite na vprašanje: Na kakšne načine so v delu uporabljeni klicaji, kdo jih izreka? Kako klicaji vplivajo na karakterizacijo protagonista romana in na pomen danega odlomka?

Slika: Josip Stritar; vir: javna domena, Wikipedia

V Parizu, 4.
julija

Danes popoldne bil sem v Louvru. Dolgo že sem želel videti to glasovito zbirališče najiminitnejših del človeških rok, človeškega duha. Veliko sem pričakoval, še več sem našel! Nov, neznan svet se mi je odkril v tem velikanskem poslopju, kamor je poslal ves izobraženi svet, kar je ustvaril

najlepšega v stoletjih. Kar sem poznal prej le po imenu, po slabih popisih in podobah, to stoji zdaj živo in resnično pred mano.

Kako bi ti mogel popisati, kaj sem čutil, ko se zagledam nenadoma sredi slavnih starogrških kipov! Kar sem videl prej v Monakovski gliptoteki, to je proti temu, kar gledam zdaj, dejal bi, le predgovor, zarja, ki naznanja sonce. Tam sem bil v veži, zdaj stojim v samem svetišču. Kaka resničnost, kaka lepota, kak mir in kako življenje! Kako naravno, preprosto vse, in kako vzvišeno, umetno! Kako telo, kak duh, kake posameznosti, kaka celota, kaka ideja!

Tukaj šele sem spoznal pravi pomen tiste pripovedke o Pigmalionu, ki je oživil mrzli, trdi kamen na svojih vročih prsah.

Zamaknjen sem stal pred tem Apolonom, kateri, če ga gledaš nekaj časa, se ti zdi, kakor da bi se ti jel bližati – bog in kralj – v svojem mirnem, tihem veličastvu in móra te skoraj, da nehote pripogneš koleno pred njim. Zdaj to Artemido s košuto, ki je plaha pribežala v zavetje deviške lovke! In pa naposled to nebeško Afrodito (Miloško). Božja oblast! Je li mogoče, da je žena tako lepa! In ta ponos, ta plemenitost! Kar pa je največje čudo – obe roki sta ji pri ramah odlomljeni, in vendar človek, ki jo zamaknjen gleda, ne pogrša ničesar! V duhu vidi, kako jih ima v deviški sramežljivosti sklenjeni nad prsmi, s katerih ji je ravno obleka zdrsnila na kolena! — Ko se siloma odtrgaš tem, čaka te vrsta drugih podob, ki se jim ne moreš prečuditi.

Zaboga! Velike, neumrjoče imenujemo ter slavimo v pesmih in podobah samopašne, surove divjake, ki so vihrali od zemlje do zemlje s svojo divjo drhaljo – pred njimi strah in trepet, za njimi požar in jok in stok umirajočih! Duh pa božji se ne bliža v grmečem viharju, bliža se v rahlem

dihanju pohlevnega vetra. Velik, neumrjoč, stvarniku enak je v resnici umetnik, kateri iz svojega osrčja stvari podobe, ki jih je v svetem trenutku spočela njegova duša, ki jim je življenje dala njegova roka; podobe – kakor njihov stvarnik, neumrjoče, katere bodo živele, dokler bodo veljale večne postave lepote in resnice!²⁸

7 Nekaj kritičnih pogledov na literarne pripovedi z vidika digitalnih raziskav

Zgoraj omenjena raziskava širjenja modernističnega toka zavesti v svojih zaključkih opozarja na napačno razumevanje humanistov, češ da skušajo s sodobnim statističnim modeliranjem in nasploh s kvantitativnimi modeli razložiti vse vidike določenega pojava. To bi pomenilo, da pri tem ni prostora za različne interpretacije rezultatov ali za odprta mesta, ki še čakajo na odgovore. Menijo, da drži nasprotno: ključna značilnost vsakega statističnega modela je, da izpostavi svoje omejitve, tj. tisto, česar model ne more razložiti.²⁹ Digitalne raziskave literarnih pripovedi se torej osredotočajo le na določene vidike sicer kompleksnih naracij, njihove interpretacije pa pogosto vodijo k bližnjemu branju in odkrivanju literarnozgodovinskih dejstev, tj. k ustaljenim metodam literarne vede.

28. Josip Stritar, Wikivir

29. Long in Jean, 2016

7.

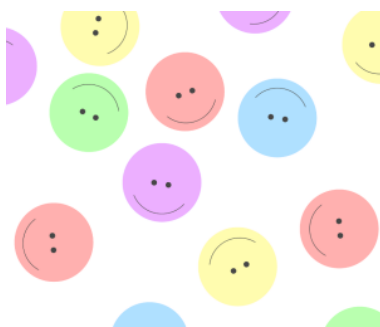
ANALIZA ČUSTEV V KNJIŽEVNOSTI

Po t. i. emocionalnem obratu v literarni vedi so čustva, ki so eden najpomembnejših delov izkušnje branja književnosti, pogosto središče zanimanja literarnih raziskovalcev, tudi tistih s področja digitalne humanistike. Kakšna čustvena stanja se odražajo na primer v romanih ali pa v literarnih kritikah, je mogoče preveriti z digitalnimi orodji za analizo čustev, ki lahko zajamejo obsežne korpuse besedil. Na ta način dobijo raziskovalci na primer vpogled v emocionalne dimenzije literarnih pripovedi, razvoja zgodbe in delovanja likov.

Tradicionalno je literarna veda področje čustev pogosto zapostavljala, čeprav gre za enega ključnih vidikov ustvarjanja in

sprejemanja književnosti. Sodobne metode literarne vede dajejo večji poudarek na čustvih tako v procesu literarne produkcije kot tudi v procesu branja književnosti, vendar se z digitalnimi metodami raziskovalke in raziskovalke ukvarjajo predvsem s čustvi v literarnih besedilih.

Z odpiranjem novih pogledov na dogajanja v okviru literarne zgodovine pa se tudi na tem polju pojavlja vprašanje čustvenih procesov, ki jih lahko raziskujemo z digitalno analizo besedil iz zgodovine, na primer pisem.



1 Občutje vs. čustvo

Orodja za analize čustev postajajo vse bolj natančna in izhajajo na primer iz obsežnih seznamov izrazov, ki imajo v posameznem jeziku določen emocionalni naboj. Algoritmi lahko razlikujejo med subjektivnimi in objektivnimi izjavami v literarnih besedilih in analizirajo občutja ter čustva. Občutje (ang. *sentiment*) je pozitivno ali negativno in spremlja določeno mnenje.¹ Analize s tega področja preverjajo, ali je občutje v besedilu pretežno

1. Liu 2015: 2

pozitivno ali negativno.² Spet druge raziskave preverjajo izključno intenzivnost čustvenega naboja v besedilih, ne glede na občutenje.³ Čustva so po drugi strani stanje, ki vključuje psihološke spremembe in motorični ter kognitivni odziv, obsega pa tudi osebne izkušnje posameznika v določeni situaciji.⁴

3 Analiza čustev v pripovedi

Analiziramo lahko dano književno delo ali literarni korpus, z digitalnimi orodji pa je v analizo mogoče zajeti zgolj posamezne čustvene vsebine, ki so povezane izključno z določenim vidikom danega besedila. Pri analizi tem književnih besedil v povezavi z občutji raziskave pogosto uporabljajo delitev na negativne, nevtralne in pozitivne, lahko pa tudi lestvico, je sestavljena iz vrednosti od 1 do 5, tj. najbolj pozitivna (5), zelo pozitivna (4), nevtralna (3), zelo negativna (2) in najbolj negativna (1). Poleg tega med tipe digitalne analize čustev uvrščamo prepoznavanje čustev glede na zaznavanje izrazov, ki se povezujejo z določenim čustvom v t. i. čustvenem leksikonu, zbirki besed, ki se povezujejo z izbranimi čustvi. Primer leksikona čustev v angleščini je [NRC Emotion](#)

2. Prim. Kim in Klinger 2021

3. Bei Yu 2008: 327-343

4. Mayer idr. 2008: 2

[Lexicon](#), seznam angleških besed in njihovih povezav z osmimi osnovnimi čustvi (jeza, strah, pričakovanje, zaupanje, presenečenje, žalost, veselje in gnus) in dvema občutkoma (negativno in pozitivno).

Kakšna čustva v zgodbi kaže pripovedovalka ali pripovedovalec – posebej če je prvoosebni/-a? Kakšne čustvene vsebine posredujejo posamezni liki ali pa kar besedilo kot celota? Z različnimi računalniškimi orodji, ki



omogočajo analizo čustev, lahko literarna dela pogledamo “od daleč” in tudi za obsežna besedila ugotovimo, kateri so čustveni vrhunci zgodb – ne z vidika bralčevega doživljanja, ampak z vidika pogostosti uporabe emocionalno nabitih besed.

Z menjavami različnih čustvenih stanj pisateljice in pisatelji oblikujejo napetosti v literarnih pripovedih in jih stopnjujejo.⁵ Z računalniškimi metodami lahko analiziramo prav te dinamike notranjega dogajanja v literarnih delih in opazujemo, kako se čustvena obarvanost besedila spreminja tekom zgodbe: je pripoved, ki se odvija, vse bolj negativno ali vse bolj pozitivno obarvana? Jo morda zaznamujejo nihanja od pozitivnih do negativnih čustev?

Primer: Stopnjevanje pozitivnih čustev med Romeom in Julijo

Eno ključnih gibal znane tragedije Williama Shakespearea *Romeo in Julija* je sklop pozitivnih čustev, ki jih zaljubljenca čutita. Vizualno mapiranje teh čustev je v rezultatih študije predstavljeno v hitro naraščajočem grafu: na horizontalni osi je označen razvoj dramskega besedila skozi verze, vrednosti na vertikalni osi pa kažejo valenco čustev. Linija se tekom tragedije izrazito vzpenja navzgor, kar pomeni stalno rast pozitivnih čustev v besedilu. Jasno je, da prehod v tragediji sega od sprva nevtralnih do izrazito pozitivnih čustev med Julijo (njena čustva označuje črna črta) in Romeom (njegova čustva so označena z belo črto, ki zaključuje sivino). Računalniška analiza je pokazala, da je rast njunih pozitivnih čustev nenehno vzajemna, saj sta njuna prikaza simetrična.⁶

4 Emocije v dialogih

6. Nalisnick in Baird 2013

Odkar je koncept dialošnosti v literaturi utemeljil vplivni raziskovalec Mihail Bahtin in pokazal, da so literarna dela pretežno dialoška, gre za enega ključnih konceptov literarne vede. Na dialošnost se lahko v okviru digitalnih metod literarne vede osredotočimo z vidika napetosti, ki jo v književnosti ustvarjajo medsebojne besedne izmenjave med književnimi liki.



Gre za to, da čustvene vsebine strojno analiziramo izključno v dialogih, ki jih obsega dano besedilo in ki kažejo na konflikte ter čustvene vezi med literarnimi osebami. Za avtomatizirano analizo čustev se vse pogosteje uporabljajo tehnike strojnega učenja in umetne inteligence, ki omogočajo razvoj modelov za prepoznavanje in interpretacijo čustev v dialogih.

Diahrone raziskave razvoja dialoga kot literarnega dispozitiva skozi različna literarnozgodovinska obdobja kako se je uporaba dialoga spreminjala in razvijala skozi čas v literaturi.⁷ Kako računalniški programi zaznajo, kdaj gre za dialog in kdaj za druge dele literarnega besedila? Muzny in drugi⁸ so na primer razvili metodo merjenja tako imenovane dialošnosti tekstov, ki jo zaznavajo v tistih

7. Muzny idr. 2017

8. Muzny idr. 2017

delih besedil, katerih strukture so podobne tistim, ki jih najdemo v govorjenem jeziku. Glede na abstraktne lastnosti posameznih besedil, na primer rabo zaimkov, strukturiranje povedi ali izražanje občutij, skupina raziskovalcev ugotavlja, do kolikšne mere je neko literarno delo dialoško.

5 Čustvene ocene knjižnih bralcev na spletu

Analizo čustev uporabljajo tudi različne raziskave drugih tipov besedil, od spletnih ocen potrošnikov do objav na družbenih omrežjih, kar ima potencial zaznavanja sprememb razpoloženja velike populacije ljudi. Na področju literature lahko raziskujemo tudi vtise o knjigah, ki jih bralci objavljajo na spletu.



