

# Primjena i problematika umjetne inteligencije u književnosti

---

**Tomiša, Maja**

**Master's thesis / Diplomski rad**

**2024**

*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:131:786070>

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

*Download date / Datum preuzimanja:* **2024-10-19**



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 BIBLIOTEKARSTVO  
Ak. god. 2023./2024.

Maja Tomiša

# **Primjena i problematika umjetne inteligencije u književnosti**

Diplomski rad

Mentor: izv. prof. dr. sc. Tomislav Ivanjko

Zagreb, rujan 2024.

## **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

<b>1. Uvod.....</b>	<b>1</b>
<b>2. Generativna umjetna inteligencija.....</b>	<b>3</b>
2.1 Kratki pregled nastanka i razvoja umjetne inteligencije.....	3
2.2 Data mining.....	6
2.2.1 Text mining.....	9
2.3 Umjetne neuronske mreže.....	13
<b>3. Obrada i generiranje prirodnog jezika.....</b>	<b>18</b>
3.1 Računalno jezikoslovlje.....	18
3.2 Postupci i mogućnosti generiranja teksta.....	21
3.2.1 Prevoditeljske sposobnosti.....	25
3.2.2 Stil generiranoga teksta.....	27
3.3 Generiranje književnoga teksta.....	29
3.3.1 Stvaranje i kreativnost.....	32
<b>4. Autorstvo.....</b>	<b>37</b>
4.1 Pitanje autorskoga prava na skupljene podatke.....	37
4.2 Pitanje autorskoga prava na generirane tekstove.....	40
4.2.1 Umjetna inteligencija kao autor, suautor ili pomoćni alat.....	41
4.3 Budućnost književnosti.....	44
<b>5. Etika umjetne inteligencije.....</b>	<b>47</b>
5.1 Etika ideja.....	47
5.2 Etika razvoja.....	49
5.3 Etika korištenja.....	49
<b>6. Zaključak.....</b>	<b>51</b>
<b>7. Literatura.....</b>	<b>53</b>
<b>Popis slika.....</b>	<b>59</b>
<b>Sažetak.....</b>	<b>60</b>
<b>Summary.....</b>	<b>61</b>

# 1. Uvod

Posljednje je desetljeće u razvoju tehnologija obilježila primjena umjetne inteligencije, a posebno velik korak u odnosu na prethodne principe izgradnje modela umjetne inteligencije i njihove primjene donijela je vrsta umjetne inteligencije koja se naziva generativnom. Otkako je 2018. godine tvrtka *OpenAI* predstavila prvu verziju *ChatGPT*-ja, *chata* koji može razumjeti prirodni jezik i ponuditi odgovore na njemu te pamtili povijest razgovora s korisnikom, mišljenja da će umjetna inteligencija promijeniti naš svijet do neprepoznatljivosti gotovo su se ukorijenila u sveopćemu mišljenju društva. Osim promjena u tehnikama pretraživanja informacija, počelo se raspravljati i o promjenama koje bi kao posljedicu mogle imati smanjenje potrebe za ljudskim angažmanom, stvaranjem i radom općenito. Generativna umjetna inteligencija već je nemalo uzdrimala i industriju zabave te se pokazala kao potencijalna prijateljica kreativnome ljudskome radu.

Kako bi se izbjeglo izvođenje ekstremnih zaključaka o mogućnostima i potencijalima generativne umjetne inteligencije potrebno je rasvijetliti temeljne tehnologije na kojima se taj oblik umjetne inteligencije zasniva. Prvi dio ovoga rada posvećen je općenitom razvoju umjetne inteligencije kao pojma koji je nastao unutar računalne tehnologije još polovinom prošloga stoljeća. Kratak pregled toga razvoja prikazuje različite polazišne točke računalnih znanstvenika i njihove ciljeve koji su s više ili manje značaja utjecali na napredak tehnologija do onoga što danas poznajemo kao generativna umjetna inteligencija. Također se u prvome dijelu obrađuju tehnike *data mininga* i *text mininga* u smislu sadržajne osnove bez koje niti jedan model generativne umjetne inteligencije ne bi mogao funkcionirati te uloga umjetnih neuronskih mreža u obrađivanju tih podataka.

Drugi dio rada sužava fokus na postupke obrade i generiranja prirodnoga jezika, odnosno mogućnosti računalnoga algoritma da razumije jezik i koristi se njime na načine na koje ga koriste ljudi, bilo u razgovoru, prevođenju, analizi ili pisanju teksta. Poseban se naglasak stavlja na složenost književnoga teksta u smislu kreativne proizvodnje, konteksta u kojemu nastaje i podteksta koji se u njemu može nalaziti.

Treći i četvrti dio rada razmatraju posebno problematično pitanje autorskoga prava na podatke koje tvrtke koriste kako bi trenirale svoje modele umjetne inteligencije, pitanje autorskoga prava na ideju, tekst ili dio teksta koji model proizvede i ponudi korisniku te, uključujući sve prethodne spoznaje, kako se sve zapravo može koristiti generativna umjetna inteligencija, obraća li se pritom pažnja na etičku komponentu same tehnologije i njezina korištenja, koja su njezina ograničenja i potencijalni smjerovi razvoja te utjecaja na ljudsko stvaralaštvo.

Cilj je rada usporediti načine učenja i stvaranja koje koristi čovjek s onima koje koriste algoritmi umjetne inteligencije te odrediti kakve rezultate oni mogu ponuditi. Analiza procesa stvaranja novih ideja i djela u književnosti mogla bi postaviti temelje za razumijevanje budućnosti u kojoj je autorska djelatnost ugrožena ili pak potpomognuta i unaprijeđena korištenjem umjetne inteligencije.

## 2. Generativna umjetna inteligencija

O umjetnoj inteligenciji govori se već desetljećima, bilo to u znanstvenim ili umjetničkim sferama. Sam pojam umjetne inteligencije nastaje i razvija se primarno u okvirima računalne znanosti, a kako računalna širi svoje utjecaje na ostale znanosti, pa čak i na aspekte svakodnevice, tako se širi i područje primjene umjetne inteligencije ili barem potiče razgovor o njezinim potencijalima i prijetnjama. Umjetnici su s vremena na vrijeme posezali za temama distopijske budućnosti u kojima neka vrsta umjetne inteligencije kontrolira ljudske živote ili roboti nasilno preuzimaju kontrolu nad svijetom. Najčešće su to književnost i film, a zbog svoje usmjerenosti na masovnu publiku, nije čudno što se takva razmišljanja pojavljuju i u nefikcionalnim diskursima koji se dotaknu teme umjetne inteligencije.

Dakako, u razvoju svake znanosti postoji dugačak niz velikih i malih koraka u napretku, od kojih je svaki na svoj način značajan. U sljedećemu pregledu spomenuti su oni koraci koji su potrebni za razumijevanje tehnologija koje se koriste u vrijeme pisanja ovoga rada te njihovih budućih tendencija.

### 2.1 Kratki pregled nastanka i razvoja umjetne inteligencije

Još je 1948. američki znanstvenik Arthur Samuel počeo raditi na razvijanju programa koji će kasnije postati poznat kao *The Samuel Checkers-playing Program*. To je bio prvi program koji je uključivao tehniku strojnog učenja (engl. *machine learning*) te popularizirao sam termin. Konkretnije, radi se o programu za igranje dame, koji će s vremenom poslužiti kao osnova za razvijanje onoga što se danas naziva generativnom umjetnom inteligencijom (Weiss, 1992). Strojno učenje koristi algoritme koji nisu namijenjeni samo poduzimanju određenih radnji, već su u mogućnosti pohraniti svaki dobiveni rezultat i uključiti ga u novi krug rada kao poznatu vrijednost. Time program pri svakom novom pokretanju posjeduje veći broj parametara s kojima može raditi. Što više puta program ponovi svoj rad, to je precizniji jer nove zadatke može usporediti s većim brojem već riješenih zadataka. Pojednostavljeni procesa izgledat će drukčije ovisno o tome kakvi su ulazni podaci te kakav se rezultat očekuje od programa u cijelosti. Ovaj postupak ponavljanja rada s podacima naziva se treniranje algoritma. Kada je riječ o jeziku, strojno učenje analizira "statistiku jezičnih uzoraka u podacima na kojima se trenira. U tim



podacima trebaju biti dokumenti za svaku rubriku kako bi se stvorio skup značajki (statistički potpis) pojedine rubrike koji će se kasnije koristiti za klasifikaciju novih dokumenata<sup>1</sup> (Barsegjan, Kuprijanov, Stepanenko i Holod, 2007a, str. 207). Drugim riječima, program analizira jezik tako što raščlanjuje riječi, analizira njihove veze s drugim riječima i gradi strukturu koja će mu pomoći da brže i preciznije analizira i poveže nove podatke u odnosu na stare (Holdsworth, 2024).

Već 1970. godine finski matematičar Seppo Linnainmaa razradio je metodu povratnoga širenja za ispravljanje pogrešaka (engl. *backpropagation*) koja omogućuje analiziranje stanja veza između pojedinih koraka od rješenja, odnosno izlazne vrijednosti, do početne ulazne vrijednosti (Griewank, 2012). Ova metoda omogućava programima da analiziraju “slabe” veze (one veze koje dovode do netočnih rješenja) i pomoću razrađenog niza matematičkih funkcija računaju promjene vrijednosti varijabli do kojih treba doći kako bi rezultat bio što točniji. Drugim riječima, zbog primjene ove metode program teži tomu da uči na svojim pogreškama. U nadolazećemu desetljeću ova se metoda koristila pri računalnomu prepoznavanju slika te se značajnije razvila u sklopu umjetnih neuronskih mreža. Sposobnost povratnoga širenja u mnogim smjerovima, a ne samo linearno, omogućilo je povezivanje mnogo međuslojeva rada te povezivanje i korištenje međurezultata na putu od ulazne do izlazne vrijednosti. Programi koji koriste takve iznimno složene strukture algoritama izdvajaju se kao posebna vrsta strojnog učenja nazvana duboko učenje (engl. *deep learning*). Iza tehnologija koje u svojim specifikacijama ili nazivima sadrže sintagmu umjetna inteligencija najčešće su upravo sustavi bazirani na dubokome učenju.

Od kraja prošloga stoljeća umjetna se inteligencija počela širiti u mnoge sfere primjene, od mehanike i medicine sve do humanističkih i društvenih znanosti, a njihovu okosnicu čini razumijevanje i korištenje prirodnoga jezika. Tijekom 2010-ih u sklopu internetskih platformi i operativnih sustava mobilnih telefona pojavljuju se *chatbotovi* – neslužbeno i opisno prevedeni na hrvatski kao roboti za čavrljanje. *Chatbotovi* se pojavljuju u *chatovima*, obliku neformalne komunikacije koji je korisnicima poznat, kako bi im pružili potrebne informacije. Komunikacija se temelji na principu pitanje-odgovor. Svako pitanje zaseban je zadatak, a ako *chatbot* ne razumije pitanje, jednostavno neće moći ponuditi odgovor. Američka tvrtka *OpenAI* predstavila

---

<sup>1</sup> Svi citati iz engleskih i ruskih izvora navedeni su prema slobodnome prijevodu autora.

je svoj proizvod *ChatGPT*, koji ne samo da može odgovarati na pitanja, već pamti odgovore i tijekom komunikacije, može odgovarati na potpitanja, priznavati svoje greške i odbijati odgovoriti na neprikladna pitanja (*OpenAI*, 2022). Svojom su naprednošću u mogućnostima *ChatGPT* i njemu slični programi postali sinonim za umjetnu inteligenciju. Akronim GPT dolazi od engleskoga *Generative Pre-trained Transformer*. Kao što mu sam naziv govori, radi se o pretvaraču koji je bio prethodno treniran na određenom korpusu i ima sposobnost obrade i stvaranja (generiranja) prirodnoga jezika. Osim već opisanih tehnika (strojno učenje, *backpropagation*, duboko učenje), ova generacija *chatbotova* osim treniranja prolazi i kroz fazu finog podešavanja (engl. *fine-tuning*). U ovoj se fazi modelu daje novi set posebno odabranih podataka specifičnih za neku temu, područje ili vrstu zadatka, kojima se model zatim posvećuje kako bi mogao dati preciznije rezultate (*OpenAI*, 2022).

No što je sa samim pojmom umjetne inteligencije? Kroz ovaj skraćeni pregled povijesti razvoja onoga što korisnici danas koriste kako ne bi sami morali pisati elektroničke poruke, školske eseje ili čak sami odlučivati koga će od prijavljenih kandidata zaposliti na otvoreno radno mjesto, možemo vidjeti da se i dalje radi o skupini algoritama koji obrađuju velike količine podataka kako bi ponudili odgovarajući rezultat.

Prijedlog Akta o umjetnoj inteligenciji Europske unije prema Uredbi Europskog parlamenta i Vijeća o utvrđivanju usklađenih pravila o umjetnoj inteligenciji (2024) nudi definiciju:

“‘UI sustav’ znači strojni sustav dizajniran za rad s promjenjivim razinama autonomije i koji nakon uvođenja može pokazati prilagodljivost te koji, za eksplicitne ili implicitne ciljeve, iz ulaznih vrijednosti koje prima, zaključuje kako generirati izlazne vrijednosti kao što su predviđanja, sadržaj, preporuke ili odluke koji mogu utjecati na fizička ili virtualna okruženja” (str. 165).

S druge strane Uredba predsjednika Ruske Federacije O razvoju umjetne inteligencije u Ruskoj Federaciji (dalje: Uredba predsjednika RF; 2019/2024) definira umjetnu inteligenciju kao kompleks tehnoloških rješenja koja su u mogućnosti oponašati ljudske kognitivne funkcije i davati sukladne rezultate. Posebno se ističe da se radi o informacijsko-komunikacijskoj strukturi, softveru koji koristi metode strojnog učenja te servisima za obradu podataka. Moglo bi se

zaključiti da se Europska unija ogradila od rasprave o tome što je točno inteligencija i može li ona biti prirodna i umjetna, ljudska ili računalna, ali iza pojma prilagodljivost svejedno stoji svojevrsna ruska paralela – oponašanje ljudskih kognitivnih funkcija. Još je Alan Turing naglašavao da stroj može postati inteligentan samo ako se na njega prenesu antropomorfne vrijednosti s ciljem da taj stroj zamijeni čovjeka (prema Schwartz, 2018). Takav pristup razvoju Oscar Schwartz naziva “*the automative vision of AI*” kojemu je suprotstavljen model J. C. R. Licklidera “*the hybrid vision of AI*” (Schwartz, 2018, str. 89) koji pretpostavlja zajednički rad čovjeka i stroja u postizanju krajnjih rezultata.

Obje definicije naznačuju da se radi o promjenjivu i višenamjenskomu sustavu. U definiciji Prijedloga Akta upotrijebljen je glagol generirati koji ukazuje na to da se pod pojmom sustava umjetne inteligencije prvenstveno podrazumijeva posebna skupina generativnih modela (poput *ChatGPT*-ja). Radi se, dakle, o modelima koji ne obavljaju samo zadane i unaprijed programirane radnje, već stvaraju (generiraju) nešto novo. Iako se generirati mogu različite vrste sadržaja (kod, slika, zvuk, glas itd.), u ovome radu naglasak će biti na postupcima i mogućnostima generiranja teksta.

## **2.2 Data mining**

Prije nego što algoritam postane dovoljno sposoban i razvijen da na upit odgovori prikladnim rješenjem, on prvo mora moći razumjeti taj upit. Kratko ćemo komentirati jedan od jednostavnijih primjera primjene strojnog i dubokog učenja kako bismo rasvijetlili korake koji u konačnici trebaju dovesti do generiranja kvalitetnoga odgovora. Programi za optičko prepoznavanje znakova (engl. *OCR – optical character recognition*) bili su trenirani na slikama svakoga pojedinog znaka svakoga fonta kako bi prema rasporedu i karakteristikama točaka u slici mogli prepoznati o kojemu se znaku radi. Taj se postupak zatim primjenjuje na rukopise i tako dalje. Vidjevši mnoštvo različitih rukopisa, program u njima pronalazi zajedničke karakteristike, naprimjer znak malog pisanog *a*, pa koliko god ono možda bilo okruglo, iskrivljeno ili nakošeno, program nauči da je to znak *a*. Ovdje, dakako, osim umjetne inteligencije, sudjeluju i tehnologije računalnoga vida i prepoznavanja uzorka, ali osim pitanja kako to program uči, na koje smo vrlo pojednostavljeno odgovorili objašnjenim primjerom,

postavlja se pitanje koliko je ulaznih podataka potrebno da bi program kvalitetno učio i odakle dolaze ti podaci. Odgovor na to pitanje pruža tehnologija *data mining*.

