

# Računalno potpomognuta analiza likova u odabranim tekstovima hrvatske književnosti

---

**Kasunić, Lorena**

**Master's thesis / Diplomski rad**

**2021**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:131:190822>

*Rights / Prava:* [In copyright](#) / [Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-06-29**



Sveučilište u Zagrebu  
Filozofski fakultet  
University of Zagreb  
Faculty of Humanities  
and Social Sciences

*Repository / Repozitorij:*

[ODRAZ - open repository of the University of Zagreb  
Faculty of Humanities and Social Sciences](#)



SVEUČILIŠTE U ZAGREBU  
FILOZOFSKI FAKULTET  
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI  
SMJER NASTAVNIČKA INFORMATIKA  
ODSJEK ZA KOMPARATIVNU KNJIŽEVNOST  
Ak. god. 2020./2021.

Lorena Kasunić

**Računalno potpomognuta analiza likova u odabranim tekstovima  
hrvatske književnosti**

Diplomski rad

Mentori: dr. sc. Petra Bago, doc. i dr. sc. Lovro Škopljanac, doc.

Zagreb, srpanj 2021.

## **Izjava o akademskoj čestitosti**

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenoj i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

---

(potpis)

## Sadržaj

1. Uvod .....	1
2. Digitalna humanistika kao znanstvena disciplina .....	2
2.1. Istraživačka građa digitalne humanistike .....	4
2.2. Negativne strane digitalne humanistike.....	5
3. Digitalna humanistika i informacijske i komunikacijske znanosti .....	8
4. Autor - tekst - čitatelj .....	10
4.1. Koje književne tekstove proučavati .....	12
4.2. Kako računalne metode mijenjanju način interpretacije .....	12
5. Naratologija.....	14
5.1. Naratologija i književni likovi .....	14
5.1.1. Elementi karakterizacije likova .....	16
5.1.2. Aktancijalni model .....	17
5.1.3. Primjena aktancijalnog modela .....	18
6. Povezana istraživanja.....	20
7. Obrada prirodnog jezika.....	21
7.1. Osnovni pojmovi i postupci.....	23
7.2. <i>Sketch Engine</i> .....	24
8. Istraživanje likova u korpusu odabranih hrvatskih književnih tekstova .....	27
8.1. Metode.....	28
8.2. Rezultati.....	29
8.3. Diskusija.....	35
9. Zaključak.....	45
Literatura .....	46
Popis slika.....	49
Popis tablica .....	49
Sažetak.....	50

## 1. Uvod

Tekst je od svojih samih početaka budio interes čovjeka za istraživanjem. Pojava pisane riječi otvorila je prostor još jednom načinu komuniciranja – usmenost više nije bila jedina mogućnost izražavanja i uspostavljanja ljudske interakcije. Razvojem društva i tehnologije, i proučavanje teksta doživjelo je određene promjene i novosti. Poseban je utjecaj na to u nedavno vrijeme imao proces digitalizacije i znanstveno-tehnološka revolucija koje su omogućile da se, osim u tiskanom, tekstu pristupi i u digitalnom mediju.

Književnost kao specifičan oblik izražavanja i stvaranja tekstova također je zahvaćena spomenutim mijenama. Spoj humanistike i tehnologije doveo je do razvoja digitalne humanistike koja se, između ostalog, bavi tekstualnim podacima. Među njima su, u kontekstu ovoga rada, posebno važni književni tekstovi. Za istraživanje i analiziranje objekata književnosti bitna je i za usko vezana obrada prirodnog jezika, domena i informacijskih i komunikacijskih znanosti, što se pokazuje u istraživačkome dijelu ovoga rada.

Rad se usredotočuje na odabrane primjere pripovjednih tekstova hrvatske književnosti, a teorijsko uporište nalazi u naratologiji. Fokus je na računalno potpomognutoj analizi likova, u kojoj su upotrijebljeni naratološki elementi karakterizacije i aktancijalni model. Cilj rada je pokazati kako se uporabom računala mogu realizirati određene književnoteorijske postavke te dobiti kvalitetni podaci korisni u tumačenju likova.

Prvi, teorijski dio rada, govori o položaju digitalne humanistike s posebnim naglaskom na njezin odnos sa znanosti o književnosti i informacijskim i komunikacijskim znanostima. Iznosi se i dio teorije književnosti koji se odnosi na tretiranje teksta, naratološku analizu likova i aktancijalni model. Nadalje, objašnjavaju se najbitniji pojmovi obrade prirodnog jezika i opcije programa *Sketch Engine*, koje su se koristile u nastavku rada. Praktični dio rada opimjeruje objašnjene teorije i pristupe na odabranim književnim tekstovima – iznose se metode, zatim rezultati, koji se nakon toga tumače. Na samome kraju prikazuju se glavni zaključci cijeloga rada.

## 2. Digitalna humanistika kao znanstvena disciplina

Digitalna humanistika relativno je nova znanstvena disciplina koja predstavlja onaj dio znanosti koji teži za povezivanjem različitih znanstvenih područja, što rezultira interdisciplinarnošću. Većina teoretičara i praktičara digitalne humanistike složiti će se da se radi o interdisciplinarnom znanstvenom području, što podrazumijeva društvene, humanističke i tehničke znanosti. Neslužbeni manifest digitalne humanistike (*The Digital Humanities Manifesto 2.0*, u nastavku: *Manifest*) navodi da je digitalna humanistika „niz konvergentnih praksi” koje djeluju u svijetu u kojem tiskani mediji više nemaju prvenstvo u diseminaciji znanja, nego se isprepliću s ostalim medijima. Također, sve zastupljeniji digitalni alati i tehnologije promijenili su načine širenja znanja iz područja umjetnosti, humanistike i društvenih znanosti (*Manifest*, n.d.). Digitalna humanistika proučava različita polja humanističkih znanosti (muzikologiju, teoriju likovnih umjetnosti, arheologiju, filologiju itd.), stoga je njezino područje interesa vrlo široko. Međutim, u ovome se radu pod terminom digitalne humanistike neće pokriti sve što je navedeno da zanima ovo znanstveno polje, već će se digitalna humanistika promatrati u jednom užem smislu – kao disciplina čiji objekt proučavanja jest književnost, ili preciznije – književni tekst. Stoga bi se moglo reći da se rad orijentira na potpolje digitalne filologije. Pojam humanistike također će se suziti u značenju i odnositi samo na humaniste koji se bave proučavanjem književnosti. Isto tako, interdisciplinarnost područja bit će prvenstveno sagledana iz kuta informacijskih i komunikacijskih znanosti (društvene znanosti) i znanosti o književnosti (humanističke znanosti).

Povijest umjetnosti, kulturalni studiji, znanost o književnosti, područja su čiji su stručnjaci oduvijek bili okrenuti pomnom čitanju, hermeneutici i proučavanju jedinstvenog autorstva (Clement i Carter, 2017: 1). Međutim, razvojem digitalne tehnologije, digitalna humanistika kao potpodručje (kako ga definiraju Tanya E. Clement i Daniel Carter) navedenih područja počela je mijenjati načine istraživanja humanističkih objekata. Znanstvenici koji se bave digitalnom humanistikom većinom su tradicionalno „obučeni” humanisti koji svoja znanja i vještine koriste kako bi objekte svojih proučavanja istraživali metodama i pristupima digitalne humanistike (Clement i Carter, 2017: 2). Dok bi to moglo vrijediti za anglosaksonsko govorno područje, situacija u Hrvatskoj ponešto je drugačija. Ovdje sâm naziv „digitalni humanist” nije raširena pojava, a oni koji se bave digitalnom humanistikom češće „potječu” iz disciplina kao što su

lingvistika, kroatistika, informacijske i komunikacijske znanosti. Isti je slučaj i s obrazovnim sustavom i ponudom obrazovnih programa – SAD i Ujedinjeno Kraljevstvo prednjače u razvijenosti institucija, ponudi smjerova i kolegija iz digitalne humanistike, dok je u Hrvatskoj akademska mreža područja još uvijek poprilično nerazvijena.

Davor Nikolić (2016) upozorava da u hrvatskom visokoobrazovnom sustavu nema dovoljno implementiranja kvantitativnih i računalnih pristupa u području filologije, a isti zaključak može se izvesti i za studije književnosti. Vrijedi istaknuti da su humanisti oduvijek radili s informacijskim izvorima – knjižnicama, arhivima, knjigama, ali digitalna okolina i drugačiji načini pohrane, pa i stvaranja informacija, stavljaju pred humaniste nove i njima dosad nepoznate aktivnosti u tradicionalnoj humanistici (Clement i Carter, 2017: 2). Nadalje, humanistici se nerijetko spočitava što nema precizno definirane istraživačke metode te što se buduće stručnjake sustavno ne podučava kako i koje metode koristiti u svojim istraživanjima. S druge strane, digitalnoj humanistici se zamjera što se previše orijentira na metode i ne povezuje metodologiju s humanističkim teorijama koje postoje već duže vrijeme (Clement i Carter, 2017: 3).

Nepodudarnost u načinu na koji se gleda na važnost metoda u znanosti o književnost i digitalnoj humanistici proizlazi iz različitih temelja na kojima su discipline nastale. Proučavanje književnosti oduvijek je dio humanističkih znanosti gdje je gotovo cijelo vrijeme glavna i/ili jedina metoda interpretacija. Njome su književni kritičari i teoretičari dolazili do vlastitih teorija o književnosti. Dakle, da bi se došlo do neke teorijske postavke, nužno je krenuti od interpretacije. S ostalim istraživačkim metodama humanisti možda neće biti upoznati tijekom svoga preddiplomskog i diplomskog obrazovanja jer te metode nisu one koje su potrebne da bi se razumjela teorija književnosti. Interpretacija tu ima prvenstvo jer nije slična „krutim” metodama koje se koriste u drugim znanostima (društvenim, tehničkim, prirodoslovnim) i uvelike ovisi o teoriji. Prigovor o prevelikoj orijentiranosti na metode koji se daje digitalnoj humanistici isto se tako može pripisati njezinoj interdisciplinarnoj naravi, gdje bi metodološki dio bio povezan s društvenim i tehničkim znanostima, a teorijski s humanističkim. Iz navedenih razlika proizlazi ono što se može primijetiti kad se čitaju znanstveni članci kojima je tema digitalna humanistika, a to je da o digitalnoj humanistici teorijski uglavnom promišljaju „tradicionalni” humanisti (u kontekstu ovoga rada preciznije bi se moglo reći proučavatelji književnosti). Čini se da digitalni humanisti praktičari nisu toliko zaokupljeni teorijskim određivanjem svoga područja, već se fokusiraju na provođenje konkretnih istraživanja.

## 2. 1. Istraživačka građa digitalne humanistike

Izuzevši numeričke podatke, tekst je najpogodniji tip podatka kojim računala mogu uspješno rukovati. Tome u prilog ide i činjenica da postoji duga tradicija obrade podataka koja se temelji na tekstovima, za razliku od ostalih humanističkih objekata kao što su slike, audio i videozapisi (Kirschenbaum, 2010: 6). Tekstualni podaci služili su za istraživanja u statistici, lingvistici, pripisivanju autorstva, a s obzirom na njihovu sveprisutnost i visoku razinu iskoristivosti, razvilo se i područje obrade prirodnog jezika. Kad Alan Liu (2011: 6) govori o tekstualno orijentiranoj strani digitalne humanistike, njezina ishodišta pronalazi u novoj tekstologiji, sociološkoj bibliografiji i kontekstualnoj ili situacijskoj tekstualnoj kritici.

Rečeno je da humanisti koriste informacije koje su pohranjene u različitim oblicima. Prije pojave i razvitka digitalnih tehnologija, okretali su se knjižnicama, arhivima i ostalim informacijsko-dokumentacijskim centrima. Internetski izvori i *online* baze podataka omogućili su puno veću dostupnost građe koju humanisti mogu proučavati. Osim povećanja količine podataka do kojih se može doći i obraditi, digitalni informacijski izvori doveli su do novih pristupa, kao što je, primjerice, estetika podatkovnih baza (Vesna, 2007 navedeno u Liu, 2011: 12). Liu (2011: 13) predviđa da će se digitalna humanistika u budućnosti još više usmjeriti prema velikim količinama podataka (engl. *big data*).

Digitalna humanistika pomiče se od empirijske paradigme (koja se umnogome odnosila na tekstualno orijentiranu stranu područja) prema paradigmi interpretacije koja počiva na ideji igre (Liu, 2011: 14). Eksperimentiranje s građom, što se sve može dogoditi ako se podvrgne određenoj vrsti obrade ili metodi, otvara mogućnost preispitivanja ustaljenih shvaćanja i stvaranja novih načina čitanja tekstova, a onda i njihova tumačenja.

Digitalna humanistika, kao izrazito interdisciplinarno područje, ima tendenciju problematizirati i propitivati disciplinarne granice, pretpostavke koje se podrazumijevaju u radu proučavatelja književnosti (primjerice, pomno čitanje, očuvanje postojećeg kanona, žanrovske i periodizacijske podjele itd.) (Nikolić, 2016: 5). Davor Nikolić (2016: 6) zalaže se da se fokus stavi na načine primjene programa u humanističkim istraživanjima, odnosno: „potrebno je računalni kôd razumijevati (i) humanistički.”



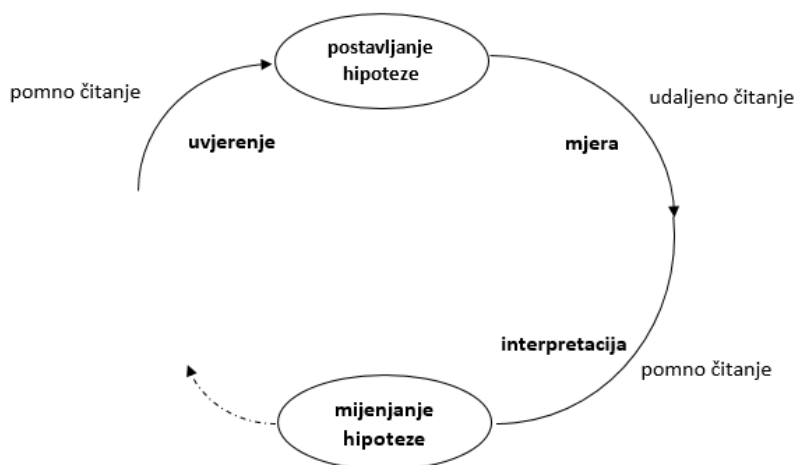
## 2.2. Negativne strane digitalne humanistike

Koji su nedostaci digitalne humanistike? Alan Liu (2011: 17) ističe da je jedan od njih manjak razlikovanja forme novih medija (e-knjige, wikiji, internetske stranice i sl.), bilo na razini žanra, retorike ili stila. Svaki medij posjeduje specifične zakonitosti i one bi se trebale uzeti u obzir prilikom istraživanja, tj. te specifičnosti mogle bi postati i sâm predmet istraživanja. Žanrovskim i stilskim razlikovanjima tradicionalnih vrsta u književnosti bavi se, primjerice, *The Stanford Literary Lab*, u kojem se provode istraživanja na području korpusne (ili kvantitativne) stilistike. Njima se pokazuje kako jezične jedinice, za koje se misli da ne nose nikakvo bitno značenje (kao što su naslovi), mogu pružiti širu sliku dužega vremenskog (a onda i književnog) razdoblja i promjena koje su se zbivale. Time se bave i oni koji proučavaju nove digitalne medije, iako Liu smatra da bi na tome trebalo poraditi, odnosno da još uvijek nema dovoljno istraživanja u tome području.

Kad se govori o odnosu digitalne humanistike i proučavanja književnosti, jedan od „kamena spoticanja” jest makroanaliza ili ono što se popularno naziva udaljenim čitanjem (usp. Jockers, 2013 i Moretti, 2013). Uobičajen rad književnih stručnjaka sastoji se od interpretacije manjeg broja tekstova ili autora jer se teži za dubinskim i detaljnim proučavanjem. To je ono što se naziva pomnim čitanjem ili mikroanalizom.

Liu (2011: 20) se zalaže da se udaljeno i pomno čitanje pomire, odnosno da se međusobno isprepliću i nadopunjuju. Andrew Piper (2018: 25) na sličnome tragu navodi kako se udaljeno i pomno čitanje, tijekom istraživanja jedne hipoteze koju je proučavatelj postavio, izmjenjuju. Slika 1 zornije prikazuje taj proces.