Katera čustva bralci izražajo pri svojih ocenah literarnih del na spletnih platformah, kakršni sta Amazon in Goodreads, je znak za to, katere knjige so se jih najbolj dotaknile. Večina bralskih ocen na teh platformah, ki bralcem ponujajo knjige, kakršni sta Goodreads in Amazon, je pozitivnih⁹

6 Nekaj kritičnih pogledov na analizo čustev v književnosti z digitalnimi orodji

Kako z natančnimi računalniškimi algoritmi ujeti vedno izmuzljiva čustva v književnih besedilih? Analiza čustev v književnosti je za računalniške programe zahtevna naloga, saj so čustva v tem tipu besedil pogosto izražena zelo subtilno.¹⁰ Zaradi tega so nekateri raziskovalci prepričani, da umetna inteligenca ne more doseči kompleksnosti človeških bralnih izkušenj.¹¹ Računalniška analiza čustev je zahtevna tudi zato, ker predvideva identifikacijo posameznih čustev, meje med njimi pa so v realnosti pogosto zabrisane.¹² Literatura je namreč sredstvo za izpovedovanje najkompleksnejših človeških občutij in vseh različnih čustvenih nians. Še ena težava analize čustev v književnosti z digitalnimi orodji pa je ironija, ki jo z orodji za analizo čustev težko zajamemo, zato je besedilo, ki temelji na ironiji, neprimeren kandidat za računalniške analize čustev.¹³

10. Nalisnick in Baird 2013: 479

11. Ramsay 2011

12. Kim in Klingler 2021

13. Elkins in Chun 2019

8.

RAČUNALNIŠKE ANALIZE LITERATURE IN SPOL

V literarna dela so vpisane določene kulturne norme in sociale okoliščine časa, v katerem so nastala, tudi norme v povezavi s spolom. Te vidike književnosti je mogoče na ožjem izboru literarnih besedil opazovati s tesnim branjem, z digitalnimi metodami pa lahko raziskave zajamejo obširnejše korpuse besedil.

Feministična digitalna humanistika z računalniškimi metodami pristopa k vprašanju spola in predstavlja kritičen glas, ki opozarja na primere, ko študije digitalne humanistike nadaljujejo patriarhalne vzorce. Poleg tega je področje digitalne humanistike, ki se povezuje s spolom, namenjeno raziskovanju sloga s tega vidika, ukvarja pa se tudi z analizami prisotnosti virov avtoric v digitalnih arhivih.

“Ženskost oziroma ženstvenost, ki jo dojemajo v besedilu, nista samodejni posledici avtorskega spola, ampak /.../ gre za druge procese, ki jih moramo tisti, ki se s tem ukvarjamo, natančno presoјati. Nezavedni stereotipi, povezani s pomenom spolnih razlik v slog pisanja image večji vpliv, kot bi si nekateri radi priznali.”¹

1 “Ženski” teksti?

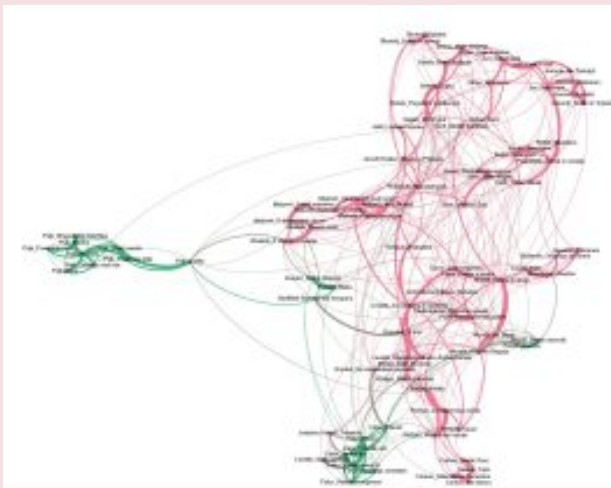
Stilometrične analize zaznavajo različne signale, ki predstavljajo kategorije, po katerih se literarna dela najbolj razlikujejo med seboj. Najmočnejši signal je avtorski, kar pomeni, da se najbolj razlikujejo dela različnih avtorjev. V stilometričnih analizah pa se pojavljajo tudi vprašanja o spolu, saj je posebej v delih starejše književnosti spol avtoric oziroma avtorjev povezan z različnimi žanri pisanja. Kot opozarja Matthew Jockers, je stilometriji in oddaljenem branju spol najmanj vpliven dejavnik na literarni slog v 19. stoletju v primerjavi z drugimi kategorijami slogovnih razlik – npr. avtor, zvrst, literarno obdobje, besedilne variacije (ločila, raba zaimkov itd.).² V stilometričnih analizah del sodobnih pisk in piscev je spolni signal še manj izrazit.

1. Koolen 2018: 251

2. Jockers 2013: 96-8

Stilometrična analiza 90 slovenskih romanov: je spol očiten signal?

Računalniška stilometrična analiza korpusa 90 slovenskih romanov iz 19. in začetka 20. stoletja iz literarnozgodovinskih obdobj slovenskega realizma in moderne, ki smo jo izvedli s paketom R studio, je pokazala, da se romani stilno ne delijo na tiste, ki so jih napisali avtorji (na spodnji sliki označeni z rdečo bravo), in tiste, ki so jih napisale avtorice (na spodnji sliki označeni z zeleno bravo). Najbolj očiten signal je avtorski, saj so se besedila istega avtorja ali avtorice glede na stilno podobnost razvrstila skupaj, vendar se ob tem dela pise in pisk niso razporuedila ločeno. Iz spodnje vizualizacije je očitno, da s svojim avtorskim slogov izstopa avtorica Pavlina Pajk, njeno pisanje pa se po podobnosti povezuje tako z romani moških kot z romani žensk. Avtorici Marica Nadlišek Bartol in Zofka Kveder sta sicer razporejeni skupaj, prav tako Josipina Urbančič Turnograjska in Lea Fatur, ki pa se jima v skupini glede na stilno podobnost pridružuje tudi Vladimir Levstik z romanom *Gadje gnezdo*, njihova dela pa so stilno bližje različnim romanom Ivana Cankarja kakor delom drugih avtoric.



vir: lasten

2 Analize rabe besedišča in kategorija spola

Presek med spolom in avtorskimi besedili zanima tudi raziskovalke in raziskovalce na področju računalniškega jezikoslovja, ki temelji na avtomatiziranem štetju pogostosti posameznih izrazov ter multivariatni frekvenčni analizi besed, ki preučuje več spremenljivk hkrati. S temi raziskavami so na podlagi posameznih korpusov besedil ugotovili, da se pojavljajo razlike med avtoricami in avtorji glede na spol. Pokazalo se je, da so najbolj tipični za posameznike oziroma skupine posameznikov ravno izbori besed, na katere sicer hitro pozabimo, tj. zaimki in predlogi.

Vaja: besedišče in spol

Rešite test, ki ga je pripravil James W. Pennebaker³ in se nanaša na nate dojemanje razlik v pisanju pri moških oziroma ženskah.

Kdo v vseh najpogostejših vsakodnevnih tipih pisanja in govorjenja bolj uporablja naslednje dele govora, moški ali ženske? Pri vsakem od vprašanj obkrožite pravilen odgovor

1. **Prva oseba ednine (npr. jaz, jaz, moj):**
 - a. ženske uporabljajo več
 - b. moški uporabljajo več
 - c. ni razlike med ženskami in moškimi
2. **Prva oseba množine (npr. mi, nas, naš):**
 - a. ženske uporabljajo več
 - b. moški uporabljajo več
 - c. ni razlike med ženskami in moškimi
3. **Členi (v angleščini: a, an, the):**
 - a. ženske uporabljajo več
 - b. moški uporabljajo več
 - c. ni razlike med ženskami in moškimi
4. **Besede, povezane s pozitivnimi čustvi (npr. ljubezen, zabava, dobro):**
 - a. ženske uporabljajo več
 - b. moški uporabljajo več
 - c. ni razlike med ženskami in moškimi
5. **Besede, ki se nanašajo na kognicijo (npr. misliti, sklepati, verjeti):**

3. Pennebaker 2011: 74–75

- a. ženske uporabljajo več
- b. moški uporabljajo več
- c. ni razlike med ženskami in moškimi

6. **Družabne besede (npr. oni, prijatelj, starš):**

- a. ženske uporabljajo več
- b. moški uporabljajo več
- c. ni razlike med ženskami in moškimi

Kot pravi avtor testa, so ga mnogi ugledni raziskovalci rešili narobe. Rešitve se glasijo:

1. a; 2. c; 3. b; 4. c; 5. a; 6. a⁴

Ste izbrali pravilne rešitve ali vas je kaj presenetilo?
Zakaj?

Ko literati ali literatke ustvarjajo svoja dela, v njih pogosto združujejo glasove različnih književnih likov s specifičnimi osebnostmi. Če v književnih besedilih nastopajo književne osebe različnih spolov, bi pričakovali, da se to odraža v stilnih lastnostih dela. Vendar raziskave kažejo, da kljub temu v malih podrobnostih pisanja, ki so predstavljene zgoraj, nad slogom posameznih likov ali pripovedovalcev prevlada avtorski slog, celo v dramatiki, ki jo večinoma sestavljajo replike med liki. Frekvenčne analize sloga, ki raziskujejo pogostost pojavljanja določenega tipa besed, pa večinoma pokažejo spol avtorja ali avtorice, ne pa spola literarnih oseb, ki v delih nastopajo.

Primer: Romeo in Julija izpod peresa Williama Shakespearea

Glede na jezikovne lastnosti besedil je Pennebaker⁵ oblikoval algoritem, ki posamezno besedilo uvršča na lestvico, sestavljeno iz devetih točk, ki na eni strani označuje maskulinito (s skrajno vrednostjo ena), na drugi strani pa femininito (s skrajno vrednostjo devet). Ko je preveril govor znanih zaljubljenca Romea in Julije v istoimenski tragediji Williama Shakespearea, je ugotovil, da sta oba lika ocenjena z oceno tri na lestvici kar pomeni, da sta obe dramski osebi glede na rabo besedišča maskulini.



Na sliki desno: Ford Madox Brown: Romeo in Julija: vir: javna domena, Wikipedia

3 Izrazi za ženske in moške like v književnosti 19. stoletja

- (2011). The secret life of pronouns: what our words say about us. 1st U.S. ed. New York, Bloomsbury Press. str. 93

Študije uporabe besedišča v literarnih delih se osredotočajo tudi na izraze, ki so v besedilih uporabljeni v povezavi z moškimi liki na eni strani in z ženskimi liki na drugi, ob tem pa preverjajo, ali gre za stereotipne predstave o spolnih družbenih vzorcih. Kateri glagoli se torej uporabljajo za opis ženskih literarnih oseb?