Rosana Ducato i Alain Strowel (2021) definiraju pojam *data mining* kao “osnovni alat podatkovnog gospodarstva (koji) koristi tehnike obrade prirodnoga jezika, strojnoga učenja, pretraživanja informacija i upravljanja znanjem za automatiziranu analizu tekstova i digitalnoga sadržaja (strukturiranih i nestrukturiranih podataka) radi izdvajanja informacija, identifikacije uzoraka, otkrivanja novih trendova, spoznaja ili poveznica” (str. 1). Ova je definicija sveobuhvatna jer *data mining* može se primijeniti na bilo koju vrstu podataka i redovito se primjenjuje u dubokom učenju. Zapravo, bez ovog dijela procesa učenja, koji analizira podatke, zatim traži poveznice i ponovno analizira podatke dobivene novim vezama, ne bismo mogli dobiti odgovarajuće izlazne vrijednosti, a da one nisu bile unaprijed zadane modelu s kojim radimo.

Slično se određenje nalazi u popisu korištenih termina u spomenutoj Uredbi predsjednika RF (2019/2024): “korak obrade strukturiranih i nestrukturiranih podataka u kojemu podacima (što uključuje tekstualne dokumente, slikovne i videomaterijale) bivaju dodijeljeni identifikatori koji odražavaju vrstu podatka”. Kao dodatak definiciji u Uredbi se navodi da se podaci mogu i interpretirati sukladno pojedinom zadatku, a cijeli proces naziva se označavanje podataka (rus. *разметка данных*, transliterirano *razmetka dannyh*). Iako je službeni naziv u Uredbi označavanje podataka, u literaturi se nerijetko koristi upravo engleski termin *data mining*, koji se pak slobodnije opisuje kao “otkrivanje skrivenih znanja u sirovim podacima” (Barsegjan, Kuprijanov, Stepanenko i Holod, 2007b, str. 58). Primjećujemo da je prvima dvjema navedenim definicijama zajedničko to da se mogu obrađivati strukturirani i nestrukturirani podaci. *Data mining* ne zahtijeva prethodno strukturiranje podataka jer će ih on sam analizirati, povezivati i iz njih zaključivati već onako kako najbolje odgovara pojedinomu slučaju. To je također važno jer tijekom dubinskog učenja u većem broju slučajeva model radi upravo s nestrukturiranim podacima. Drugim riječima, *data mining* ne pretražuje prema kriterijima, već sam tijekom procesa određuje kriterije, a zatim prema otkrivenim karakteristikama organizira podatke.

Barsegjan i sur. (2007b) opisuju postupak *data mininga* kao proces klasifikacije ili regresije koji se provode na osnovi dviju varijabli – zavisne i nezavisne, te se provodi u dva koraka. Zavisna varijabla označava karakteristiku podatka koja stoji u opreci s ostalim,

nezavisnim varijablama. Sukladno tomu svaka nezavisna varijabla može postati zavisna ako se u procesu promijeni fokus. Prvi je korak obuka modela na podacima kojima su varijable poznate. Model prema poznatim varijablama stvara određenu klasifikaciju prema kojoj će u drugome koraku svrstavati sve nove podatke koji su mu ponuđeni. Naravno, ako model među novim podacima pronade nove karakteristike ili posebna odstupanja, on će od njih stvoriti nove varijable koje mogu u budućnosti biti uračunate u proces analize te se time proširuje klasifikacija. Joyce Jackson (2002) dijeli setove postupaka *data mininga* u dvije vrste, ovisno o tome je li konačni cilj stvaranje novog modela ili isključivo pronalazak uzoraka u podacima. Prvi se set postupaka odnosi upravo na spomenutu klasifikaciju i regresiju, dok drugi obuhvaća prepoznavanje uzoraka i detektiranje odstupanja od norme koju su uzorci postavili (Jackson, 2002). Još je jedna sposobnost *data mininga* da na osnovi analiziranih podataka i dobivenih veza predvidi ishode analize još neobrađenih podataka. Ovaj dio procesa temelji se na vjerojatnostima pa se dobivene pretpostavke ne moraju nužno u potpunosti poklapati s krajnjim rezultatima analize, ali prepoznavanje uzoraka i predviđanje mogu ubrzati proces. Zbog uzročno-posljedičnoga povezivanja karakteristika podataka, ovakav se postupak naziva sekvencijskom analizom (rus. *сиквенциональный анализ*, transliterirano *sikvencional'nyj analiz*). Grafički prikaz ovoga postupka prikazan je na slici 1. Obuka modela podrazumijeva analizu poznatih podataka (*data sets* i *pre-processing*) te raspoređivanje podataka prema određenim i novostečenim kriterijima (*database systems* i *statistics*). Posljednje dvije sastavnice procesa (*analytics* i *evaluation*) odnose se na spomenuto prepoznavanje uzoraka i mogućnost rada s novim podacima kao što to objašnjava Jackson (2002).



Slika 1. Sastavnice procesa data mining (*Data Mining Process*, 11. 7. 2024.)

Autori navode nekoliko primjera korištenja svih navedenih mogućnosti *data mininga*, od internetskih trgovina do postavljanja dijagnoze i predlaganja liječenja (Barsegjan i sur., 2007b,

str. 71). *Data mining* analizira podatke (proizvodi stavljeni u košaricu, simptomi bolesti i osobine pacijenta), određuje njihove karakteristike i s obzirom na kategorije u kojima se nalaze predviđaju buduće postupke (koje su proizvode ostali kupci kupovali uz već odabrane, koji su lijekovi prikladni za određenu bolest, ali naprimjer isključi one koji nisu pogodni ako je pacijent alergičan na neki sastojak). Kao što je već istaknuto u ovome radu, statistika ne mora točno odrediti ishod (neće svaki kupac svaki put kupiti isti broj proizvoda u istoj kombinaciji) te je ljudski faktor prijeko potreban za provjeru dobivenih rezultata (liječnička dijagnoza i određivanje terapije naspram pretraživanja simptoma na internetu).

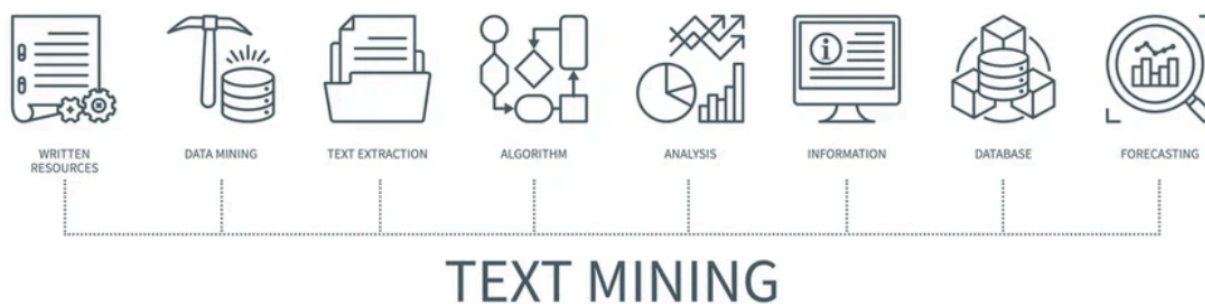
### 2.2.1 *Text mining*

Postavlja se pitanje kako *data mining* funkcionira kada su podaci koje obrađuje tekstualnoga tipa. Da bi neki model generativne umjetne inteligencije mogao stvoriti tekst, on prije svega mora naučiti analizirati, razumjeti, a tek potom koristiti jezik. *Chatbotovi* prošloga desetljeća mogli su odgovarati na upite gotovim tekstovima, ali svejedno su morali razumjeti upit, odnosno u upitu registrirati riječ ili sintagmu koja služi kao okidač ili svojevrsni putokaz za to u kojemu skupu rješenja *chatbot* može pronaći odgovor najbliži očekivanomu. Pojedini se upiti korisnika mogu, ovisno o kontekstu u kojemu se koristi *chatbot*, predvidjeti te sukladno tomu može se isprogramirati i odgovor. Ali situacija je daleko previše da bi se sve predvidjele, a ispisivanje svake moguće varijante upita i povratne informacije oduzimalo bi previše vremena. Generativni modeli nemaju zadane odgovore ili obrasce, već uče koristiti prirodni jezik i sami stvaraju obrasce po kojima mogu korisnicima ponuditi odgovore.

Ranije je spomenuto da *data mining* ima mogućnost obrade strukturiranih i nestrukturiranih podataka. Jedna od mogućih vrsta nestrukturiranih podataka su upravo tekstualni podaci, to jest dokumenti koji sadržavaju tekst. Osim što je unutarnja struktura teksta jednako složena koliko i jezik na kojemu je tekst pisan, toliko je i strukturirana vanjska forma u kojoj se tekst pojavljuje. Drugim riječima prirodni su jezici živi i gramatike ih više opisuju nego što ih propisuju, a tekst dostupan tehnologijama *data mininga* može biti u obliku dokumenta ili internetske stranice, sekcije s komentarima ili nečega drugog. Kako bi se mogli obraditi ovakvi izvori i podaci u njima osim *data mininga* potrebno je primijeniti i tehnologije “obrade prirodnih jezika, pretraživanja informacija, prikupljanja informacija i upravljanja znanjem” (Barsegjan i

sur., 2007a, str. 194). Kada se ove tehnologije primjenjuju zajedno radi obrade prirodnog jezika, taj se proces naziva *text mining*.

Nakon što se utvrdi je li pojedini izvor pogodan za primjenu *text mininga* (postoji li uopće tekst koji bi se mogao analizirati) potrebno je izvršiti preliminarnu obradu teksta (rus. *предварительная обработка*, transliterirano *predvaritel'naja obrabotka*). Ova obrada uključuje: zanemarivanje zaustavnih riječi, korjenovanje ili koji drugi način morfološke analize te ujednačavanje oblika zapisa (Barsegjan i sur., 2007a, str. 196–197). Ovaj dio postupka obavlja se u sklopu prvih triju sastavnica na grafičkom prikazu na slici 2. Zaustavne su riječi (engl. *stopwords*) one riječi koje imaju samo gramatičko značenje i u tekstovima se pojavljuju često. Radi se o nepunoznačnim riječima (prijedlozi, veznici, čestice, uskllici, članovi i sl.), odnosno riječima koje nemaju leksičko značenje pa njihovim izuzimanjem iz preliminarne analize smanjujemo opseg podataka koje treba obraditi, a time i vrijeme obrade. Gaikwad, Chaugule i Patil (2014) opisuju preliminarnu analizu kao proces analize dokumenta u kojoj se provjerava format i set znakova koji su iskorišteni u dokumentu kako bi on bio spreman za sljedeći korak – tekstualnu analizu (četvrta i peta sastavnica prikaza na slici 2).



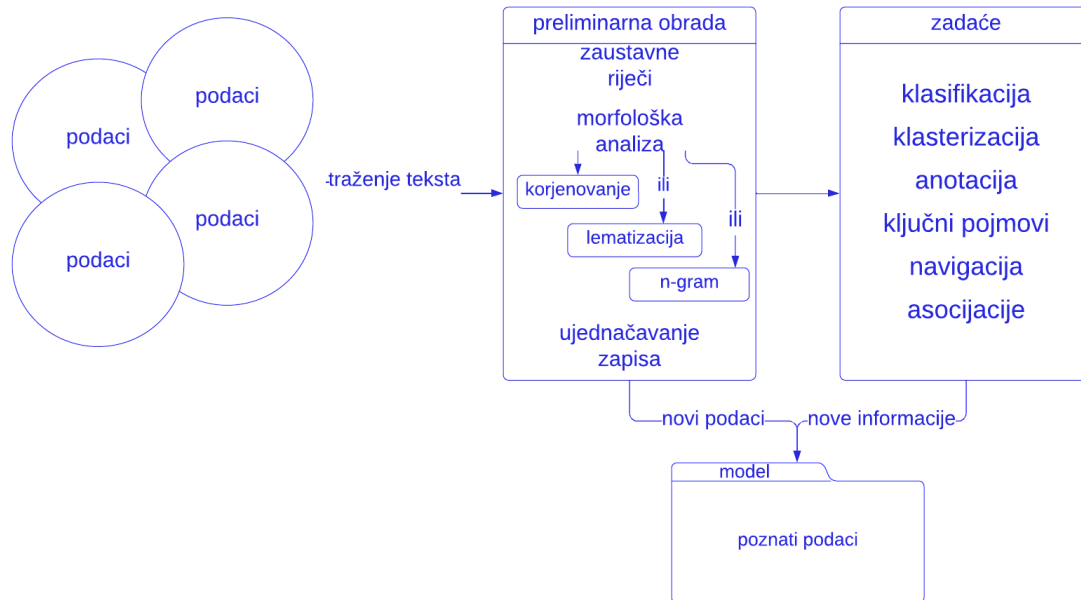
Slika 2. Sastavnice procesa text mining (*Text Mining Concept*, 11. 7. 2024.)

Korjenovanje (engl. *stemming*) vrsta je morfološke analize kojom se određuje korijen riječi. Taj korijen Šnajder (kao što citira Pandžić) preciznije naziva pseudokorijenom jer se ne radi o pravome morfološkom korijenu. Nadalje autori navode da se time pokušava dobiti “normalan oblik riječi” (Barsegjan i sur., 2007a, str. 196), a normalan oblik opisuju kao oblik riječi koji nije sklanjani ili u množini. Prema njihovoj definiciji, ono na što su autori mislili jest kanonski oblik riječi (rus. *исходная форма*, transliterirano *ishodnaja forma*), što bi podrazumijevalo nominativ jednine za imenice, infinitiv za glagole, nominativ jednine muškoga

roda za pridjeve i tako dalje. Taj se proces zapravo naziva lematizacijom. Jacob Murel i Eda Kavlakoglu (2023) u članku posvećenom upravo korjenovanju i lematizaciji pojašnjavaju razliku između ovih dvaju termina: “Literatura uglavnom opisuje korjenovanje kao proces uklanjanja afikasa s riječi kako bi se dobili nizovi korijena, a lematizaciju kao obuhvatniju aktivnost uklanjanja morfoloških varijacija kako bi se dobio temeljni rječnički oblik”. Još je jedna mogućnost analiza n-grama, odnosno nizova znakova neovisno o vrsti riječi, značenju i afiksima. Posljednji korak preliminarne analize jest ujednačavanje oblika zapisa koje se odnosi na povezivanje riječi s obzirom na njihov grafički izgled – riječ može biti napisana malim slovima, velikim slovima ili njihovom kombinacijom, ali dokle god analiza iz prethodnog koraka daje isti rezultat, one će sve biti smještene u istu jedinicu.

Tijekom analize izvora *text mining* stremi rješavanju nekoliko zadataka koje mogu biti ili njegov krajnji cilj ili tek jedan u nizu koraka kojima će se riješiti neka druga, složenija zadataka. Osnovne su zadatke: klasifikacija, klasterizacija, automatska anotacija, određivanje ključnih pojmova, navigacija i traženje asocijacija (Barsegjan i sur., 2007a, str. 197–198). Možemo primijetiti da ove zadatke u sebi već podrazumijevaju razumijevanje različitih razina prirodnoga jezika. Osim toga ove su zadatke međusobno usko povezane. Bez razumijevanja jezika odabrani model ne može podijeliti tekstove u klase, odrediti koje su njihove glavne karakteristike te sukladno tomu po čemu se razlikuju ili po čemu su slični i tako ih grupirati (klasterizacija). Određivanje ključnih pojmova omogućuje navigaciju po tekstu i traženje asocijacija, odnosno veza između ključnih pojmova, a automatska anotacija nastaje povezivanjem konteksta s ključnim pojmovima i asocijacijama. Autori ističu složenost ovoga procesa jer model prolazi korake svojevrstne morfološke i leksičke analize u kojoj određuje gramatičke karakteristike riječi i pretražuje postojeće rječnike kao pomoć u povezivanju značenja; zatim prolazi sintaktičku analizu u kojoj se sintagme ili riječi koje se najčešće nalaze u neposrednoj blizini analiziraju kao cjelina u odnosu na kontekst u kojemu se nalaze; na koncu se analizira odnos teme i upotrijebljenih riječi kako bi se odredila veza između klasifikacije teksta i iskorištenih termina (Barsegjan i sur., 2007a, str. 199). Sve informacije i nova znanja dobiveni iz jednoga teksta ulaze u radni fond modela te on time dobiva iskustvo za obradu svih sljedećih tekstova koje će dobiti ili pronaći kako bi ih obradio (posljednje dvije sastavnice procesa prikazane na slici 2). Sumiran cijeli proces analize koju prolazi *text mining* prikazan je na slici 3.