Slika 1. Prikaz procesa istraživanja hipoteze



Liu (2011: 23) daje opasku digitalnoj humanistici kazujući da ona još uvijek ne posjeduje kritičku svijest o društvenim, političkim, kulturnim i ekonomskim aspektima. Humanističke analize i interpretacije, bilo likovnih, glazbenih, književnih i drugih djela, sadržavaju barem neku od navedenih kritika, tj. djela se tumače uzimajući u obzir različite segmente ljudskog djelovanja. Autor promatra digitalnu humanistiku kroz međuodnos tri pojma: tehnologije, medija i informacije. Potonji pojam dobio je poseban značaj razvojem informacijski orijentiranih polja kao što su kibernetika, računalna lingvistika, korpusna analiza, rudarenje podataka i sl. (Liu, 2011: 6). Ono što Liu (2011: 10) primjećuje jest izostanak pojma komunikacije, koji je u tehničkim (te u nešto manjem razmjeru u društvenim) znanostima naznačen uporabom termina informacijske i komunikacijske tehnologije. Jednim od potencijalnih razloga izostanka pojma komunikacije u radu digitalnih humanista može se smatrati naglašavanje lingvističkih i formalističkih značajki teksta. To se posebice zapaža u književnosti, odnosno književnoj teoriji koja je tijekom 20. stoljeća iznjedrila teorije kao što su ruski formalizam, nova kritika, strukturalizam i poststrukturalizam. Dolazi do udaljavanja od značenja, komuniciranja ideja s čitateljima, a naglašava se forma, struktura jezika, njegova zvučnost. Do pomaka je ipak došlo zahvaljujući razvoju *web-a* koji je otvorio prostor teoriji društvenih mreža i u čijoj implementaciji u području književnosti prednjače A. Piper, F. Moretti i Z. Frank (Liu, 2011: 11).

Za proučavatelje književnosti može se reći da zastupaju dva pola. Prvi se odnosi na zagovornike digitalne humanistike (među kojima je vjerojatno najzvučnije ime Franco Moretti), koji se služe digitalnim alatima i programima te metodama digitalne humanistike kako bi čitali „iskosa“, pristupajući književnosti na drugačiji način od onog tradicionalnog. Nerijetko se zanimaju za istraživanje onih književnih elemenata koji su dosad bili zanemareni (ponovno se prizivaju naslovi tekstova), žele dati širu sliku, „čitati“ veliku količinu tekstova i autora kako bi se uočio određen uzorak. S druge strane, postoje oni koji se oštro protive digitalnoj humanistici i novostima koje je ona donijela u proučavanje književnosti. Jedan od najgorljivijih kritičara jest Stanley Fish. On u digitalnoj humanistici vidi opasnost brisanja svih granica linearnosti i sistematičnosti te potkopavanja autoriteta i disciplinarne snage (Fish, 2012). Linearno iskustvo čitanja, koje organizira autor time što izlaže informacije u poretku koji on smatra ispravnim, digitalna tehnologija zamjenjuje višesmjernim iskustvom u kojem autor postaje samo jedan od sudionika

(Fish, 2012). Ono što Fish (2012) također kritizira jest ideja „teksta u procesu”. Naime, riječ je o oksimoronu zbog toga što je proces nužno usmjeren na dovršeni artefakt (tekst), a to je ono što nestaje shvaćanjem teksta kao nečega što se stalno mijenja, kako ga shvaćaju zagovornici digitalnih medija. Digitalna humanistika promiče suradnju između autora i čitatelja, njihovo međusobno nadopunjavanje, otvoreni pristup istraživanjima i publikacijama koje ne bi trebale biti dostupne samo članovima akademskih krugova. Nadalje, Fish (2012) je skeptičan oko optimističnih ideja digitalne humanistike koje prognoziraju da će se intenzivnijim uključivanjem šire društvene zajednice humanistika pozicionirati kao poželjna i zanimljiva znanstvena disciplina te da će digitalna humanistika pružiti studentima vještine uz pomoć kojih će biti konkurentniji na tržištu rada, unutar i izvan akademske zajednice. Međutim, ključno pitanje koje Fish postavlja jest nudi li digitalna humanistika nove i bolje načine o shvaćanju ciljeva humanistike, ili želi u potpunosti promijeniti što ciljevi humanistike i rad humanista jesu.

David Golumbia (2014: 2) također izražava negativan stav prema digitalnoj humanistici, točnije, digitalnim humanistima. Digitalna humanistika nije samo skup metoda, nego izravno interveniranje u znanost o književnosti. Digitalni humanisti pokušavaju umanjiti autoritet „tradicionalnih” humanista i odrediti što jest a što nije dio proučavanja književnosti, kako napominje Golumbia (2014: 2). Golumbia zamjera što se digitalni humanisti upliću u rad proučavatelja književnosti jer smatra da znanost o književnost i digitalna humanistika nisu jednaka područja. Jedan od primjera kojim Golumbia (2014: 6) to nastoji dokazati jest to što su se unutar znanosti o književnosti književne teorije vrlo brzo prihvaćale i postajale dio znanstvene prakse (primjerice poststrukturalizam), dok teorija digitalne humanistike nije zaživjela u proučavanju književnosti na taj način (spominje automatsko prepoznavanje žanra, modeliranje teme). Sljedeće, digitalni humanisti koji su vrlo aktivni u svome području rijetko se bave književnošću, a slično je i kod proučavatelja književnosti – mali broj njih uključuje teorije digitalne humanistike u svoj rad (Golumbia, 2014: 7). Golumbia nema ništa protiv toga da se digitalna humanistika razvija kao znanstvena disciplina, ali ne i kao dio znanosti o književnosti. Ako se digitalni humanist želi baviti proučavanjem književnosti, on to mora činiti poštujući postojeću znanstvenu praksu, odnosno pristupe i načela studija književnosti. Dakle, autor želi da digitalni humanisti jasno definiraju svoj odnos prema znanosti o književnosti.

### **3. Digitalna humanistika i informacijske i komunikacijske znanosti**

Digitalna humanistika, kao što je ranije već navedeno, interdisciplinarno je područje u kojem se upotreba digitalne tehnologije i informacijskih izvora često ističe. To je samo jedna od mnogih poveznica, odnosno dodirnih točaka digitalne humanistike s područjem koje se u hrvatskoj akademskoj zajednici najčešće naziva informacijskim i komunikacijskim znanostima. Obje discipline imaju poprilično nedefiniran položaj i status unutar znanstvene zajednice. Naime, i digitalna humanistika i informacijske i komunikacijske znanosti zbog svoje interdisciplinarnosti nemaju čvrstu poziciju u klasifikaciji znanstvenih područja. Nepostojanje čvrste pozicije u klasifikaciji znanosti odnosi se na međunarodnu klasifikaciju jer su u Hrvatskoj informacijske i komunikacijske znanosti regulirane Pravilnikom o znanstvenim i umjetničkim područjima, poljima i granama.

Fangli Su, Yin Zhang i Zachary Immel (2020) proveli su zanimljivo istraživanje interdisciplinarnih suradnji u digitalnoj humanistici proučavajući članke koji sadrže istraživanja iz digitalne humanistike. Istraživanje je pokazalo svu raskoš disciplina s kojima digitalna humanistika „suraduje”, a one koje su kroz vrijeme konstantne su: računalna znanost, informacijske i komunikacijske znanosti, lingvistika i književnost (Su, Zhang i dr., 2020: 1). Još jedna činjenica koju ističe Tibor Koltay (2016: 3) je ta da su digitalna humanistika i informacijske i komunikacijske znanosti proizašle iz uslužne funkcije jer su znanstvenoj zajednici nudile izvore zabilježenih informacija koji su se upotrebljavali u različitim vrstama istraživanja.

Lyn Robinson, Ernesto Priego i David Bawden (2015) nabrajaju sedam poveznica između informacijskih i komunikacijskih znanosti i digitalne humanistike. Prije nego se navedu, valja napomenuti da autori informacijske i komunikacijske znanosti (tj. ono što se u literaturi zna nazivati i LIS - *Library and information science*) definiraju kao disciplinu koja se bavi proučavanjem „komunikacijskog lanca zabilježenih informacija” (ibid: 1, vlastiti prijevod), koji se temelji na praksama iz knjižničarstva, informacijskog menadžmenta i sl. Ipak, čini se da, na temelju niže navedenih poveznica, autori ne uključuju i komunikacijske znanosti, nego samo informacijske.

Digitalnu humanistiku i informacijske znanosti povezuje:

1. što se mnoga istraživanja digitalne humanistike ostvaruju u suradnji s informacijsko-dokumentacijskim institucijama, primjerice: knjižnicama, arhivima, muzejima

2. podrijetlo u uslužnim funkcijama unutar akademske zajednice (ono što je napomenuo i Tibor Koltay)
3. praksa korištenja dokumenata u širem smislu riječi, kako bi „začahurile” sve zabilježeno znanje
4. zanimanje za teme: pretraživanje, metapodaci i opisivanje izvora, ontologija, klasifikacija, taksonomija, kultura baštine, vizualizacija podataka, rudarenje podataka itd.
5. smještanje u iste akademske jedinice, odnosno institucije
6. obostrano korištenje materijala iz druge discipline (digitalna humanistika se u svom edukacijskom programu služi materijalima povezanim s informacijskim znanostima, i obratno)
7. objavljivanje radova digitalne humanistike u časopisima i ostalim publikacijama koji su primarno specijalizirani u informacijskim znanostima, iako postoji sve više publikacija koje su isključivo orijentirane na digitalnu humanistiku (Robinson, Priego i dr., 2015: 2).

Ono u čemu se digitalna humanistika i informacijske i komunikacijske znanosti djelomično podudaraju, ali i donekle razilaze jest odnos prema teorijskom i praktičnom dijelu svojih područja. Koltay (2016) navodi da je za informacijske i komunikacijske znanosti karakteristično to što su u svome razvoju polazile od prakse prema teoriji, dakle induktivnim putem. Digitalna pak humanistika, zahvaljujući tradicionalnoj humanistici, ima čvrstu teorijsku pozadinu na koju se može osloniti, međutim često je više zanimaju metodološke rasprave jer, kako ističe Koltay (2016), njih je jednostavnije razriješiti, nego one koje se tiču teorije. Nadalje, odnos disciplina prema tehnologiji vrlo je sličan – tehnologija služi kako bi se informacijski objekti mogli sačuvati, tumačiti, pretraživati. U središtu interesa je kako tehnologija mijenja čovjekov odnos prema izvorima informacija i kako se ti izvori mijenjaju. Dakle, digitalna tehnologija nužno je vezana uz informacije, tj. dokumente, objekte koji se proučavaju, a nije izdvojeni predmet istraživanja (kao što bi se moglo reći za, primjerice, računarstvo).

Robinson, Priego i Bawden (2015: 5), promatrajući odnos digitalne humanistike i informacijskih i komunikacijskih znanosti kroz prizmu komunikacijskog lanca dolaze do zaključka da se digitalna humanistika većinski orijentira na rubne dijelove lanca – nastanak i uporabu informacija, a informacijske znanosti na središnji dio komunikacijskog lanca – organizaciju, pretraživanje i upravljanje informacijama.

Na kraju ovog kratkog pregleda odnosa digitalne humanistike i informacijskih i komunikacijskih znanosti, korisno je spomenuti kakva se budućnost predviđa ovim disciplinama. I u tom se pogledu uočava povezanost – nalaze se u stanju „egzistencijalne krize” (ibid: 5). Informacijskim i komunikacijskim znanostima prijeti da postanu suvišne, a razlog tomu je u intenzivnoj promjeni komunikacijskih tehnologija, diseminacijskih metoda i izdavaštva (Robinson, Priego i dr., 2015: 5). Digitalnoj humanistici moglo bi se dogoditi da aktivnosti kojima se bavi preuzmu ostala znanstvena područja koja se pritom ne bi izjašnjavala oznakom „digitalna humanistika”, čime bi se disciplina „izgubila” unutar drugih znanstvenih polja (Robinson, Priego i dr., 2015: 6). Može se primijetiti da ovo nije stvar nekog budućeg vremena, nego nešto što se događa od samih početaka. Ono što bi moglo pružiti svjetliju budućnost ovim disciplinama jest to da se udruže u onim temama koje su im zajedničke. Digitalni humanist Andrew Piper (2018: 20) u svojoj knjizi *Enumerations: Data and Literary Study* ističe upravo informacijske znanosti kao jednu od glavnih disciplina koja je statističkim i algoritamskim pogledom na tekst otvorila prostor računalnom proučavanju književnosti. To je jedan od mogućih putova ostvarivanja suradnje između digitalne humanistike i informacijskih i komunikacijskih znanosti.

#### **4. Autor - tekst - čitatelj**

Književni teorijski trokut, načinjen od autora, teksta i čitatelja, bio je predmet interesa mnogih književnih teoretičara i kritičara. Tijekom povijesti razvoja znanosti o književnosti, naglasak na tim elementima konstantno se mijenjao. Afirmacijom i autonomijom autora (posebice od novovjekovlja), sve se više govorilo o autorovom životu, cijenili su se autobiografski elementi u književnim djelima. Takav pristup unutar književnosti naziva se pozitivizam: „Pozitivizam tako karakterizira *biografizam*, tj. iscrpan rad na proučavanju biografija pojedinih autora, *historizam*, tj. uvjerenje da su književna djela objašnjiva kao spomenici svog vremena te se moraju razmatrati i tumačiti u okviru povijesnog istraživanja o uvjetima njihova nastanka i *psihologizam*, tj. uvjerenje da je potrebno proučavati psihologiju autora kako bi se razumjele njegove tvorevine: umjetnička književna djela.” (Solar, 1984: 216). Sve je to dovelo do udaljavanja od samog teksta u zadnjoj trećini 19. stoljeća. Tekst se nije shvaćao kao objekt koji je neovisan o svome autoru i koji se treba proučavati bez „potrage” za autorom. Uvidjevši nedostatke smjera u kojem se književna kritika uputila, početak dvadesetog stoljeća vraća pozornost na tekst. Jedna od prvih književnih teorija koja se htjela odmaknuti od pozitivističkog proučavanja književnosti bila je ona ruskih formalista

(pretežito 20-e i 30-e godine prošlog stoljeća). Oni su se zalagali da se književnost promatra u svom užem smislu, tj. da se znanost o književnosti bavi samo literarnošću književnosti (Solar, 1984: 219). U središtu interesa jest tekst kao umjetničko djelo, odnosno kako od jezičnog materijala nastaje književni. Kao što se može zaključiti iz naziva, ruski se formalisti najviše zanimaju za formalna određenja književnog teksta. Na engleskom govornom području, točnije, u Sjedinjenim Američkim Državama, razvija se „nova kritika” (između 1920-ih i 1950-ih). Novi kritičari također se okreću književnom tekstu i ne pridaju važnost autoru i njegovoj biografiji ili ostalim izvanknjiževnim kontekstima. Upravo su novi kritičari skovali pojam pomnog čitanja: „Time su upozorili na potrebu da čitalac savjesno i koncentrirano pazi na sve detalje izraza, odnosno na sve one elemente koje tekst sadrži, a koji lako izbjegnu pažnji u čitanju orijentiranom isključivo prema razumijevanju riječi u smislu svakodnevnog govora ili jezika znanosti.” (ibid: 225). Sljedeća književna teorija koja se fokusirala na književni tekst je strukturalizam (razvijao se od 30-ih pa sve do 60-ih godina 20. stoljeća) koji svoje korijene vuče iz strukturalne lingvistike. Ugledajući se na lingvističku teoriju Ferdinanda de Saussurea, strukturalisti se često bave označiteljem i označenim te arbitrarnosti jezičnog znaka (Solar, 1984: 236). Osim strukturalne lingvistike, Milivoj Solar (1984: 236) napominje da je na strukturaliste utjecala i teorija informacije od koje su preuzeli ideju komunikacije kao razmjene poruke između pošiljatelja i primatelja putem kanala. Već se iz navedene sheme prijenosa informacija od ishodišta do odredišta može nazrijeti interes za trećim dijelom književnog trokuta – čitateljem. Proučavanje književnosti koje se usredotočuje na čitatelja, odnosno publiku donijela je teorija recepcije u 60-im godinama 20. stoljeća. Teoretičare koji pripadaju tome pravcu zanima kako čitatelji razumiju i prihvaćaju književnost (Solar, 1984: 242). Uz njih se vezuje pojam „horizont očekivanja” koji se odnosi na „spremnost publike da na temelju svog prethodnog poznavanja književnosti prihvati i novo, originalno književno djelo.” (ibid: 243).

Današnje proučavanje književnosti većinski se orijentira na tekst i čitatelja. U tom smjeru ide i računalno potpomognuto tumačenje književnosti. Međutim, u većoj mjeri radi se o analizi književnih tekstova. Razlog tomu je to što unutar digitalne humanistike postoji tradicija tekstualne analize, tekst kao izvor informacija i humanistički objekt podoban je za podvrgavanje računalnim metodama i alatima. Istraživanje čitatelja i/ili publike više počiva na metodama koje se koriste u društvenim znanostima, primjerice sociologiji. Riječ je o elementu književnosti koji se ne nalazi eksplicitno u samom tekstu. Ipak, tijekom posljednjih godina sve se više analizira i recepcija

čitatelja. Kao primjer može poslužiti studija Federica Pianzole i suradnika (2020) koji su proučavali na koji način mladi čitatelji komentiraju i stvaraju fikciju na *Wattpadu*, platformi za društveno čitanje.