V študiji leposlovja iz 19. stoletja sta Jockers in Kirilloff⁶ proučevala upodobitev in dejanja moških in ženskih likov. Njuna metoda je bila, da sta v 3329 različnih literarnih delih z avtomatskim računalniškim zaznavanjem preučevala te like prek zaimkov in glagolskih oblik, ki so bili uporabljeni v zvezi s temi liki, na primer: ona je hodila ali on je oblečen. Raziskava je pokazala, da so bili v leposlovnem korpusu moški in ženske upovedani z različnimi besednjaki v skladu z normami 19. stoletja. Glagoli, ki označujejo čustva in občutke (kot sta jokati, ljubiti), so bili v izrazito večji meri povezani z ženskimi liki, medtem ko so bili glagoli, ki označujejo dejanje in gibanje (kot so napredovati, pristopiti, jezdit), močneje povezani z moškimi liki. "Zdi se, da ta rezultat podpira delo literarnih in kulturnih kritikov, ki so opazili težnjo 19. stoletja po vrednotenju pasivnih žensk in aktivnih moških," sta poudarila avtorja.⁷ Navajamo nekaj značilnih besed, ki so jih zaznali med najpogostejšimi 50 besedami, povezanimi z moškimi liki: vzel, hodil, jezdil, prišel, odšel, naredil, poklical, ponudil, vstopil, se vrnil, plačal, vedel, dal, uspel, poslal, mislil, želel, sledil, govoril. Med najpogostejšimi 50 besedami, povezanimi z ženskimi liki, pa najdemo na primer naslednje značilne izraze:

6. Jockers in Kirilloff 2016

7. Jockers in Kirilloff 2016: 21

sedela, jokala, čutila, vedela, izbruhnila, odgovorila, vstala, ljubila, poslušala, se smejala, zašepetala, sprejela, zavrnila, slišala, se bala, ozdravela, gledala, brala, vrgla, ostala, poljubila, sprejela, razumela, gledala, trpela.⁸

“Tradicionalne družbene vrednote se ne odražajo le na ravni zgodbe, temveč na ravni stavkov in pravzaprav tudi na ravni posameznih besed.”⁹

Kaj pa fizični opisi ženskih in moških likov v književnih besedilih? Raziskave vzorcev opisovanja ženskih likov v književnih delih, ki povezujejo lastnosti dobrote in lepote, kažejo, da so ti zelo pogosti v delih po vsem svetu.¹⁰ Tudi študija lastnosti likov, povezanih s spolom, v sodobnih slovenskih romanih,¹¹ je pokazala, da se ta povezava pojavlja, do podobnih sklepov pa je prišla tudi analiza izbranih del slovenskih avtoric in avtorjev z obdobja modernizma ter realizma.¹² Raziskava opisov telesa ženskih in moških likov v angleški literaturi je pokazala, da se pri karakterizaciji

8. Jockers in Kirilloff 2016: 21

9. Kirilloff idr. 2018: 842

10. Gottschall 2005

11. Žunkovič 2017

12. Pribožič in Jančič 2022

žensk telesni opisi pojavljajo pogosteje tako v starejši kakor tudi v sodobni književnosti.¹³ Pokazala je, da so liki v starejši književnosti izrazito binarno razdeljeni na moške in ženske književne osebe, kar pa v sodobnejših delih ni več tako izrazito.¹⁴

3 Literarna kritika in spol

Analiza besedišča recenzij knjig ameriškega časopisa New York Times, ki ima dolgo tradicijo recenziranja književnih besedil, je pokazala razlike med ocenami del, ki so jih napisale osebe, ki so se opredelile kot moški oziroma ženske. Študija kritik, ki so izšle med letoma 1905 in 1925, je pokazala naslednje: o delu avtoric se je razpravljalo znotraj redukativnega, domačijskega, zasebno/javnega besednjaka.¹⁵

Kaj pa literarna kritika v današnjem času? Analiza 1000 recenzij iz sodobne revije New York Times Book Review je pokazala, da kljub družbenim premikom nekatere razlike ostajajo ter da ločevanja vztrajajo. V recenzijah se mestoma še vedno pojavlja stereotipno razlikovanje med moškimi in ženskami in izpostavljanje te binarnosti kot nasprotje. S tem literarne kritike reproducirajo "razcep javno/zasebno, ki smo ga podedovali iz 19. stoletja."¹⁶

4 Nekaj kritičnih pogledov na računalniške analize literature in spol

13. Cheng 2020

14. Cheng 2020

15. Lavin 2020

16. Piper in So 2016

Digitalne raziskave književnosti predpostavljajo, da je gradivo v danem korpusu razporejeno glede na določene kategorije, med njimi so pogosto tudi spolne. To pomeni, da dobijo določena dela oznake avtorstva glede na spol, enako velja za književne osebe. Vendar Johanna Drucker opozarja, da so spolne identitete pogosto tudi zamolčane in vnaprej določene kategorije v digitalni humanistiki lahko vplivajo na to, da si posamezne slogovne lastnosti književnih besedil razlagamo premalo kompleksno.¹⁷ To Laure Mandell¹⁸ imenuje stereotipiziranje in opozarja, da je bila v nekaterih kvantitativnih analizah kulture kategorija spola biologizirana in predstavljena kot nespremenljiva binarna opozicija med moškimi in ženskami.

“Če so bili literarni podatki zbrani, razvrščeni in celo izdelani v skladu s kategorijama moški/ženska, bomo našli statistično pomembne razlike, ne glede na to, ali so te razlike ključne ali ne.”¹⁹

Pomembno je torej, da tudi digitalne raziskave, ki na kakršen koli način vključujejo kategorijo spola, to storijo na utemeljen način. Občutljivost za vprašanja spola

17. Drucker 2017: 632

18. 2016

19. Mandell 2016

pokaže razlaga kategorije spola kot konvencionalne vloge, ki naj bi jo posamezniki (literarne avtorice ali avtorji) prevzeli, da bi se umestili v določen družbeni kontekst.²⁰ Raziskava s področja digitalne humanistike je za razporejanje književnih besedil v korpusu uporabila shemo, ki je avtorski spol razdelila na tri različne možnosti: ženski, moški ali neznan. To je bilo utemeljeno z dejstvom, da je ta shema razdelitve spolnih kategorij organizirala večino javne reprezentacije spola v obdobju, ki je predmet raziskovanja, tj. v viktorijanski dobi angleškega leposlovja. Ob tem je v študiji izpostavljeno naslednje: “/s/pol je zagotovo lahko bolj zapleten, kot kažejo te kategorije /.../ toda ta raziskava se sprašuje o zgodovini konvencionalnih vlog, ne o resnici osebne identitete ali temeljnih procesih, ki vplivajo na spolno usmerjenost.”²¹

Avtorice so bile dolga stoletja na področju književnosti zapostavljene, saj je bilo pisanje v posameznih zgodovinskih obdobjih domena moških in je bila to za žensko nenavadna odločitev. Vendar se tudi v sodobnem času lahko zgodi, da se vzorci v preteklosti ponovijo. Primer je pojavljanje avtoric na spletni strani Wikipedija, ki lahko odraža podrejeno vlogo, ki so jo imele v literarnem sistemu predvsem v preteklosti. V Združenih državah Amerike je tako odmeval primer, ko je Wikipedia premaknila znana imena pisateljic iz kategorije “Ameriški pisatelji” v kategorijo “Ameriške pisateljice romanov”, ob tem pa se je izkazalo, da se kar 90 odstotkov urednikov tega spletišča identificira kot moški.

20. Underwood idr. 2018

21. Underwood idr. 2018

9.

SPLETNE PODATKOVNE ZBIRKE IN LITERARNO(VEDNO) GRADIVO NA SPLETU

Spletne podatkovne zbirke in digitalne knjižnice zajemajo bodisi primarno digitalne vire bodisi digitalizirane vire, ki so sprva nastali v tiskani ali celo rokopisni obliki. Gre za prepise besedil (na primer romanov, pisem, zgodovinskih dokumentov) in/ali digitalizacijo v slikovni obliki (na primer rokopisov ali zemljevidov).

Tako različne spletne zbirke podatkov kakor tudi digitalne knjižnice igrajo ključno vlogo pri razširjanju znanja in podpirajo humanistične raziskave, tudi digitalne. Med drugim gre za digitalne knjižnice univerz, arhivske ustanove in komercialne platforme za e-knjige.

Spletne zbirke in digitalne knjižnice so v sodobnem času civilizacijsko pomembne, saj družbi omogočajo ohranjanje kulturne dediščine, ki z digitalizacijo dobiva novo življenje na spletu za prihodnje generacije, kar je še posebej ključno pri arhiviranju dragocenih starejših dokumentov.

1 Zbirke podatkov

Spletne podatkovne zbirke so zbirke digitaliziranih podatkov, ki so zbrani in na posebne načine organizirani ter dostopni prek spleta za namene raziskovanja v humanistiki. Spletne podatkovne zbirke na svetovnem omrežju na področju digitalne humanistike obsegajo različne tipe virov, žanrov in tematik, od ožjih do najširših. Gre na primer za enciklopedične ali biografske zbirke, zbirke znanstvenih člankov, literarnih del in raznih digitaliziranih zgodovinskih virov. Tovrstne platforme raziskovalcem in raziskovalkam, pa tudi širši javnosti omogočajo, da dostopajo do virov znanja in pridobivajo podatke za svoje nadaljnje raziskave. Podatki v tovrstnih zbirkah so strukturirani na način, da obiskovalce podpirajo pri raziskovanju vsebin, lahko jim omogočajo tudi, da podatke ali metapodatke prenesejo na svoje naprave in jih obdelujejo, spet druge platforme pa poleg samih podatkovnih zbirk obiskovalcem ponujajo tudi digitalna orodja, s katerimi lahko digitalne vsebine analizirajo, vizualizirajo in interpretirajo. V nadaljevanju predstavljamo ključne slovenske zbirke podatkov na področju literarne in kulturne zgodovine, pa tudi nekaj zanimivih tujih zbirk podatkov.

1.1 Elektronska zbirka Pisma

Pomemben dokument preteklosti so pisma, ki so si jih ljudje nekdanj medsebojno pošiljali in zgodovinsko dogajanje prikazujejo z vidika posameznikov ter posameznic in njihovih intimnih življenjskih zgodb. Množica pisem



slovenskih avtorjev in avtoric – tako znanih osebnosti kakor tudi drugih ljudi, od kmetov do primorskih aleksandrink – so zbrana v [elektronski zbirki Pisma](#), ki javnosti predstavlja ta pomemben del kulturne dediščine.

Znamka; vir: Elektronska zbirka Pisma

Spletna platforma predstavlja tudi biografije piscev in pisk, ki jih vključuje, ter njihove fotografije. Med pismi lahko iščemo glede na naslednje kategorije: pošiljatelj oziroma pošiljateljica, prejemnik oziroma prejemnica in obdobje korespondence. Poleg tega lahko med pismi iščemo na podlagi ključnih besed, ki jih napišemo v brskalnik, pisma pa so poleg tega razdeljena glede na najrazličnejše tematike, ki nam omogočajo iskanje, na primer cenzura, domoljubje, erotika, gledališče, hrana, ljubezen, oblačila, politika, vojaščina in zdravje.



Slika: Naslovna stran Elektronske zbirke Pisma z možnostjo iskanja med gradivom; vir: Mihurko: Elektronska zbirka Pisma

Izziv: Elektronska zbirka Pisma

V spletni zbirki Pisma z iskalnikom tematik, ki ga spletna stran ponuja, poiščite zadetke s ključno besedo “žensko prijateljstvo”. Preberite pisma, ki ustrezajo tej ključni besedi, in odgovorite na naslednja vprašanja:

- na kakšne načine so si ženske medsebojno izkazovale prijateljstvo? Poiščite nekaj zanimivih izrazov.
- Med zadetki najdete tudi pisma znanih pisateljic in urednic iz časa, ko se so prvič uveljavljale v slovenskem literarnem prostoru. Iz pisem skušajte ugotoviti, kako so bila njihova

prijateljstva povezana z njihovim poklicnim udejstvovanjem.

- Na podlagi lastnih izkušenj primerjajte nekdanje izkazovanje prijateljstva med ženskami v korespondencah s sodobnim izkazovanjem prijateljstva prek digitalnih medijev.

1.2 Gledališka cenzura in uprizoritve Dramatičnega društva v Ljubljani (1873–1914)

V spletnem repozitoriju [Gledališka cenzura in uprizoritve Dramatičnega društva v Ljubljani \(1873–1914\)](#) najdemo podatke in gradivo iz različnih virov, ki pričajo o cenzuri na gledaliških odrih v omenjenem obdobju. Repozitorij se nanaša na slovenske gledališke uprizoritve Dramatičnega društva v Deželnem gledališču v Ljubljani. Repozitorij daje s kronološkim prikazom gledaliških del, ki so jih v Dramatičnem društvu v Ljubljani poslali v cenzorsko presojo, pomemben vpogled v cenzorske postopke iz preteklosti. Skenirane dokumente si obiskovalci spletne strani lahko tudi поблиže pogledajo.



Slika: Cenzuriran rokopis drame Nora Henrika Ibsena;
vir: Mihurko: Gledališka cenzura in uprizoritve
Dramatičnega društva v Ljubljani (1873–1914)

**Izziv: Gledališka cenzura in uprizoritve Dramatičnega društva
v Ljubljani (1873–1914)**

V spletni zbirki Gledališka cenzura in uprizoritve
Dramatičnega društva v Ljubljani (1873–1914) pogledite

dokumente, povezane z uprizoritvijo [drame Henrika Ibsena Nora](#), ki jo je ljubljansko Dramatično društvo uprizorilo leta 1892. Odrska uprizoritev dela je bila s strani cenzorstva dovoljena pogojno. Iz digitaliziranega zgodovinskega dokumenta skušajte razbrati pogoj, ki so ga postavili ustvarjalcem.