Slika 3. Slijed obrade podataka tijekom text mininga

Pohranjivanje znanja, proces razvoja i iskustvo koji su dosad bili spominjani u ovome radu teško je nazivati ikako drukčije s namjerom da se izbjegnu poveznice s načinom na koji uče ljudi. Proces po kojemu modeli generativne umjetne inteligencije uče nalikuju procesima ljudskoga učenja upravo zato što je zamišljen po uzoru na ljudsko učenje, ali treba biti oprezan s poistovjećivanjem ovih dvaju nizova radnji. Djeca uče jezik na osnovi izvanjezične stvarnosti u kojoj se nalaze, povezuju označitelje i označeno, a to kako se jezik uistinu uči i koristi poznato je uglavnom u okvirima poteškoća koje nastaju uslijed oštećenja nekoga od dijelova mozga. Tako je primjerice poznato da je Wernickeovo područje “dio sljepoočnoga moždanog režnja koji je odgovoran za jezično razumijevanje” (*Wernickeovo područje*, bez dat.). Pod pojmom jezično razumijevanje primarno podrazumijevamo poznavanje značenja riječi (odnos označitelja i označenoga) te smisleno povezivanje i upotrebu riječi (Moawad, 2023). Primjećujemo da su to osnovne funkcije na kojima se temelje dobro razvijeni modeli generativne umjetne inteligencije. Iskander Mil’zigin i Ljudmila Baeva (2017) u uvodu svojega članka ističu da je baratanje jezikom kreativni postupak koji zahtijeva poznavanje strukture jezika, ali i mogućnost snalaženja u fleksibilnosti te strukture. Podaci pretraživani i analizirani tehnikama *text mininga* prolaze kroz i nadograđuju strukturu zamišljenu i nazvanu prema načinu funkcioniranja mozga, konkretnije neuronske mreže.

## 2.3 Umjetne neuronske mreže

Mogućnost učenja i (samostalnog) razvoja zajednička je mozgu i modelima generativne umjetne inteligencije, što znači da oboje imaju neku ishodišnu točku ili barem početno stanje od kojega zatim prolaze kroz ranije navedene procese. Ti procesi ne moraju nužno voditi do neke zamišljene krajnje točke jer, kao što logika nalaže, ljudski mozak ne doseže točku nakon koje više ništa ne može naučiti ili u kojoj posjeduje konačnu zalihu informacija kojom raspolaže. Istim slijedom razmišljanja pretpostavljamo da se tako razvijaju i umjetne neuronske mreže – od početne točke do bilo kojega zadovoljavajućeg stanja koje se uvijek može promijeniti ovisno o potrebama korisnika.

Pokušat ćemo razjasniti analogiju koja je nastala između prirodne i umjetne neuronske mreže, odnosno zašto je određeni skup algoritama prozvan upravo neuronskom mrežom. Kora mozga sastavljena je od milijuna živčanih stanica (neurona) koje se međusobno povezuju u svojevrsnu mrežu. Kako objašnjavaju Liane Gabora i Scott B. Kaufman (2010) ono što živčane stanice raspoređene po cijelome tijelu impulsima prenose do mozga u njemu stvara sjećanja, stvara ili jača veze između pojedinačnih neurona u mozgu. U razvijenomu mozgu *Homo sapiens*a osim sjećanja na vanjske podražaje postoje i apstraktne misli te sjećanja na njih, a ona tvore cjelokupno pamćenje. Pojednim dijelovima pamćenja može se pristupiti prema podražaju ili sadržaju, a u prizivanju pojedinog sjećanja sudjeluje više različitih neurona te, sukladno tomu, jedan neuron sudjeluje u aktivaciji više sjećanja (Gabora i Kaufman, 2010, str. 5). Tako će podražaj aktivirati neuron povezan sa sjećanjem na takav ili njemu sličan prijašnji podražaj, ali i sve povezane neurone koji ne moraju nužno imati konkretnu sadržajnu povezanost s izvornim podražajem. Apstraktna misao u ovome slučaju igra značajnu ulogu jer označava sposobnost aktivacije neurona koja nije uzrokovana vanjskim podražajem, već isključivo sjećanjem samim, odnosno mišlju – misao rađa misao. Jedna od promjena koja će biti značajna za daljnju diskusiju u ovome radu jest ona do koje dolazi u principu reagiranja neurona.

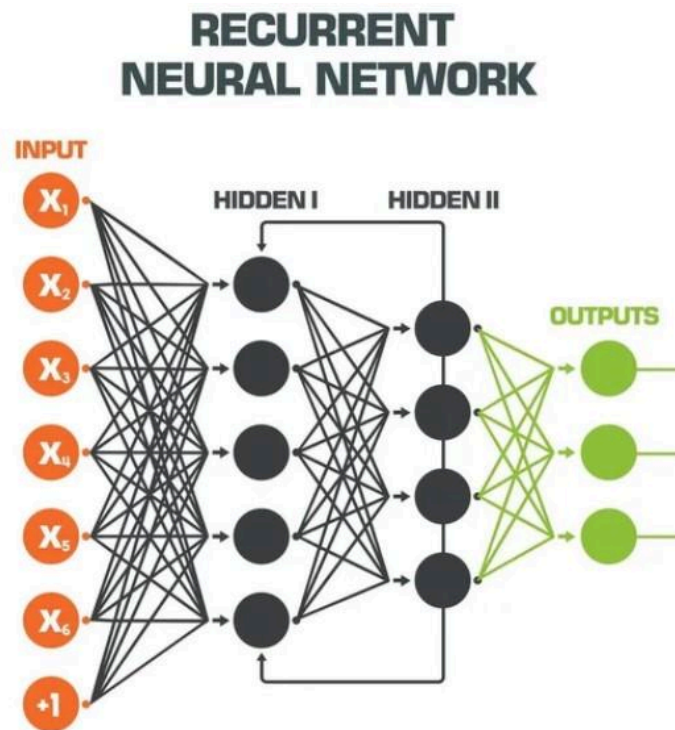
Kako sjećanja, misli i veze između neurona postaju kompleksnije, tako neuroni postaju osjetljiviji na fragmente podražaja, što znači da je i sama naznaka podražaja dovoljna da se aktivira cijeli jedan sklop u neuronskoj mreži i širi se u bilo kojemu smjeru. Neuroni koji su bliži i jače povezani reagiraju na fragmente koje povezuje neka sličnost, a oni udaljeniji reagiraju na

ono što ih razlikuje. Drugim riječima aktivacija i reagiranje neurona ne temelje se isključivo na bliskosti, sličnosti i najvišim vjerojatnostima povezivanja, već i na vrsti veze, koja može odabrati najslabiji ili najudaljeniji put kojim bi neka misao mogla proteći. Gabora i Kaufman (2010) nazivaju ovaj proces stvaranjem kreativnih veza.

Barsegian i sur. (2007b, str. 79) opisuju (umjetnu) neuronsku mrežu kao “razred modela temeljen na biološkoj analogiji s ljudskim mozgom koji je namijenjen različitim zadacima analize podataka nakon takozvanoga stadija obuke na postojećim podacima”. Ova definicija govori nam da se radi o sklopu sličnih sastavnica koje rade zajedno kako bi postigle neki rezultat kroz već spomenuti korak treniranja na unaprijed zadanim podacima. Početni je podatak u ovome slučaju prvotni podražaj. On ulazi u mrežu neurona (koji se jednako nazivaju i na ruskome i na engleskome jeziku), koji su pak posloženi u slojeve (rus. *слои*, transliterirano *sloi*, engl. *layers*). Andrej Karpathy, jedan od osnivača tvrtke *OpenAI*, opisuje slojeve kao računske korake (engl. *computational steps*), vraćajući time fokus s biološke na računalnu stranu tehnologije. Naglasivši da se radi o računanju, postavlja se pitanje što se to računa unutar neuronske mreže. Pojednostavljeno rečeno računa se vrijednost, odnosno stanje podatka od ulazne do izlazne točke, i to u obliku vektora. Karpathy ističe da veliki napredak u radu neuronskih mreža predstavlja mogućnost promjene vrijednosti vektora. Uklanjanje ograničenja koje traži da ulazna vrijednost bude određene veličine ili oblika kako bi izlazna vrijednost bila unaprijed točno zadanog oblika omogućava neuronskoj mreži da slobodno radi s podacima i pronalazi optimalan način za njihovu obradu te njihov konačni oblik u rezultatu (Karpathy, 2015).

Posebna vrsta neuronskih mreža koje se koriste u svrhe koje će za ovaj rad biti važne su povratne neuronske mreže (engl. *recurrent neural network*, *RNN*). povratne neuronske mreže “kombiniraju ulazni vektor s vektorom stanja i fiksiranom (ali naučenom) funkcijom kako bi stvorile novi vektor stanja” (Karpathy, 2015). Neuronska mreža prati kako se obrađuju i mijenjaju podaci tijekom obrade, što se događa s njima prilikom prolaska kroz svaki sloj mreže, u kojim se smjerovima kreću (proces nije linearan od ulaza do izlaza) i kako se cjelokupno stanje mreže mijenja ovisno o zaključcima izvedenima iz obrađenih podataka. Naime povratna neuronska mreža osim podataka analizira i nadzire svoj napredak. Liza S. Ardenij opisuje ovu mogućnost neuronske mreže:

“na osnovi pokušaja i pogrešaka neuronska mreža akumulira informacije o svojem radu i dobivenomu rezultatu te tako usavršava svoj rad. Program ili sustav koji se zasniva na neuronskoj mreži s vremenom daje kvalitetnije rezultate zadanih zadataka za razliku od običnih programa. Akumulira se programski kod u kojemu je fiksirano ‘iskustvo’ takvoga programa” (2019, str. 77).



Slika 4. Povratna neuronska mreža (Neural Network Algorithm, 11. 7. 2024.)

Radi se o mogućnosti koja se nalazi u pozadini onoga što Karpathy naziva “varljivo jednostavno aplikacijsko programsko sučelje” (2015). Na površini, odnosno u svojem kodu, ono uistinu prihvaća ulazne vrijednosti i nudi neke druge izlazne vrijednosti, ali te se izlazne vrijednosti ne formiraju samo na osnovi vidljivih ulaznih podataka, već i na svim prijašnjim podacima i njihovim stanjima koje model pamti i koristi za svaku novu analizu. Ilustracija procesa skrivenoga između ulaznih i izlaznih podataka prikazana je na slici 4. Kada je riječ o obradi prirodnih jezika, model prvo uči i pretpostavlja znakove i nizove znakova, a među njima prepoznaje bjeline, koje isprva dolaze nasumično, ali u daljnjem procesu pomognu modelu da nauči raspoznavati riječi. Logičnim slijedom analize model nauči kratke, zatim duže riječi, pa kao što je bilo spomenuto u prethodnome poglavlju, sintagme ili riječi koje se često pojavljuju zajedno ili u nekom određenom slijedu, nakon čega slijede cijele rečenice i tako dalje (Karpathy,

2015). Možemo reći da za razliku od čovjeka koji materinski jezik uči slušanjem, umjetna neuronska mreža jezik uči gledanjem. Izračunat će vjerojatnosti kada se pojavljuje koji znak (uključujući bjeline), naprimjer znak *a* pojavljuje se u kombinaciji s bjelinom (dva znaka: *a\_*), a nakon toga može uslijediti bilo što (uči da *a\_b* nije povezani skup znakova, mogao bi biti dio niza *lijepa\_boja*, *on\_igra\_biljar* ili štogod drugo). U ovoj fazi model prepoznaje pseudorečenice, to jest nizove znakova na pozicijama izračunatima po vjerojatnosti.

Govoreći o hrvatskome jeziku, model bi mogao primijetiti da često u kombinaciji s bjelinom dolaze znakovi *a*, *e*, *i*, *u*, i *m* (posljednji glasovi oblikotvornih sufiksa za e-deklinaciju imenica). S više ponavljanja ili više ponuđenog teksta model primjećuje da se određeni skupovi znakova uvijek pojavljuju unutar dviju bjelina, pohranjuje tu informaciju i počinje shvaćati kako izgleda riječ te da ona ima svoje unutarnje ustrojstvo: kada niz počinje s *\_knj*, najveća je vjerojatnost da će se nastaviti znakom *i*, odnosno kombinacijom *ig* ili *iž* (knjiga, knjigom, knjižnica, književni itd.).

Naravno, dani primjeri pojednostavljena su shema rada vrlo kompleksnih modela i sklopova algoritama potrebna radi lakše vizualizacije procesa. Ovisno o tome koji je glavni fokus rada, bila to tehnička strana procesa ili samo njegovi rezultati, autori zauzimaju različita stajališta o danoj temi. Tako Luke Tredinnick i Claire Laybats (2023) u svome članku upozoravaju na proces koji se događa između ulaznih i izlaznih podataka koji smo upravo opisali:

“Nema ničega fundamentalno drukčijega u načinu rada simuliranih neuronskih mreža u usporedbi s tradicionalnim algoritmima. Zapravo, koliko god izvanredne bile izlazne vrijednosti koje one stvore i koliko je god ta intrinzična priroda skrivena iza režima treniranja, simulirane neuronske mreže ostaju algoritamski procesi. Generativna umjetna inteligencija jednako je obvezana pravilima kao i bilo koji algoritamski proces. Razlika je u tome što su pravila izvedena iz treniranja na podacima *post hoc*” (str. 99).

Isto ističe i Anna M. Plotnikova (2023) te nudi dva opisa pojma neuronska mreža prema kojima se mogu podijeliti i pogledi onih koji se bave ovom temom. Prvo se shvaćanje odnosi na sagledavanje umjetne neuronske mreže isključivo kao algoritma (kako to rade Tredinnick i Laybats), a drugo se odnosi na čovjekoliki sustav, s time da taj sustav nije samo nastao po uzoru na ljudski mozak, već ima potencijal jednak ljudskom mozgu (Plotnikova, 2023, str. 49). S

obzirom na to da se ne zna što se točno događa u svakomu pojedinom koraku kroz koji podatak prolazi u neuronskoj mreži i koje je sve zaključke model u tom procesu izvukao, Tredinnick i Laybats (2023) proces treniranja i svake sljedeće obrade podataka u modelima generativne umjetne inteligencije nazivaju crnom kutijom (engl. *black box processes*, str. 100). Isticanje nerazumijevanja detalja procesa dovodi nas do upotrebe glagola u kontekstu generativne umjetne inteligencije i zapažanja da umjetna neuronska mreža može razumjeti na bilo kojemu jeziku, ali da ne može “sanjati, zamišljati, maštati”, bez obzira na to što može “izmišljati” (Plotnikova, 2023, str. 49). U Uredbi predsjednika RF (2019/2024) stoji da su veliki generativni modeli (rus. *большие генеративные модели*, transliterirano *bol'shie generativnye modeli*) sposobni interpretirati i stvarati podatke u različitim medijima. Interpretiranje podrazumijeva pružanje potrebnih informacija na osnovi upita, a uz različite medije navedeni su tekstovi, slike, videomaterijali te sve slično tomu uz naglasak da se rezultati mogu proizvoditi “na razini koja je usporediva s rezultatima intelektualne djelatnosti čovjeka ili čak na višoj razini” (2019/2024). Jednako tako jasnoća pogleda na tehnologiju gubi se kada se ona ne prikazuje kao inspirirana prirodom (neuronska mreža inspirirana radom ljudskoga mozga), već se pokušava nametnuti razmišljanje da priroda radi poput stroja, odnosno da su procesi u mozgu jednaki onima kroz koje prolazi neki niz algoritama (Elam, 2023).

Može se zaključiti da stručnjaci koji se bave proučavanjem rezultata umjetnih neuronskih mreža imaju konfliktan stav prema napretku. S jedne strane smatraju da tehnologija može i svakako hoće postati bolja od čovjeka, što je pozitivno jer unaprjeđenje tehnologije je ljudsko postignuće. Ali s druge strane gledaju na taj napredak negativno jer ljudi moraju pod svaku cijenu održati nadmoć u nekoj neodređenoj kategoriji nad nečime za što niti ne znaju kako točno funkcionira. Iako su povratne neuronske mreže velik korak u razvoju umjetne inteligencije, uvijek postoji mogućnost da se model preuči, što ga ne bi odvelo ni u kojemu korisnom smjeru. Drugim riječima ako se model stalno uči na istim podacima, on počinje nuditi ista rješenja za sve slične vrste upita, što nikako ne možemo nazvati opasnim po buduće korištenje ljudskih potencijala.