#### **4.1. Koje književne tekstove proučavati**

Bavljenje proučavanjem književnosti neminovno je vezano uz pitanje književnog kanona i autorskog opusa, odnosno uz pitanje o tome što i koliko toga se čita. Dakle, kvantiteta je nešto čega su se književni kritičari i teoretičari morali dotaknuti tijekom svoga profesionalnog rada. Problem o kojem se često diskutira jest po kojim kriterijima određena književna djela postanu dio kanona, a druga ne. Uz to se nadovezuje i pitanje o tome koliko bi kanon smio ili trebao biti velik. Očito je da se vrši selekcija i da unutar samog kanona postoje autorski opusi koji su „razvikaniji” i učestalo tumačeni, a postoje i oni tekstovi kojima prijete da iziđu iz kanonskih granica. Pitanje kvantitete pročitanoa i što se sve čita (odnosno interpretira) Andrew Piper (2018: 17) postavlja na nov način, tako da problem stavlja u područje digitalne humanistike, a inspirira se obradom prirodnog jezika, strojnim učenjem, rudarenjem podataka. Točnije, u središtu njegove pozornosti su ponavljanja u jeziku, ono što je učestalo. Upravo je u tome različitost spram tradicionalnog tumačenja književnih tekstova gdje prevladava interes za ono neobično, singularno, jedinstveno, što se također često proučava u nekim računalnim analizama teksta, primjerice forenzičkoj lingvistici, detekciji plagijata i autorstva i sl. Piper (2018: 19) u ponavljanjima i onome što je zajedničko vidi srž proučavanja književnosti. Matthew L. Jockers (2013: 15), koji u svom izučavanju književnosti koristi makroanalizu, također smatra da se današnji proučavatelji književnosti više ne mogu i ne bi trebali zadovoljiti anegdotalnim dokazima, probranim tekstovima za koje se tvrdi da su reprezentativni, već iskorištavati ono što digitalna humanistika nudi, a to je mnoštvo digitaliziranih tekstova i računalnih alata, pomoću čega se mogu analizirati i inicijalno „nezanimljivi” književni tekstovi.

#### **4.2. Kako računalne metode mijenjanju način interpretacije**

Kad se tumači neki književni tekst, većinom se polazi od određene ideje koja se zatim potvrđuje i dokazuje analizom teksta. Dakle, autor s namjerom odabire tekst koji će biti egzemplan za ono što želi istaknuti kao zanimljivu ili bitnu karakteristiku proučavanog književnog teksta. Piper



(2018: 26) kritizira takav pristup jer se njime ne dobiva informacija o tome koliko tekstova ne podupire početnu ideju i autor *a priori* zaključuje da je njegova ideja ispravna. Metodologija računalno potpomognute analize je drugačija i preuzeta je iz prirodnih znanosti. Polazi se od neke hipoteze, ona se testira na većem broju tekstova i ne isključuje se mogućnost da će hipoteza biti pobijena. Osim toga, računalne metode pomažu u postavljanju pravila, sistematiziranosti istraživanja i time se omogućava da se isti procesi provedu nad nekim drugim tekstovima. Nadalje, algoritmi, koji su sastavni dio izgradnje nekog računalnog alata, tjeraju istraživače da detaljno razrade svoje ideje, da odaberu što žele da računalo učini s analiziranim tekstom, a sve to podrazumijeva samorefleksiju (Piper, 2018: 26).

Računalni programi koje digitalni humanisti koriste prilikom proučavanja književnosti često su već gotovi proizvodi, alati koji su dostupni na internetu, a u koje su implementirane računalne metode. Budući da su pitanja koja proučavatelji književnosti postavljaju vrlo raznolika, teško je da općeniti i alati namijenjeni širokoj upotrebi mogu na zadovoljavajući način poduprijeti analizu književnih tekstova. Na tome tragu Mathew L. Jockers (2013: 4) navodi da su računalne metode za tekstualnu analizu primitivne: „Lista frekventnosti riječi, konkordancije i lista ključnih riječi u kontekstu (KWIC) korisni su za određenu vrstu analiza, ali ovi zalogajčići teško mogu zadovoljiti apetit digitalnih humanista za više.”

Pitanje koje se postavlja jest trebaju li humanisti posjedovati znanje o programiranju i računalnim tehnologijama kako bi mogli aktivnije sudjelovati i/ili samostalno izrađivati alate koji su im potrebni za njihova istraživanja. Pisanje programskih skripti i razvijanje alata odvija se u obliku suradnje više stručnjaka iz različitih područja, što podupire ukorijenjenost digitalne humanistike u interdisciplinarnosti. Ako bi se proučavatelji književnosti bavili onime što spada u domenu tehničkih znanosti, ne bi se mogli dovoljno usredotočiti na sâmo tumačenje književnih tekstova i pojava. Jockers (2013: 17) uspoređuje metode prisutne u prirodnim i humanističkim znanostima (preciznije, znanosti o književnosti) i pritom dolazi do zaključka da bi humanističke metode (od kojih je u proučavanju književnosti najvažnije pomno čitanje) od tzv. tvrdih znanosti trebale preuzeti elemente pouzdanosti i testiranja.

## **5. Naratologija**

Unutar znanosti o književnosti postoje različite književne teorije koje na temelju svojih specifičnih postavki nastoje tumačiti književne tekstove. Jedna od njih je naratologija, najkraće definirana kao

opća teorija pripovjednoga teksta. Tvorac pojma je Tzvetan Todorov, francuski strukturalist (Grdešić, 2015). Naratologija se bavi svim elementima pripovjednoga teksta: pripovjedačem, vremenom, prostorom, likovima, ali i samim činom pripovijedanja, odnosom fabule i sižea, tj. priče i diskurza itd. Strukturalistička naratologija fluidnije je shvaćala pojam pripovijedanja te je tako naratologija počela proučavati i druge oblike u kojima se javlja određena pripovjedna struktura – usmenu književnost, kazališne predstave, filmove, stripove itd.

Budući da je fokus ovog rada na književnim likovima te se njihova analiza temelji na naratološkoj teoriji, u nastavku slijedi pregled određenih pristupa unutar naratologije koji su osnova za kasniju računalno potpomognutu analizu likova iz hrvatske književnosti.

### **5.1. Naratologija i književni likovi**

Za početak, što je (ili tko) književni lik? Fotis Jannidis (2012), kako je navedeno u Hühn i drugi (2014: 30), definira lika kao sudjelovatelja u svijetu priče, koja može nastati u različitim medijskim oblicima. Dakle, ova se definicija ne odnosi samo na likove koji su dio književnoga teksta, nego na bilo koje likove koji se nalaze unutar određene priče – filmske, kazališne, folklorne i sl. Kad god se pokušava definirati pojam lika, postavlja se pitanje je li lik isključivo proizvod jezika i jedino u njemu postoji, odnosno je li on diskurzni element, ili je lik entitet koji nastaje jezikom, ali može postojati i izvan teksta i traži od čitatelja da uključi znanje o ljudskim bićima. Rimmon-Kenan (2002), kako je navedeno u Grdešić (2015: 64), ističe dvije struje u teoriji likova – realističku i purističku. Realistička odgovara shvaćanju likova kao stvarnih ljudi, što znači da se mogu izdvojiti iz teksta i da njihovo postojanje nije vezano uz narativne događaje. Puristička strana poima likove kao jedan od motiva u tekstu, čime se naglašava njihova jezična i nereferencijalna narav.

Ako je lik sudionik u radnji, znači da je vezan uz događaje. Takvo razmišljanje ima svoje korijene još iz antičke Grčke, gdje je Aristotel u *Poetici* istaknuo prvenstvo radnje nad karakterima. Karakterizacija je u službi radnje, a likovi nisu nužan element teksta, kao što je slučaj s radnjom: „Fabula je, dakle, temelj i kao neka duša tragičkog umijeća, a karakteri su na drugome mjestu.” (2005: 17). Time su se vodili strukturalisti kad su u naratologiju uveli pojam aktera, naglašavajući da je lik nositelj radnje. Nadalje, Vladimir Propp i Algirdas Julien Greimas osmislili su aktancijalne modele kako bi kategorizirali likove na temelju radnji koje vrše. Proppov (1982), kako je navedeno u Grdešić (2015: 62), aktancijalni model (koji je razvijen pomoću ruskih bajki) sastoji

se od: zlikovca, davatelja, pomoćnika, princeze (tražene osobe) i njezina oca, pošiljatelja, junaka i lažnog junaka. Greimasov (1989), kako je navedeno u Grdešić (2015: 62), model sadrži subjekta, objekta, pošiljatelja, primatelja, pomoćnika i protivnika. Ono što se navodi kao zamjerka ovakvom pristupu likovima jest redukcija teksta na jedan model i pitanje što se događa u primjerima gdje su likovi toliko složeni da se ne može odrediti jasna funkcija ili u tekstovima koji gotovo da ni nemaju radnje, a radnja je u aktancijalnim modelima glavno određenje lika (Grdešić, 2015: 63).

Bez obzira poima li se lika realistički ili puristički, književni kritičari i teoretičari, ali i neprofesionalni čitatelji, zanimaju se za njihovo tumačenje. Likovi se žele klasificirati i okarakterizirati. Kategorizacija likova znači da postoje određene kategorije na temelju kojih se likovi smještaju u skupine. Budući da su književni likovi često vrlo kompleksni, skloni mijenjaju svojih obilježja tijekom radnje, ponekad ih je teško jednoznačno svrstati u jednu kategoriju ili je kategorija nedovoljno ekstenzivna. Jedan od najpoznatijih sustava kategorizacije fikcionalnih likova jest onaj E. M. Forstera koji nudi dvije kategorije: plošni (engl. *flat*) i zaokruženi (engl. *round*) likovi. Plošni likovi su jednostavniji, ne razvijaju se tijekom radnje već ostaju isti, dok su zaokruženi likovi kompleksniji, detaljnije su razrađeni. Jannidis (2012), kako je navedeno u Hühn i drugi (2014: 34), navodi Schneiderovu teoriju načina na koji se odvija proces kategorizacije lika. Može se ići odozgo prema dolje, što znači da već unaprijed postoji neka kategorija koja se zatim primjenjuje na lika. Druga mogućnost je da se ide odozdo prema gore što znači da se lik integrira u kategoriju, odnosno da se iz teksta povlače obilježja lika pomoću kojih se stvara kategorija. Maša Grdešić (2015: 64) pravi distinkciju između kategorizacije (autorica koristi pojam klasifikacije) i karakterizacije, gdje je kategorizacija analiza priče, a karakterizacija analiza teksta. U svojoj knjizi *Uvod u naratologiju* Grdešić (2015: 66), prije nego što navede elemente karakterizacije likova, osvrće se na vrste narativnih figura koje iznosi Gajo Peleš. Prema Pelešu, postoje tri narativne figure - psihemska (ili figura osobnosti), sociemska (ili figura skupnosti) i ontemska (ili figura opstojnosti). Psihemska narativna figura predstavlja zbroj svojstava osobnosti lika čiju jezgru čini vlastito ime. Sociemskom narativnom figurom ocrtava se odnos pojedinca i skupine kojoj pripada (najčešće se radi o obitelji, društvenoj klasi, zanimanju). Od svih navedenih, ontemska narativna figura je najopćenitija. Njome se označava opća tema djela koju lik utjelovljuje, tj. lik se promatra kao nositelj određene ideje. Pelešove narativne figure dobra su osnova na koju se mogu nadovezati postupci koje Grdešić (2015) ističe kao elemente karakterizacije likova a to su: izravna definicija, neizravna prezentacija i karakterizacija analogijom.

### 5.1.1. Elementi karakterizacije likova

Izravna definicija lika znači da pripovjedač ili drugi lik iznosi određena obilježja lika čime se može stvoriti slika o tome kakav je opisani lik. Kod ovakvog načina karakterizacije javlja se pitanje pouzdanosti jer različiti tipovi pripovjedača posjeduju različite razine pouzdanosti, a isto vrijedi i za likove koji su rijetko nepristrani dok govore o drugome liku (Grdešić, 2015: 69). Udaljavanjem od devetnaestostoljetne ideje romana i dolaskom modernizma izravna definicija kao postupak karakterizacije može čak imati i negativan predznak, kako ističe Grdešić (2015: 70). Sljedeći postupak karakterizacije jest neizravna prezentacija. Ona obuhvaća radnje (rutinske i nerutinske), govor, vanjski izgled i okolinu (Grdešić, 2015). Važnost radnji lika prepoznata je još u Aristotelovo vrijeme, što je bilo prethodno istaknuto. I rutinske i nerutinske radnje mogu pružiti dobar uvid u svojstva lika. Rutinske radnje pokazuju onu stalnu, relativno stabilnu i nepromjenjivu stranu lika, a naglašavanje rutine lika može biti okosnica njegove izgradnje. Nerutinske ili iznimne radnje također puno govore o liku. Vjerojatno najpoznatiji primjer koliko jedna iznimna radnja može obilježiti lika jest Raskoljnikov, glavni lik *Zločina i kazne*, koji ubojstvom stare lihvarice postaje prepoznatljiv i neprestano iznova tumačen lik u svjetskoj književnosti.

Vanjski izgled može se shvatiti kao nešto što lik nije u mogućnosti kontrolirati i to je točno za pojedine elemente kao što su boja očiju, visina, crte lica i sl. Međutim, dijelovi vanjskog izgleda kao što su odjeća, frizura ili šminka odaju osobine lika jer su to stvari na koje lik može svjesno utjecati. Nadalje, Roland Barthes (1992: 78), kako je navedeno u Grdešić (2015), govori o informantima. To su informacije pomoću kojih se tekst smješta u prostor i vrijeme, a kad je riječ o likovima u informante spada primjerice dob, spol, omiljena hrana, knjige, jutarnje navike, boja očiju itd. Još jedan element karakterizacije koji daje korisne informacije o liku je okolina - fizička ili ljudska (Grdešić, 2015: 78). Fizičko okruženje se uglavnom odnosi na prostor u kojem lik boravi – stan, kuća, priroda itd. Ljudsko okruženje je zapravo ono što Peleš naziva socijalskom narativnom figurom, dakle obitelj, društvena skupina, prijatelji, kolege na poslu i sl. Okolina je bila posebno važna i isticana u romantizmu (priroda kao prostor refleksije, zrcalo unutarnjih osjećaja) i realizmu (unutarnji prostori koji odražavaju duševno i ekonomsko stanje lika).

Posljednji oblik karakterizacije je karakterizacija analogijom u što Grdešić (2015: 80) svrstava: analogna imena, analogan krajolik i analogiju između likova. Imena često ukazuju na neku vanjsku ili unutarnju osobinu lika. To se može realizirati na razini forme i na razini značenja. Analogan krajolik je zapravo ono što je objašnjeno na primjeru romantizma i realizma, gdje je često prostor

odražavao kako se lik osjeća, što ga brine i sl. Analogija između likova naziva se još i recipročna karakterizacija, a označava trenutke u pripovjednom tekstu kada se dva ili više likova nađe u sličnoj situaciji u kojoj njihov način ponašanja (koji može biti sličan ili različit) naglašava osobine koje posjeduju (Grdešić, 2015: 83).

### **5.1.2. Aktancijalni model**

Algirdas Julien Greimas književni je teoretičar koji je u šezdesetim godinama prošlog stoljeća osmislio aktancijalni model. To je svojevrsni teorijski alat kojim se može analizirati bilo koja stvarna ili fikcionalna radnja (Hébert, 2011: 1). Ipak, model se pretežito primjenjuje na književne pripovjedne tekstove.

Aktancijalni model sastoji se od šest aktanata raspoređenih u tri opozicije, od kojih svaka predstavlja po jednu os. Aktanti su: subjekt, objekt, protivnik, pomagač, pošiljatelj i primatelj, a osi su sljedeće: os želje, os moći i os prijenosa. Na osi želje su subjekt i objekt, na osi moći protivnik i pomagač, a na osi prijenosa pošiljatelj i primatelj. Svaki od aktanata može se realizirati kroz konkretni lik u djelu, ali se može odnositi i na neke apstraktnije kategorije (primjeri će se moći vidjeti na tekstovima koji se analiziraju u ovome radu).

Subjekt je onaj ili ono čiji je predmet interesa objekt i često je glavni lik djela. Objekt je onaj ili ono za čime se žudi. To može biti neka ideja, cilj (nešto apstraktno), a može biti i neki lik (primjerice, u bajkama princ kao subjekt žudi za princezom kao objektom). Protivnik pak želi spriječiti subjekt da dođe do svojeg objekta. Pomagač, s druge strane, želi da subjekt dođe do objekta i zato mu pomaže u tome. Pošiljatelj je onaj/ono što daje objekt nekome (taj netko ne mora biti subjekt), dok je primatelj konačni dobivatelj objekta.

Na osi želje prikazuje se odnos između subjekta i objekta. Louis Hébert (2011: 1) navodi dvije mogućnosti ostvarivanja veze ova dva aktanta: može doći do konjunkcije (subjekt i objekt su povezani) ili disjunkcije (subjekt se uspješno rješava objekta). Pomagač, koji se s protivnikom nalazi na osi moći, pomaže u spajanju subjekta i objekta, tj. želi da dođe do konjunkcije. Protivnik pak ometa subjekta u dolaženju do objekta. Treća os je os prijenosa gdje pošiljatelj daje objekt s namjerom da se uspostavi veza između subjekta i objekta, a primatelj je element za koji se pronalazi objekt.