1.3 Zgodovina ženskega tiska

Projekt [Women's Print History](#) (Zgodovina ženskega tiska) je obsežna angleška zbirka podatkov o prispevkih žensk k tisku v 18. stoletju. Vsebuje podrobne informacije o knjigah, ki so jih ustvarile avtorice, tiskarke, urednice, prevajalke in založnice, pa tudi o delih, ki so povezana z ilustratorkami in skladateljicami ter s prodajalkami in graverkami. Po repozitoriju je mogoče iskati na podlagi naslovov, avtoric, imen tiskarskih družb, formatih, žanrih in virih, zadetke pa je mogoče posebej razvrstiti glede na različne funkcije žensk, od tiskark do literatk, in glede na različne funkcije tiskarskih podjetij, od knjigarjev do založnikov.

1.4 Zbirka bralnih izkušenj v Združenem kraljestvu

Na proces literarne recepcije in zbiranje izkušenj bralcev je osredotočena zbirka z imenom [UK Reading Experience Database](#) (slo.: Zbirka bralnih izkušenj v Združenem kraljestvu), ki vsebuje 30.000 bralnih vtisov ljudi v sedanjem Združenem kraljestvu med letoma 1450 in 1945. Bralci, ki jih najdemo v tej zbirki z odprtim

dostopom, so tako znane osebe kakor tudi anonimneži, na primer predniki prostovoljcev, ki so v svojih družinskih arhivih našli zapise o branju.

2 Digitalne knjižnice

Digitalne knjižnice so spletni viri prosto dostopnih besedil fizično obstoječih rokopisov, knjig, revij, časopisov in drugih gradiv v elektronski obliki. Uporabnikom



navadno omogočajo prost dostop do vsebin prek njihovih spletnih brskalnikov ali za to namenjenih aplikacij, gradivo pa lahko prenesejo na svoje digitalne naprave.

Digitalne knjižnice svoje gradivo ponujajo bodisi brezplačno bodisi z naročnino; med gradivom je mogoče iskati po avtorjih in naslovih ali z uporabo ključnih besed ter drugih kriterijev.

2.1 Digitalna knjižnica Slovenije

[Digitalna knjižnica Slovenije](http://dLib.si) (dLib.si) vsebuje digitalizirano gradivo s področij slovenske znanosti, umetnosti in kulture, na portalu pa lahko najdemo knjige, periodiko, rokopise, glasbo, slike in zemljevide. Gradivo za portal temelji na zbirkah Narodne in univerzitetne knjižnice, pa tudi njenih partnerjev, od knjižnic do dediščinskih ustanov. Nekatere od vsebin so tudi izvorno digitalne.



One or more interactive elements has been excluded from this version of the text. You can view them online here: <https://books.ung.si/digitalnahumanistika/?p=277#oembed-1>

Video: Predstavitev Digitalne knjižnice Slovenije; vir: Digitalna knjižnica Slovenije – dLib v novi podobi, Youtube

2.2 Project Gutenberg

Project Gutenberg je bil prvi portal, ki je nudil brezplačne e-knjige. Ustanovitelj projekta Gutenberg je Michael Hart, ki je spletno platformo zagnal leta 1971. Hart je znan tudi kot izumitelj e-knjig, ki je svoje poslanstvo opisal z naslednjimi besedami: “Literatura mora biti brezplačna tako kot zrak, ki ga dihamo.”¹

3 Izbrana svetovna spletišča s knjigami

3.1 Internet Archive

[Internet Archive](#) je digitalna knjižnica, kjer lahko najdemo brezplačne knjige, filme, programsko opremo, glasbo, spletne strani in druge vsebine. [Google Books](#) je

1. 50 years of eBooks: 1971-2021. Project Gutenberg

spletišče, ki v celoti ponuja knjige, ki so jim avtorske pravice potekle, za dela, katerih avtorske pravice so še v veljavi, pa ponuja kratek opis in prikaz vzorčnih strani. [Google Books Ngram Viewer](#) je spletni iskalnik, kjer lahko preverimo pogostost pojavljanja besed, besednih zvez ipd. v tiskanih virih, ki so bili v različnih jezikih objavljeni med letoma 1500 in 2019.

3.2 Orlando

Zbirka z naslovom [Orlando: Women's Writing in the British Isles from the Beginnings to the Present](#) skuša uporabnikom literarno zgodovino predstaviti drugače. Ni knjiga, čeprav je po dolžini enakovredna več kot 80 znanstvenim knjigam, in ni digitalna izdaja obstoječega besedila, je bogato iskana zbirka besedil digitalno rojenega izvirnega pisanja. Poln je interpretativnih informacij o ženskah, literaturi in kulturi, z več kot 8 milijoni besed besedila v dokumentih o življenju in pisanju več kot 1400 avtorjev, skupaj z veliko kontekstualnega zgodovinskega gradiva o pomembnih temah, kot je izobraževanje, politika, znanost, pravo in ekonomija.

3.3 Women writers project

Projekt pisateljic z imenom Women writers project je dolgoročni raziskovalni projekt, posvečen literarnih

besedilom angleških ženskih avtoric iz predviktorijanskega časa in iz obdobja zgodnjega modernizma. Cilj projekta je arhivirati to literarno dediščino, ki je prosto dostopna javnosti, ob tem pa s funkcijami spletne strani podpirati raziskave o avtoricah. Spletna stran vsebuje tudi prikaze raziskav, ki so bile izvedene na tem korpusu, na primer [analizo govora moških in žensk](#) v posameznih dramah in [zemljevid toponimov](#), ki so omenjeni v korpusu.

4 Literarnovedno znanje na spletu: od Wikipedije do “živih” priročnikov

Znanje se hitro spreminja in dopolnjuje, zato so spletne izdaje publikacij s področja literarne vede primerno okolje za dinamične vsebine s tega področja. Na ta način se znanje širi v strokovni javnosti, bralci pa niso le pasivni sprejemniki vsebin, ampak aktivno sodelujejo pri njihovem oblikovanju. V slovenskem prostoru je zelo pomemben vir nastajajo pa tudi spletni ubčeniki, monografije in gradiva na literarnovedne tematike.

Primeri: Nova pisarija in Izhodišča slovenske pripovedne proze

- [Nova pisarija: Strokovno pisanje na spletu](#) Mirana Hladnika je objavljena na spletišču Wikibooks in obsega najrazličnejše tematike s področja strokovnega pisanja s poudarkom na

literarnovednem raziskovanju. Tam najdemo praktične napotke za pisanje z digitalnimi orodji tako z vidika avtorstva kot z vidika bralstva. Dotakne se različnih tematik procesa priprave besedila in objave, od citiranja in recenziranja prek raznovrstnih sodobnih medijev, pa tudi žanrskih in slogovnih razlik. Poglavlji sta namenjeni tudi področjem iskanja virov in digitalne humanistike.²

- Monografija [Izhodišča slovenske pripovedne proze](#) Andrejke Žejn prikaže razvoj starejše slovenske pripovedne proze, in sicer se osredotoča na njen zgodnji razvoj. Delo predstavlja nova odkritja s tega področja in zasleduje literarnozgodovinsko obdobje od časa, ko so v slovenskem prostoru obstajali zgolj prevodi pripovedne proze iz tujine do izida prve slovenske izvirne pripovedi *Sreča v nesreči* leta 1836. Tudi Žejn uporablja poleg ustaljenih literarnovednih metod tudi digitalne metode raziskovanja književnosti. Z uporabo računalniške stilometrije avtorica analizira besedila od prve polovice 17. stoletja do Ciglerjeve *Sreče v nesreči* in pokaže slogovne povezave v korpusu besedil. Med drugim se pokaže slogovna podobnost med deli Christopha Schmidta in Janeza Ciglerja, ki se je po Schmidu zgledoval.³

2. Hladnik 2014/2016

3. Žejn 2021

Tako da potrebe bralcev v najršiirši javnosti kot tudi za potrebe šolstva so pomembne objave o slovenskih literarnih avtoricah in avtorjih, pa tudi literarnih zgodovinarkeh ter zgodovinarjih, literarnizgodovinskih obdobjih in drugih temah, povezanih s književnostjo, na spletni enciklopediji Wikipedija. Tu najdemo relevantne informacije o omenjenih temah, ki se sproti posodabljaajo in je spremenila tudi odnos skupnosti do strokovne verodostojnosti, kot opozarja Miran Hladnik: “Te ne jamčita več doktorat in objavljena monografija, relevantna postajajo tudi argumentirana stališča ljudi iz drugih strok in splošne publike.”⁴

Status člankov na Wikipediji je urejen z licenco *creative commons* (cc), kar dobesedno pomeni ‘licenca ustvarjalnega ljudstva’, za katero se nadejamo, da bo nekoč nadomestila tradicionalno avtorsko zakonodajo, znano pod imenom *copyright* ©. Stara je podobno kot Wikipedija in izhaja iz sorodnih kulturnih izhodišč, kar dokumentira že z izbiro podobnega ključnega izrazja: pri Wikipediji se govori o skupnosti ali občestvu (z rubriko Pod

lipo (angl. *village pump*) namiguje na z gledovanje pri tradiciji neposredne demokracije v okviru lokalnih skupnosti), *creative commons* pa prevajajo z »ustvarjalna gmajna«, kar z asociacijo na slovensko uporniško pesem »Leukhup, leukhup, leukhup, leukhup woga gmaina« iz šestnajstega stoletja opozarja na svojo uporniško dimenzijo, uporniško vsaj v razmerju do omejujočega *copyrighta*.⁵

Wikimedijina spletišča poleg spletne enciklopedije zajemajo tudi Wikiknjige, kjer so objavljeni priročniki in knjige, Wikiverzo, kjer lahko univerze objavljajo gradiva za seminarje in projekte, pa tudi predavanja in podobno, poleg tega pa tu najdemo še Wikivir, ki obsega stara besedila v javni lasti. Gre za obsežno zbirko digitaliziranih besedil, ki je pomembna tako za bralke in bralce kakor tudi za literarnovedno raziskovanje z digitalnimi orodji.

Izziv: Cvetje v jeseni

Na Wikiviru odprite prozno besedilo Ivana Tavčarja

[Cvetje v jeseni](#). V kateri reviji je zgodba najprej izšla? Če besedila ne poznate dobro, ga (ponovno) preberite v celoti. Nato z iskalnikom besed v digitaliziranem besedilu poiščite vse primere, ko se v pripovedi omenja “cvetje”, besedo, ki jo najdemo že v naslovu pripovedi. Preberite te dele besedila in oblikujte kratek zapis, v katerem utemeljite, kako se omembe cvetja v delu povezujejo z njegovim naslovom in kaj cvetje v tej zgodbi simbolizira. Med drugim boste našli naslednjo omembo, ki jo izreče protagonist zgodbe:

“Zasmehujte me, kolikor vas je volja! Da, prijateljice, kar vam ni bilo dano, dano je bilo meni! Pretresla me je ljubezen, enaka viharju, enaka nevihti! Prišla je pozno, prav tako je prišla, kakor pride časih kako cvetje v pozni jeseni. Sadu ni rodila, kakor ga ne rodi jesenski cvet. Če me hočete poslušati, vam razgrnem vse. Pri tem tudi izveste, čemu nisem hotel zakonske družice, katera bi me bila ravnotako okrog prsta obvijala, kakor obvijate ve svoje može. Pri poslušanju se bode morda obilo smejale! Vprašanje je torej: me li hočete poslušati?”⁶

Na spletu lahko najdemo tudi številne slovenske in tuje znanstvene članke z literarnovednega področja, ki jih za prosto branje objavljajo znanstvene revije. Nekatere znanstvene revije v tujini na svojih spletnih straneh objavljajo tudi članke, ki so še v recenziji in niso še sprejeti v objavo, ob tem pa povabijo svetovno strokovno javnost, da jih komentira. Spet drugič raziskovalci povabijo predstavnike javnosti, da preberejo njihove najnovejše

6. Tavčar: Cvetje v jeseni; Wikivir

zamisli in jim posredujejo svoje komentarje. Nove oblike znanstvenih in poljudno znanstvenih ter strokovnih del med drugim izkoriščajo tudi možnost interakcije s svojimi bralci, ki želijo sodelovati v razpravi prek komentarjev. Tovrstne spletne strani bralcem omogočajo, da spreminjajo vsebino, s sledenjem sprememb pa lahko vidijo vse modifikacije besedila.