### 3. Obrada i generiranje prirodnog jezika

Model generativne umjetne inteligencije dobio je podatke koje može obrađivati, proučavati i analizirati te je u mogućnosti ponuditi rezultate varijabilne kvalitete. S obzirom na to da podaci kojima se ovaj rad bavi predstavljaju tekstove ili tekstualne upite na prirodnome jeziku, razumljivo je da računalne znanosti ne mogu same ponuditi sva objašnjenja koja bi se ticala pojedinih postupaka “razumijevanja” i strukturiranja jezika od strane modela. Nakon *text mininga* i obrade podataka u umjetnoj neuronskoj mreži, model se može posvetiti koherenciji teksta koji korisnik dobiva kao rezultat svoga upita. U ovome dijelu procesa glavnu ulogu u nadzoru rada preuzima računalno jezikoslovlje. Kako se tijekom razvoja umjetnih neuronskih mreža za nit vodilju uzimao ljudski mozak i možebitni načini na koje ljudi uče jezik, tako se u procesu generiranja teksta oslanja na procese ljudskoga stvaranja i kreativnosti kako bi se pratio napredak razvoja modela.

#### 3.1 Računalno jezikoslovlje

Računalno jezikoslovlje (također: računalna lingvistika) grana je općega jezikoslovlja, odnosno “interdisciplinarno znanstveno područje povezano s računalnim modeliranjem prirodnoga jezika utemeljenim na statistici ili pravilima te s proučavanjem računalnih pristupa jezičnim pitanjima” (*Računalno jezikoslovlje*, bez dat.). Ova se grana jezikoslovlja fokusira na jezičnu analizu danoga teksta, ali u isto vrijeme uzima u obzir da je tekst produkt nekoga računalnog procesa. Zbog te specifičnosti izvora teksta, jezik se pokušava približiti računalnim procesima koji se temelje na matematičkim operacijama te mu se dodjeljuju brožčane vrijednosti kako bi se on mogao statistički analizirati, stvarajući time svojevrsnoga posrednika između računalnoga koda i prirodnoga jezika, odnosno “istraživački su naponi kod CL-a<sup>2</sup> usmjereni na otkrivanje kako su nizovi riječi povezani s njihovim značenjima te na razvoj formalizama pomoću kojih su opisane te povezanosti. Fokus je kod CL-a na jeziku, a računalni su algoritmi u funkciji potpore” (Šuman, 2021, str. 373).

Područje istraživanja računalnoga jezikoslovlja vrlo je široko. Strojno prevođenje, pretraživanje informacija, sažimanje i anotiranje teksta glavna su područja kojima se računalno

---

<sup>2</sup> engl. *Computational linguistics*

jezikoslovlje bavi na sadržajnoj razini teksta. Jedna od prvih uspješnije izvršenih zadaća bilo je formiranje modela za automatsku provjeru i ispravljanje teksta (Bol'šakova, 2017). Provjera pravopisa jedna je od jednostavnijih provjera koje posebni alati mogu vrlo točno izvršiti ako im se jasno zadaju pravopisna pravila jezika na kojemu se radi. S druge strane gramatika je ipak nešto drugačija. Naprimjer ponekad nije dovoljno provjeriti je li riječ točno napisana, već i je li ona u dobrome obliku s obzirom na kontekst. Tako kod flektivnih jezika neki alati za provjeru još uvijek neće prepoznati da se u sintagmi "pisati knjigom" radi o pogrešnome padežu, odnosno da riječ *knjigom* postoji u tome obliku, ali kao izravni objekt uz glagol *pisati* prema njegovoj rekciji mora doći u akuzativu. Osim navedenoga, u sklopu računalnoga jezikoslovlja proučavaju se i kategorizacija i klasifikacija teksta te, ono najzahtjevnije, mogućnost održavanja dijaloga na prirodnome jeziku, formiranje odgovora prema upitu na prirodnome jeziku, generiranje teksta te prepoznavanje i sinteza govora (Ibid, str. 13).

Neovisno o tome koja je od navedenih funkcija (ili više njih) cilj pojedinoga programa, svaki se sastoji od nekoliko modula od kojih je svaki zadužen za jednu razinu jezične analize. Elena I. Bol'šakova (2017) tako navodi module od najniže razine raspoznavanja riječi pa sve do značenja cjelokupnoga teksta: modul za gramatičku analizu, koja zapravo obuhvaća određivanje granica riječi; modul za morfološku analizu; modul za sintaktičku analizu te modul za semantičku i pragmatičku analizu, koji su ključni za funkcionalnu komunikaciju s modelom.

Modul koji je zadužen za gramatičku analizu većim je dijelom već objašnjen u sklopu tehnologije *text mining* i procesa umjetnih neuronskih mreža koje počinju između skupova znakova raspoznavati riječi. Modul za morfološku analizu teksta onaj je u kojemu nastaju prve veće razlike koje ovise o tome na kojemu se od prirodnih jezika radi, odnosno koje su njegove karakteristike. Ovisno o razini fleksije, upotrebljava se korjenovanje za manje flektivne jezike ili lematizacija za jezike s razvijenom fleksijom. Neovisno o tome koji se od ovih dvaju postupaka koristi, modul za morfološku analizu može odrediti vrstu riječi, gramatičke kategorije riječi te morfove koji tvore pojedinu riječ (Klyšinskij i Sanin, 2017).

Bol'šakova (2017) dijeli module za morfološku analizu prema rezultatu koji daju, prema već spomenutim metodama same analize i prema načinu rada s rječnikom. Rezultati tako mogu biti kanonski oblik riječi koji se povezuje s rječnikom ili osnova potkrijepljena gramatičkim kategorijama pojedine riječi. Šuman značajno pojednostavljuje i generalizira morfološku analizu



kao određivanje *stema* i afikasa, odnosno račšlambu riječi na primarne i dodatne dijelove riječi (2021, str. 374). Što se tiče rječnika, za početne se analize mogu koristiti gotovi rječnici danoga jezika ili model može pristupiti analizi bez njega. S druge strane, rezultati morfološke analize pohranjuju se u posebne rječnike kako bi im model mogao ponovno pristupiti prilikom sljedeće analize. Rječnik se nadopunjuje i pomaže pri novim analizama kako se cijeli postupak ne bi morao svakoga puta u potpunosti ponavljati. Najčešći formati koji se koriste za izradu rječnika su XML, TEI i CoNLL (Klyšinskij i Sanin, 2017).

Posebna je situacija primjećivanje jezičnih specifičnosti poput homonimije i njezino rješavanje. Za takve specijalne zadatke upotrebljavat će se i posebne metode unutar modula, što znači da čak i ako modul radi na principu lematizacije, može koristiti n-grame koji obuhvaćaju više riječi kako bi se dobio kontekst u kojemu je homonim iskorišten. Na taj način model može naučiti razlikovati značenja homonima (Bol'shakova, 2017).

Moduli za sintaktičku, semantičku i pragmatičku analizu nisu toliko samostalni kao onaj za morfološku analizu. S obzirom na to da ova tri modula trebaju tekst veći od jedne riječi kako bi analiza bila izvršena, oni ispunjavaju različite zadatke koje ih mjestimično udružuju na različitim razinama. Dvije zadatke na koje se u današnje vrijeme stavlja poseban fokus su izdvajanje informacija i određivanje tona teksta. Treniranje modela provodi se uglavnom na novinskim člancima i komentarima na društvenim mrežama, ali pretpostavlja se da će se ove analize moći koristiti i na drugim vrstama tekstualnih podataka. Izdvajanje informacija može uključiti jedan tekst ili veći broj tekstova. Postupak može biti točniji ako se radi na većem broju tekstova jer model može uspoređivati sve informacije koje pronađe te prema učestalosti njihova ponavljanja odrediti koje su najvažnije. U vrijeme bezbrojnih portala s vijestima i kada se članci ili njihovi dijelovi recikliraju iz stranice u stranicu, ovako istrenirani model može sastaviti skup važnih informacija koje je pronašao te tako korisnicima uštedjeti vrijeme koje bi utrošili na pretraživanje (Bol'shakova i Efremova, 2017).

Zadaca određivanja tona teksta ponešto je složenija jer se odvija ponajprije na semantičkoj razini. Ponekada su to jednostavni slučajevi u kojima se pronalaze riječi pozitivna ili negativna značenja, koje onda mogu ukazati modelu na to što autor misli o temi, odnosno kakav stav zauzima. S druge strane komentar na društvenoj mreži ili komentar u smislu recenzije nekoga proizvoda ne mora uopće sadržavati jasno pozitivne ili negativne riječi poput *odlično*, *u*

*redu* ili *loše, grozno*. Tako na internetskim stranicama poput *Goodreads* možemo pronaći komentar iz 2020. godine: “Čemu služi globalna pandemija ako ne čitanju ovakvih knjiga?” S obzirom na to da se radi o komentaru na *Zločin i kaznu* možemo pretpostaviti da korisnik želi ukazati na to da je roman dugačak te da čitatelji tijekom prisilnog boravka kod kuće konačno imaju dovoljno vremena posvetiti se takvome djelu, ali ne saznajemo je li se ovome konkretnome korisniku roman svidio ili ne. U ovakvim slučajevima pomaže sustav ocjenjivanja zvjezdicama ili ocjenama (od 1 do 5 ili od 1 do 10) koje modelu mogu ukazati na to da je ovo bio pozitivan komentar ako je ocijenjen s pet zvjezdica.

Natal'ja V. Lukašević (2017) osvrće se upravo na problem razumijevanja teksta, navodeći da se model može trenirati na unaprijed zadanim pozitivnim i negativnim tekstovima, ali sastavljanje korpusa oduzima mnogo vremena. Zatim navodi da istrenirani modeli nisu lako prenosivi, naprimjer model istreniran na komentarima knjiga neće dobro funkcionirati na komentarima koji se odnose na najnoviji model nekoga kućanskog aparata. Posljednje, ali ne i manje važno, jest to da je teško ispraviti rezultate koje model daje jer je ponekad teško objasniti u čemu je bila pogreška, odnosno zašto je model krivo opredijelio pojedini tekst (Lukašević, 2017). U članku se također spominju ironija i sarkazam kao izrazito problematične pojave jer pozitivan iskaz zapravo ima negativno značenje i obrnuto, ali autorica ne nudi rješenje koje bi poboljšalo analizu.

### **3.2 Postupci i mogućnosti generiranja teksta**

Spomenuvši mogućnost izdvajanja informacija iz teksta i razlikovanja važnih od nevažnih informacija, može se zaključiti da je situacija poprilično jasna, barem što se tiče novinskih članaka. Tekst koji ima specifičnu strukturu, odnosno sastavljen je od određenog broja i redoslijeda informacija koje nose neki ključni pojmovi, nije osobito složen za reproducirati. Modeli generativne umjetne inteligencije trenirani na vijestima ili s pristupom portalima s novostima vrlo će brzo početi nuditi slične tekstove jer jedan od zadataka semantičke analize jest prezentiranje važnih informacija korisniku. No već u samim počecima razvoja generativne umjetne inteligencije stručnjaci su se upuštali u generiranje književnih tekstova različitih vrsta, što je zahtijevalo sasvim drugačiji pristup cijelome procesu. Ponekad se tomu procesu pristupalo s potpuno tehničke strane, a to je podrazumijevalo sastavljanje velikog fonda za treniranje te

naprosto čekanje da model sam shvati što mu je činiti s podacima. Tek u posljednjem desetljeću s razvojem dubokog učenja možemo vidjeti značajniji pomak pri prepoznavanju, razumijevanju i generiranju prirodnoga jezika (Schwartz, 2018).

Okretanje ka književnosti može se smatrati ispitivanjem granica mogućnosti tehnologije. Licklider (kao što citira Schwartz) razrađuje principe rada u kojima čovjek i stroj rade ono u čemu su najbolji te ostvaruju vrlo efikasnu suradnju koja može dovesti do “poboljšanja kapaciteta i ljudske i računalne inteligencije” (Schwartz, 2018, str. 94). Računalu su u ovome slučaju namijenjeni rutinirani zadaci koje može izvršiti mnogo brže od čovjeka, svakoga puta s jednakom preciznošću, a na čovjeku je da određuje svrhu i cilj rada.

U slučaju korištenja generativne umjetne inteligencije kakvu poznajemo danas, možemo reći da je čovjekova zadaća ostala osmišljavanje i zadavanje upita koji predstavlja smjernice za cilj koji računalo može postići, ali s druge strane neobično je shvaćanje da je stvaranje književnoga teksta rutinirana radnja poput matematičkih izračuna ili izdvajanja informacija iz zadanoga teksta. Zato pristupamo ovoj temi kao eksperimentu.

Proučavanje rezultata rada, odnosno generiranih tekstova vrlo jasno pokazuje koje su (trenutne) mogućnosti pojedinoga modela generativne umjetne inteligencije. Očekivani ili neočekivani odgovori na upite korisnika te jezična analiza tekstova mogu otkriti kako model koristi podatke na kojima je istreniran. Nakon postavljanja upita na prirodnome jeziku, pokreću se svi ranije objašnjeni mehanizmi pretraživanja i analize podataka te obrnuti proces sklapanja odgovarajućih podataka u odgovor koji je korisniku razumljiv.

Zanimljiv je stadij u kojemu se modeli trenutno nalaze kada je u pitanju razumijevanje podataka i razumijevanje književnoga teksta. Za kratak eksperiment odabrano je pitanje *Kako je kažnjen Raskoljnikov?*, zadano 25. travnja 2024. na hrvatskom, engleskom i ruskom, dva ChatGPT-ju verzije 3.5 i dva Microsoftovu Copilotu (prikazi na slikama 5, 6 i 7).

как был наказан раскольников?



Раскольников, главный герой романа "Преступление и наказание" Фёдора Достоевского, был наказан внутренним моральным страданием и сожалением за совершенное им преступление. Физический наказательный акт со стороны закона не произошел в романе.

Однако Раскольников самостоятельно претерпел мучительное внутреннее переживание и моральное разочарование из-за своих действий. Он начинает испытывать угрызения совести и мучительные сны, в которых его преступление преследует его. С каждым днем он погружается все глубже в душевный кризис, осознавая тяжесть своего греха и невозможность жить нормальной жизнью после этого.

Кульминацией этой моральной кары становится его встреча с Соней Мармеладовой, женщиной, испытавшей множество жизненных испытаний, и ее способность прощать и видеть в человеке добро даже после совершения ошибок. Эта встреча приводит Раскольникова к внутреннему обращению и пониманию своих истинных чувств.

### *Slika 5. Snimka zaslona; ChatGPT.*

kako je raskoljnikov kažnjen?

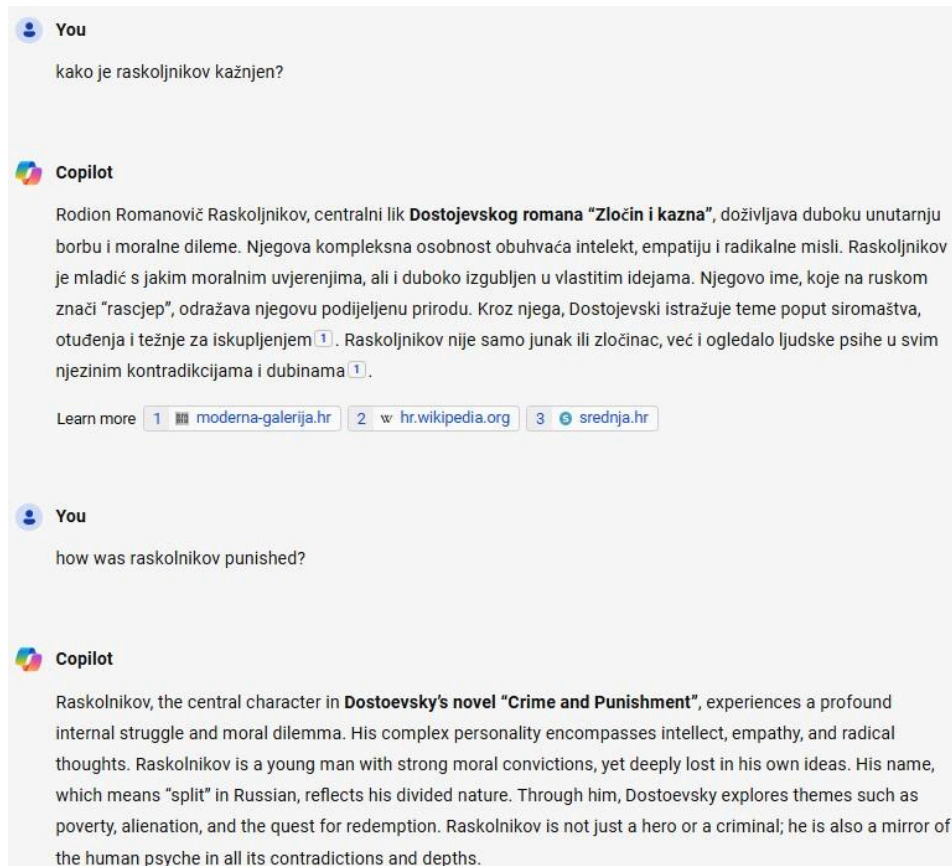


Raskoljnikov, glavni lik romana "Zločin i kazna" Fjodora Dostojevskog, kažnjen je na nekoliko različitih razina:

1. **Unutarnja kazna:** Raskoljnikov prolazi kroz intenzivnu unutarnju borbu i duševnu patnju nakon što počinu ubojstvo starice Alione Ivanovne. Osjećaj krivnje i unutarnji konflikti dovode ga do emocionalnog sloma i dubokog moralnog propitkivanja.
2. **Fizička kazna - izgnanstvo u Sibir:** Kao posljedica svojeg zločina, Raskoljnikov biva osuđen na izgnanstvo u Sibir. Iako se to ne prikazuje izravno u romanu, na kraju se saznaje da je Raskoljnikov poslan u Sibir, što je teška fizička kazna koja predstavlja gubitak slobode i udaljenje od doma.
3. **Mentalna i emocionalna transformacija:** Tijekom romana, Raskoljnikov prolazi kroz duboku unutarnju transformaciju. Pod utjecajem svojih iskustava, osobito susreta s Sonjom Marmeladovom, koja simbolizira suosjećanje i iskupljenje, Raskoljnikov postupno počinje prihvaćati svoju krivicu i tražiti put prema moralnom obnovi.