### 5.1.3. Primjena aktancijalnog modela

Kako funkcionira aktancijalni model u praksi, prikazat će se na književnim tekstovima koji se koriste u istraživačkome dijelu ovoga rada.

U *Šumi Striborovoj* autorice Ivane Brlić-Mažuranić mogu se izdvojiti dvije glavne narativne linije: ljubavni zaplet sina i snahe te protjerivanje stare majke iz šume. U prvoj narativnoj liniji postoje dva subjekta – stara majka i snaha – koje žude za objektom – sinom. Snaha ujedno ima i ulogu protivnika jer želi zapriječiti ostvarivanje veze između majke i sina. Kao pomagač javlja se lik ubogog djevojčeta, koji je i primatelj jer na kraju bajke ona i sin završe zajedno. Stara majka, osim što je subjekt, također je i pošiljatelj jer daje objekt (svog sina). U drugoj narativnoj liniji majka je ponovno subjekt, a ostanak u šumi objekt jer majka žudi za životom u čarobnoj šumi, sa svojim sinom. U tome joj pomažu Domaći (posebice lik Malika Tintilinića) i Stribor. I ovdje je protivnik snaha koja nahuška supruga da istjera staricu iz kuće. Šuma kao čaroban i idealan svijet je pošiljatelj, a selo kao simbol stvarnosti primatelj.

Novela *Oprava* Josipa Kozarca ima jednu glavnu narativnu liniju, stoga je aktancijalni model nešto jednostavniji nego u *Šumi Striborovoj*. Subjekt je Đuro Vučetić, a objekt njegova supruga Jelena koju on želi zadržati za sebe. Tajnik Barić javlja se kao protivnik, kao i sin Dušan. Međutim, ova dvojica likova nisu tipičan primjer protivnika jer ne djeluju izravno u sprječavanju subjekta da dobije objekt, već neposredno utječu na subjektovu (Đurinu) težnju da zadrži vlastitu suprugu. Barić to čini time što se bio upustio u preljub s Jelenom, a Dušan je plod te nevjere i time, kako sam subjekt smatra, neprijatelj u odnosu supružnika. Pomagač zapravo ne postoji jer sve okolnosti idu nauštrb glavnog lika Đure. Kao pošiljatelj bi se mogla shvatiti Đurina reproduktivna mana (neplodnost) koja je „gurnula” objekt (Jelenu) u naručje drugog muškarca. Time Barić postaje i primatelj objekta.

Putopis *Oko Lobora* Antuna Gustava Matoša, s obzirom na to da ima samo jednog lika (sam pripovjedač), aktante ostvaruje preko nešto apstraktnijih elemenata. Pripovjedač je istovremeno i subjekt i primatelj i pošiljatelj. Objekt je put na koji se pripovjedač otisnuo. Dvojnu ulogu pomagača i protivnika imala bi prisjećanja i promišljanja pripovjedača o krajevima kroz koje prolazi jer mu s jedne strane pomažu da nastavi put, a s druge mu odmažu jer je putnik zaokupljen refleksijama o minulim vremenima.

Kao i *Šuma Striborova*, pripovijetka *Krvavi most u Zagrebu* Mirka Bogovića ima dvije narativne linije: borba za Ružičinu ljubav i sukob Garduna i biskupa Mihalja oko Medved- i Garić-grad. Budući da je pripovijetka književna vrsta koja se svojom složenošću i opsegom nalazi između novele i romana, Bogovićev tekst ima dvije narativne linije, razgranatiju fabulu i više likova u odnosu na novelu *Oprava* koja se koncentrira na jednu radnju i praktički jednog lika (Đuru Vučetića). U prvoj narativnoj liniji subjekt je Pavao Slavinić čiji predmet žudnje (dakle objekt) je Ružica. Protivnik mu je Gardun jer mu on sprječava da ostvari vezu s Ružicom, ali je Gardun istovremeno i subjekt jer također žudi za objektom. Pošiljatelji su Ružičina braća Ivan i Hinko Gusingovac koji daju svoju sestru namijenjenom primatelju, a to je Gardun. Međutim, objekt na kraju završava u rukama Pavla Slavinića, stoga je on stvarni primatelj. Kćerka Margitica također bi se mogla smatrati primateljem jer uživa plodove ljubavi svojih roditelja, Pavla i Ružice. Druga narativna linija za subjekta ima Garduna, čiji je objekt preuzimanje vlasti nad oduzetim teritorijem. Protivnik mu je biskup Mihalj, ali i Pavao Slavinić koji se hrabro uz njega bori, a pomagači su mu braća Gusingovci. Osveta prema biskupu Mihalju mogla bi se kategorizirati kao pošiljatelj jer to je ono što dovodi do žudnje za preuzimanjem vlasti nad oduzetim teritorijem. Međutim, Gardunova želja ostaje neostvarena jer primatelj postaje biskup Mihalj, budući da je on krajnji pobjednik sukoba.

Najkompleksniji aktancijalni model nalazi se u romanu *Posljednji Stipančići* Vjenceslava Novaka. Mogu se pratiti tri narativne linije: propast patricijske obitelji Stipančić, Lucijin položaj u obitelji i ljubavni odnos između Lucije i Alfreda. Prva narativna linija usko je vezana uz likove koji čine društveno-politički svijet romana. Subjekt je Ante Stipančić, glava obitelji, a svi njezini članovi (Ante, Valpurga, Lucija i Juraj) su objekti jer Ante želi da njegova obitelj ide u točno određenom smjeru. Konkretnih pomagača i protivnika u obliku likova nema jer splet društveno-političkih okolnosti utječe na sudbinu obitelji Stipančić. Hrvatski politički interesi bili bi pošiljatelj jer oni „guraju” obitelj Stipančić prema primatelju, a to su strani politički interesi. Kao akter pošiljatelja funkcionira lik kanonika Vukasovića, dok se kao akteri primatelja javljaju major Benetti, Fabijan Martinčić i Čolić. Druga narativna linija tiče se Lucijinog nepovoljnog položaja u obitelji. Ona je subjekt, a njezin objekt je želja za školovanjem i slobodom u odlučivanju. Majka Valpurga glavni joj je pomagač jer pokušava uvjeriti Antu da bi se i Luciji trebalo dopustiti da se obrazuje. Lucijini protivnici su Ante i Juraj – otac ima aktivnije i neposrednije djelovanje i ovdje isto sprječava Lucijino stvaranje veze s objektom, a Juraj je posredni protivnik jer se samo njemu daju financijska

sredstva i mogućnosti koje on na kraju ni ne iskorištava na ispravan način. Pošiljatelj bi bio Ante Stipančić, a primatelj sin Juraj jer je on taj koji dobiva mogućnost obrazovanja i slobode. U trećoj narativnoj liniji (koja se odnosi na ljubavni dio radnje) subjekt je Lucija, a objekt Jurjev prijatelj Alfred. Ponovno je majka ona koja ima ulogu pomagača i nastoji da se ostvari veza između Lucije i Alfreda. U tome joj posredno pomaže i Martin Tintor koji piše pisma u Alfredovo ime. Ono što sprječava ostvarivanje ljubavne veze jesu Alfredove neiskrene namjere i njegovo zavlacenje Lucije. Kao pošiljatelj objekta mogao bi se shvatiti Juraj jer je on taj koji je doveo Alfreda u kuću, a primatelj je „sretan brak” koji je sklopio jer se time nije razvio ljubavni odnos kako je subjekt (Lucija) priželjkivao.

## 6. Povezana istraživanja

Postoje mnogi elementi književnog teksta koji se mogu analizirati računalnom metodom – prostor, vrijeme, forma, stil itd. Naravno, što će se proučavati ovisi o književnom rodu kojem tekst pripada. Epika, lirika i drama imaju svoje specifične zakonitosti kojima se stvara okvir za potencijalne hipoteze koje se mogu ispitati. Ovaj rad bavi se analizom pripovjednih prozih tekstova, stoga su kao glavni književni element koji se proučava odabrani likovi. Unutar digitalne humanistike postoje slična istraživanja na tu temu, a ovo poglavlje donosi kratki pregled nekih od njih.

Hardik Vala i drugi (2015: 1) u svome radu na početku ističu jedan problem kad se radi računalno potpomognuta analiza likova, a to je prevlast upotrebe NER (engl. *Named Entity Recognition*) alata. Naime, kako ističu autori, korištenjem isključivo NER-a ne dolazi se do detaljne i ekstenzivne slike likova u književnom tekstu. Razlog tomu je što postoje mnogi neimenovani likovi ili likovi koje se ne oslovljava imenom i/ili prezimenom, a bitni su za građenje društvenog konteksta i osvjetljavanja imenovanih (najčešće glavnih) likova (Vala, Jurgens i dr., 2015: 1). Problemu bi se moglo doskočiti upotrebom univerzalnih ovisnosti (engl. *universal dependencies*) koje prave gramatičku anotaciju uzimajući u obzir nadrećenične odnose (Mrežnik, 2021). U svome istraživanju autori su pokušali detektirati što više inačica naziva likova kako bi se stvorila potpunija slika. Pritom su različitim pravilima nastojali povezati različite nazive i izraze koji se zapravo odnose na jedan lik. Kao što i navode u radu, njihovo istraživanje fokusiralo se na detekciju likova, ne i opsežniju književnoteorijsku analizu (Vala, Jurgens i dr., 2015: 4). Valja napomenuti da su u istraživanju koristili prozne tekstove, točnije 58 romana.



Joseph Bullard i Cecilia Ovesdotter Alm (2014) svoje su istraživanje proveli na dramskim tekstovima Augusta Strindberga i Henrika Ibsena. Analizirali su likove na temelju tri kategorije: spol, dob (stari nasuprot mladi) i društveno-ekonomski položaj (srednja klasa naspram niže klase). Autori su se fokusirali na stvaranje označenog skupa podataka i računalnog modela kojim bi se predvidjele društveno-biografske značajke književnih likova u dramama (Bullard i Ovesdotter Alm, 2014: 1).

Još jedan zanimljiv pristup književnim likovima jest onaj Lucie Flekove i Iryne Gurevych (2015) koje se služe praksama iz psihologije, preciznije petfaktorskim modelom osobnosti. Autorice pomoću NLP-a (engl. *Natural Language Processing*) i profiliranja osobnosti (engl. *personality profiling*) predviđaju osobnosti protagonista u književnim tekstovima, odnosno romanima. Nadalje, konstruiranje osobnosti likova povezuje se s osobnostima čitatelja – autorice postavljaju hipotezu kako čitatelji preferiraju likove koji su im karakterno slični (Flekova i Gurevych, 2015: 1). Flekova i Gurevych (2015: 2) temeljile su istraživanje na analizi triju elemenata: izravan govor, radnje lika te predikati i prilozi. Svaki element imao je svoj odvojeni sustav koji se trebao napraviti. Još jedno istraživanje koje u sebi ima psihologijski aspekt jest ono Arthura M. Jacobsa (2019) koji radi analizu sentimenta (engl. *sentiment analysis*) likova iz knjiga o Harryju Potteru autorice J. K. Rowling. Autor se također oslanja na teoriju velikih pet dimenzija ličnosti i zanima se za razinu dopadljivosti likova čitateljima.

## 7. Obrada prirodnog jezika

Prirodni jezik je svaki oblik jezika koji se kod čovjeka razvio na prirodan način kroz njegovu upotrebu i služi kao sredstvo komunikacije. Prirodni jezik ne podrazumijeva samo pisani tekst, nego i usmeni govor. Pojam oprečan prirodnom jeziku bio bi formalni jezik čiji je nastanak unaprijed planiran, a nije se razvio spontano kao što je slučaj s prirodnim jezikom. Primjer formalnog jezika bili bi programski jezici. Osnovna razlika između ove dvije kategorije jezika je višeznačnost. Prirodni jezik je promjenjiv, fleksibilan i „živa tvar”, čime se želi reći da se riječima neprestano mijenjanju i pridodaju nova značenja, da jedna riječ može nositi više od jednog značenja. Višeznačnost se ostvaruje i prenesenim značenjima, figurama ironije, paradoksa itd. S druge strane, formalni jezik odlikuje se jednoznačnošću, što proizlazi iz njegove stroge sintakse i pravila slaganja elemenata te korištenja simbola i čestog oslanjanja na matematičke i logičke zakonitosti.

Obrada prirodnog jezika ne bavi se samo tekstem jer bi to značilo da se analizira niz znakova (jer je tekst u svojoj biti zapravo sekvenca znakova). Obrada jezika je obrada niza znakova koji nose jezičnu poruku, dakle koji imaju neko značenje i to značenje je bitno za analizu. Lucja M. Iwanska i Stuart C. Shapiro (2000: 8) kažu da je prirodni jezik reprezentacijski sustav znanja koji se može računalno dobro kontrolirati zbog svoje strukturiranosti i bliske veze sintakse i semantike. Zašto je onda proučavanje prirodnog jezika toliko zahtjevno, posebice onog koji se realizira u književnim tekstovima?

Alati koji se koriste u obradi prirodnog jezika većinom su nastali tako što su trenirani na elektronički najdostupnijim tekstovima – internetskim člancima, stranicama, blogovima i sl. Riječ je o suvremenim tekstovima koji imaju određene zakonitosti oblikovanja, vokabular koji upotrebljavaju, način strukturiranja rečenica. Ostale vrste tekstova kao što su znanstveni radovi, povijesni tekstovi (u koje se mogu svrstati i nesuvremeni književni tekstovi), stručne knjige, puno su manje zastupljeni u korpusu kojim se trenira određeni alat za obradu teksta. Budući da vrijeme nastanka i vrsta teksta (ili žanr) utječu na to kako se alat razvija, za očekivati je da je književne tekstove teže obraditi jer se prilikom obrade koriste alati koji su namijenjeni analizi drugih tekstova. Jedan od glavnih problema koje ističe Michael Piotrowski (2012: 3) jest pitanje standardizirane ortografije. Uzme li se za primjer hrvatski jezik, može se uočiti da se tijekom povijesti njegov pravopis mijenjao i da i u današnje vrijeme postoje neriješena pitanja vezana uz hrvatsku ortografiju. Sve te mijene otežavaju obradu prirodnog jezika. Isto tako, za razliku od formalnog jezika, prirodni jezik ima fluidnu sintaksu koja otežava stvaranje strogih i nepromjenjivih pravila za računalnu obradu. Tijekom same analize povijesnih tekstova poteškoće se mogu javiti i zbog promjene u značenju riječi, drugačijem kontekstu u kojem se određena riječ koristila, i naravno, kako se zapisivala. Kada se govori u kontekstu književnih tekstova, implicitni elementi, metaforičko izražavanje i višeznačnost dodatno otežavaju obradu i tumačenje.

## **7.1. Osnovni pojmovi i postupci**

Kako bi se jednostavnije pratio istraživački dio ovoga rada u kojem se koriste termini važni za obradu prirodnog jezika, slijede kratka obrazloženja korištenih pojmova.

1. *Pojavnica*. *Hrvatski mrežni rječnik – Mrežnik* (2021) definira pojavnicu na sljedeći način: „sve što se nalazi između dva znaka koja služe kao graničnici (svako individualno pojavljivanje); svaka pojava jezične jedinice u korpusu, na razini riječi svaki oblik uključen u leksem”. Kad

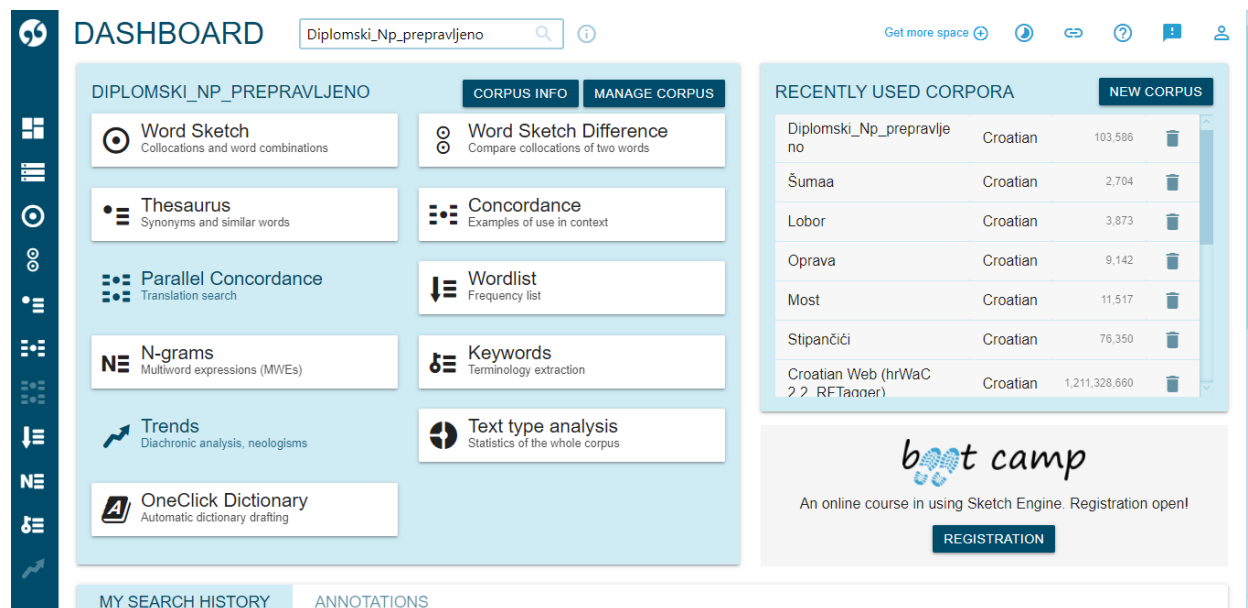
se govori o veličini jezičnog korpusa, redovito se govori o broju pojavnica, a ne riječi jer u pojavnice spadaju i interpunkcijski znakovi.