Primer: Deset let oblikovanja priročnika o naratologiji

[The living handbook of narratology](#) (*Živi priročnik naratologije*) temelji na priročniku, ki je sprva izšel v tiskani obliki, pozneje pa so ga objavili na spletu in uporabnikom omogočili, da posegajo v besedilo. Tako so se oblikovale mnoge zanimive razprave o posameznih področjih naratologije, vsebino priročnika pa je preverila množica bralcev. Ta primer pa kaže na moč skupinskega oblikovanja znanja, saj gre danes za enega najbolj citiranih virov s področja naratologije. Priročnik je na spletu nastajal deset let, v tem času pa so ga različni strokovnjaki redno dopolnjevali z novimi prispevki in koncepti o naratologiji, bralci pa so ga dopolnjevali, dokler ga niso v njegovi zadnji različici leta 2021 začasno zaprli za spremembe. Priročnik bralcem omogoča različne funkcije, med njimi iskanje po celotnem besedilu, izvoz podatkov in integracijo digitalnih orodij za besedilno analizo.

5 Branje z digitalnih medijev

Digitalna sodobnost spreminja tudi načine, kako je mogoče brati literarna dela, saj bralci prehajajo s tradicionalnih tiskanih medijev na digitalne platforme, kot so e-knjige, spletni članki, blogi in drugi digitalni viri. Vse bolj razširjeno branje z zaslonov sproža vprašanja, ali gre za enakovreden proces branju s tiskanega medija. Digitalni format namreč ob dolgotrajnem branju utruja oči, poleg tega pa nudijo digitalni mediji izrazito interaktivnost (npr. povezave do različnih spletnih strani s povezanimi vsebinami v obliki teksta, slik, videa, podcastov ipd.). To lahko bodisi obogati bralno izkušnjo bodisi vodi do manjše zbranosti, kar velja tudi za označevanje, podčrtavanje in ustvarjanje zaznamkov pri digitalnem branju. Nekateri raziskave kažejo, da bralci, ki berejo z zaslona, kompleksne informacije dojemajo manj natančno od tistih, ki berejo s tiskanih medijev.

Vaja: anketa o branju z digitalnih medijev med znanci

Na eni od spletnih platform, na primer slovenski [1Ka](#), oblikujte kratko anketo na temo branja z digitalnih medijev. Delite jo med drugimi študenti, družinskimi člani, prijatelji, znanci in drugimi, ki bodo med anketiranjem anonimizirani. Zberite tudi podatke o njihovi starosti, spolu, izobrazbi, kraju bivanja. Vprašanja naj obsegajo na primer: kako pogosto berejo knjige v digitalni obliki (npr. e-knjige) in na katerih napravah (e-bralnik, tablica, telefon ipd.), kateri tip digitalnih medijev najpogosteje uporabljajo za branje (e-knjige, spletne časopise, bloge ipd.), ali raje

berejo tiskane knjige in časopise ali raje digitalne medije ter zakaj. Preverite, do kakšnih rezultatov ste prišli, in jih predstavite v skupinah.

6 Nekaj kritičnih pogledov na spletne podatkovne zbirke in digitalne knjižnice

Z digitalizacijo obsežnega gradiva z vseh področij človekovega življenja v preteklosti in sedanjosti, pa tudi ožje na polju literarnih virov se širijo možnosti dostopa do znanja, pa tudi novih računalniških raziskav. Vendar se ob tem porajajo vprašanja, na kakšne načine izbirati gradivo, ki se digitalizira, in kako te izbire argumentirati. Kritiki opozarjajo, da digitalizacija pogosto zaobide vire avtorjev in avtoric, ki so zunaj klasičnega kanona in so že tako izključeni/-e. Da se v okviru digitalizacije besedil književnih avtorjev oziroma avtoric pojavljajo spolne neenakosti, je pokazala na primer raziskava, ki je preverjala korpus literarnih del v angleščini od 30. let 19. stoletja do sodobnosti. Izkazalo se je, da je bilo bistveno manj digitaliziranih romanov žensk v primerjavi z deležem romanov avtorjev.⁷

7. Riddell in Bassett 2020

“Gotovo tudi knjižnice, tako kot druge institucije, reflektirajo marginalizacijo in izključevanje v družbi, ki ji služijo.”⁸

Ne preseneča, da je ena ključnih tem, na katero opozarjajo raziskovalke in raziskovalci s področja digitalne humanistike, prost in enostaven dostop do digitaliziranih virov. Mnogi digitalni humanisti zagovarjajo, da je javni dostop do podatkov ključen, in poudarjajo medsebojno sodelovanje raziskovalnih institucij ter prosto uporabo digitalnih orodij za analize gradiva. Posledica omejenih korpusov, ki so digitalizirani, je dejstvo, da že na izbiro tematik, s katerimi se tovrstne makroanalize ukvarjajo, vplivajo možnosti dostopa do virov.⁹ To predstavlja pomembno omejitev raziskav, ki se na primer pogosto osredotočajo na dela, ki so digitalizirana in so jim že potekle avtorske pravice. Ena od težav pri računalniških raziskavah književnosti je nedostopnost tistih besedil, ki imajo veljavne avtorske pravice in jih ni mogoče dobiti v digitalizirani obliki, do katere lahko dostopajo programi za digitalne raziskave literarnih in drugih virov. Možna rešitev, ki se je uveljavila v tujini, je tip spletnega arhiva, kakršen je arhiv z imenom [HathiTrust Research Center](#), ki vsebuje skoraj 17 milijonov del in je za namene digitalnih raziskav omogočil dostop

8. Hope 2001

9. Weidman in Pastor 2021

tudi do besedil, ki so še zaščiteni z avtorskimi pravicami. Poleg tega za raziskave ponuja zbirko metapodatkov o književnih delih, ki jih lahko raziskovalci in raziskovalke uporabljajo pri svojih analizah, pa tudi nekaj orodij za besedilne analize, ki jih lahko uporabniki ter uporabnice prosto preizkusijo.

Čeprav sta skoraj vse založništvo in razširjanje književnosti v 21. stoletju odvisna od računalniške tehnologije, po mnenju kritikov in kritičark digitalni pristopi zaradi zaprtega dostopa skrbništvo nad besedilnimi arhivi omogočajo le izbranim in v svojem prizadevanju za oblikovanje programskih orodij delujejo neoliberalno.¹⁰ Po mnenju drugih raziskovalcev in raziskovalk se digitalna humanistika obrača stran od teorij, kot so jih tradicionalno poznali v humanistiki, saj prinaša poudarek na empiričnih pristopih, ki temeljijo na obdelavi podatkov,¹¹ njen cilj pa naj bi bilo ustvarjanje novih orodij, repozitorijev ipd., ne pa oblikovanje teoretičnih spoznanj. Kritični pogledi na digitalne zbirke podatkov in elektronske knjižnice opozarjajo, da se zaradi tega razloga izpodriva politično progresivna humanistična misel v korist izdelave digitalnih orodij in arhivov.¹² V tem Daniel Allington, Sarah Brouillette in David Golumbia prepoznavajo neoliberalne težnje po prestrukturiranju univerze po vzoru korporativnega sveta, ki bi služila gospodarstvu. Tudi področje humanistike naj bi se po mnenju teh raziskovalcev in raziskovalke torej z digitalnim preobratom k

10. Hui Kyong Chun idr. 2016

11. prim. Kleymann, 2020

12. Allington idr. 2016

praktičnim orodjem obrnilo v neoliberalno
smer.¹³

10.

DIGITALNE KRITIČNE IZDAJE KNJIŽEVNE KLASIKE

Tekstna kritika se natančno ukvarja literarnimi besedili in drugimi dokumenti iz zgodovine, pri čemer išče okoliščine njihovega nastanka ter recepcije, pa tudi njihove različice v času. Zaradi oddaljenosti njihovega nastanka se ukvarja predvsem s starejšimi deli. Z digitalnimi orodji so se možnosti za oblikovanje tekstnokritičnih spletnih izdaj še razširile in nastajajo sodobne digitalne kritične izdaje, ki vključujejo različne interaktivne funkcionalnosti. Digitalne izdaje lahko vsebujejo grafike in interaktivne vizualizacije, ki lahko prikažejo različne vidike besedila, na primer povezave med liki, dogodki, temami in drugimi elementi.

Tekstna kritika predpostavlja urejanje, analiziranje in

komentiranje literarnih in drugih besedil iz preteklosti. Tradicionalno so tako v slovenskem prostoru na primer izhajale tekstnokritične izdaje zbranih del klasičnih slovenskih pisateljev in pisateljic, ki so bile za bralko oziroma bralca informativne zaradi bogatih pojasnjevalnih opomb. Tovrstne izdaje poleg besedila bralcu ponujajo tudi informacije o zgodovinskem kontekstu dela in njegovem izvoru ter raznih posegih v tekst (uredniških, cenzorskih, avtorskih idr.), razlage posameznih manj jasnih delov, okrajšav, starinskih poimenovanj, vsebuje pa tudi prikaz različic besedila in povezave do dodatnih virov.

1 Različne metode tekstne kritike, tudi digitalne

Sodobna digitalna tekstna kritika, pa tudi tekstna kritika pred digitalnim obratom je lahko, kot ugotavljata Odd Einar Haugen in Daniel Apollon,¹ oblikovana s treh različnih zornih kotov, ki pa se medsebojno ne izključujejo, ampak najpogosteje dopolnjujejo:

- pogled nazaj: gre za iskanje izvirov danega besedila, iskanje različic in sprememb originalnega dela ter analizo različnih variant. Poleg tega gre lahko za raziskovanje avtorskih zapiskov in njihov razvoj do končnega rezultata;
- pogled od zunaj: tj. razumevanje besedila z vidika njegovega družbenega in zgodovinskega konteksta, kar pomeni, da se analiza ne osredotoča toliko na izvor besedila in njegovo materialnost, ampak predvsem na socialno-

kulturno realnost, ki je zaznamovala nastanek besedila;

- pogled naznoter: gre za izhodišče, ki besedilo razume kot samostojni dokument in ga ne povezuje z zunajbesedilnimi dejstvi, marveč išče različne pomene ter interpretacije.²

Primer: Foglarjev rokopis

Foglarjev rokopis oziroma Foglarjeva pesmarica (cca. 1757-1762) je najstarejša znana rokopisna pesmarica, ki so jo napisali na Spodnjem Štajerskem. Elektronska kritična izdaja, ki jo je uredila Nina Ditmajer, rokopis predstavlja z digitalnim faksimilom in opombami, znanstvenim komentarjem z razlago okoliščin nastanka rokopisa, jezikovnih in literarnih posebnosti ter uredniških načel prepisa ter študijo. Predstavljene so zgodovinske posebnosti besedila



in variante nekaterih verzij pesmi iz osmih rokopisnih ter tiskanih slovenskih pesmaric iz 18. in 19. stoletja.³

Slika desno: Foglarjev rokopis; vir: Javna domena

2 Oblika digitalne izdaje

Če vsebujejo tradicionalne tiskane kritiške izdaje kazalo, pojmovnik in/ali seznam omenjenih oseb ter druge vsebine, ki omogočajo iskanje po knjigi, pa morajo spletne izdaje možnosti iskanja oblikovati drugače. Digitalna orodja pogosto ne prinašajo novih vprašanj, ampak so učinkovitejša in omogočajo hitre ter zanesljive raziskave, ki so bile nekoč preobsežne, da bi si jih kdo zadal in jih tudi izpeljal.⁴ Zelo pomembno je, da snovalci in snovalke digitalne izdaje upoštevajo potrebe uporabnic in uporabnikov, hkrati pa oblikujejo vmesnik, ki je preprost za uporabo. Med besedili v digitalni kritični izdaji lahko iščemo na primer prek ključnih besed, poimenovanj oseb, krajev idr., iskanje pa je lahko tudi prosto: uporabnica ali uporabnik ne izbira ključne besede, ampak v iskalnik vpiše poljubne izraze.

Digitalna izdaja mora biti označena s podatki o strukturi in vsebini dokumenta, pri čemer upoštevamo standard za označevanje XML (ang. *Extensible Markup Language*), ki določa način označevanja dokumentov. Poleg tega

3. Ditmajer: Foglarjeva pesmarica; Ditmajer idr. 2019

4. Haugen in Apollon 2014

upoštevamo Inicijativo za zapis besedil TEI (ang. *Text Encoding Initiative*), ki določa smernice za kodiranje besedila na podlagi številnih elementov, ki jih lahko uporabimo.