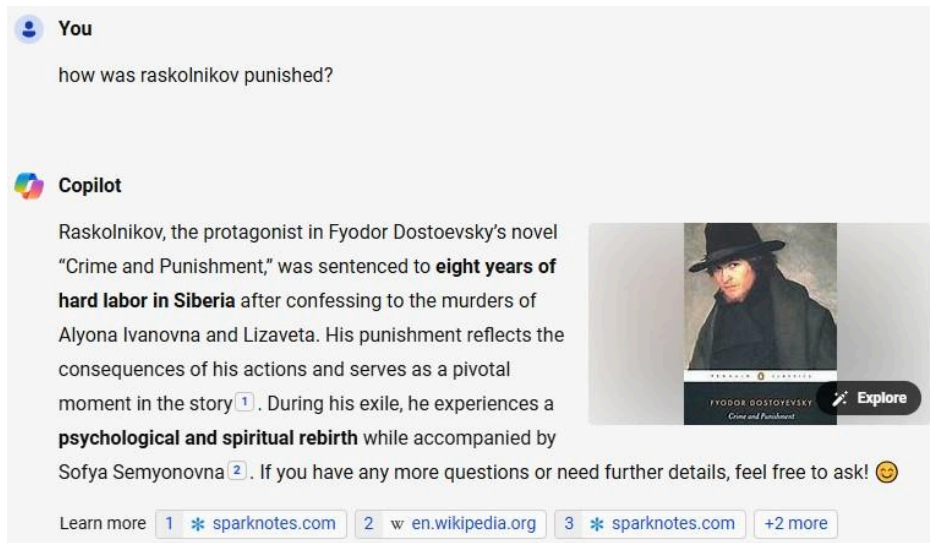
Kazna koju Raskoljnikov doživljava nije samo vanjska, već prije svega unutarnja. On se suočava s vlastitim demonskim stranama, moralnim propadanjem i unutarnjim sukobima te postupno prepoznaje važnost empatije, suosjećanja i etičkih vrijednosti. Njegovo putovanje kroz romaneskni svijet Dostojevskog simbolizira složenost ljudske prirode i potrebu za duhovnom obnovom i iskupljenjem.

### *Slika 6. Snimka zaslona; ChatGPT.*



Slika 7. Snimka zaslona; Microsoft Copilot.

Pitanje na koje bez problema može odgovoriti svatko tko je pročitao roman *Zločin i kazna*, *chatovima* je predstavljalo problem u tri od četiri slučaja. Logično je bilo postaviti pitanje na ruskom s obzirom na to da se radi o ruskome klasičnom djelu, ali *ChatGPT* ne samo da odgovara da je Raskolnikov kažnjen unutarnjim moralnim patnjama i žaljenjem, već odlazi toliko daleko da naglašava da nije bilo fizičke kazne. *Copilot* na hrvatski i engleski upit odgovara gotovo jednako, ali za hrvatski odgovor nudi korištene izvore u kojima se odgovor može potvrditi, dok za engleski to ne nudi. Prilikom čitanja primjećujemo da je engleski odgovor prijevod hrvatskoga pa zato niti ne može ponuditi izvore. Bilo je potrebno otvoriti novi razgovor i ponoviti pitanje da bi *Copilot* ponudio originalniji odgovor (slika 8). S druge strane zanimljivo je što novi odgovor sadržava isključivo kaznu koju je odredio sud, a ne spominje "unutarnje kazne", a prva dva odgovora, odnosno hrvatski odgovor i njegov prijevod na engleski, uopće ne odgovaraju točno na pitanje.



Slika 8. Snimka zaslona; Microsoft Copilot.

Jedini točan i potpun odgovor iz prvog pokušaja ponudio je *ChatGPT* i to na upit na hrvatskome jeziku. Tragikomičnost situacije proizlazi iz toga što je hrvatski jezik zasigurno jezik s manje resursa od ruskog i engleskog, ali s daleko većom količinom dostupnih analiza književnih djela, sastavljenih i objavljivanih kako učenici ne bi morali čitati lektire.

Prvo što izravno utječe na odgovor jest jezik na kojemu je postavljen upit. Isti upit može biti postavljen na više različitih jezika i svaki će put dati drugačiji odgovor jer će se on bazirati na materijalima na tome jeziku. Tako naprimjer upit na hrvatskome neće nužno dati odgovor s informacijom koja je modelu dostupna samo na engleskome jeziku, čak i ako model ima mogućnost prevođenja teksta.

### 3.2.1 Prevoditeljske sposobnosti

Prije generiranja književnog teksta, a nakon analize podataka za treniranje, može se smjestiti još jedan prijelazni korak u razumijevanju generativnih modela umjetne inteligencije, a to je vanjska analiza teksta. Model razumije prirodni jezik i na njemu nudi odgovore te analizira dostupne podatke, no postavlja se pitanje može li transparentno analizirati tekst, odnosno prirodnim jezikom objasniti što je naučio iz svoje "računalne" analize. U osnovne vrste prijevoda ubrajaju se ne samo prijevod teksta, već i prijevod slike, odnosno prepoznavanje znakova u slikovnim datotekama i prijevod govora (Wang, Wu, He, Huang, Church, 2022). Kvaliteta

prijevida ovisi o kvaliteti podataka dostupnih o jezicima s kojih i na koje se prevodi, što bi značilo da model mora istovremeno baratati dvjema bazama podataka i posredovati između njih svojim stečenim znanjem. Wang i sur. (2022) dijele jezike na one bogate i siromašne izvorima te ističu kako nedovoljna količina podataka u jednome jeziku automatski uzrokuje lošu kvalitetu prijevida jer model ne može razviti osjetljivost za vokabular i kontekst upotrebe riječi. Čak i ovakvi modeli mogu djelomično biti od pomoći pri generalnom razumijevanju nekoga teksta jer model je svakako u mogućnosti ponuditi nekakvu izlaznu vrijednost, ali može li, osim povezivanja riječi u različitim jezicima, obavljati i druge prevoditeljske zadaće?

O zapažanjima iz ove sfere govore Marija V. Artamonova, Akbar A. Mambetov i Ekaterina V. Tulina (2023) s Katedre za lingvistiku i prevođenje Instituta humanističkog obrazovanja u svome radu u kojemu su generiranim tekstovima pristupili s pozicije prevoditelja. Lingvisti su isprobali tri *chatbota* zasnovana na *ChatGPT*-ju i zadali mu uobičajene prevoditeljske zadatke: određivanje učestalosti pojedinih riječi; traženje kolokacija; usporedbu paronima; traženje sinonima; odabir rime te preliminarnu prevoditeljsku analizu teksta.

Prvi zadatak nije izvršio niti jedan *chatbot*, iako je svaki ponudio rješenje. Postoje razni korpusi jezika u kojima se ovakve informacije mogu provjeriti, ali pokazalo se da *chatbot* ne može koristiti funkciju stranice koja pretražuje korpus. Rješenja koja su *chatbotovi* ponudili bila su izmišljena i neprovjerljiva (Artamonova i sur, 2023). U ovakvim slučajevima uobičajeno je reći da model halucinira jer ne zna se točno što ga je nagnalo da ponudi odgovor koji ne postoji nigdje u njemu dostupnim izvorima.

Kolokacije, paronime i sinonime *chatbotovi* su uglavnom pronalazili, uspoređivali i ispisivali bez problema, ali kada ih se zatraži da potkrijepe svoje odgovore izvorima, ponudili bi izmišljene poveznice ili poveznice na tekstove u kojima se ponuđene riječi uopće ne pojavljuju. Rimu također mogu ponuditi jer se u principu radi o pretraživanju riječi koje sadržavaju jednake nizove znakova na svojim krajevima, ali autori su primijetili problem koji bi mogao imati svoju varijantu u svakome jeziku, a to je izgovor. Tako je jedan od *chatbotova* ponudio par *peremen–vreden* (rus. *перемён–вреден*; *promjena* u genitivu množine, *štetan* u nominativu jednine muškoga roda). Naglasak u imenici je na posljednjemu slogu, a u pridjevu na prvome, zbog čega se riječi ne rimuju u potpunosti (Artamonova i sur, 2023, str. 245). S druge strane možemo dati jedan jednostavan primjer kao što je rima u paru *ruk–krug* (rus. *рук–круг*, *ruka* u

genitivu množine, *krug* u nominativu jednine) jer se u ruskome provodi obezvučenje na kraju riječi pa u izgovoru par glasi *ruk–kruk*, ali model to ne može prepoznati jer provjerava samo grafeme.

Posljednji su zadatak *chatbotovi* odradili najbolje. Na osnovi ulomka iz teksta kojemu nisu pridodani nikakvi podaci o izvoru, *chatbotovi* su sasvim točno pružali informacije o tome o kojoj se knjizi radi, tko je autor, kojemu žanru knjiga pripada i slično. Ta informacija ne čudi jer ta sposobnost spada u zadaće izdvajanja informacija te model generativne umjetne inteligencije ne mora s ponuđenim tekstom raditi ništa do li usporediti ga s dostupnim podacima i odrediti o kojoj se knjizi radi.

U svakome slučaju, strojni prijevod konačno može nakon mnogo godina razvoja dati dobre rezultate, ali jednako tako kako model generativne umjetne inteligencije može ponuditi bilo koji (točan) odgovor, ne može objelodaniti i objasniti svoje postupke. Prevođenje može biti potpomognuto i ubrzano dobro istreniranim modelom te u tome možemo vidjeti pozitivne strane poput smanjenja pritiska na prevoditelja i brže širenje tekstova. Ovo bi trebao biti pozitivan primjer suradnje čovjeka i tehnologije kako je prevoditelj i dalje taj koji bi morao pregledati rezultate koje je model ponudio. Jedina prava prijetnja prevoditeljskoj struci ne nalazi se u napretku tehnologije, već u nerazumijevanju mogućnosti i u nebrizi za kvalitetu rezultata.

### 3.2.2 Stil generiranoga teksta

Osim klasičnoga odgovora na neko pitanje ili zadatka koji uključuje pretraživanje, sumiranje ili prezentiranje informacija, s obzirom na to da se može služiti prirodnim jezikom, model generativne umjetne inteligencije trebao bi biti u mogućnosti sastaviti smislen tekst bilo koje tematike i opsega. Upit može biti samo tema teksta, ponekad i vrsta, naprimjer izvještaj, obavijest ili nešto općenito kao *Napiši kratki tekst o*. Dosad u svim slučajevima, bez obzira na temu i vrstu teksta, pa čak i u “običnim” odgovorima na upite, modeli pokazuju iste karakteristike koje gotovo možemo nazvati stilskima.

U slučaju u kojemu se modelu zadaje tema na koju mora napisati tekst, on će, po svemu sudeći, pronaći neke informacije o toj temi ili tekstove koji su slični po vrsti kako bi pokušao oponašati njihov stil. U istraživanju Romana E. Tel’pova i Stanislave V. Larcine (2023) s Katedre



za opću i rusku lingvistiku Državnog instituta ruskoga jezika A. S. Puškina u Moskvi ispitanici su bili ljudi i *ChatGPT*. Zadane su četiri teme na koje je napisano 160 tekstova, 20 tekstova po temi, ali polovinu su pisali ljudi, a polovinu *ChatGPT*. S obzirom na to da model bazira svoj tekst na podacima koji su vezani uz temu, on sadržava znatno više ponavljanja ključnih riječi, dok ispitanici ponekad u cijelome tekstu nisu iskoristili ključnu riječ koja je određivala temu, ali tekst je temi ipak odgovarao. Također leksička raznovrsnost u tekstovima koje su napisali ispitanici znatno je viša od one u tekstovima koje je napisao *ChatGPT*. Model koristi sinonime kako se ne bi očito ponavljao, ali ponavljanje postaje jasno kada se u obzir uzme informacija koju je model ponudio – informacija je samo jedna, ali izrečena je na različite načine, i to ne drukčijim jezičnim konstrukcijama, već samo sinonimima.

Primjere nije teško pronaći. Dok su se u sklopu nekoga drugog aspekta generiranja teksta ispitivale mogućnosti analize književnog teksta, *ChatGPT* ponudio je sljedeće opise dvaju likova:



Willie Beier is a character in Sven Hassel's novels who appears as one of the soldiers fighting on the Eastern Front during World War II. He is often depicted as a member of the main group of protagonists alongside characters like Sven, Porta, Tiny, and others.

Beier's character is typically portrayed as a loyal and courageous soldier, deeply entrenched in the brutal realities of war. He participates in the group's missions and battles, facing the horrors of combat alongside his comrades. Like the other characters in Hassel's novels, Beier confronts the challenges of survival, camaraderie, and the moral dilemmas inherent in wartime situations.

While Beier may not be as prominently featured as some of the other characters like Sven, Porta, or Tiny, he still plays a significant role in the narratives, contributing to the overall themes and dynamics within the group of soldiers.

### *Slika 9. Snimka zaslona; ChatGPT.*



Alfred Kalb is a character in Sven Hassel's novels who typically appears as one of the soldiers fighting on the Eastern Front during World War II. Like many of the characters in Hassel's novels, Kalb is portrayed as a member of the main group of protagonists, alongside characters such as Sven, Porta, Tiny, and others.

Kalb's character is usually depicted as a loyal and brave soldier, facing the harsh realities of war alongside his comrades. He participates in their missions and battles, enduring the hardships and dangers of combat on the Eastern Front.

While Kalb may not be as prominently featured as some of the other characters like Sven or Porta, he still plays a significant role within the group dynamics of the soldiers. His presence contributes to the camaraderie, conflicts, and shared experiences depicted throughout the novels.

### *Slika 10. Snimka zaslona; ChatGPT.*

Prilikom čitanja veće količine teksta koju je *ChatGPT* ponudio kao popis i karakterizaciju likova, lako je previdjeti ovakvo ponavljanje, pogotovo ako uistinu tražimo odgovor na pitanje na koje ni sami ne znamo odgovoriti. U ovome slučaju o Beieru i Kalbu ne saznajemo mnogo osim toga da su vojnici, ali primjećujemo da su konstrukcije *typically portrayed* i *usually depicted* parovi sinonima, kao što su *courageous* i *brave*, a cijeli prvi dio prvih rečenica ima identičnu konstrukciju: subjekt – predikatni skup od predikata i jednog priloga – priložna oznaka usporedbe od dva pridjeva i imenice. Jednaka je situacija s parovima *brutal* i *harsh*. Tekstovi se razlikuju po tome kada je spomenuta jedinica (u prvoj ili drugoj rečenici), a zatim se odgovor počinje očitije ponavljati (*missions and battles*). Na kraju *ChatGPT* pribjegava prvoj opciji korištenja iste konstrukcije: predikat – objekt s varijacijama u sinonimima (*facing the horrors* i *enduring the hardships*). Iako jezično ovi odgovori nisu problematični, oni zapravo nisu informativni. Usporedbe radi, opis prvoga lika sa službene stranice glasi ovako: “a carpenter by occupation; hates the war and wants to go back home as fast as possible to his workshop, his wife and children. Found calmly smoking his pipe and viewed as the father figure of the group” (*Characters*, bez dat.). Iako ovaj opis nije ništa duži od onoga koji je ponudio *ChatGPT*, on sadrži mnogo više informacija kao što su zanimanje lika, njegov odnos prema situaciji, njegove želje i karakter. *ChatGPT* zapravo nam je samo rekao da je Beier lik koji sudjeluje u radnji.