2. *Različnica*. „Pojedinačna riječ koja se razlikuje od druge riječi” (*Mrežnik*, 2021). Različnica se dakle odnosi na jedinstveno pojavljivanje riječi, čime se razlikuje od pojavnice koja predstavlja svako pojavljivanje riječi. Na primjer, izraz „jedan po jedan” ima tri pojavnice i dvije različnice jer riječ „jedan” se pojavljuje dva puta u istome obliku.
3. *Lema i lematizacija*. Postupak određivanja leme svakoj pojavnici u tekstu zove se lematizacija, a lema je natuknički (kanonski) oblik riječi (Tadić, 2003: 32). Često je informacija o vrsti riječi bitna za pravilno određivanje leme. Primjerice, hrvatska riječ „oko” može po vrsti biti prijedlog, a može biti i imenica (oko kao organ vida). To uvelike mijenja daljnje označavanje.
4. *Vertikalizacija*. Potupak razdvajanja svake pojavnice teksta i smještanje u zaseban red. Svi slojevi informacije vezani uz neku pojavnicu dodaju se u isti red, zatim se određuje koji stupac nosi koju informaciju što omogućava nadopunjavanje korpusa i širenje dodatnim slojevima informacija. Na primjer, POS oznake temelj su razvijanja MSD označivača, a pomoću MSD oznaka dolazi do nastanka NERC označivača. Dakle, sustav se proširuje i uči na prethodnim podacima.
5. *Označavanje vrsta riječi*. Jedan od osnovnih postupaka u obradi prirodnog jezika je POS (engl. *Part-of-speech*) označavanje. Tim se postupkom označuju vrste riječi na način da se svakoj pojavnici u tekstu pridružuje informacija o kojoj vrsti riječi se radi (Tadić, 2003: 32). Alat kojim se označavanje vrši naziva se POS označivač (engl. *Part-of-speech tagger*). Kao što je ranije spomenuto, alati za obradu prirodnog jezika većinski su trenirani na suvremenim tekstovima i zato je i POS označavanje zahtjevnije izvoditi na povijesnim (književnim) tekstovima. Ipak, POS označavanje može dati korisne informacije za književne tekstove jer pruža uvid u raznolikost vrsta riječi, s obzirom na književnu vrstu ili žanr, period u kojem je nastalo djelo, autora koji ga je napisao (što bi se moglo iskoristiti za pripisivanje autorstva), ali i na općenitoj razini jer neki književni tekstovi sigurno u većoj količini koriste neke vrste riječi, nego što je to uobičajeno u ostalim tipovima neknjiževnih tekstova.
6. *Morfosintaktičko označavanje*. Još jedan postupak sličan označavanju vrsta riječi je morfosintaktičko (MSD) označavanje. Razlika se krije u tome što MSD označavanje ide dublje u analizu riječi, odnosno daje više informacija o riječima – uz vrstu riječi, dobije se i podatak

o rodu, broju, padežu, licu i sl. To je posebno korisno za hrvatski budući da je on flektivni jezik.

## 7.2. Sketch Engine

*Sketch Engine* je internetski dostupan program koji služi istraživanju načina funkcioniranja jezika (*Sketch Engine*, 2021). Nudi različite alate za više od 90 jezika, uključujući i hrvatski. *Sketch Engine* alati dostupni za hrvatski jezik su: skica riječi (*Word Sketch*), usporedba skica riječi (*Word Sketch Difference*), tezaurus (*Thesaurus*), konkordancija (*Concordance*), lista riječi (*Wordlist*), n-gramovi (*N-grams*) i ključne riječi (*Keywords*). Programom se služe leksikografi, prevoditelji, lingvisti, studenti i svi koji se bave nekim oblikom proučavanja jezika. Korisnici imaju pristup korpusima, a za hrvatski jezik dostupni su sljedeći korpusi: *Croatian Web (hrWaC)*, *CHILDES Croatian Corpus*, *DGT*, *EUR-Lex judgements Croatian*, *OPUS2 Croatian*, *Riznica v0.1*, *Croatian Medical Corpus*. Osim unaprijed dostupnih korpusa za pojedini jezik, program omogućuje da korisnik stvori svoj korpus koristeći se vlastitim tekstovima. Slika 2 prikazuje korisničko sučelje *Sketch Enginea*.



Slika 2. Sučelje internetskog programa *Sketch Engine*

U nastavku će se objasniti opcije koje program nudi a koje su bile korištene prilikom istraživačkog dijela ovoga rada.

Opcija koja je najviše koristila za istraživanje korpusa jesu konkordancije, a pojmovnik Hrvatskog mrežnog rječnika (2021) definira ih na sljedeći način: „konkordancija (engl. *concordance*) – popis riječi sa svim oblicima u kojima se pojavljuju s kontekstom i oznakom izvora, koji se nalaze u nekome korpusu.” Dakle, unese se upit (u ovome radu bila je riječ o CQL<sup>1</sup> izrazima) i program pronađe sve primjere koji odgovaraju zadanom upitu.

Jedna od mogućnosti dodatnog sortiranja dobivenih konkordancija je GDEX (*Good Dictionary EXamples*). Kao što sâm naziv govori, GDEX se koristi kako bi se pronašli dobri primjeri upotrebe riječi (čije konkordancije se traže) koji mogu poslužiti za učenje jezika ili kao rječnički primjeri. Rečenice se vrednuju na temelju duljine, vokabulara, prikladnosti konteksta, zamjenica koje povezuju rečenicu s okružujućim tekstom i sl. (*Sketch Engine*, 2021). Postoji generički GDEX sustav, ali ono što je važno je da se sustav može specifično konfigurirati za svaki jezik, dakle ovisan je o jeziku.

Lista riječi (*Wordlist*) još je jedna opcija unutar programa *Sketch Engine*. Ona stvara frekvencijsku listu lema, vrsta riječi, POS ili MSD oznaka i sl. Moguće je definirati minimalnu i maksimalnu frekvenciju, a u istraživačkome dijelu ovoga rada lista riječi se koristila u svrhu dobivanja frekvencije likova koji su dio aktancijalnih modela i kako bi se dobio broj pojavljivanja likova koji su se detaljnije analizirali.

---

<sup>1</sup> CQL (*Corpus Query Language*) je oblik izraza kojim se mogu pretraživati kompleksni gramatički uzorci u korpusu ili nešto što se ne može pretražiti jednostavnim korisničkim upitom (*Sketch Engine*, 2021). Tako se mogu raditi pretraživanja pomoću POS ili MSD oznaka, lematiziranog oblika riječi, također se mogu stvarati upiti kojima se traže izrazi od većeg broja riječi.

## 8. Istraživanje likova u korpusu odabranih hrvatskih književnih tekstova

Kao dio istraživačkog dijela ovoga rada, napravljeno je istraživanje o klasifikaciji i karakterizaciji likova u pet proznih tekstova hrvatske književnosti. Istraživanje se sastoji od kombiniranja naratoloških teorijskih postavki i metoda i alata koji se koriste u obradi prirodnog jezika. Korpus se sastoji od sljedećih književnih tekstova: *Šuma Striborova* Ivane Brlić-Mažuranić, *Oprava* Josipa Kozarca, *Oko Lobora* Antuna Gustava Matoša, *Krvavi most u Zagrebu* Mirka Bogovića i *Posljednji Stipančići* Vjenceslava Novaka. Podaci o veličini korpusa dostupni su u Tablici 1.

Tablica 1. Veličina korpusa

	<i>Šuma Striborova</i>	<i>Oprava</i>	<i>Oko Lobora</i>	<i>Krvavi most u Zagrebu</i>	<i>Posljednji Stipančići</i>	ukupan korpus
broj pojava (postotak u ukupnom korpusu)	3421 (2,83 %)	10696 (8,86 %)	4656 (3,86 %)	13046 (10,80 %)	88948 (73,65 %)	120767

Izabrani su tekstovi koji su napisani na hrvatskom jeziku jer to omogućava da se analiziraju tekstovi u svom izvornom obliku, da se ne radi s književnim prijevodima. Poznato je da svaki prijevod, bilo kojeg teksta, a posebno književnog, neminovno podrazumijeva određena odstupanja od svoga originala. Budući da je u književnosti bitno ne samo što je napisano, nego i kako je napisano, rad s tekstovima na izvornom jeziku činilo se kao najbolji izbor. Nadalje, postoji malo znanstvenih radova i istraživanja koji se bave računalno potpomognutom analizom hrvatskih tekstova koji su ujedno i književna djela, stoga je ovaj rad doprinos tom trenutno slabo razvijenom području digitalne filologije. Svaki od pet odabranih književnih tekstova pripada različitoj književnoj vrsti – bajci (*Šuma Striborova*), noveli (*Oprava*), putopisu (*Oko Lobora*), pripovijetki (*Krvavi most u Zagrebu*) i romanu (*Posljednji Stipančići*). Raznolikošću pripovjednih žanrova htjelo se ispitati mogu li se uočiti razlike između načina prikaza likova, nameće li žanr neke zakonitosti po tom pitanju. U istraživačkom dijelu rada fokus je na naratološkoj analizi likova.

Glavna hipoteza je da se računalnom potpomognutom analizom mogu dobiti kvalitetni podaci o određenim naratološkim elementima karakterizacije likova. Istraživanje polazi od Greimasovog aktancijalnog modela klasifikacije likova i naratoloških elemenata karakterizacije. Pritom treba naglasiti da su izabrani oni elementi karakterizacije za koje se procijenilo da ih je u ovoj početnoj fazi najlakše računalno identificirati.

Za istraživanje se koristi internetski alat *Sketch Engine*<sup>2</sup>, a tekstovi su preuzeti s internetske stranice *e-lektire*<sup>3</sup>. Najprije su otvoreni u internetskom formatu kako bi se mogli kopirati u tekstualne datoteke koje su upotrebljavane u nastavku istraživanja.

## 8.1. Metode

Istraživanje književnog korpusa odvijalo se pomoću statističkih metoda i onih koje su specifične za obradu prirodnog jezika. Nakon što su se tekstovi u internetskom formatu kopirali u tekstualne datoteke, napravila se normalizacija. Normalizacija tekstova podrazumijeva da se izbace svi izvantekstualni podaci koji nisu dio „čistog” književnog teksta, već su dodani kao dio nakladničkog uređivanja. Tako su se trebale ukloniti fusnote, nazivi tekstova i imena njihovih autora, rječnici manje poznatih riječi. Sljedeće što je trebalo napraviti je vertikalizirati tekstove i svakoj pojavnici pripisati odgovarajuću morfosintaktičku (MSD) oznaku. Vertikalizacija i MSD označavanje napravilo se pomoću CLARIN.SI<sup>4</sup> internetske usluge i u alatu *Sketch Engine* koji taj proces čini za svaki korpus koji se u njemu stvori. Korištenjem dva internetska alata bilo je moguće učiniti usporedbu rezultata. Zatim je uslijedila provjera točnosti klasifikatorâ. Uzeo se nasumičan segment jezično najkompleksnijeg književnog teksta koji je dio korpusa i ručno se označavao, a potom su se ručne oznake usporedile s onima *ReLDI* i *Sketch Engine* klasifikatora.

Sljedeći korak bio je ručno ispravljanje onih pogrešaka koje su se češće ponavljale i onih koje su bile vezane uz nazive likova. Ispravljeni vertikalizirani korpus potom se postavio na *Sketch Engine* kako bi se mogli poduzeti daljnji koraci u analizi.

Izračunat je udio osobnih zamjenica koje se odnose na književne likove u nasumičnom uzorku osobnih zamjenica iz svakog teksta koji se dobio opcijom *random sample* u alatu za konkordancije,

---

<sup>2</sup> Dostupno na: <https://www.sketchengine.eu/>

<sup>3</sup> Dostupno na: <https://lektire.skole.hr/>

<sup>4</sup> Dostupno na: <http://clarin.si/services/web/query>

a isti postupak je napravljen za udio osobnih imena likova u nasumičnom uzorku osobnih imenica. Naredni koraci odnosili su se na upotrebu različitih mogućnosti u *Sketch Engine* kako bi se mogla napraviti karakterizacija likova i analiza aktanata. Vodeći se razrađenim aktancijalnim modelima za svaki tekst odabrani su sljedeći likovi koji su se detaljnije analizirali: majka (baka) i Domaći (te Malik Tintilinić kao glavni predstavnik te skupine) iz *Šume Striborove*, Đuro Vučetić iz *Oprave*, Gardun i Pavao Slavinić iz *Krvavog mosta u Zagrebu*, Lucija Stipančić i kanonik Vukasović iz *Posljednjih Stipančića* te pripovjedač (koji je ujedno glavni i jedini lik) iz *Oko Lobora*. Za njihovo tumačenje koristile su se konkordancije, GDEX, lista riječi te distribucija kroz tekst.

Prvo se napravila analiza rada GDEX opcije na odabranim likovima. Gledalo se na koji način GDEX vrednuje, tj. rangira određene primjere rečenica s obzirom na to koliko su one informativne za karakterizaciju likova, posebice kao izravne definicije likova. Uz to, napravljena je i analiza primjera neizravnih prezentacija s naglaskom na glagole koji se pojavljuju uz određeni lik, a do njih se došlo analizom GDEX-a. Nadalje, zadavanjem CQL izraza dobile su se rečenice koje odražavaju povezanost odabranog lika s ostalim likovima u tekstu. Na razini svih likova koji su dio aktancijalnih modela izvršena je distribucija ostvarenih aktancijalnih uloga kroz tekst (koliko se određena uloga pojavljuje u određenim dijelovima teksta – na početku, u sredini i na kraju), listom riječi se dobio prikaz koji su likovi više, a koji manje zastupljeni.

## 8.2. Rezultati

Provjeravanje točnosti *ReLDI* i *Sketch Engine* klasifikatora dalo je različite rezultate. Službena internetska stranica *CLARIN.SI*-a navodi podatak da točnost klasifikatora za hrvatski jezik iznosi 92,53 % posto (Kasunić, 2019). Točnost klasifikatora na razini vrsta riječi za književni korpus koji se koristi u ovome radu izračunala se na uzorku od 603 pojavnica (uvršteni su i interpunkcijski znakovi). Ručnim provjeravanjem došlo se do točnosti od 90,38 %. *Sketch Engine* klasifikator dao je nešto slabiji rezultat. Koristio se identičan uzorak od 603 pojavnice, a točnost je iznosila 88,06 %. Budući da se istraživanje fokusiralo na književne likove, napravljene su matrice pogreške za vlastite imenice, osobne zamjenice i glagole (zasebno za svaki rod – muški, ženski i srednji). U Tablici 1 mogu se vidjeti rezultati koji su dobiveni za *ReLDI* klasifikator, a Tablica 2 prikazuje matricu pogreške za klasifikator koji koristi *Sketch Engine*.



Tablica 2. Matrica pogreške za *ReLDI* klasifikator

<i>ReLDI</i> oznaka	Točna oznaka					
		Np	Pp	V(m)	V(f)	V(n)
	Np	21	0	0	0	0
	Pp	0	17	0	0	0
	V(m)	0	0	14	0	0
	V(f)	0	0	0	6	0
	V(n)	0	0	0	0	6
	Nc	1	0	0	0	0
	A	1	1	1	0	0

Legenda:

Np – vlastite imenice

Pp – osobne zamjenice

V (m) – glagoli u muškom rodu

V(f) – glagoli u ženskom rodu

V(n) – glagoli u srednjem rodu

Nc – opće imenice

A – pridjevi<sup>5</sup>

Matrica pogreške pokazuje da su dvije vlastite imenice označene pogrešnom oznakom za vrstu riječi („Omiri” su označeni kao opća imenica, a „Ugarska” kao pridjev). Sve osobne zamjenice ispravno su označene, kao i glagoli u ženskom i srednjem rodu. Samo je jedan glagol u muškom rodu označen pogrešno – „slušo” je *ReLDI* klasifikator identificirao kao pridjev.