Kljub drugemu mediju in novim možnostim spleta pa imajo digitalne kritične izdaje pogosto podobno obliko kot klasične tiskane izdaje, kar “ni nujno znak konservativnosti ali pomanjkanja iznajdljivosti, gre bolj za to, da pustijo, da je besedilo v ospredju in prikazano tako, da ne pritegne pozornosti na digitalni medij.”⁵

Primer: Wittgensteinovi rokopisi



Ko je avstrijsko-britanski filozof Ludwig Wittgenstein leta 1951 umrl, je

zapustil obsežno zbirko še neobjavljenih in v času njegovega življenja neznanih zapisov, ki je obsegala približno 20.000 strani. Gre za t. i. Wittgensteinovo zapuščino, ki vsebuje avtorjeve rokopise in tipkopise: filozofske zvezke, rokopise, tipkopise in zapise po nareku.⁶

Ker je avtor v času svojega življenja objavil le malo zapisov, so imeli njegovi neobjavljeni pripevki nenavadno veliko vlogo pri recepciji njegovega dela. Elektronska izdaja, ki so jo pripravili na Univerzi v Bergnu, ni omejena s pravili tiskane objave in ponuja podrobne informacije o vsaki strani rokopisov in tipkopisov, ki sestavljajo Wittgensteinove dokumente.⁷ [Spletna stran](#) med drugim ponuja barvne fotografije vsake strani izhodiščnega besedila, podrobno izdajo tega besedila, ki vsebuje številne informacije o delu, in izdaja, ki ne prikazuje raznih različic besedila in drugih informacij.⁸

Ludwig Wittgenstein; vir: javna domena

3 Kritični pogledi na digitalne kritične izdaje literarne klasike

Digitalni znanstveni urednik ali urednica je vedno razpet med potrebo po ohranitvi izvirnega videza dela ali skupine del, ki jih obravnava, tudi pri objavi na spletu, in potrebo po čim bolj učinkovitem izkoristku možnosti spletnega okolja. Slednje pomeni rabo različnih orodij in multimedijskih možnosti. Če skuša biti tipičen prikaz kritične izdaje na spletu v čim večji meri podoben tradicionalni tiskani izdaji, po drugi strani premalo izkoristi možnosti, ki jih imajo digitalni mediji.⁹

6. The Wittgenstein Archives

7. Stern 2010

8. Stern 2010

9. Haugen in Apollon 2014

11.

KRITIČNA DIGITALNA HUMANISTIKA

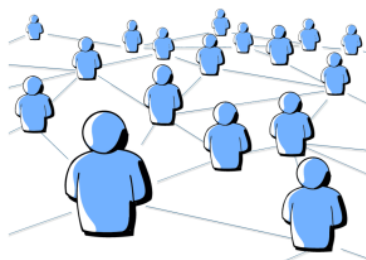
Kritična digitalna humanistika se ukvarja z vprašanji družbene pravičnosti na področju digitalnega sveta. Z vidika kritične teorije naslavlja teme, povezane s hierarhijami in neenakovrednim položajem določenih posameznikov, odpraviti pa skuša tudi primere algoritmov, ki izvajajo krivične prakse z vidika rase, etničnosti, spola in drugih kategorij, ter opozarja na stereotipe.

Kritične težnje znotraj digitalne humanistike izpostavljajo potrebo po natančni refleksiji sodobnih digitalnih orodij, algoritmov, strukture raziskav, načina zbiranja podatkov, postopkov digitalizacije gradiva in drugih vidikov. Izhajajo iz predpostavke, da sodobni mediji nudijo mnoge možnosti, pa tudi omejitve za prikazovanje družbene realnosti, ob tem pa niso le

sredstvo za prikazovanje resničnosti, ampak resničnost tudi na novo ustvarjajo.

1 Sodelovanje s širšo skupnostjo

Kritični pogledi na digitalno humanistiko med drugim opozarjajo na stereotipne predstave, ki se iz družbenih struktur selijo v digitalne prostore in celo na področje raziskav.



Poleg tega opozarjajo na potrebo, da digitalna humanistika izstopi iz akademskega polja in ne sodeluje le z univerzo, ampak tudi s širšo skupnostjo ter drugačnimi profili posameznikov: “/p/otrebujemo nove hibridne praktike: umetnike-teoretike, humaniste programerje, aktiviste-raziskovalce, arhiviste teorij, programerje z občutljivostjo za kritično rasno teorijo.”¹ Pojavljajo se skupine, ki opozarjajo na pristranskost digitalnega okolja in delujejo v preseku med aktivizmom, digitalno humanistiko in umetnostjo, na primer [Queer technologies](#), ki je z ironičnimi digitalnimi produkti in videi opozarjala na zapostavljeno mesto kvir oseb v digitalnem okolju.

Primer: Spremenimo digitalno humanistiko

Gibanje med digitalnimi humanisti in humanistkami z imenom #transformDH (spremenimo digitalno humanistiko), ki se je začelo na družbenem omrežju Twitter z istoimenskim ključnikom, si prizadeva za kritični prostor, ki bi reflektiral raziskovalne prakse, digitalna orodja, zbirke gradiva in druge pojave v polju sodobne digitalne humanistike. Gibanje nagovarja naslednje ključne ideje:²

1. Vprašanja rase, razreda, spola, spolnosti in hendikepiranosti morajo biti v središču digitalne humanistike ter študija digitalnih medijev.

2. Feministični, queerovski in antirasistični aktivistke ter aktivisti, umetnice in umetniki, ustvarjalke ter ustvarjalci medijev izven akademskega kroga opravljajo delo, ki prispeva k digitalnim študijem v vseh oblikah. To delo produktivno destabilizira norme in standarde institucionalno priznanega akademskega dela.

3. Fokus digitalne humanistike bi morali preusmeriti s tehničnih procesov na politične in si vedno prizadevati razumeti družbeni, intelektualni, ekonomski, politični ter osebni vpliv digitalnih praks, ko jih razvijamo. Potrebujemo digitalno humanistiko, ki bo osredotočena na presečišče digitalne produkcije in družbene transformacije skozi raziskovanje, pedagogiko in aktivizem ter ne bo omejena na institucionalne akademske prostore.³

Nekateri digitalni humanisti oziroma humanistke so prepričani/-e, da lahko to novo področje spremeni globoko zakoreninjene tradicije akademskega sveta, ki ga

2. Bailey idr. 2016

3. Bailey idr. 2016

po njihovih ocenah zaznamuje kolonialistični pogled na kulturo in zahodnocentrični privilegiran položaj.⁴

2 Vpliv digitalne tehnologije na življenje ljudi: humanistični pogled

Sodobni avtomatizirani programi so danes prisotni na različnih področjih življenja, informacije pa pridobivajo iz množice virov, ki lahko včasih odražajo neenakovredna razmerja med ljudmi.

Programi na primer iz zbirke življenjepisov avtomatsko izberejo najboljše kandidate za določeno delovno mesto, pri čemer lahko zapostavljajo določene skupine posameznikov. Algoritmi med drugim sprejemajo odločitve na polju zdravstvene varnosti in na drugih pomembnih področjih, ki vplivajo na vsakdan ljudi. Za raziskovalce, ki se ukvarjajo z digitalno humanistiko, je zato pomembno, da razumejo, kako delujejo algoritmi, ki se skrivajo za določenimi programi. To je ključno, da se lahko izognejo morebitnim pristranskim interpretacijam gradiva. Digitalna humanistika opozarja, da odločitve v zvezi s tehnologijo niso nujno nevtralne, saj imajo lahko velike posledice za življenje posameznikov. Sodobni razvoj velikih podatkov in umetne inteligence je prinesel mnogo novih možnosti, vendar je ob teh velikih spremembah bistveno spodbujati kritične ter družbeno ozaveščene pristope k zbiranju in analizi podatkov.⁵ To pa otežuje



4. Lothian in Phillips 2013

5. Prescott 2023

dejstvo, da je večina sodobnega razvoja umetne inteligence v rokah komercialnih podjetij, predvsem korporacij iz Silicijeve doline, lastniki algoritmov pa nočejo izdati svojih poslovnih skrivnosti, zato je težko ugotoviti obseg in naravo pristranskosti njihovih produktov.⁶

“Največja grožnja človeštvu s strani umetne inteligence ne izvira iz avtonomnih robotov morilcev, temveč iz družbenih motenj in krivic, ki jih povzročajo pretirano zanašanje na slabo zasnovane in slabo dokumentirane komercialne črne skrinjice za upravljanje vsega, od zdravstvene oskrbe do javnega reda in kriminala.”⁷

“Privzeto stališče bi morda moralo biti, da se vsi komercialni paketi umetne inteligence, ki niso v celoti dokumentirani, obravnavajo kot pristranski do določenih skupin.”⁸

6. Prescott 2023

7. Prescott 2023

8. Prescott 2023



3 Ena od težav: računalniški algoritmi in kritična rasna teorija

Kritična rasna teorija analizira oblikovanje in uporabo računalniških algoritmov, s tem pa prepoznava potencialne pasti, na primer nezavedne predsodke. Tudi spletni brskalniki, ki jih uporabljajo široke skupine ljudi, lahko odražajo nepravilnost do določenih, največkrat marginaliziranih posameznikov, ki se kaže v družbi, kar še utrjuje različne stereotipe, tudi rasne.⁹ Poleg tega so iskalniki neposredno povezani s komercialnimi interesi korporacij in s težnjami podjetij, ki pri njih oglašujejo svoje izdelke ter storitve.

“Kaj se zgodi, ko o drugačnosti ne razmišljamo

9. prim. Benthall in Haynes 2019

več kot o primanjkljaju, ki ga je treba obvladovati in popraviti, ampak jo razumemo kot sistem, znotraj katerega operiramo, kot del naših hipotez, našega navdiha, naših ciljev?”¹⁰

Sporni zadetki v spletnih brskalnikih

Kako lahko računalniški algoritmi krepijo rasno zatiranje, se je pokazalo v raziskavi Safiye Umoja Noble¹¹. Avtorica je izpostavila, da so napačni rezultati v iskalnikih zaradi tako imenovanih algoritmov zatiranja etično sporni in imajo velike družbene posledice, zato poziva k večji osveščenosti v javnosti. Safiya Umoja Noble je med drugim opozorila, da so bile med Googlovimi fotografijami slike temnopoltih moških označene z izrazom “gorila”.¹²

5 Jezikovna pristranskost

Na svetu je več kot 7000 jezikov, a približno 40 % ljudi se sporazumeva v katerem od osmih najpogostejših jezikov.

10. Barnett 2014

11. 2018

12. Noble 2018

Mnogi jeziki niso le redki, ampak tudi popolnoma izključeni iz procesa digitalizacije.¹³ Le 5 % jezikov je živih na spletu in po nekaterih ocenah večina svetovnih jezikov do konca 21. stoletja na spletu ne bo preživela.¹⁴ Eden od ciljev digitalne humanistike je večji poudarek na egalitarnosti jezikov.

Izziv: Večja pravičnost algoritmov

Kaj menite, kako bi lahko v družbi zagotovili, da bi bili računalniški algoritmi pravični in ne bi diskriminirali določenih posameznikov ali skupin? Zamislite si nekaj rešitev brez uporabe spleta in drugih virov ter jo zapišite. Zatem v spletnih in drugih virih poiščite možne rešitve. V skupini študentje in študentke drug drugemu predstavite svoje zamisli in jih komentirajte z vidika lastnih idej in rešitev, ki ste jih našli na spletu ter v drugih virih.

13. Mikami in Kodama 2012

14. Kornai 2013

VIRI IN LITERATURA

Literatura

1. Adwetewa-Badu, Bemma, 2020: Poetry from Afar: Distant Reading, Global Poetics, and the Digital Humanities. *Modernism / Modernity* 5/1.
2. Alexander, Sam, 2019: Social Network Analysis and the Scale of Modernist Fiction, *Modernism, modernity* 3/4. Dostopno na: <https://modernismmodernity.org/forums/posts/social-network-analysis>.
3. Allington, Daniel, Brouillette, Sarah, and Golumbia, David, 2016: Neoliberal Tools (and Archives): A Political History of Digital Humanities. *Los Angeles Review of Books*. Dostopno na: <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.
4. Allison, Sarah, Gemma, Marissa, Heuser, Ryan, Moretti, Franco, Tevel, Amir, and Yamboliev, Irena, 2013: Style at the Scale of the Sentence. Pamphlets of the Stanford Literary Lab 5.