S obzirom na to da se površinski jezično gledano generirani tekstovi ne ponavljaju, oni su u svome izričaju originalni. Model ne kopira ni tekstove na kojima je treniran ni izvore koje je koristio prilikom sastavljanja novog odgovora na korisnikov upit. To izvrsno pokazuje i činjenica da kada je upitan za izvor informacije ne može točno navesti odakle ju je saznao. Prilikom provjere teksta u programima za otkrivanje plagijata generirani tekstovi dobivaju procjenu 98–100% originalnosti (Ajdagulova, 2023). Neki su se programi već prilagodili novonastaloj situaciji i tekstove koji sadržavaju stilske karakteristike generiranih tekstova (ponavljanje ključnih riječi, ponavljanje konstrukcija sa sinonimima) označavaju kao sumnjive bez obzira na postotak originalnosti.

### 3.3 Generiranje književnoga teksta

Već je ranije u ovome radu istaknuto da se generiranje prirodnoga jezika odvija na osnovi statističke vjerojatnosti pojavljivanja riječi, sintagmi ili čak nizova riječi. Što se književnih

tekstova tiče, uobičajeno je da ljudi poeziju gledaju na drukčiji način u odnosu na prozu, no umjetnoj neuronskoj mreži ona je i dalje jezični predložak koji se statistički obrađuje, ona dijeli pjesme na riječi i specifične strukture te može prepoznati čak i pojavljivanje ritma (Schwartz, 2018). Nema ničega izrazito specifičnoga u čitanju i pisanju, odnosno analiziranju i generiranju stihova kada model nauči percipirati ih kao razmake ili pojedinačne riječi. Svaka pojava koja može biti primijećena i zatim matematički izražena može se ponovno reproducirati u varijacijama, kao što je to bio slučaj s dvama atributima uz imenicu u priložnoj oznaci koja se pojavila u opisu lika iz prethodnog poglavlja. Da se modelu pruži veća količina tekstova u kojima svaka imenica uz sebe ima tri atributa, model bi nakon nekog vremena treniranja na tim tekstovima počeo prilagati tri pridjeva uz imenicu. Time on ne pridodaje samome sadržaju dodatno značenje, vrijednost ili oznaku stila u smislu u kojemu to radi autor dok stvara tekst (Tredinnick i Laybats, 2023). Model generativne umjetne inteligencije svakako posjeduje mogućnost odabira riječi koje odražavaju stil pojedinog autora, odnosno takvo se oponašanje može shvatiti kao posjedovanje jednake razine kreativne inteligencije (Schwartz, 2018). S druge strane u svome istraživanju Imke van Heerden i Anil Bas primjećuju mogućnost zavaravanja čitatelja od strane modela generativne umjetne inteligencije prilikom predlaganja generiranog teksta u Shakespeareovu stilu, ali stručnjaci su brzo i lako prepoznavali generirane tekstove zbog “niske razine emocionalnoga utjecaja i čitljivosti” (2021, str. 177). Može se reći da pitanje nije *Je li tekst kvalitetan? nego Tko može prepoznati kvalitetan tekst?* Autori također naglašavaju da je jedan od izazova razvoja generativnih modela ostvarivanje uspješne suradnje između računalnih stručnjaka i teoretičara književnosti. S obzirom na to da stručnjaci u jednome području najčešće nisu upoznati s onim drugim, pojednostavljivanje materije dovodi do nerazumijevanja problema i sprečavanja napretka. U istome radu autori zagovaraju kvalitetnu suradnju i razvoj metoda koje bi preformulirale apstraktne književne koncepte u računalno razumljive registre koje bi modeli mogli koristiti, iako se nigdje ne komentira koji je krajnji cilj razvoja takvih modela i kome bi takvo nešto bilo potrebno i korisno.

No postavlja se pitanje možemo li posve isključiti namjeru korištenja točno određene riječi ili kombinacija riječi prilikom stvaranja književnog teksta. Već je samo razumijevanje prirodnoga jezika veliko tehnološko postignuće, ali treba uzeti u obzir da umjetna neuronska mreža može proizvesti jezično smislenu, ali površnu književnost. Elam (2023) ističe da obrada i predviđanje prirodnoga jezika nisu nužno glavni problem, nego da je moguće da se radi o

namjernom ograničavanju modela na najpoželjnije oblike odgovora (ponovno je fokus stavljen na statistiku), što nužno smanjuje raznolikost generiranih odgovora. Modeli koji generiraju slike proizvode klišeizirane prikaze, a modeli specijalizirani za glazbu proizvode melodije izvedene iz postojećih i tako dalje. Tredinnick i Laybats (2023, str. 99) ističu da se ne radi nužno o nedostacima pojedinoga modela, već da su oni izrađeni tako da ne budu u potpunosti originalni kako bi se spriječilo generiranje nepodobnih odgovora. Ova tvrdnja može naći svoje uporište u generiranju humora, koji je sam po sebi specifičan što u proizvodnji, što u tumačenju, a posebno zahtjevan u postupcima prevođenja.

U istraživanju Eduarda V. Lihačeva (2023) pokazalo se da šale koje je generirao *ChatGPT* imaju puniju sintaksu od onih koje su smislili ljudi te da izbjegava stereotipni i crni humor. Moguće je da su modeli generativne umjetne inteligencije posebno uređeni kako bi izbjegavali određene teme, što je neizravno navedeno i u njihovim pravilnicima za korištenje i uvjetima korištenja. Kada se s *chatbotom* zapodjene razgovor u kojemu se spomene neka od diskutabilnih tema koje bi mogle dovesti do nepodobnih odgovora ili potaknuti korisnike na neželjene reakcije, u *chatu* će se pojaviti upozorenja na temu koja se ne slaže s politikom vlasnika modela. S obzirom na to da stereotipi i teme kojih se dotiče crni humor izvan okvira humorističnoga konteksta mogu biti interpretirane kao uvredljive, može se pretpostaviti da nije riječ o nesposobnosti modela da generira takav humor, već da ga uistinu svjesno izbjegava.

Pretpostavka da model može izbjeći određenu temu ili iskaz znači da bi model svejedno mogao generirati odgovor unutar te kategorije kada bi zabrana bila uklonjena. Također model nije sam zaključio koje su teme diskutabilne ili problematične već mu je to odredio čovjek, a nije potrebno naglašavati da procjene ovakvoga tipa mogu biti u određenoj mjeri subjektivne. Zbog kulturološke neosjetljivosti i pokušaja programiranja iste generirani tekstovi mogu ponavljati određene pristranosti ili općenitosti te smanjiti raznovrsnost stila i tematike. Kada bi se modeli počeli učiti na tekstovima koje su generirali drugi modeli, proces bi počeo produbljivati ove karakteristike (Epstein i sur., 2023). S druge strane subjektivnost kao takva neophodna je za stvaranje književnoga djela bilo koje vrste, što pak otvara mnoga pitanja određivanja tema, ograničenja i smjerova u kojima se modeli generativne umjetne inteligencije uopće mogu razvijati, odnosno što im ljudi mogu dopustiti ili što je izvan ičije kontrole te kakva bi se ponašanja trebala korigirati kada se primijete u odgovorima.

### 3.3.1 Stvaranje i kreativnost

Generativna umjetna inteligencija, iako ne mora nužno imati izvore u načinima i procesima stvaranja kojima se koriste ljudi, uvijek se s njima uspoređuje. Kako je kod razvoja umjetnih neuronskih mreža bilo potrebno razumjeti osnove funkcioniranja neuronske mreže u ljudskome mozgu, tako je za razumijevanje kreativnoga procesa generativne umjetne inteligencije potrebno razumjeti na koji način kreativno djeluju ljudi. U mnogim se komentarima recentnoga stanja tehnologije polazi od pretpostavke da umjetna i ljudska inteligencija funkcioniraju po istim principima, što neminovno dovodi do više ili manje netočnih zaključaka i predviđanja budućega razvoja umjetne inteligencije.

U teoriji evolucije antropolozi pretpostavljaju da je do značajnije promjene u moždanome sklopu došlo tijekom prijelaza vrste *Homo habilis* u *Homo erectus*. *Homo habilis* imao je mogućnost pamćenja situacija i pristupanja tome pamćenju kada bi se ponovno našao u istoj situaciji, ali vjeruje se da je *Homo erectus* mogao pristupiti sjećanjima na pojedinačne situacije neovisno o vanjskim okolnostima, odnosno on je mogao “svojevoljno pristupiti sjećanju neovisno o poticajima i na taj način odglumiti događaj koji se dogodio u prošlosti ili koji bi se mogao dogoditi u budućnosti” (Gabora i Kaufman, 2010, str. 4). Francuski filozof Henri Bergson ističe da se sjećanje sastoji od svega što se dogodilo, a što je ostalo svjesno i nesvjesno zapamćeno: “moždani mehanizam stvoren je kako bi u nesvjesno potisnuo ukupnost prošlosti, a u svijest doveo samo ono što može rasvijetliti danu situaciju” (2001, str. 42). Ponovno je glavni razlog prizivanja sjećanja konkretna korisnost u danoj situaciji, ali sve što ostaje pohranjeno u nesvjesnom utječe na svaku sljedeću interpretaciju situacije jer um nikada ne može dvaput proći kroz isto stanje (Bergson, 2001). Sposobnost razmišljanja o prošlim i budućim događajima postaje svojevrsan oblik maštanja jer pojedinac nije, naravno, mogao znati što će se dogoditi u budućnosti, ali mogao je smisliti varijacije ishoda pojedinih situacija. Vanjski podražaj time gubi na važnosti u svojoj ulozi okidača misaonih procesa i čovjek stvara unutarnji svijet u kojemu se misao može stvarati, kretati i prekidati bez posebno korisnoga cilja.

U misaonome se procesu stvara nova sposobnost koja utječe na razvoj kreativnoga mišljenja, a to je simboličko razmišljanje, odnosno simbolički prikaz predmeta, ljudi ili bilo kojih drugih aktera pojedine situacije. Simboličko razmišljanje u obzir uzima sav potencijal pojedinog objekta o kojemu se razmišlja i smješta ga se u priču prema svim varijantama u kojima

se taj potencijal može iskoristiti (Gabora i Kaufman, 2010). Procesom simbolizacije obična grana u rukama djeteta u igri postaje mač kao što protok vremena i uzročno-posljedične veze stvaraju paralelne svjetove, pa čak i svemire, u idejama filmskih scenarista.

Nadalje, nakon što prestaje biti ograničen fizičkim svijetom i izravnim djelovanjem, ljudski mozak postaje sposoban usmjeriti pažnju na detalj ili širu sliku te kombinirati detalje na različite načine. Misaoni procesi postaju divergentni ili konvergentni, odnosno fokus misli može se raspršiti na širu sliku i oprečne ili nepovezane objekte ili usredotočiti se na jedan objekt i konkretizirati ga. Gabora i Kaufman (2010, str. 9) kao primjer divergentnoga procesa navode razmišljanje o poeziji, a kao primjer konvergentnoga procesa navode matematičke operacije. Sintagma kreativni proces označava postupke širenja i sužavanja fokusa te procese kombinacije detalja kako bi se stvorila nova ideja ili objekt. Osnovnih pet strategija kreativnoga procesa su: kombinacija suprotnih svojstava u jednome elementu; dopuna elementa novim svojstvom; miješanje elemenata različitih karakteristika; postavljanje očekivanoga elementa u neočekivani kontekst te postavljanje neočekivanoga elementa u poznati kontekst (Mil'gizin i Baeva, 2017, str. 67). Bergson, s druge strane, definira intelekt kao "sposobnost izraditi umjetni predmet, osobito oruđe za izradu oruđa, te beskrajno izmjenjivati njihovu proizvodnju" (2001, str. 153). U suštini ova definicija obuhvaća strategije kreativnoga procesa jer podrazumijeva različito postupanje s poznatim elementima, njihovo kombiniranje te korištenje za proizvodnju novih elemenata.

Bergson uz intelekt neraskidivo veže instinkt kao nadopunjujuću karakteristiku svih živih bića, odnosno tvrdi: "nema intelekta u kojemu se ne mogu pronaći tragovi instinkta, nema instinkta koji nije okružen intelektom kao izmaglicom" (2001, str. 150). Kako je instinkt karakteristika isključivo živih bića, logično je da se on ne spominje u proučavanju umjetne inteligencije, ali ako su mogućnost spoznaje i instinkt nerazdvojni, mogu li se ljudska i umjetna inteligencija uopće uspoređivati kao sposobnosti različitih razmjera, ali istih temelja? Prema Bergsonovim određenjima instinkt je zapravo sposobnost korištenja organiziranog oruđa, a intelekt neorganiziranoga. Ako podatke koje model generativne umjetne inteligencije obrađuje kako bi korisniku ponudio odgovor shvatimo kao organizirano oruđe, možemo reći da je sposobnost modela da odredi vjerojatnost pojavljivanja neke riječi instinkt koji model posjeduje. Intelekt koji okružuje taj instinkt odnosi se na mogućnost modela da razumije prirodni jezik, odnosno razumije veze i odnose pojedinih elemenata, bilo da se radi o jednakosti oblika, sadržaja

ili uzroku i posljedici. Umjetna inteligencija radi s prirodnim jezikom, ali intelekt bi se trebao odnositi i na sposobnost stvaranja novih ideja. Angel Hsing-Chi Hwang (2022) kao jednu od faza kreativnoga procesa izdvaja osmišljavanje ideja koje nisu nužno korisne ili izvedive, ali s pomoću njih mozak dolazi do novih koncepata koji mogu pokrenuti cjelokupni kreativni proces. Ponekad taj proces zahtijeva prekid razmišljanja, ostavljanje ideje da “odleži” jer mozak je u stanju procesirati je i u podsvijesti. Posebna grana tehnologija umjetne inteligencije bavi se procesom donošenja odluka, koji je iznimno važan pri odabiru najbolje osmišljene ili generirane ideje, a ponekad odluka ne može biti matematički prikazana kako bi se algoritmu pokazalo na koji način ju je trebao donijeti.

Profesor Olivier Toubia (2023) sa Sveučilišta Columbia, čija je specijalnost ideacija, upozorava da moramo razlikovati božansko stvaranje i kreativnost, odnosno stvaranje nečega novog ni iz čega i stvaranje nečega od postojećih elemenata. Na prvi se pogled čini kao da to upozorenje ne ide na ruku umjetnoj inteligenciji, ali zapravo se odnosi na preuveličavanje ljudskih sposobnosti. Svaka je ljudska ideja nastala jednom od pet strategija kreativnoga procesa ili njihovim kombinacijama, a najveći sporovi koji se tiču ljudskih sposobnosti i sposobnosti umjetne inteligencije zasnivaju se na tome da se jednoj od tih dviju pridaju prevelike zasluge. Čak se jedna točka Uredbe predsjednika RF osvrće na ovaj problem:

“algoritme rada neuronske mreže izrazito je složeno interpretirati, sukladno tome rezultati njihova rada mogu se podvrgnuti sumnji ili ih može izmijeniti čovjek. Nedostatak razumijevanja postupka kojim umjetna inteligencija dobiva rezultate jedan je od razloga niske razine povjerenja u suvremene tehnologije umjetne inteligencije i može postati preprekom u njihovom razvoju” (2019/2024).

Potvrdu toga da je najčešći oblik kreativnosti druga strategija – dopuna poznatoga elementa novim svojstvom – pronalazimo i u razmatranjima Tredinnicka i Laybats (2023) koji tvrde da kreativnost nije sama po sebi jednaka originalnosti, već u sebi sadržava tragove tradicije i postojeće kulture kojima se pridonosi nešto novo. Autori uspoređuju ljudsku i umjetnu inteligenciju kao dva sustava sposobna stvarati uvjetno nove elemente, ali ne s razlikom u kvaliteti već kvantiteti, barem što se trenutnog stanja tehnologije tiče. Procjenu jednaku umjetnoj inteligenciji iz Uredbe predsjednika RF Tredinnick i Laybats daju za ljudski mozak: “Ali ljudski je um također crna kutija i dok proizvode ljudske kreativnosti određujemo kao originalne radove,

ne znamo odakle oni proizlaze ili kako su točno nastali” (2023, str. 100). S druge strane trebalo bi uzimati u obzir o kojoj se razini kreativnosti radi, odnosno kakav kreativni napor se u konačnici zahtijeva te koji bi trebali biti njegovi produkti. Kreativnost se tako može odnositi na svakodnevne kreativne aktivnosti koje su možebitno podložne automatizaciji ili se može odnositi na kreativne doprinose na višim razinama intervencije, odnosno potpune promjene u odnosu na postojeće stanje (Rafner, Beaty, Kaufman, Laubart, Sherson, 2023).