Tablica 3. Matrica pogreške za *Sketch Engine* klasifikator

---

<sup>5</sup> MULTEXT-East oznake dostupne na: <http://nlp.ffzg.hr/data/tagging/msd-hr.html>

Sketch Engine oznaka	Točna oznaka					
		Np	Pp	V(m)	V(f)	V(n)
	Np	22	0	0	0	0
	Pp	0	18	0	0	0
	V(m)	0	0	14	0	0
	V(f)	0	0	0	5	0
	V(n)	0	0	0	1	5
	Nc	0	0	0	0	0
	A	1	0	0	0	0
	R	0	0	1	0	0
Q	0	0	0	0	1	

Legenda:

Np – vlastite imenice

Pp – osobne zamjenice

V (m) – glagoli u muškom rodu

V(f) – glagoli u ženskom rodu

V(n) – glagoli u srednjem rodu

Nc – opće imenice

A – pridjevi

R – prilozi

Q – čestice

Klasifikator *Sketch Enginea* pogrešno je označio jednu vlastitu imenicu. Radi se o imenici „Ugarska” koja je (kao i kod prethodnog klasifikatora) označena kao pridjev. Ponovno su sve osobne zamjenice točno označene. Kod glagola je za svaki rod jedna riječ neispravno označena: „slušo” je dobio oznaku priloga, „nastojala” glagola srednjeg roda i „bilo” čestice.

Nakon što se ispitala točnost klasifikatorâ, izračunao se udio likova u nasumičnom uzorku vlastitih imenica. Za svaki od pet tekstova u *Sketch Engineu* generiralo se 100 nasumičnih vlastitih imenica

opcijom *random sample*. Valja napomenuti da se nisu gledali samo likovi koji su uvršteni u naratološku analizu, nego svi likovi u tekstu, a isto je napravljeno i za zamjenice.

Bajka *Šuma Striborova* na uzorku od 100 vlastitih imenica imala je 88 imenica (88 %) koje su imena likova (međutim, treba uzeti u obzir da su imena likova ručno ispravljena budući da se zapravo radi o općim imenicama koje je klasifikator početno tako i označio – npr. „baka”, „snaha”, „djevojče”). Novela *Oprava* imala je udio od 77 % imena likova u vlastitim imenicama, pripovijetka *Krvavi most u Zagrebu* 41 %, putopis *Oko Lobora* 0 % (s obzirom na to da postoji samo jedan lik koji je ujedno i pripovjedač), a *Posljednji Stipančići* 84 %. Osim što su izračunati udjeli, gledalo se na što se odnose ostale vlastite imenice, stoga se u Tablici 3 za svaki tekst može vidjeti u kojim kategorijama su se pojavile vlastite imenice, a da to nisu samo imena likova.

Tablica 4. Broj različitih kategorija vlastitih imenica

	<i>Šuma Striborova</i>	<i>Oprava</i>	<i>Krvavi most u Zagrebu</i>	<i>Oko Lobora</i>	<i>Posljednji Stipančići</i>
Np (ukupno u tekstu)	188	117	543	154	3001
Np (uzorak)	100	100	100	100	100
Np (likovi)	88	77	41	0	84
Np (lokacije)	0	8	25	42	8
Np (povijesne osobe)	0	2	20	41	5
Np (narodi)	0	0	1	6	1
Np (institucije)	0	0	0	1	0
Np (ostalo)	7	1	1	6	0
Pogrešno označene riječi	5	12	12	4	2

Sličan postupak napravljen je i za osobne zamjenice – uzeo se nasumičan uzorak od 101 osobne zamjenice za svaki tekst, izuzev *Oko Lobora* jer u njemu ima samo 96 osobnih zamjenica (zato je i uzet broj 101 kako bi ukupan broj iznosio 500). Postotak osobnih zamjenica koje se odnose na likove u *Šumi Striborovoj* iznosi 86,14 %, u *Opravi* također 86,14 %, u *Krvavom mostu u Zagrebu* 87,13 %, u *Oko Lobora* 16,67 % i u *Posljednjim Stipančićima* 89,11 %. Ukupni brojevi osobnih zamjenica u cijelim tekstovima su sljedeći: *Šuma Striborova* – 136 (3,98 % ukupnog broja pojava u tom tekstu), *Oprava* – 730 (6,82 %), *Krvavi most u Zagrebu* – 391 (3 %), *Oko Lobora* – 96 (2,06 %), *Posljednji Stipančići* – 3870 (4,35 %).

Rezultati opcije GDEX u *Sketch Engineu*, koji su vidljivi u Tablici 4, pokazuju koliko je za svaki lik bilo ukupno pronađeno primjera i ako je taj broj bio manji od 100, analizirali su se svi primjeri, a u slučaju lika Lucije Stipančić i kanonika Vukasovića te su brojke bile veće, stoga je uzeto prvih 100 primjera za svaki lik. U tablici se također vidi koliko izravnih definicija je pronađeno te koji GDEX rezultat im je pripisan (moguće su vrijednosti u rasponu od 0 do 1). Valja spomenuti da za putopis *Oko Lobora* GDEX nije funkcionirao jer ne postoji konkretan naziv lika, već se morala gledati osobna zamjenica „ja” za koju GDEX nije dao nijedan primjer.

Tablica 4. GDEX izravne definicije odabranih likova

	baka (majka)	Domaći	Malik Tintilinić	Lucija Stipančić	Vukasović	Đuro Vučetić	Pavao Slavinić	Gardun
broj GDEX primjera	66	18	11	463 (uzorak: 100)	196 (uzorak: 100)	34	56	78
broj izravnih definicija	3	1	0	2	12	3	5	6
GDEX vrednovanj	0,425; 0,374;	0,149	/	0,449; 0,339	0,875; 0,4; 0,375; 0,374;	0,174; 0	0	0,104; 0

e izravnih definicija	0,12				0,281; 0,112; 0,0097; 0			
-----------------------	------	--	--	--	-------------------------------	--	--	--

U konkordancijama (ujedno i GDEX primjerima) koje je *Sketch Engine* pronašao za svakog lika izvukli su se glagoli tako što se gledalo da se pojavljuju u istoj rečenici u kojoj se navodi ime lika čiji GDEX primjeri su se tražili, i da se glagoli odnose na taj lik, a ne na neki drugi. Razvrstani su u četiri kategorije: glagoli koji se odnose na dinamične radnje (npr. *poći, potpaliti, navaliti, predvoditi*), glagoli kojima se izražava statična radnja (npr. *sjesti, gledati, upitati, slušati, govoriti*), glagoli koji su vezani uz emocije i promišljanja lika (npr. *ražalostiti se, plakati, klonuti, pocrvenjeti, namrštititi se*) i glagoli koji izriču neko zbivanje (npr. *mladovati, umrijeti, živjeti, odrastati*). Tablica 5 donosi podatke o navedenim kategorijama za odabrane likove.

Tablica 5. Prikaz radnji za odabrane likove

	baka (majka)	Domaći i Malik Tintilinić	Lucija Stipančić	Vukasović	Pavao Slavinić	Gardun	Đuro Vučetić
zbivanje	2 (5 %)	1 (3,22 %)	6 (7,23 %)	3 (4,84 %)	3 (4,83 %)	1 (1,05 %)	1 (2,32 %)
emocije i promišljanja	10 (25 %)	4 (12,90 %)	26 (31,33 %)	10 (16,13 %)	13 (20,97 %)	22 (23,16 %)	17 (39,53 %)
dinamične radnje	15 (37,5 %)	19 (61,29 %)	7 (8,43 %)	12 (19,35 %)	31 (50 %)	28 (29,47 %)	0 (0 %)
statične radnje	13 (32,5 %)	7 (22,58 %)	44 (53,01 %)	37 (59,68 %)	15 (24,19 %)	44 (46,31 %)	25 (58,14 %)

ukupan broj glagola	40	31	83	62	62	95	43
------------------------	----	----	----	----	----	----	----

Osim što su se tražili podaci na temelju kojih bi se interpretirali pojedini likovi, i aktancijalne uloge na razini cijelog korpusa također su bile predmet interesa. Zato se u nastavku donosi broj pojedinih aktancijalnih uloga s obzirom na broj pojavnica likova koji preuzimaju određenu ulogu (subjekta, objekta, pomagača, protivnika, pošiljatelja i primatelja):

- N (subjekti): 1058
- N (objekti): 1462
- N (pomagači): 483
- N (protivnici): 704
- N (pošiljatelji): 793
- N (primatelji): 399.

Uz to, za svaki tekst se izračunalo koja aktancijalna uloga je najčešća u njemu:

- *Šuma Striborova*: subjekt (40,31 %)
- *Oprava*: protivnik (48,35 %)
- *Posljednji Stipančići*: objekt (34,40 %)
- *Oko Lobora*: subjekt/pošiljatelj/primatelj (100 %)
- *Krvavi most u Zagrebu*: subjekt (41,31 %).

### 8.3. Diskusija

Rezultati točnosti oba klasifikatora pokazuju da iako su alati većinski trenirani na neknjiževnim tekstovima, njihova točnost je na zadovoljavajućoj razini. Posebno ako se uzme u obzir da je uzorak na kojem se ispitivala točnost ulomak iz jezično najkompleksnijeg književnog teksta koji se koristio u ovome radu – *Krvavi most u Zagrebu*. Ta složenost prvenstveno se odnosi na velik broj zastarjelica i arhaizama te padežnih nastavaka koji se više ne upotrebljavaju u suvremenom hrvatskom jeziku. Iz matrica pogreški za *ReLDI* i *Sketch Engine* klasifikator vidi se da nije bilo

problema s ispravnim označavanjem vlastitih imenica, osobnih zamjenica i rodova za glagole. Ono što se pokazalo kao problem za daljnji rad s označenim riječima je neispravno označavanje likova kao općih imenica. To se ručno ispravljalo jer je točna MSD oznaka (ne nužno cijela oznaka, nego dio koji se odnosi na to je li imenica opća ili vlastita) bila važna za CQL izraze koji su se u nastavku koristili. Tekst *Šuma Striborova* u tom je kontekstu predstavljala specifičan problem. Budući da pojedini bitni likovi u toj bajci nemaju vlastito ime, nego su imenovani općom imenicom (npr. „baka”, „sin”, „snaha”), za potrebe analize likova te su imenice označene kao vlastite imenice jer se odnose na konkretni lik u tekstu. Nije došlo do šuma u ostalim tekstovima jer je svaki od pet tekstova činio jedan potkorpus, pa promjene u jednom tekstu nisu utjecale na ostale.

Vlastitim imenicama ne označavaju se samo imena osoba (u ovome slučaju likova), nego i neke druge kategorije – mjesta, institucije, geografski pojmovi, božanstva, nazivi publikacija itd. U književnim se tekstovima očekuje da su neke od navedenih kategorija dominantne, a neke da se pojavljuju puno rjeđe. Kad se interpretiraju pripovjedni književni tekstovi, često se, primjerice, spominju mjesta radnji iz čega se može zaključiti da se osim likova u fikcionalnim narativnim tekstovima redovito pojavljuju i određene lokacije. Kako su po kategorijama raspoređene vlastite imenice u književnim tekstovima koji se analiziraju u ovome radu, može se vidjeti u Tablici 3. U *Šumi Striborovoj*, uz nazive likova, najviše se pojavljuju nazivi božanstva, konkretnije imenica „Bog”. Ostalih vlastitih imenica gotovo da i nema. Takva raspodjela vlastitih imenica može se povezati s književnom vrstom kojoj tekst pripada – umjetničkom bajkom. Visoka razina fikcionalnosti i odmicanje od bilo kakvog referiranja na stvarna mjesta, događaje ili osobe dovodi do toga da su čak i nazivi likova zapravo opće imenice. Bajka je i čest predložak za tumačenje aktancijalnih uloga, samo što Ivana Brlić-Mažuranić aktere tipične za bajke „prevodi” u aktere koji su bliski hrvatskoj kulturi – snaha, baka, sin, ubogo djevojčice i sl. Zazivanje Boga u radosti i nevolji isto se tako može protumačiti kao nešto čime ova bajka odražava pripadnost hrvatskom mentalitetu i tradiciji. Teorija književnosti razlikuje narodnu i umjetničku bajku pri čemu je narodna bajka sklonija formulaičnom izražavanju, ponavljanju motiva, nepromjenjivim likovima, jasnim oprekama dobra i zla te stoga pripada jednostavnim proznim oblicima (Solar, 1984: 169). S druge strane, umjetnička bajka je složeni prozni oblik koji nije podložan čvrstim okvirima izražavanja i konvencijama kao narodna bajka te dopušta autoru više slobode. Ana Pintarić (2008: 12) navodi da je narodna bajka vezana uz mitska, neodređena vremena, dok se za umjetničku (ili autorsku) bajku zna vrijeme i mjesto nastanka (francuski klasicizam). Ivana Brlić-Mažuranić

pisala je umjetničke bajke u kojima je kombinirala elemente klasične bajke i hrvatskog folklora te kršćanski svjetonazor (Pintarić, 2008: 166).

Novela *Oprava* također nema previše drugih vlastitih imenica osim imena likova. Mogle bi se jedino istaknuti pojavnice koje su lokacije, točnije riječ je o Gleichbergu koji se pojavljuje 7 puta, i Zagrebu koji se u analiziranom uzorku spominje samo jednom. Gleichberg je mjesto gdje glavni lik odlazi na liječenje i odakle se vraća kući u Slavoniju. Novela kao književna vrsta fokusirana je na jednu radnju i manji broj likova, stoga ne čudi da je i lokacija koja se pretežito spominje samo jedna. Time se uočava statičnost radnje i naglasak na emocionalno i psihičko stanje likova, zbog čega se *Oprava* i smatra psihološkom novelom. Pripovijetka *Krvavi most u Zagrebu* daje drugačiju sliku u odnosu na prethodna dva teksta. Uz imena likova, brojem pojavljivanja dominiraju lokacije i povijesne osobe. Razlog tomu može se iščitati iz činjenice da pripovijetka ima dvije narativne linije od kojih se jedna odnosi na povijesna zbivanja borbe za vlast između biskupa Mihalja i župana Garduna. Dijelovi teksta koji se odnose na tu narativnu liniju spominju različite vladare, osobe iz hrvatske prošlosti, mjesta njihova podrijetla. Ova pripovijetka jasno pokazuje jedan problem s kojim se proučavatelji književnosti mogu susresti kada rade istraživanje o likovima koristeći se računalnim alatima a to je da MSD, NER ili bilo koje slično označavanje ne može razlikovati fiktivne likove od povijesnih osoba na automatskoj razini. Jedino osoba koja je upoznata sa sadržajem teksta može odrediti što jesu a što nisu književni likovi. U putopisu *Oko Lobora* postoji samo jedan glavni lik koji je istovremeno i pripovjedač, stoga nije eksplicitno imenovan. Njegova prisutnost u tekstu na jezičnoj razini može se uočiti jedino preko osobne zamjenice „ja”. Zbog toga tekst nema nijednu vlastitu imenicu koja bi označavala lika. Kao i kod *Krvavog mosta u Zagrebu*, najveći dio vlastitih imenica odnosi se na lokacije i povijesne osobe. Matošev putopis opisuje putovanje glavnog lika do Lobora, tijekom čega se prisjeća hrvatske prošlosti, objašnjava razlike između Hrvata i Srba i navodi asocijacije koje u njemu bude krajevi kroz koje prolazi. Iz navedenog se može zaključiti da čim se tekst izmjesti iz sadašnjosti u prošlost pripovijedanja, bilo u obliku događaja ili sjećanja, nužno se povećava broj lokacija i povijesnih osoba. Taj bi se podatak mogao iskoristiti kao neki oblik detektiranja tematike ili čak žanra književnih tekstova. U Novakovom romanu *Posljednji Stipančići* najviše se vlastitih imenica odnosi na likove; lokacija i povijesnih osoba je svega nekoliko. Iako se i u ovome romanu dio radnje bavi društveno-političkim pitanjima, situacija je drugačija nego u Bogovićevoj pripovijetki.



*Posljednji Stipančići* kao likove imaju aktere senjske i hrvatske suvremenosti i zato nije potrebno spominjanje vanknjiževnih povijesnih referenci u obliku stvarnih osoba.