Stanford, California: Stanford Literary Lab.

5. Anderson, Chris 2008: The End of Theory: the Data Deluge Makes the Scientific Method Obsolete. *Wired*. Dostopno na: <https://www.wired.com/2008/06/pb-theory/>.
6. Anhert, Ruth, Anhert, Sebastian, Coleman, Catherine, and Weingart, Scott, 2020: *The network turn*. Cambridge: Cambridge University Press.
7. Apollon, Daniel, Bélisle, Claire, and Régnier, Philippe (ur.), 2014: *Digital Critical Editions*. In Haugen, Odd Einar, and Apollon, Daniel: *The Digital Turn in Textual Scholarship: Historical and Typological Perspectives*. Illinois: University of Illinois Press. 35–57.
8. Bailey, Moya, Cong-Huyen, Anne, Lothian, Alexis, and Phillips, Amanda, 2016: *Reflections on a Movement: #transformDH*, *Growing Up. Debates in the Digital Humanities*. Dostopno na: <https://dhdebates.gc.cuny.edu/read/untitled/section/9cf90340-7aae-4eae-bdda-45b8b4540b6b>.
9. Bajt, Jan in Robnik-Šikonja, Marko, 2022: *Strojna analiza tematik in sentimenta*. *Uporabna informatika* XXX/1. 19–36.
10. Baldha, Rohan, 2021: *A Study on Sentimental Analysis of Novels and Shakespeare's Play*. *International Journal for Research in Applied Science & Engineering Technology* 9/11. 2088–2099.

11. Barnett, Fiona, 2014: The Brave Side of Digital Humanities. *Differences* 25/1. 64–78.
12. Ben Blatt, 2017: Nabokov's favorite word is Mauve. New York: Simon & Schuster.
13. Biemann, Chris, Crane, Gregory R., Fellbaum, Christiane D., and Mehler, Alexander, 2014: Computational Humanities – bridging the gap between Computer Science and Digital Humanities. *Dagstuhl Reports* 4/7.
14. Beveridge, Andrew and Shan, Jie, 2016: Network of Thrones. *Math Horizons* 23/4. 18–22.
15. Benthall, Sebastian, and Haynes, Bruce D., 2019: Racial categories in machine learning. *Proceedings of the Conference on Fairness, Accountability, and Transparency*. Atlanta, USA. Dostopno na: <https://arxiv.org/pdf/1811.11668.pdf>.
16. Boyd, Ryan L., Blackburn, Kate G., and Pennebaker, James W., 2020: The narrative arc: Revealing core narrative structures through text analysis. *Science advances* 6/32. Dostopno na: <https://www.science.org/doi/10.1126/sciadv.aba2196>.
17. Caroline Levine, 2013: From Nation to Network. *Victorian studies* 55/4. 647–666.
18. Cheng, Jonathan Y., 2020: Fleshing Out Models of Gender in English-Language Novels (1850-2000). *Journal of cultural analytics* 5/1. Dostopno na: <https://culturalanalytics.org/>

- article/11652-fleshing-out-models-of-gender-in-english-language-novels-1850-2000.
19. Christen, Alessio and Spadini, Elena, 2019: Modeling genetic networks. Gustave Roud's œuvre, from diary to poetry collections. *Umanistica Digitale* 3/7. 77–104.
 20. Culler, Jonathan, 2010: Introduction: Critical Paradigms. *PMLA: Special Topic Literary Criticism for the Twenty-First Century* 125/4. 905–915.
 21. Ditmajer, Nina, Ogrin, Matija, and Erjavec, Tomaž, 2019: Encoding textual variants of the early modern Slovenian poetic texts in TEI. *Prispevki za novejšo zgodovino = Contributions to the contemporary history; Contributions á l'histoire contemporaine; Beiträge zur Zeitgeschichte* 59/1. 10–28.
 22. Dotson, Daniel, 2009: Portrayal of Physicists in Fictional Works. *Comparative Literature and Culture* 11/2.
 23. Drucker, Johanna, 2017: Why Distant Reading Isn't. *PMLA/Publications of the Modern Language Association of America* 132/3. 628–635.
 24. Eder, Maciej, Piasecki, Maciej, and Walkowiak, Tomasz, 2017: An open stylometric system based on multilevel text analysis. *Cognitive Studies/Études Cognitives* 17. 1–26.
 25. Eder, Maciej, 2018: Elena Ferrante: a virtual

- author. In Tuzzi, A. and Cortelazzo, M. A. (Eds.): *Drawing Elena Ferrante's Profile*. Padova: Padova University Press. 31–45.
26. Edelstein, Dan, Findlen, Paula, Ceserani, Giovanna, Winterer, Caroline in Coleman, Nicole, 2017: *AHR Forum Historical Research in a Digital Age: Reflections from the Mapping the Republic of Letters Project*. *American historical review* 122/2. 400–424.
 27. Efstathios Stamatatos, 2008: *A Survey of Modern Authorship Attribution Method*. *Journal of the American Society for Information Science and Technology* 60/3. 538–561.
 28. Elkins, Katherine and Chun, Jon, 2019: *Can sentiment analysis reveal structure in a “plotless” novel?* Cornell University. Dostopno na: <https://arxiv.org/pdf/1910.01441.pdf>.
 29. Elson, David, Dames, Nicholas, and McKeown, Kathleen, 2010: *Extracting Social Networks from Literary Fiction*. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Uppsala: Association for Computational Linguistics. 138–147.
 30. Eve, Martin Paul, 2022: *The digital humanities and literary studies*. Oxford: Oxford University Press.
 31. Eve, Martin Paul, 2019: *Close Reading with Computers Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*. Stanford: Stanford University Press.

32. Fish, Stanley, 2012: Mind Your P's and B's: The Digital Humanities and Interpretation. New York Times. Dostopno na: <https://archive.nytimes.com/opinionator.blogs.nytimes.com/2012/01/23/mind-your-ps-and-bs-the-digital-humanities-and-interpretation/>.
33. Fitzpatrick, Kathleen, 2010: Reporting from the Digital Humanities 2010 Conference. The Chronicle of Higher Education. Dostopno na: <https://web.archive.org/web/20190829004943/https://www.chronicle.com/blogs/profhacker/reporting-from-the-digital-humanities-2010-conference/25473>.
34. Flores, Leonardo, 2021: Third-Generation Electronic Literature. In O'Sullivan, James (ur.): Electronic Literature as Digital Humanities: Contexts, Forms, & Practices. New York: Bloomsbury Academic. 26–43.
35. Habjan, Jernej, 2012: The Bestseller as the Black Box of Distant Reading: The Case of Sherlock Holmes. *Primerjalna književnost* 35/1. 91–105.
36. Heuser, Ryan in Le-Khac, Long, 2012: A quantitative literary history of 2,958 nineteenth-century British novels: The semantic cohort method. Pamphlets of the Stanford Literary Lab

4. Dostopno na: <https://litlab.stanford.edu/LiteraryLabPamphlet4.pdf>.
37. Hladnik, Miran, 1993: Preštevna določila slovenske povesti. Slavistična revija 41/1. 65–75.
38. Hladnik, Miran, 1998: Regionalizem in slovenska književnost. Ljubljana: Filozofska fakulteta, Oddelek za slovanske jezike in književnosti.
39. Hladnik, Miran, 2016: Vključevanje drugega in drugačnega v slovensko literarno zgodovino. Zupan Sosič, Alojzija (ur.): Drugačno v slovenskem jeziku, literaturi in kulturi: 52. SSJLK: Zbornik predavanj. Ljubljana: Znanstvena založba Filozofske fakultete. 46–53.
40. Hladnik, Miran, 2014/2016: Nova pisarija. Dostopno na: <http://lit.ijs.si/pokrajin.html>. Hladnik, Miran, 2014/2016: Nova pisarija: Strokovno pisanje po spletu. Ljubljana: Znanstvena založba Filozofske fakultete.
41. Hockey, Susan, 2000: The history of humanities computing. Schriebman, Susan, Siemens, Ray in Unsworth, John (ur.): A Companion to Digital Humanities. Blackwell publishing. Dostopno na: https://companions.digitalhumanities.org/DH/?chapter=content/9781405103213_chapter_1.html.
42. Hui Kyong Chun, Wendy idr., 2016: The Dark Side of the Digital Humanities. Gold, Matthew K.

- in Klein, Lauren F. (ur.): *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press. 493–509.
43. Grandjean, Martin, 2015: Network visualization: mapping Shakespeare's tragedies. Dostopno na: <http://www.martingrandjean.ch/network-visualization-shakespeare/>.
 44. Greenfield, Patricia M., 2013: The Changing Psychology of Culture From 1800 Through 2000. *Psychological Science* 24/9. 1722–1731.
 45. Hachmann, Gundela, 2023: Network analysis in literature and the arts: rethinking agency and creativity. *JLT* 17/2. 221–240.
 46. Elkins, Katherine in Chun, Jon, 2019: Can sentiment analysis reveal structure in a “plotless” novel? *arXiv*. Dostopno na: <https://arxiv.org/pdf/1910.01441.pdf>
 47. Fox, Neal P. and Ehmoda, Omran in Charniak, Eugene, 2012: Statistical Stylometrics and the Marlowe-Shakespeare Authorship Debate. *Proceedings of the Georgetown University on Language and Linguistics*. Washington: Georgetown University. Dostopno na: <https://cs.brown.edu/research/pubs/theses/masters/2012/ehmoda.pdf>.
 48. Gavin, Michael, 2016: Historical Text Networks: The Sociology of Early English Criticism. *Eighteenth-Century Studies* 50/1. 53–80.
 49. Gavin, Michael, 2016: Historical Text Networks:

- The Sociology of Early English Criticism. *Eighteenth-Century Studies* 50/1. 53–80.
50. Gessey-Jones, Thomas, Connaughton, Colm, Dunbar, Robin, and Yose, Joseph, 2020: Narrative structure of *A Song of Ice and Fire* creates a fictional world with realistic measures of social complexity. *PNAS* 117/46. 28582–28588.
 51. Gottschall, Jonathan, 2008b: The "Beauty Myth" is No Myth. *Human Nature* 19. 174–188.
 52. Haugen, Odd Einar, and Apollon, Daniel, 2014: *The Digital Turn in Textual Scholarship: Historical and Typological Perspectives*. In Apollon, Daniel, Bélisle, Claire, and Régnier, Philippe (ur.): *Digital Critical Editions*. Illinois: University of Illinois Press. 35–57.
 53. Jančič, Nika in Pribožič, Katarina, 2022: **Obravnava mita lepote v delih Govekarja, Kersnika, Kvedrove in Pajkove. Jezik in slovstvo** 67/1–2. 197–207.
 54. Jagoda, Patrick, 2017: *Literature: Networks*. Dostopno na: <https://oxfordre.com/literature/display/10.1093/acrefore/9780190201098.001.0001/acrefore-9780190201098-e-135>.
 55. Jockers, Matthew L. in Mimno, David, 2013: Significant themes in 19th-century literature. *Poetics* 41/6. 750–769.

56. Jockers, Matthew, 2013: *Macroanalysis*. Urbana, Chicago, Springfield: University of Illinois Press.
57. Jockers, Matthew L. in Kirilloff, Gabi, 2016: Understanding Gender and Character Agency in the 19th Century Novel. *Journal of Cultural Analytics* 2/2. Dostopno na: <https://culturalanalytics.org/article/11066-understanding-gender-and-character-agency-in-the-19th-century-novel>.
58. Juola, Patrick, 2008: Authorship attribution. *Information Retrieval* 1/3. 233–334.
59. Juola, Patrick, 2015: The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions. *Digital Scholarship in the Humanities*. 30/1. 100–113.
60. Juvan, Marko, Šorli, Mojca in Žejn, Andrejka, 2021: Interpretiranje literature v zmanjšanem merilu: “oddaljeno branje” korpusa “dolgega leta 1968”. *Jezik in slovstvo* 66/4. 55–76, 140.
61. Kazantseva in Szpakowicz, 2010: Summarizing Short Stories. *Computational Linguistics* 36/1. 71–109.
62. Kearns, Edward, 2020: Annotating and quantifying narrative time disruptions in modernist and hypertext fiction. *Proceedings of the 1st Joint Workshop on Narrative Understanding, Storylines, and Events*.