Uzevši sve navedeno u obzir, može se zaključiti da svako pohranjeno iskustvo mijenja stanje ljudskoga uma i time izmjenjuje kompleksne nizove varijacija koje on može proizvesti, čak i kada se radi o nesvjesnome jer ono također utječe na razmišljanja, stavove i ponašanja. Umjetna inteligencija nema iskustvo, ne živi život u kojemu je izložena raznoraznim podražajima koji uzrokuju preslagivanje spoznaja i koji mogu uzrokovati nastanak ideje. Generativna umjetna inteligencija može isključivo reflektirati skup podataka na kojima se trenira i ne može reagirati na događaje koji se odvijaju u stvarnome vremenu (Toubia, 2023).

Nezanemariv čimbenik u kreativnome procesu također je i potreba. Ideje nastaju kao odgovor na određeno traženje, zadovoljenje potrebe ili estetsko uređenje nekoga elementa te ne moraju nužno proizlaziti iz prirodne selekcije (odabir onoga što je najkorisnije), nego se samoorganizacijom i razmjenom dijele, zamiru ili preoblikuju (Gabora i Kaufman, 2010). I Bergson (2001) govori o tome kako intelekt djeluje u svakome području u kojemu želi jer on može uzeti oblik bilo kojega elementa i napuniti ga bilo kakvim značenjem, čak i onime koji ničemu ne služi. Ideje koje se tiču književnosti ujedno su dio kulture i njezin odraz. Za Toubiju (2023) zabava prenosi kulturu i čovjek stvara priče kako bi objasnio svijet samome sebi, poučio druge o njemu i pokušao prenijeti određene vrijednosti ili zabavom stvara nove vrijednosti. Postavlja se pitanje može li stvaranje kulture biti prepušteno računalu. Martin Senftleben na to pitanje pruža vrlo izravan odgovor:

“Sustav umjetne inteligencije može oponašati ljudsku kreativnost i generirati usporedive književne i umjetničke rezultate. Ali nije u mogućnosti probiti se ispod površine ljudskoga umjetničkog djela, nadići njegov pukli oblik i izgled te procijeniti njegovu poruku i značenje u svjetlu trenutnih društvenih uvjeta” (2023, str. 1538).



Sva se razmatranja ponovno svode na fond na kojemu se trenira pojedini model generativne umjetne inteligencije. Čak i kada bi se fond konstantno obnavljao recentnim podacima, književnim tekstovima, člancima, čak i književnim kritikama, on ne bi mogao reagirati već samo reflektirati dobivene podatke. U generiranim tekstovima ne bi moglo biti ničega avangardnoga, a za ljude ne bi bilo ničega ispunjavajućega u tome da svoje ideje zadaju računalu i pričekaju da im ono preda rezultate svoga rada (Senfleben, 2023, str. 1539–1541).

## 4. Autorstvo

Razvoj generativne umjetne inteligencije i njezina sve veća primjena u raznim kontekstima u posljednjih su godinu dana uzdrмали svijet svojim napretkom i mogućnostima, kako na pozitivan, tako i na negativan način. Osmišljavanje potencijalnih primjena i zamišljanje savršenih, brzih i jeftinih rezultata rada generativnih modela potaknulo je niz reakcija javnosti, a i sami su stručnjaci podijeljeni u stavovima kako bismo se trebali nositi s novim tehnologijama. U razdoblju kada tvrtka *Meta* objavljuje promjenu Pravila o privatnosti u kojima navodi da će se svi sadržaji stvoreni i dijeljeni na njihovim platformama koristiti za treniranje umjetne inteligencije, više nego ikada raspravlja se o autorskom pravu, krađi podataka i intelektualnoga vlasništva. Posebno su glasni na društvenim mrežama postali umjetnici koji pozivaju na ulaganje službenih pritužbi s namjerom da se njihovi sadržaji izuzmu iz fonda za treniranje umjetne inteligencije. Naime *Meta* je korektno ponudila mogućnost pritužbe, ali uz upozorenje da neće sve pritužbe biti uvažene.

S druge strane problema zaštite podataka i sadržaja nalazi se problem njihovoga korištenja. Ovogodišnja dobitnica Nagrade “Akutagawa”, japanske nagrade za književnost, spisateljica Rie Kudan obznanila je da je prilikom pisanja svoga nagrađenoga romana koristila *chatbot* koji je napisao oko pet posto teksta njezina romana (*Kyodo*, 2024). Iako se ispostavilo da se u romanu radi o odgovorima koje nudi *chatbot* AI-built, koji je autorica predstavila kao jednog od likova u romanu, rasprave o uvođenju posebnih pravila o korištenju generativne umjetne inteligencije prilikom pisanja te o tome koliko je uopće pošteno koristiti pomagala ove vrste ne jenjavaju.

### 4.1 Pitanje autorskoga prava na skupljene podatke

Da bi model generativne umjetne inteligencije funkcionirao, mora biti istreniran na određenome skupu odgovarajućih podataka, a da bi se mogao razvijati, mora mu biti osiguran stalni priljev novih podataka u fond za treniranje. Ovisno o tome koja se tehnika *data* i *text mininga* koristi i s kojim ciljem, tijekom procesa može doći do umnažanja sadržaja zaštićenog autorskim pravom, izdvajanja dijela neke baze podataka ili kopiranja i prilagođavanja dijela računalnoga programa (Ducato i Strowel, 2021). Međutim, neovisno o korištenoj tehnici,

vlasnici modela za koji se *data* i *text mining* provode mora imati zakonit pristup podacima koji će ući u fond, što bi u slučaju umjetničkih djela, posebice književnih, značilo da bi autor trebao dati neku vrstu privole za korištenje svoga rada ili biti na to obvezan nekim drugim ugovorom.

Senftleben (2023) objašnjava koja su to četiri glavna problema vezana uz principe prikupljanja podataka, dozvola za prikupljanje i eventualne sustave pružanja naknade autorima za korištenje njihovih djela za treniranje generativne umjetne inteligencije. Prvi se problem odnosi na općenito ujednačavanje pravila o pristupu i korištenju podataka konkretno na području Europske unije, ali problem može biti razmatran i na globalnoj razini. Provjeravanje uvjeta korištenja i uvjeta naplate naknade nositeljima autorskih prava za svako pojedino djelo koje se želi unijeti u fond značajno bi usporavalo svako treniranje modela i na koncu ne bi bilo isplativo, tvrdi Senftleben (2023). Ovakvo usporavanje procesa dovelo bi do zaostataka u cjelokupnome napretku novih tehnologija jer dok bi se druge organizacije bavile treniranjem modela sa slobodnim pristupom podacima, Europska unija bavila bi se papirologijom o autorskom pravu, što bi je postavilo u lošu poziciju na globalnome tržištu (Ibid, str. 1548). Sljedeća se dva problema odnose na sam proces isplaćivanja naknade autoru: s obzirom na količine podataka koje modeli obrađuju moglo bi biti zahtjevno uskladiti i pratiti odnose iskorištenih i plaćenih djela; u cijelome bi se procesu moglo dogoditi da sam autor uopće ne dobije naknadu za korištenje rada, već ona bude isplaćena organizacijama koje upravljaju repozitorijima (Ibid).

Nadalje kada bi se i postavila jasna pravila o korištenju podataka, neki autori smatraju (Ducato i Strowel, 2021) da bi trebale postojati iznimke, odnosno situacije u kojima bi se određenim institucijama trebao pružiti pristup podacima neovisno o ograničenjima i zabranama. Takve bi institucije prije svega uključivale ustanove visokoškolskoga obrazovanja na kojima se provode znanstvena istraživanja te kulturne ustanove. U trenutno aktivnim propisima takve iznimke postoje, ali nejasnoće u njihovu definiranju mogu pak dovesti do nemogućnosti njihove primjene. Tako neovisna istraživanja ne bi mogla dobiti pravo na *data mining* jer se ne provode u sklopu nekoga sveučilišta; čak i ako se iznimka odnosi na sadržaj, ne mora se nužno odnositi na kod pa bi on bio nepristupačan odjelima koji se bave računalnom znanosti (Ibid, str. 11, 14). Također, postoje izvori koji već u sebi imaju implementirane zabrane za *data* i *text mining*, a povećanje broja neograničenoga pristupa moglo bi potaknuti povećanje naknade za pretplatu postojećim korisnicima pojedinoga izvora (Ibid, str. 19).

Drugi pristup pravu na korištenje podataka ne kreće od podataka samih, već od tekstova koje generira umjetna inteligencija:

“fokus bi mogao biti na ponudi proizvoda i usluga generativne umjetne inteligencije na tržištu (ovo je izlazna dimenzija). Umjesto zahtijevanja plaćanja naknade na razini treniranja, mogao bi biti nametnut porez na konačni sustav umjetne inteligencije koji je u mogućnosti ponuditi književne i umjetničke rezultate. Ovaj porez na umjetnu inteligenciju mogao bi biti iskorišten u svrhu financijske potpore, obuke i novih književnih i umjetničkih projekata za autore od krvi i mesa” (Senftleben, 2023, str. 1537).

Prijedlog je svakako zanimljiv, iako u kapitalističkome društvu u kojemu autori od svojih nakladnika često ne dobivaju odgovarajuće naknade teško je zamisliti da bi ih dobivali za korištenje svojih djela u treniranju modela od treće strane koja najčešće prikuplja podatke bez ikakvih prethodnih obavijesti. No ovaj se prijedlog dotiče nove razine autorskoga djelovanja, a to je primjena umjetne inteligencije u komercijalne svrhe. Ako se pojedini model trenira s namjerom da zamijeni ljudskoga autora ili da mu u nekome trenutku postane ozbiljna konkurencija, onda se autorima izravno nanosi šteta time što se njihova djela koriste bez ikakve naknade, posebice ako oni sami nisu pristali da njihovo djelo uđe u fond za treniranje.

S druge strane tvrtke koje razvijaju generativnu umjetnu inteligenciju ovise o podacima koje stvaraju ljudi. Kako je već bilo rečeno, modelu se mora pružiti stalni priljev novih podataka kako on ne bi ušao u zatvoreni krug ponavljanja obrade i time počeo nuditi jednake i sve lošije rezultate. Senftleben (2023) naglašava kako bi porez na umjetnu inteligenciju zapravo koristio i autorima i tvrtkama jer bi se (moguće na godišnjoj bazi) osigurala isplata naknade autorima, a tvrtke bi dobile pristup potrebnim podacima bez otežavajuće papirologije i analiziranja svakoga pojedinog djela i njegovih uvjeta korištenja. Autor dodaje kako bi moguća opcija bila i isplaćivanje naknade tek nakon što dovoljno istrenirani model “konačno ubije potražnju za istom ljudskom kreativnošću koja je osnažila sustav umjetne inteligencije” (Ibid, str. 1538). Po nekima neizbježan, trenutak ubijanja potražnje ipak jest diskutabilan i o njemu će još biti riječi.

Prema nekim gledištima problem razmotren u ovome poglavlju uopće ne postoji jer “*text* i *data mining* ne utječe nužno na autorska i srodna prava. *Text* i *data mining* može uključivati procesiranje pukih činjenica i podataka koji nisu sami po sebi zaštićeni autorskim pravom”

(Ducato i Strowel, 2021, str. 8). Generativna umjetna inteligencija ne kopira podatke koje obrađuje, već iz njih uči i pokušava ostvariti rezultate statistički srodne podacima na kojima se trenira. Pojedino je djelo samo sredstvo prijenosa informacija i ostaje zapamćeno u povijesti modela i on ga može reproducirati jednako koliko i čovjek koji je jednom pročitao knjigu pa je pokušava recitirati napamet.

## 4.2 Pitanje autorskoga prava na generirane tekstove

O autorskome pravu na generirane tekstove mnogo se manje raspravljalo u odnosu na autorsko pravo na podatke koji ulaze u fond za treniranje. Uzroci tomu najvjerojatnije su trenutne mogućnosti generativne umjetne inteligencije i nešto jednostavnija situacija promatrana sa stajališta stvaranja proizvoda bilo koje vrste. Generativna umjetna inteligencija u trenutku pisanja ovoga rada u književnosti nije uspjela ponuditi cjelovita djela najtraženije vrste, a to su romani. Na primjerima generiranih tekstova jasno je da je umjetna inteligencija još uvijek u procesu svladavanja stilistike, a pitanje je hoće li ikada ovladati relevantnošću koju nude književna djela koja su napisali ljudi. Isto tako bez prevelikoga udubljivanja u tematiku može se pretpostaviti da autorska prava na generirano djelo pripadaju tvrtki koja posjeduje model, odnosno bilo kojoj drugoj privatnoj ili pravnoj osobi koja s tvrtkom ugovorom uredi uvjete korištenja njihova modela.

Prema Zakonu o autorskom pravu i srodnim pravima (NN 111/21) autorsko pravo na neko djelo pripada fizičkoj osobi koja ga je stvorila, a na temelju ugovora nositelj prava može biti fizička ili pravna osoba koja ima pravo na udio u prihodima. Jednaka je situacija i u Ruskoj Federaciji za čije zakone Artenij (2019) navodi da je autor znanstvenog, književnog ili umjetničkog djela čovjek čijim je kreativnim radom to djelo nastalo. Europska se unija uglavnom ograđuje od djelovanja umjetne inteligencije usmjeravajući odgovornost za njezin rad na čovjeka (Ibid), ali pitanje zapravo nikada nije je li model generativne umjetne inteligencije autor književnoga djela, već je li autor čovjek koji je zadao upit ili vlasnik programa modela.

Odredbe koje bi se odnosile na kreiranje književnog djela s pomoću generativne umjetne inteligencije situacijskoga su karaktera i moguće je da na pojedine rezultate koje model ponudi ne može biti primijenjen zakon o autorskome pravu jer on zahtijeva značajan doprinos čovjeka

(Lawton, 2023). Ipak, značajan je doprinos teško definirati. Ako korisnik zada upit *Napiši kratku priču o zecu koji živi u šumi.*, moglo bi se reći da određivanje teme teksta nije dovoljan doprinos. Ako korisnik razradi temu, likove, glavni slijed događaja i zatraži od modela da izradi cjelovitu priču, hoće li situacija postati diskutabilna? Ako kažemo da je tada korisnik autor djela, sveli smo autora na razinu pukoga generatora ideja i motiva, a opet nije nepoznato da su ljudi znali krasti ideje koje se, kao osnova za književno djelo, mogu smatrati intelektualnim vlasništvom.

Čim autorsko pravo pripadne korisniku, možemo se pitati kako bi na to gledala tvrtka koja je vlasnik modela, pogotovo ako korisnik iskoristi generirani tekst u komercijalne svrhe. Sljedeći je problematični element generirani tekst koji je odraz iskustva umjetne inteligencije, a ne samoga korisnika, što znači da bi informacijski jednaki upiti različitih korisnika rezultirali istim tekstom, dok više autora na istu temu nikako ne bi napisalo iste tekstove. Korisnik je tako odgovoran za rezultate rada generativnoga modela dokle god se uspijeva uz njegovu pomoć izraziti i dok rezultati odražavaju njegove namjere i osobni izražajni potpis (Epstein i sur., 2023). Za pripadnost autorskih prava korisniku koji je koristio alate umjetne inteligencije zauzimaju se i McGibbon i Solomon (2024), ali uz jasno određivanje granica između radova koje je stvorio pojedini model i onih koje su stvorili ljudi uz pomoć modela.

#### **4.2.1 Umjetna inteligencija kao autor, suautor ili pomoćni alat**

Umjetna inteligencija u ovome vremenu ne može se smatrati autorom u punome smislu te riječi. Tekstovi koje može generirati uglavnom nisu ni po čemu kvalitetni, a zanimljivi su isključivo sa stajališta promatranja razvoja tehnologije. Donekle istreniran model generativne umjetne inteligencije može ponuditi tekst u kojemu postoji osnovna radnja i beskarakterni likovi. Model može ponuditi solidnu analizu likova kada mu je ponuđen potpuni književni tekst, ali spoznaju o karakteru pojedinoga lika ne može uključiti u generiranje novoga teksta na temelju polaznoga. Također ako se radi o književnome tekstu koji je tek u nastanku i model nema mnogo sadržaja koji bi mogao analizirati, neće nikako moći uhvatiti bit likova i radnje kojima je autor htio prožeti tekst. U svakome slučaju, barem zasad, uvijek mora postojati čovjek koji će inicirati ideju i povremeno je usmjeravati u željenome pravcu.