Upotreba GDEX (*Good Dictionary EXamples*) opcije u *Sketch Engineu* bila je inicijalno zamišljena na način da bi se pomoću nje izvlačile izravne definicije književnih likova, budući da GDEX vrednuje koliko je neka rečenica dobar rječnički primjer ili materijal za učenje jezika. Međutim, korištenje GDEX-a je pokazalo da ono što se smatra kvalitetnim rječničkim primjerom većinom nije izravna definicija lika. Štoviše, pokazalo se da su primjeri s najnižim GDEX rezultatom najinformativniji oblici izravne definicije, što se odnosi i na primjer kanonika Vukasovića u nastavku. Nadalje, može se primijetiti da su kraće i jednostavnije rečenice ostvarivale bolji rezultat (npr. *Ali još nije došlo vrijeme, da baka umre.*), nego one koje su puno složenije i koje su primjer izravne definicije lika (npr. *Glava pokreta, kanonik Vukasović, bio je uman čovjek, od prirode bogato nadaren, a naukom prosvijetljen, te ne samo porijeklom i imenom nego i duhom svojim pravi narodni aristokrat.*). U Tablici 4 navedene su brojevi onoga što se može smatrati izravnim definicijama pojedinih likova. Pritom treba napomenuti da je broj GDEX primjera jednak broju konkordancija za pojedinog lika. Ono što GDEX daje kao dodatnu informaciju naspram obične konkordancije je vrednovanje svake rečenice na skali od 0 do 1. Moglo bi se kao svojevrsno pravilo uzeti da što je niža mjera (broj bliži nuli), to je izravna definicija bolja i detaljnija. Zanimljivo je primijetiti da je lik Vukasovića iz *Posljednjih Stipančića* imao najviše primjera izravnih definicija. One dolaze i od pripovjedača i od drugih likova (redom Vukasovićevih protivnika). Zbog njegove važnosti u kontekstu društveno-političkih previranja u romanu i utjecaja koji ima na druge likove, razumljivo je da postoji potreba da ga se definira i prikaže u određenom svjetlu, ali i zbog socijskih i onemogućavajućih okolnosti koje onemogućavaju izravnu definiciju njegovih interesa. Ostatak primjera koji nisu izravne definicije pretežito se odnosi na radnje likova i upravni govor. Vrlo je malo primjera gdje se opisuje vanjski izgled lika. Može se zaključiti da je tomu tako zato što opisi slijede nakon što se navede ime lika i u njima se eksplicitno ne navodi kako se lik zove, nego se zamjenicama ukazuje na to da je riječ o liku koji je u prethodnoj rečenici bio oslovljen svojim imenom. Budući da se upit za GDEX primjere sastojao isključivo od naziva lika, očekivano je da su primjeri opisa izostali. Posebno informativnima su se pokazali primjeri u kojima likovi nešto rade, osjećaju ili razmišljaju. Tablica 5 pokazuje koliki je udio pojedinih kategorija glagola zastupljen za pojedinog lika. Pomoću tih

podataka može se dobiti slika o tome je li lik više orijentiran na iskazivanje emocija i refleksiju, je li dinamičan ili statičan.

Majka (tj. baka) te Domaći i Malik Tintilinić najviše se ostvaruju kroz dinamične radnje. Neki od primjera glagola kojima se izražava aktivnost majke su: *otići, donesti, poskočiti, potpaliti, poletjeti*. Majka je stalno u pokretu jer ju snaha tjera na obavljanje različitih zadataka (ili zamki), a uz to, majka čini sve one poslove koje su u njezinoj domeni kao domaćice i skrbiteljice obitelji. U nešto manjem postotku od dinamičnih javljaju se i statične radnje (razlika u postotku je 5 %). Majka često nešto pripovijeda, govori, sjedi i gleda. Time se ukazuje na to da je majka u brizi, da ima neki problem koji dijeli s drugima (u ovome slučaju to su Domaći koji slušaju njezinu priču i nevolju oko snahe). Ujedno se tako razvija sociemska figura majke jer se iz odnosa s drugim likovima i količine interakcije stvara slika o liku, a također i psihemskom narativnom figurom koja se ostvaruje statičnim radnjama. Domaći, i Malik Tintilinić kao njihov vođa, vrlo su dinamični likovi što pokazuje udio dinamičnih radnji, odnosno glagola kojima se one ostvaruju, u iznosu od 61,29 %. Primjeri glagola su: *izaći, nanizati se, skočiti, topnuti, podmetnuti, baciti*. Domaći su nestašna bića koja pokušavaju razveseliti majku, ali joj i pomoći na konkretan način i zato kod njih prednjače dinamične radnje – dijelom se njima odražava njihova dječja osobnost, a dijelom radnje kojima nastoje razotkriti narav snahe-guje (u tome se ističe Malik Tintilinić).

Lucija Stipančić i kanonik Vukasović najviše se izražavaju kroz statične radnje (Lucija 53,01 %, Vukasović 59,68 %). Međutim, kad se pogledaju konkretni glagoli, mogu se uočiti nepodudaranja. Lucija često gleda, sjedi, promatra, okreće (glavu, oči), dok Vukasović govori, pripovijeda, saziva, opire se, huška. Daje se naslutiti da, iako su navedeni glagoli svrstani u statične radnje, lik Lucije je više okrenut prema promatranju događaja i introspekciji. Vukasović pak često vodi glavnu riječ, obraća se protivnicima i pristašama, riječima nastoji politički utjecati na one koji ga slušaju. Njihove psihemske figure su različite. Možda je bilo za očekivati da će kod Lucije prevladavati glagoli koji su usko vezani uz emocionalna stanja i promišljanje i zapravo se to donekle ostvarilo jer su statične radnje ipak više povezane s njezinim unutarnjim životom te glagoli kojima se izražavaju emocije i misli čine oko jednu trećinu analiziranih glagola (31,33 %). Tako Lucija često porumeni, klone, plače, osjeća, smije se. Jedini odabrani lik iz korpusa koji se može usporediti po razini „emocionalnosti” je Đuro Vučetić, međutim emocije i promišljanja koje on proživljava drugačije su od Lucijinih, što se može vidjeti u nastavku ovog potpoglavlja.

Likovi Pavla Slavinića i župana Garduna iz *Krvavog mosta u Zagrebu* imaju izražene statične i dinamične radnje, ali u različitim odnosima. Dinamične radnje čine 50 % analiziranih glagola Pavla Slavinića – on predvodi, navaljuje, prodire, povrati se, odbija, brani se itd. Te radnje su povezane s njegovom ulogom u sukobu sa županom Gardunom. Slavinić predvodi pristaše biskupa Mihalja kako bi zaustavio Gardunovu vojsku. Samo iščitavajući glagole koji se vežu uz ovaj lik a bez poznavanja književnog teksta, Slavinić se može okarakterizirati kao hrabar, poduzetan i ratoboran pojedinac. Međutim, ono što bi se tada moglo zaključiti je da je Pavao Slavinić nasilnik i onaj koji napada, dok je on zapravo branitelj i pozitivan lik. Slični glagoli javljaju se i uz lik župana Garduna – on odbija, navaljuje, sukobljava se, harači itd. Jedino se po glagolu *haračiti* može naslutiti da je riječ o negativnom liku. Zanimljivo je da iako je Gardun također aktivan sudionik borbe i predvodi svoje pristaše, kod njega prednjače statične radnje (46,31 %, a dinamične 29,47 %). Glagoli kojima se najčešće izražavaju statične radnje tiču se govorenja: *zapovjediti, sazvati, reći, hrabriti, savjetovati, vikati* i sl. Može se primijetiti sličnost s likom kanonika Vukasovića iz *Posljednjih Stipančića* – obojica govorenjem utječu i pokazuju svoju moć.

Đuro Vučetić, glavni lik *Oprave*, nema nijedan glagol kojim bi se iskazivala neka dinamična radnja, pojavljuje se samo jedan glagol zbivanja, glagoli kojima se izriču statične čine 58,14 % ukupnog broja analiziranih glagola, a preostalih 39,53 % čine glagoli vezani uz emocije i promišljanja. Vučetić promatra, gleda, otkriva, nazire, ili pak drhti, plače, dolazi na pomisao, ruje po uspomenama, zna, proniče itd. Ranije je rečeno da se za *Opravu* kaže da je psihološka novela, a to se potvrđuje radnjama glavnoga lika. Zaokupljen je sumnjom u ženinu nevjeru i pitanjem je li Dušan zaista njegov sin. Uz to je slaba zdravlja, duševno uznemiren i osjetljiv. Lik je orijentiran na promišljanje i introspekciju (slično kao Lucija Stipančić), što se odražava i na njegovo emocionalno stanje. Stoga je kod njega izražena psihemska narativna figura. Uvidom u glagole stvara se dobar temelj za psihološku karakterizaciju Đure Vučetića. Vrsta riječi koja se zdravorazumski nameće kao korisna za karakterizaciju likova su pridjevi. Međutim, na njih se u ovome radu nije fokusiralo jer se ne pojavljuju toliko često uz imena likova (kao glagoli) da bi ih u ovoj fazi računalne obrade bilo jednostavno pronaći.

Već je spomenuto da, budući da nema nijednog lika osim pripovjedača, putopis *Oko Lobora* nema GDEX primjera. Ipak, gledajući obične konkordancije za osobnu zamjenicu „ja” i glagole koji se uz nju javljaju, može se zaključiti da je pripovjedač rijetko aktivan lik koji vrši neku radnju. Jedini put kad se glavni lik može smatrati subjektom je kada zamišlja što bi on sve učinio da je posljednji

gospodar dvorca Lobora. U svim ostalim primjerima on je onaj kojeg se vodi, gleda, prati, na koga se škilji, koga se pita. Pritom vršitelji radnje nisu živa bića, nego sjena, noć, cesta, dvor i sl.

Neizravna prezentacija (kako ju obrazlaže naratologija) dijeli radnje likova na rutinske i nerutinske. Takva podjela ne može se u potpunosti ostvariti samo nabranjem glagola koji se vežu uz pojedinog lika. Prethodno je potrebno pomno pročitati tekst i upoznati se s radnjom i likovima. Ipak, brojnost određenih glagola može ukazati na radnje koje lik više puta obavlja, a to se vidi iz broja pojavnica glagola. Primjerice, Vukasović i Gardun nerijetko nešto govore, pripovijedaju, nekome se obraćaju. Može se reći da su to njihove rutinske radnje. Ili Lucija koja redovito plače i crveni se. Ako se ide obrnutim smjerom, glagoli koji su najmanje zastupljeni trebali bi predstavljati nerutinske radnje lika. Međutim, u primjerima likova koji se analiziraju u ovome radu nema neke radnje koja bi bila vrlo specifična i važna za određenog lika, a da se pojavljuje vrlo rijetko. Primjerice, iako se glagoli *ciknuti* i *pucati* (od smijeha) javljaju samo jednom uz Domaće i Malika Tintilinića, oni podupiru rutinske radnje i karakterizaciju – njihovu nestašnost, zaigranost i dinamičnost. Može se zaključiti da je ovakav pristup radnjama lika prikladniji i informativniji za rutinske nego za nerutinske radnje. Količina glagola mogla bi biti upotrebljiva i za razlikovanje književnih rodova i vrsta. U epici su radnje više zastupljene nego u drami i lirici, stoga bi i brojnost glagola trebala biti veća, posebice onih koji se tiču dinamičnih radnji. Unutar epike također bi se mogle napraviti analize glagola koje bi ukazivale na to kakva je tematika teksta, koliko je naglasak na radnji, a koliko na opisima, emocionalnim stanjima i sl. Na primjeru ovog istraživanja vidljivo je da su *Šuma Striborova* i *Krvavi most u Zagrebu* najviše orijentirani na dinamične radnje, što je u skladu s temama kojima se bave. S druge strane, novela *Oprava*, putopis *Oko Lobora*, pa i roman *Posljednji Stipančići*, nemaju toliko izražene glagole kojima se prikazuje aktivnost likova. Naravno, ovo su zaključci proizašli iz analize svega nekolicine likova. Za detaljniju analizu potrebno je sagledati tekst u cjelini.

Govor kao još jedan element neizravne prezentacije prisutan je u obliku glagola kojima se izriče neki oblik govorenja ili šutnje i upravo po odabiru glagola se može okarakterizirati lika. Majka je ona koja pripovijeda svoju nesretnu sudbinu, Domaći manje govore, ali zato slušaju, Lucija muči, ne želi odgovarati, ne progovara, ne odvrća, za razliku od Vukasovića koji često govori, prepire se, huška i umiruje. I Pavao Slavinić i Gardun viču, a Gardun još i zapovijeda, savjetuje, priopćava, ispipava. Vučetić pak gotovo da i ne koristi nijedan glagol (ne)govorenja. To je u skladu s njegovom tendencijom da o stvarima promišlja, a ne da ih verbalizira i ulazi u komunikaciju.

Okolina je bitan element karakterizacije lika – gdje se nalazi, čime je okružen, kako to odražava njegov karakter i kako okolina na njega utječe. U ovome radu bavilo se ljudskom okolinom – socijalskom narativnom figurom, kako kaže Peleš (1999, navedeno u Grdešić, 2015). Likovi koji su odabrani za detaljniju analizu stavljeni su u odnos s drugim likovima koji čine aktancijalne modele teksta kojem likovi pripadaju. To je učinjeno tako da su se pisali CQL izrazi kojima su se obuhvatili svi likovi aktancijalnog modela.

Za svaki lik zadana su dva CQL izraza, ovisno je li lik ispred ili iza drugog lika koji se s njime javlja u interakciji. Upit se ograničio na rečeničnu razinu i zadalo se da razmak između dva lika može biti 0, 1, 2, 3 ili 4 pojavnica. U *Šumi Striborovoj* majka je najviše u interakciji sa sinom (7 pojavljivanja), zatim s Domaćima i Malikom Tintilinićem (6 pojavljivanja), snahom (5 pojavljivanja), a sa Striborom stupa u interakciju u dva pojavljivanja. Razlog česte bliske pojave sina i majke u rečenici može se pripisati tome da subjekt žudi za objektom aktancijalnog modela, dakle majka želi uspostaviti odnos sa sinom. Domaći kao glavni pomagači potpomažu da se taj odnos realizira i zato se javljaju 6 puta, a isto se može reći za snahu koja kao protivnik želi spriječiti da majka zadobije sina na svoju stranu. Domaći (uz Malika Tintilinića kao predstavnika), shodno navedenom, najviše stupaju u interakciju s majkom, sa šumskim starješinom Striborom 2 puta te sa snahom i sinom po jedanput.

Kod Lucije Stipančić uvjerljivo dominira njezin odnos s majkom Valpurgom koja se u istoj rečenici s Lucijom pojavljuje čak 29 puta. Ponovno bi se situacija mogla povezati s Valpurginom ulogom pomagača, ali i s dijeljenjem uloge objekta jer su obje članice obitelji Stipančić (uz Jurja i Antu) sudionice propasti Stipančića zbog političke aktivnosti oca i stanja u hrvatskom političkom životu. Sljedeći likovi s kojima se Lucija javlja u puno manjoj mjeri su: Juraj (6 pojavljivanja), Ante (4 pojavljivanja) i Alfred (3 pojavljivanja). Kanonik Vukasović stupa u interakciju sa samo dva lika, i to u istoj količini – Antom Stipančićem i Čolićem (11 pojavljivanja svaki). Odnosi ova tri lika uvjetovani su njihovim političkim aktivnostima i interesima te iako su zalažu za različite politike, međusobno surađuju, prave dogovore i pokušavaju zadobiti što više pristaša.

Pavao Slavinić vrlo se rijetko pojavljuje unutar iste rečenice s drugim likom, svega 9 puta. Od toga se šest puta njegovo ime nalazi pored Ružice, predmeta njegove žudnje. Poput likova majke i sina, Slavinić kao subjekt ostvaruje socijalsku narativnu figuru s Ružicom kao objektom. Iako druga narativna linija pripovijetke prati sukobljavanja Garduna i biskupa Mihalja, čija uzdanica u borbi

je Pavao Slavinić, dvojica protivnika (Gardun i Pavao) u neposrednoj jezičnoj blizini nalaze se svega dva puta. Razlog se može pronaći u načinu pisanja tog teksta. Naime, Bogović daje dugačke rečenice koje se fokusiraju samo na lika koji je njihov subjekt, a tek se u nekoj od sljedećih rečenica javlja drugi lik, ili su unutar iste rečenice, zbog duljine, likovi previše udaljeni, tj. previše riječi se nalazi između njih. Kad bi se u upit stavilo, primjerice, 10 pojavnica umjesto 4, ne bi se dobilo značajnije više rezultata. Dakle, glavni problem su dugačke rečenice i nizovi rečenica koji govore o jednom liku. Župan Gardun najčešće se u istoj rečenici spominje s braćom Gusingovcima (5 puta), a po jedanput se javlja s Ružicom i Pavlom Slavinićem.

Zanimljivi rezultati dobiveni su za lik Đure Vučetića. Naime, nije pronađen nijedan primjer gdje se glavni lik pojavljuje u istoj rečenici s drugim likom. Uzrok toga vjerojatno leži u tome što se likovi u *Opravi* više spominju kroz zamjenice, a ne preko vlastitih imena. Nadalje, radnja novele je poprilično rudimentarna i naglasak je na promišljanjima i dvojabama Đure Vučetića, stoga on ni ne dolazi u izravnu interakciju s bilo kojim od ostalih likova. Slična stvar se događa i s glavnim likom (pripovjedačem) u putopisu *Okolobora* – budući da je on jedini lik, nema s kime stupiti u interakciju.

Bitno je napomenuti da između pojedinih likova unutar tekstova nije došlo do izravne interakcije ne samo zato što se analiza radila na razini jedne rečenice, nego i zato što se neki likovi pojavljuju kasnije u tekstu. Konkretno, to se može zaključiti za Alfreda i Martina Tintora, Stribora i Malika Tintilinića i sl. Likovi koji su vezani uz istu narativnu liniju i vrijeme radnje vjerojatno će se pojavljivati puno bliže jedan drugome.