- Association for Computational Linguistics.
72–77. Dostopno na: <https://aclanthology.org/2020.nuse-1.9.pdf>.
63. Kim, Evgeny in Klinger, Roman, 2021: A survey in sentiment and emotion analysis for computational literary studies. *Zeitschrift für digitale Geisteswissenschaften*. Erstveröffentlichung vom. Dostopno na: https://zfdg.de/2019_008.
64. Kirilloff, Gabi, Capuano, Peter J., Fredrick, Julius in Jockers, Matthew L., 2018: From a distance 'You might mistake her for a man': A closer reading of gender and character action in *Jane Eyre*, *The Law and The Lady*, and *A Brilliant Woman*. *Digital Scholarship In The Humanities* 33/4. 821–844.
65. Kleymann, Rabea, 2020: Softwareprototyp: Zur Theoriemündigkeit Experimenteller Software in Den Digital Humanities. In Huber, Martin (ur.): *Wovon sprechen wir, wenn wir von Digitalisierung sprechen? Gehalte und Revisionen zentraler Begriffe des Digitalen*. Frankfurt am Main: Universitätsbibliothek Johann Christian Senckenberg, CompaRe. 119–138.
66. Kornai, Andras, 2013: Digital language death. *PLOS ONE* 8/10. 1–11.
67. Liu, Bing, 2015: *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge: Cambridge University Press.

68. Long, Hoyt in So, Richard Jean, 2016: Turbulent Flow: A Computational Model of World Literature. *Modern Language Quarterly* 77/3. 345-367. [
69. Lothian, Alexis, and Phillips, Amanda, 2013: Can digital humanities mean transformative critique? *Journal of e-media studies* 3/1. Dostopno na: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b065de73479b1fd31b2df3d275377d112077f589>.
70. Marinković, Mila, 2023: Data visualizations in the “PISMA” collection: Cultural Heritage at the intersection of informatics and literary science. *Tipkopolis*.
71. Mandell, Laure, 2016: Gendering Digital Literary History: What Counts for Digital Literary History. In Schreibman, Susan, Siemens, Ray and Unsworth, John (ur.): *A New Companion to Digital Humanities*. Wiley-Blackwell. 511–523.
72. Mandić, Lucija, 2023: Kako meščanski je v resnici meščanski roman? Kvantitativna vsebinska analiza daljše slovenske pripovedne proze v dolgem 19. stoletju. Simpozij Obdobja. V tisku.
73. McGillivray, Barbara, Poibeau, Thierry, and Ruiz Fabo, Pablo, 2020: Digital Humanities and Natural Language Processing: “Je t’aime... Moi non plus”. *Digital Humanities Quarterly* 14/2. Dostopno na: <https://www.digitalhumanities.org/dhq/vol/14/2/000454/000454.html>.

74. McPherson, Tara, 2012: Why is the Digital Humanities So White?, or, Thinking the Histories of Race and Computation. In Gold, Matthew K. (ur.): Debates in the Digital Humanities. Minneapolis: Minnesota University Press. 138–160.
75. Moretti, Franco, 2000a: Conjectures on World Literature”. *New Left Review* 1/54.
76. Moretti, Franco, 2000b: The slaughterhouse of literature. *Modern Language Quarterly* 61/1. 207–227.
77. Muzny, Grace, Fang, Michael, Chang, Angel in Jurafsky, Dan, 2017: A Two-stage Sieve Approach for Quote Attribution. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia: Association for Computational Linguistics. 460–470.
78. Nalisnick, Eric T. in Baird, Henry S., 2013: Character-to-Character Sentiment Analysis in Shakespeare’s Plays. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Sofia: Association for Computational Linguistics. 479–483.
79. Noble, Safiya Umoja, 2018: Algorithms of Oppression: How Search Engines Reinforce Racism. New York: New York University Press.
80. Olson, Hope A., 2001: The Power to Name: Representation in Library Catalogs. *Signs* 26/3.

639–668.

81. Prescott, Andrew, 2023: Bias in Big Data, Machine Learning and AI: What Lessons for the Digital Humanities? *Digital Humanities Quarterly* 17/2. Dostopno na: <https://www.digitalhumanities.org/dhq/vol/17/2/000689/000689.html>.
82. Pennebaker, James W., 2011: *The secret life of pronouns: what our words say about us*. New York: Bloomsbury Press. 74–75.
83. Piper, Andrew, 2016: *There Will Be Numbers*. *Journal of Cultural Analytics* 1/1. Dostopno na: <https://culturalanalytics.org/article/11062-there-will-be-numbers>.
84. Piryani, Rajesh, Gupta, Vedika, Kumar Singh, Vivek, and Pinto, David, 2018: Book impact assessment: A quantitative and text-based exploratory analysis. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology* 34/5. 3101–3110.
85. Pojoga, Vlad, Neagu, Laurentiu-Marian in Dascalu, Mihai, 2020: The Character Network in Liviu Rebreanu's *Ion*: A Quantitative Analysis of Dialogue. *Metacritic Journal for Comparative Studies and Theory* 6/2. Dostopno na: <https://www.metacriticjournal.com/article/164/the-character-network-in-liviu-rebreanus-ion-a-quantitative-analysis-of-dialogue>.

86. Prescott, Andrew, 2023: Bias in Big Data, Machine Learning and AI: What Lessons for the Digital Humanities? Digital Humanities Quarterly 17/2. Dostopno na: <https://www.digitalhumanities.org/dhq/vol/17/2/000689/000689.html>
87. Rodrigo-Ruiz, Débora, 2022: Emotions in the literary text. Emotions in the literary text. A digital analysis of Cervantes' narrative. *Káñina* 46/1.117–141.
88. Slodnjak, Anton. 1961. Realizem II. Legiša, Lino (ur.): Zgodovina slovenskega slovstva III. Ljubljana: Slovenska matica.
89. Stanley, David J., 2014: Computers, visualisation and history: How new technology will transform our understanding of the past. London: Routledge.
90. Ramsay, Stephen, 2011: Reading Machines: Toward an Algorithmic Criticism. Illinois: University of Illinois Press.
91. Rommel, Thomas, 2000: Literary Studies. Schriebman, Susan, Siemens, Ray in Unsworth, John (ur.): A Companion to Digital Humanities. Blackwell publishing. Dostopno na: <https://companions.digitalhumanities.org/DH/>.
92. Sack, Graham, 2012: Character Networks for Narrative Generation. Intelligent Narrative Technologies: Papers from the 2012 AIIDE. 38.

93. Shore, Daniel, 2016: Was It for This? Influence, Archive, and Network. *Modern Philology* 113/3. 398-421.
94. Stern, David, 2010: The Bergen Electronic Edition of Wittgenstein's Nachlass. *European Journal of Philosophy* 18/3. 455-467.
95. Tanigawa, Katie, 2013: Warping the City: Joyce in a Mudbox: The Maker Lab in the Humanities (MLab). Dostopno na: <https://maker.uvic.ca/about/>
96. Yoshiki, Mikami, and Kodama, Shigeaki, 2012: Measuring linguistic diversity on the web. In Vannini, Laurent, and Le Crosiner, Herve (ur.): *Net.Lang: Towards the Multilingual Cyberspace*. Caen: C&F Editions. 121-139.
97. Vitali, Giovanni Pietro, 2018: Rethinking Rome as an Anthology: The Poeti del Trullo's Street Poetry. *Umanistica Digitale* 3. 105-136.
98. Warwick, Claire, 2016: Building Theories or Theories of Building? A Tension at the Heart of Digital Humanities. Unsworth, John idr. (ur.): *A New Companion to Digital Humanities*, edited by New York: John Wiley & Sons. 538-52.
99. Weidman, Sean in Pastor, Aaren, 2021: Modernism and Gender at the Limits of Stylometry. *Digital humanities quarterly*. 15/4. Dostopno na: <http://www.digitalhumanities.org/dhq/vol/15/4/000566/000566.html>.
100. Zehe, Albin, Becker, Martin, Hettinger, Lena, Hotho, Andreas, Reger, Isabella in Jannidis,

Fotis, 2016: Prediction of Happy Endings in German Novels based on Sentiment Information. P. Cellier, T. Charnois, A. Hotho, S. Matwin, M.-F. Moens, Y. Toussaint (ur.): Proceedings of DMNLP. Riva del Garda.

101. Ẓejn, Andrejka, 2020: Rac̣unalniṣ ko podprta stilometric̣ na analiza pripovedne literature Janeza Ciglerja in Christopha Schmidia v slovenščini. *Fluminensia* 32/2. 137–158.
102. Ẓejn, Andrejka, 2021: Izhodišča slovenske pripovedne proze. Dvestoletna tradicija slovenske pripovedne proze: od sredine 17. do sredine 19. stoletja. Ljubljana: ZRC SAZU, Inštitut za slovensko literaturo in literarne vede. Dostopno na: <https://ispp.zrc-sazu.si/kolofon/>.

Viri

1. Anderson, Chris, 2008: The End of Theory: the Data Deluge Makes the Scientific Method Obsolete. *Wired*. Dostopno na: <https://www.wired.com/2008/06/pb-theory/>.
2. Ćosić, Vuk, 2023: Nacija – Kultura. Dostopno na: <https://nacija-kultura.si/#/projekt>.
3. Daniels, Matt: The largest vocabulary in hip hop. Dostopno na: <https://pudding.cool/2017/02/vocabulary/>.

4. Digitalna knjižnica Slovenije, Dlib. Narodna in univerzitetna knjižnica Ljubljana. Dostopno na: <https://www.dlib.si>.
5. Ditmajer, Nina (ur.): Foglarjeva pesmarica. Ljubljana: Založba ZRC SAZU. Dostopno na: <https://zalozba.zrc-sazu.si/sl/publikacije/foglarjeva-pesmarica>.
6. ELO, zbirke 1–4; Electronic Literature Collection (2006–2022). Dostopno na: <https://collection.eliterature.org/>.
7. Erjavec, Tomaž, Hladnik, Miran, Juvan, Marko in Mihurko Poniž, Katja, 2021: *Slovenian Novel Corpus (ELTeC-slv): April 2021 release (v2.0.0)*. Zenodo. Dostopno na: <https://doi.org/10.5281/zenodo.4662600>.
8. European Association for Digital Humanities. Dostopno na: <https://eadh.org/publications/what-digital-humanities>.
9. Mihurko, Katja (ur.): Elektronska zbirka Pisma. Raziskovalni center za humanistko Univerze v Novi Gorici. Dostopno na: <https://pisma-rch.ung.si>.
10. Mihurko, Katja (ur.), 2023: Gledališka cenzura in uprizoritve Dramatičnega društva v Ljubljani (1873–1914). Ljubljana: Inštitut za novejšo zgodovino. Dostopno na: <https://sidih.github.io/cenzura/index.html>.
11. Purg, Peter, 2005: Literatura in digitalna kultura. Dostopno na: digilit.pleter.net.

12. Scheinfeldt T, 2008: Sunset for ideology, sunrise for methodology? Found History. Dostopno na: <https://foundhistory.org/2008/03/sunset-for-ideology-sunrise-for-methodology/>.
13. Rice, Justin, 2018: What Makes Hemingway Hemingway? The LitCharts Blog. Dostopno na: <https://www.litcharts.com/blog/analitics/what-makes-hemingway/>.
14. Stamatatos, Efstathios, 2018: A Survey of Modern Authorship Attribution Method. Journal of the American Society for Information Science and Technology 60/3. 538–556.
15. Stritar, Josip: Zorin. Wikivir. Dostopno na: <https://sl.wikisource.org/wiki/Zorin>.
16. Tavčar, Ivan: Cvetje v jeseni; Wikivir. Dostopno na: [https://sl.wikisource.org/wiki/Cvetje_v_jeseni_\(Ljubljanski_zvon\)](https://sl.wikisource.org/wiki/Cvetje_v_jeseni_(Ljubljanski_zvon)).
17. The Wittgenstein Archives. University of Bergen (WAB). Dostopno na: https://wab.uib.no/wab_nachlass.page/
18. What is digital humanities? European Association for Digital Humanities. Dostopno na: <https://eadh.org/publications/what-digital-humanities>.
19. Železnikar, Jaka, 1999: Interaktivlija. Dostopno na: <https://www.jaka.org/1997/int/int.htm>.
20. Zunković, Igor, 2017: Reprezentacija spola

pri prejemnikih nagrade kresnik. Slavistic ~ na
revija 65/2017, s ~ t. 1. 183–198.