Na razini cjelokupnog korpusa (svih 5 tekstova zajedno) gledao se broj pojavnica za svaku aktancijalnu ulogu. S time da u obzir treba uzeti da su se tako mogle detektirati jedino uloge koje su se ostvarile kroz određen lik, a ne kroz apstraktne instance (kao što su, primjerice, šuma i selo u *Šumi Striborovoj*). Iako su subjekti često likovi koje se smatra glavnima u pripovjednom književnom tekstu, oni nisu imali najveći broj pojavnica (njih je 1058). Uloga objekta je najbrojnija u količini pojavljivanja u iznosu od 1462 pojavnice. Obje se uloge ostvaruju kroz 8 likova. Ipak, moglo bi se zaključiti da se imena objekata češće spominju jer su subjekti fokusirani na ostvarivanje veze s njima, a uz to mogu biti predmet žudnje ne samo jednog subjekta, nego više njih. Time se i moraju učestalije spominjati u tekstu. Budući da su subjekti i objekti najbrojniji, os želje se može izdvojiti kao dominantna naspram preostale dvije, što se i dokazuje zbrojem objekta

i subjekta u iznosu od 2520 pojavnica. Sljedeća najbrojnija aktancijalna uloga je uloga pošiljatelja (u koje su uključeni i akteri pošiljatelja). Likovi koji imaju ulogu pošiljatelja i koji su akteri pošiljatelja većinski pripadaju romanu *Posljednji Stipančići*. Budući da je to najopsežniji tekst u korpusu, i njihova imena se ponavljaju u velikom broju. Pošiljatelji u ostalim tekstovima uglavnom su apstraktni elementi, stoga se njihova pojavnost nije mogla ubrojiti. Os prijenosa (koju čine pošiljatelji i primatelji) po brojnosti pojavnica (u iznosu od 1192) slijedi nakon osi želje. Iza pošiljatelja slijede protivnici kojih ima 704. Zanimljivo je da su brojniji od pomagača (i po broju likova [8], i po broju pojavnica), iz čega se može zaključiti da se protivnici eksplicitnije izražavaju u tekstovima nego pomagači i da općenito postoji više likova koji žele spriječiti ostvarivanje odnosa između subjekta i objekta nego onih koji žele pomoći. Ako se gleda os prijenosa, uočava se da ona ima najmanji zbroj (1187) u odnosu na ostale osi, međutim os prijenosa ima svega 5 pojavnica više, stoga su te dvije osi podjednako zastupljene. Od svih uloga aktancijalnog modela, najmanja je pojavnost primatelja (399). Kao i kod pošiljatelja, u ovaj broj su uračunati i akteri primatelja.

Navedeni podaci sugeriraju da se u tekstovima koji su odabrani za analizu u ovome radu najviše pojavljuju objekti, protivnici i pošiljatelji. Uočava se da dolazi do prevage na tzv. „negativnu” stranu jer su pošiljatelji često oni koji ne žele da se ostvari veza subjekta i objekta, nego protivnika i objekta. Pritom su „pozitivni” likovi (koji se nerijetko ostvaruju u ulogama primatelja i pomagača) u nepovoljnijem položaju.

Osim brojnosti pojedinih uloga aktancijalnog modela, htjela se napraviti i njihova distribucija kroz svaki tekst (koliko se neka uloga javlja na početku, sredini i kraju teksta, s time da su se tekstovi proporcionalno podijelili na tri dijela). Međutim, nije pronađen nijedan uzorak koji bi se mogao prikazati kao svojevrsna tendencija određene uloge (npr. da se primatelji većinski pojavljuju u zadnjoj trećini teksta). Uzrok tomu može biti višestrukost uloga koje likovi na sebe preuzimaju, ali i struktura književnog teksta koja je često složena i istovremeno se prati više narativnih linija. Ono što je distribucija pokazala jest to da visoka ili niska frekvencija pojavnosti lika može ukazati na to u kojem dijelu teksta se odvija koja narativna linija. Primjerice, Ante Stipančić pojavljuje se u prve dvije trećine teksta, a u posljednjoj izostaje, dok se Lucija i Valpurga Stipančić najviše pojavljuju u zadnjoj trećini teksta, čime se potvrđuje da smrću glave obitelji žene dobivaju više prostora i one sada moraju brinuti o budućnosti Stipančića.

## 9. Zaključak

Digitalna humanistika znanstveno je polje koje od onih koji se njime bave zahtijeva spremnost na suradnju, interdisciplinarnost i gledanje izvan naviknutih kategorija. To je temeljna snaga digitalne humanistike. Književnost je područje interesa humanističkih znanosti i kao takva ulazi u krug onoga što digitalna humanistika proučava. Prilikom analiziranja književnih tekstova, važno je da se u obzir uzima književna teorija jer time analiza postaje detaljnija i mogu se dobiti kvalitetniji rezultati. To je bila nit vodilja i u ovome radu – na temelju izložene teorije o karakterizaciji likova i aktancijalnim modelima, napravljena je analiza pojedinih likova hrvatske književnosti.

Provođenjem istraživanja uočeno je da treba biti prilagodljiv i spreman na promjene ako se neki element iz književne teorije ne može „prevesti“ u konkretnu računalnu metodu ili alat, već je potrebno kombinirati više opcija. Unatoč takvim poteškoćama, opcije programa *Sketch Engine* koje su se koristile za analizu pokazale su zanimljive rezultate vrijedne tumačenja, čime se ostvario cilj i potvrdila hipoteza rada. Iz rezultata se moglo vidjeti da radnje, frekventnost, pojavnost lika uz drugi lik, načini ostvarivanja aktancijalnih uloga mogu izroditi karakterizacijom književnoga lika. Uz to, s obzirom na oblike eksplicitnog pojavljivanja imena likova, njihov udio u vlastitim imenicama prepoznao se kao naznaka određene razlike u žanrovima.

Prostor za daljnji rad i napredak u primjeni alata svakako postoji. Jedna od stavki na kojima bi se moglo poraditi jest kako pronaći one elemente karakterizacije (npr. vanjski izgled, ambijent u kojem se lik nalazi, dijelovi tijela, namještaj ili predmeti koje lik koristi) koji se u tekstu ne vezuju uz eksplicitno imenovanje lika. Sličan je problem i s apstraktnim aktancijalnim ulogama – postavlja se pitanje o tome na koji ih način detektirati. Dakako, tu je i već spomenuto pitanje o tome kako izvući izravne definicije. Budući da trenutno ne postoji (gotovo) nikakav broj razvijenih alata ili programa za računalno potpomognutu analizu književnih tekstova na hrvatskome jeziku, brojne su mogućnosti razvijanja ovog područja digitalne humanistike.



## Literatura

Bullard, J., Ovesdotter Alm, C. (2014). Computational analysis to explore authors' depiction of characters. *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLfL)*. [internet], str. 11 – 16. Raspoloživo na: <https://www.aclweb.org/anthology/W14-0902.pdf> [pristupljeno: 2. 5. 2021.].

Clement, T. E., Carter, D. (2017). Connecting Theory and Practice in Digital Humanities Information Work. *JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY*. [internet] 68 (6), str. 1385 – 1396. Raspoloživo na: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23732> [pristupljeno: 30. 4. 2021.].

*Digital Humanities Manifesto 2.0*. [internet]. Raspoloživo na: [http://www.humanitiesblast.com/manifesto/Manifesto\\_V2.pdf](http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf)

Fish, S. (2012). *The Digital Humanities and the Transcending of Mortality*. [internet]. Raspoloživo na: <https://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/?mtrref=www.google.com&gwh=3C5859B4CF46C665158841E9BEFCD439&gwt=regi&assetType=REGIWALL> [pristupljeno: 25. 3. 2021.].

Flekova, L., Gurevych, I. (2015). Personality Profiling of Fictional Characters using Sense-Level Links between Lexical Resources. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. [internet], str. 1805 – 1816. Raspoloživo na: <https://www.aclweb.org/anthology/D15-1208.pdf> [pristupljeno: 26. 3. 2021.].

Golumbia, D. (2014). Death of a Discipline. *differences: A Journal of Feminist Cultural Studies*. [internet] 25 (1). Raspoloživo na: [https://www.academia.edu/7344839/Death\\_of\\_a\\_Discipline](https://www.academia.edu/7344839/Death_of_a_Discipline) [pristupljeno: 30. 4. 2021.].

Grdešić, M. (2015). *Uvod u naratologiju*. Zagreb: Leykam international d. o. o.

Hébert, L. (2011). THE ACTANTIAL MODEL. U: *Tools for Text and Image Analysis: An Introduction to Applied Semiotics*. [internet]. Raspoloživo na: [http://www.revue-texto.net/Parutions/Livres-E/Hebert\\_AS/5.Actantial-Model.pdf](http://www.revue-texto.net/Parutions/Livres-E/Hebert_AS/5.Actantial-Model.pdf) [pristupljeno: 20. 5. 2021.].

Hrvatski mrežni rječnik – MREŽNIK. Raspoloživo na: <http://ihjj.hr/mreznik/> [pristupljeno: 1. 4. 2021.].

Hühn, P., Meister, J. C., Pier, J., Schmid, W. (2014). *Handbook of Narratology*. Berlin; New York: W. de Gruyter.

Iwanska, L., Shapiro, S. C. (2000). Natural Language Is a Powerful Knowledge Representation System: The UNO Model. U: *Natural language processing and knowledge representation : language for knowledge and knowledge for language*. Menlo Park, Calif.: Cambridge, Mass.; London: AAAI Press; MIT Press.

Jacobs, A. M. (2019). Sentiment Analysis for Words and Fiction Characters From the Perspective of Computational (Neuro-)Poetics. *Frontiers in Robotics and AI*. [internet], 6. Raspoloživo na: <https://www.frontiersin.org/articles/10.3389/frobt.2019.00053/full> [pristupljeno: 27. 3. 2021.].

Jockers, M. L. (2013). *Microanalysis: Digital Methods and Literary History*. University of Illinois Press.

Kasunić, L. (2019). *Računalna obrada književnih tekstova na primjeru analize korpusa ruskih romantičara i realista*. [internet]. Raspoloživo na: <https://repozitorij.ffzg.unizg.hr/islandora/object/ffzg%3A179/datastream/PDF/view> [pristupljeno: 1. 5. 2021.].

Kirschenbaum, M. G. (2010). What Is Digital Humanities and What's It Doing in English Departments?. *ADE Bulletin*. [internet] 150, str. 55 - 61. Raspoloživo na: <https://mkirschenbaum.files.wordpress.com/2011/03/ade-final.pdf> [pristupljeno: 2. 4. 2021.].

Koltay, T. (2016). Library and information science and the digital humanities: Perceived and real strengths and weaknesses. *Journal of Documentation*. [internet] 72 (4), str- 781 - 792. Raspoloživo na: <https://www.emerald.com/insight/content/doi/10.1108/JDOC-01-2016-0008/full/html> [pristupljeno: 2. 4. 2021.].

Liu, A. (2011). The state of the digital humanities: A report and a critique. *Arts & Humanities in Higher Education*. [internet] 2 (1-2). Raspoloživo na: <https://journals.sagepub.com/doi/10.1177/1474022211427364> [pristupljeno: 2. 4. 2021.].

Nikolić, D. (2016). Digitalna humanistika i nacionalna filologija: o mogućim implikacijama računalnog obrata. *Croatica*. [internet] 50 (60), str. 75 - 87. Raspoloživo na: <https://hrcak.srce.hr/174512> [pristupljeno: 30. 3. 2021.].

Pintarić, A. (2008). *Umjetničke bajke – teorija, pregled i interpretacije*. Osijek: Filozofski fakultet Osijek.

Piotrowski, M.. (2012). *Natural language processing for historical texts*. University of Toronto: Graeme Hirst.

Piper, A. (2018). *Enumerations: Data and Literary Study*. The University of Chicago Press.

Robinson, L., Priego, E., Bawden, D. (2015). Library and information science and digital humanities: two disciplines, joint future?. International Symposium on Information Science , 19-21 May 2015, Zadar, Croatia. Raspoloživo na: <https://openaccess.city.ac.uk/id/eprint/11889/> [pristupljeno: 1. 4. 2021.].

Sketch Engine. Raspoloživo na: <https://www.sketchengine.eu/> [pristupljeno: 15. 5. 2021.].

Solar, M. (1984). *Teorija književnosti*. Zagreb: Školska knjiga.

Su, F., Zhang, Y., Immel, Z. (2020). Digital humanities research: interdisciplinary collaborations, themes and implications to library and information science. *Journal of Documentation*. [internet] 77 (1), str. 143 – 161. Raspoloživo na: <https://www.emerald.com/insight/content/doi/10.1108/JD-05-2020-0072/full/html> [pristupljeno: 2. 5. 2021.].

Tadić, M. (2003). *Jezične tehnologije i hrvatski jezik*. Zagreb: Exlibris.

Vala, H., Jurgens, D., Piper, A., Ruths, D. (2015). Mr. Bennet, his coachman, and the Archbishop walk into a bar but only one of them gets recognized: On The Difficulty of Detecting Characters in Literary Texts. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. [internet], str. 769 – 774. Raspoloživo na: <https://www.aclweb.org/anthology/D15-1088.pdf> [pristupljeno: 2. 5. 2021.].

## Popis slika

Slika 1. Prikaz procesa istraživanja hipoteze.....	5
Slika 2. Sučelje internetskog programa Sketch Engine.....	24

## Popis tablica

Tablica 1. Veličina korpusa.....	26
Tablica 2. Matrica pogreške za ReLDI klasifikator .....	29
Tablica 3. Matrica pogreške za Sketch Engine klasifikator .....	29
Tablica 4. Broj različitih kategorija vlastitih imenica .....	31
Tablica 5. Prikaz radnji za odabrane likove .....	33

# Računalno potpomognuta analiza likova u odabranim tekstovima hrvatske književnosti

## Sažetak

U radu se iznosi narav odnosa digitalne humanistike i tradicionalne humanistike (ponajprije iz perspektive književnosti) te poveznice između digitalne humanistike i informacijskih i komunikacijskih znanosti. S obzirom na to da je glavni fokus rada računalno potpomognuta analiza književnih likova, objašnjava se kako naratologija tumači likove, njihovu klasifikaciju i karakterizaciju, što predstavlja književnoteorijsko polazište za analizu likova u središnjem dijelu rada. Također, navode se i objašnjavaju računalne metode koje se koriste prilikom računalno potpomognute analize književnih likova. U istraživanju se upotrebljavaju sljedeći književni tekstovi: *Posljednji Stipančići* (Vjenceslav Novak), *Krvavi most u Zagrebu* (Mirko Bogović), *Šuma Striborova* (Ivana Brlić-Mažuranić), *Oprava* (Josip Kozarac) i *Oko Lobora* (Antun Gustav Matoš), koji su dostupni na internetskoj stranici *e-lectire*. Istraživanje je usredotočeno na naratološku analizu književnih likova potpomognutu računalnim metodama i alatima. Analiza rezultata pokazuje korisnost računalnih metoda i alata u naratološkoj analizi likova. Proučavanje glagola koji se vežu uz pojedini lik, traženje podataka o interakciji likova, analiza GDEX primjera izravnih definicija, statistički podaci o udjelima vlastitih imenica i osobnih zamjenica postupci su kojima se ostvaruje karakterizacija. Navode se i prijedlozi za budući rad i razvoj računalno potpomognutog proučavanja književnih likova.

**Ključne riječi:** digitalna humanistika, naratologija, informacijske i komunikacijske znanosti, računalne metode, književni tekst

# Computer-based character analysis in a selection of Croatian literature

## Summary

This paper presents the nature of the relationship between digital humanities and traditional humanities (primarily from the perspective of literature) as well as the link between digital humanities and information and communication sciences. Given that the main focus of the paper is computer-aided analysis of literary characters, it describes narratological character interpretation, character classification and characterization, which represents the literary theory starting point for character analysis in the central part of the paper. Furthermore, the computer methods used in computer-aided character analysis are listed and explained. The following literary texts are used in the research: *Posljednji Stipančići* (Vjenceslav Novak), *Krvavi most u Zagrebu* (Mirko Bogović), *Šuma Striborova* (Ivana Brlić-Mažuranić), *Oprava* (Josip Kozarac) and *Oko Lobora* (Antun Gustav Matoš) which are available on the *e-lectire* website. The research focuses on the narratological analysis of literary characters aided by computer methods and tools. The analysis of the results shows the usefulness of computer methods and tools in narratological analysis of characters. The study of verbs related to a particular character, searching for data on the interaction of characters, analysis of GDEX examples of direct definitions, statistics on the shares of proper nouns and personal pronouns are procedures by which characterization is accomplished. Suggestions for future work and development of computer-aided study of literary characters are also given.

**Key words:** digital humanities, narratology, information and communication science, computer methods, literary text