Kada se sagledaju sve trenutne mogućnosti i budući potencijali umjetne inteligencije, treba ukloniti fokus s bipolarnog razmišljanja u kojemu su postojeće mogućnosti sljedeće:

umjetna inteligencija ne može ništa, naprema umjetna inteligencija može sve. Odstupajući od razmišljanja “umjetna inteligencija ili čovjek” i prelazeći u “umjetna inteligencija i čovjek” moguće je ostvariti dobre i korisne rezultate (Khosrowshahi, 2023). Jerry W. Washington (2023) govori o suradnji pisaca s generativnom umjetnom inteligencijom s ciljem poboljšanja učinkovitosti pri pisanju te ističe da korištenje umjetne inteligencije može pozitivno utjecati na percepciju vlastitih sposobnosti u pisca. Neke od tih sposobnosti su stvaranje visokokvalitetnoga sadržaja i prevladavanje kreativne blokade. Doduše, Washington daje nespretno objašnjenje: “Izlaganjem pisaca oglednim uzorcima i inovativnim stilovima pisanja tehnologija služi kao izvor zamjenskoga učenja” (2023, str. 20). Prilikom kreativne blokade korisnik može “razgovarati” s modelom ili mu ponuditi ideje i dijelove teksta koje model može obraditi i komentirati ili na koji drugi način preusmjeriti i potaknuti razmišljanje korisnika. Takav postupak može djelovati oslobađajuće na korisnika jer mu nešto izvan njega (u ovome slučaju model) reflektira njegove misli u drugome obliku pa ih korisnik može sagledati s odmakom potrebnim za pokretanje novih vlastitih misli (Elam, 2023). Treba uzeti u obzir da ideje i stilovi koje model može ponuditi nisu sami po sebi inovativni, već mogu biti novi i nevideni pojedinomu piscu. Fond na kojemu je model treniran služi kao knjižnica ideja i ne može korisniku ponuditi nešto u srži revolucionarno, već ideju isfiltriranu iz postojećih sadržaja kojima korisnik možda nije imao pristup, nije imao priliku susresti se s njima ili im se jednostavno nije dovinuo u danome trenutku. Schwartz (2018) također miče fokus s generiranja književnosti u smislu reflektiranja ljudskih tvorevina prema suradnji s računalom koja može rezultirati novim kreativnim djelovanjem.

Generativnoj umjetnoj inteligenciji može se prepustiti stvaranje književnoga teksta, ali taj proces još uvijek ne može biti dovršen bez uplitanja čovjeka. U svojevrsnoj ulozi suautora umjetna inteligencija može ponuditi temu ili smjer kretanja, čak i više njih, no čovjek će ili morati izabrati ili zadati modelu da sam odabere temu i prepusti se generiranju odgovarajućega teksta:

“Alati generativne umjetne inteligencije omogućuju kreativcima eksperimentiranje s više izbora u smislu kreativnoga smjera. Ali u središnjoj fazi kreativci moraju potvrditi da se rezultat poklapa s njihovim namjerama. Na kraju procesa ljudi moraju urediti i preraditi

sadržaj kako bi osigurali da odgovara i primjereno odražava stvaraočevu estetiku i namjeru” (Lawton, 2023).

Svojom mogućnošću statističkog analiziranja teksta i pristupom opširnome fondu model može korisniku ponuditi komentare na korisnikov originalni tekst. Prilikom analiziranja teksta model uspoređuje podatke s onima koji su pohranjeni u njegovoj povijesti učenja te može korisniku ponuditi varijacije rečenica, dodatke koji bi tekst učinili jasnijim ili bogatijim. Velika je prednost ovoga postupka njegovo djelovanje u stvarnome vremenu pa korisnik ne mora tekst slati beta-čitatelju ili uredniku te čekati na odgovor jer umjetna inteligencija analizu odradi u roku od nekoliko sekundi do nekoliko minuta, ovisno o veličini teksta i duljini odgovora. Čak i ako komentari nisu oni koje bi dali ljudi koji su pročitali tekst, brzina odgovora ima poticajan učinak na pisca te se on može posvetiti prepravkama bez negativnoga osjećaja koji bi mogla izazvati kritika (Washington, 2023).

U ulozi pomoćnoga alata umjetna inteligencija djeluje već neko vrijeme. Predviđanje riječi tijekom tipkanja i ispravljanje pravopisnih i gramatičkih pogrešaka već je implementirano u mnoge programe. Ovi se alati također zasnivaju na statistici te zadanim pravilima i mogu značajno smanjiti vrijeme potrebno za uređivanje teksta, iako je njihova učinkovitost različita ovisno o jeziku koji se koristi. Izvan književnosti umjetna se inteligencija vrlo efikasno koristi za ponavljajuće radnje poput promjene veličine i izrezivanja slika u grafičkom dizajnu te označavanje, uređivanje videomaterijala te transkripciju (Lawton, 2023). Ako se izuzme uređivanje, tijekom stvaranja književnoga teksta pisac ne nailazi na opetovane radnje koje oduzimaju vrijeme koje je moglo biti iskorišteno za pisanje samoga teksta, tako da se uloga umjetne inteligencije ponovno fokusira na mogućnost generiranja ideja ili ulomaka teksta koji mogu inspirirati pisca, bilo tako da ih iskoristi u svojem tekstu ili da od njih stvori nešto sasvim novo (Washington, 2023).

Negativna strana korištenja generativne umjetne inteligencije prilikom stvaranja književnoga teksta jednaka je potencijalnome riziku korištenja bilo koje tehnologije, a to je stvaranje ovisnosti o njezinoj pomoći. Autori upozoravaju (Lawton, 2023; Washington, 2023) da bi prekomjerno oslanjanje na umjetnu inteligenciju moglo dovesti do gubitka autentičnosti, kulturološkog konteksta ili do nedovoljne estetske razvijenosti s obzirom na to da su stil i kulturološki kontekst još uvijek nedostižni i uglavnom nerazumljivi čak i dobro istreniranim



modelima. Također povećanje produktivnosti za koje je zaslužan model može u korisnicima izazvati nesigurnost u svoje sposobnosti. Upravo taj osjećaj glavni je pokretač sada već globalnoga straha od masovnoga gubitka posla i radnih mjesta zbog uvođenja alata umjetne inteligencije u poslovni svijet (Hwang, 2022).

### 4.3 Budućnost književnosti

Budućnost generativne umjetne inteligencije mogla bi utjecati i na budućnost književnosti, bilo u okviru načina proizvodnje tekstova, bilo u okviru nove tematike i književnih pokreta kao reakcije na tehnološku evoluciju. De Cramer, Morini Bianzino i Falk (2023) predviđaju tri mogućnosti razvoja umjetne inteligencije: (1) dobra kombinacija velikih jezičnih modela i algoritama umjetne inteligencije pomoći će ljudima da brže osmisle veću količinu ideja i rješenja s kojima mogu eksperimentirati; (2) radovi umjetne inteligencije smanjit će potrebu za ljudskim radovima, što će smanjiti njihovu proizvodnju i usporiti napredak jer umjetna inteligencija treba fond sastavljen od ljudskih kreacija kako bi se razvijala; (3) veća količina generiranih proizvoda utjecat će na porast vrijednosti ljudskih proizvoda. Autori naglašavaju da ove tri opcije ne isključuju nužno jedna drugu, što bi značilo da su mogućnosti istovremeno ostvarive u različitim područjima.

U književnosti već sada možemo vidjeti sve tri opcije, iako generativna umjetna inteligencija najvjerojatnije nije ostvarila svoj puni potencijal. Korisnik može zadati upit nekome *chatu* koji koristi generativnu umjetnu inteligenciju i unutar nekoliko sekundi dobiti niz ideja koje bi mu mogle pomoći da prebrodi kreativnu blokadu ili razriješi neko problematično mjesto u tekstu. Čak i odgovori koji ne odgovaraju korisnikovim potrebama mogu pobuditi inspiraciju. Epstein i sur. (2023) u svome radu naglašavaju da je umjetna inteligencija alat kao i mnogi drugi te da bi mogla izmijeniti mnoge industrije, u početku na štetu ljudskih radnih mjesta, ali da bi na kraju mogla naći svoju primjenu u posve novim oblicima rada. Rafner i sur. (2023) također odabiru opciju u kojoj je najefikasniji i etički način rada suradnja čovjeka i umjetne inteligencije.

Druga je inačica budućnosti ostvariva s tehničke strane ako tvrtke koje razvijaju umjetnu inteligenciju ne osiguraju priljev svježih podataka na kojima će trenirati svoje modele. Hoće li taj

priljev biti dobrovoljan ili prisilan, osiguran zakonski, transakcijski ili će se čak raditi o masovnoj krađi podataka, teško je sa sigurnošću predvidjeti.

Na kraju, već se sada može vidjeti zaziranje od generiranih radova, kako umjetničkih općenito tako i književnih. Mogućnosti generativne umjetne inteligencije nisu osporive i korisnicima su zanimljive, ali kada žele konzumirati neki sadržaj još uvijek će radije odabrati autore, a ne algoritme. To potvrđuje i negativan stav publike na spomenuto priznanje autorice Rie Kudan da je koristila umjetnu inteligenciju prilikom pisanja svoga romana. Generativna kreativnost samo je odraz ljudske kreativnosti te se na nju nastavlja, a ne zamjenjuje ju, zbog toga se mijenjaju vrijednosti intelektualnoga rada (Tredinnick i Laybats, 2023). Epstein i sur. (2023) također ističu da korisnik prilikom susreta s umjetničkim djelom bilo koje vrste u njemu traži autorovu namjeru, poticaj za stvaranje djela, što u računalnih algoritama ne postoji. Takvo gledište podupire i Aydoğdu Çelik: "(model) ne posjeduje ljudsku kreativnost i emocionalnu inteligenciju koje su bitan sastavni dio velikih književnih djela. Književna djela generirana umjetnom inteligencijom nisu stvorena kreativnim procesom, već postupkom prepoznavanja uzoraka" (2023, str. 152).

Suradnja između ljudi i umjetne inteligencije i dalje ostavlja dojam najpovoljnije opcije. Pojedina bi (poslovna) područja mogla prosperirati ako se kao krajnji cilj korištenja umjetne inteligencije zadrži kvaliteta proizvoda (Torkington, 2023). Ono što je trenutno privlačno kod umjetne inteligencije jest njezina brzina odgovora, ali malo vrijedi serijal romana generiran u pet sekundi ako ti romani nisu čitljivi. Isto ističe i Khosrowshahi (2023) skrećući pozornost na potražnju – tržište mogu preplaviti generirani proizvodi, ali od njih neće biti koristi ako ih ljudi ne žele konzumirati. Također ističe da ljudi osjećaju povezanost s autorom umjetničkoga djela koje im se sviđa te pokazuju znakove idolatrije prema njemu, što je teško zamislivo s računalnim algoritmom. Ovakav stav prema autoru primjećuju i drugi autori: "Kada strojevi zamijene ljudske autore na književnome i umjetničkom području, lišit će društvo uzora za ljudski estetski angažman" (Senftleben, 2023, str. 1540).

Doduše, Senftleben također tvrdi kako ljudi možda neće više osjećati potrebu da razvijaju vlastite estetske prakse ili potrebu da se uključuju u igru ekspresije, s čime se nikako ne možemo složiti. U proučavanju evolucije ljudske kreativnosti Gabora i Kaufman (2010, str. 17) donose zaključak da je nagon za stvaranjem usporediv s nagonom za razmnožavanjem, odnosno da se

radi o sili prirode koja može nagnati ljude da ovjekovječe svoju kulturu. Autori navode da kreativnost proizlazi iz želje za razumijevanjem, objašnjavanjem svijeta te ovladavanjem njime, a razmišljanjem o svijetu oko sebe gradimo svoj unutarnji svijet koji je uvijek promjenjiv (Ibid). Elementi kulture se prenose generacijama podložni promjenama, a ostaju očuvani u svojoj srži zbog mogućnosti samoorganizacije i samoregeneracije, odnosno mogućnosti prilagodbe okolnostima. Književni tekst kao odraz kulture i dio šire vizije pojedinca odraz je individualnih doživljaja zajedničkoga konteksta, a računalni algoritam nema doživljaje niti stavove o kontekstu u kojemu se nalazi. Također treba uzeti u obzir vrednovanje umjetnosti kao takve kada je njezina vrijednost u samome procesu stvaranja: “Simulacija ili oponašanje glasa i stila izvedbe ili djela očito zanemaruje ulaganje vremena, vještine i novca koji su potrebni za postizanje vrijednog osobnog izričaja te je zbunjujuće i varljivo za javnost” (McGibbon i Solomon, 2024, str. 3). Naravno da će konačni ishod biti važan, ali razmišljanje koje dovodi promatrača ili čitatelja do određenih stavova o djelu koje ga trenutno okupira proizlazi iz sposobnosti improvizacije, pokušaja utjecaja na okolinu, prezentiranja individualnog pogleda umjetnika, a ne iz prepoznavanja uzoraka, odnosa uloženog i dobivenog niti “tehnički svezajuće anonimnosti” (Elam, 2023, str. 284).

Toubia (2023) priznaje da generativna umjetna inteligencija možda nikada neće moći stvoriti umjetničko djelo jednake kvalitete kao ono koje je stvorio čovjek, ali da će biti dovoljno profitabilno ako može stvoriti nešto približno toliko dobro, ali brže i jeftinije. Čak i ako se ta verzija budućnosti ostvari, ljudi neće nužno prestati stvarati kao što nisu niti do sada. Društvene mreže, internetske platforme ili samostalne izložbe dobar su pokazatelj da ljudi stvaraju i žele stvarati umjetnost neovisno o tome želi li im netko za tu umjetnost platiti te jednako tako postoji publika koja će tu umjetnost gledati, čitati i podržavati neovisno o tome što ju je autor sam stavio na raspolaganje.

## 5. Etika umjetne inteligencije

Kada se u obzir uzme vrsta tehnologije, kako ona nastaje, razvija se te kako se na koncu koristi, umjetna inteligencija općenito, a posebno generativna, otvara mnoga pitanja koja se tiču etičkih principa i stavova. Neka od tih pitanja već su neposredno postavljena u ovome radu. Kao i kod ostalih ljudskih izuma, teško je osigurati da će se umjetna inteligencija koristiti isključivo za dobro i korist čovječanstva, a ako pretpostavimo da smo u današnjemu vremenu tek na početku potencijalnoga većeg razvoja, možemo vidjeti da se etička komponenta ne kreće u pozitivnome smjeru. Etička pitanja mogu se podijeliti u tri skupine: etika ideja, etika razvoja i etika korištenja. Etika ideja odnosila bi se na početne korake prije samoga stvaranja modela i baze podataka, odnosno stavove profesionalaca uključenih u razvoj pojedinoga modela i njihovo poimanje etike i etičkih problema. Etika razvoja okuplja sva pitanja koja se postavljaju tijekom treniranja modela kao što su njegov cilj i odnos prema prikupljenim podacima. Etika korištenja najjednostavnija je u svojem značenju, ali vjerojatno najsloženija u smislu reguliranja jer uključuje neograničen broj korisnika koji dolaze u doticaj s modelom i sve one kojima su radovi generativne umjetne inteligencije predstavljeni u obliku gotovoga djela.

### 5.1 Etika ideja

Postoji nekoliko različitih dokumenata koji okupljaju etičke principe korištenja umjetne inteligencije koje tvrtke koje umjetnu inteligenciju razvijaju mogu potpisati, a dokumente stvaraju same tvrtke ili su ponuđeni prema geografskom položaju ili u sklopu neke međunarodne organizacije (npr. *SDAIA AI Ethics Principles* Saudijske Arabije, *NSW Mandatory Ethical Principles* Novog Južnog Walesa, *Etičke smjernice za pouzdanu umjetnu inteligenciju* Europske unije itd.). UNESCO je također predstavio prijedlog smjernica i principa za etičko korištenje umjetne inteligencije koji sadržava sve glavne potencijalne probleme te pitanja i način procjene rizika za svaki pojedini problem s kojim se susrećemo ili možemo susresti prilikom razvoja i korištenja umjetne inteligencije. Počevši od etike ideja, UNESCO naglašava: “države članice bi uvijek trebale osigurati preuzimanje etičke i zakonske odgovornosti za bilo koji dio životnoga ciklusa sustava umjetne inteligencije” (2023, str. 33). Čak i ako ljudi nisu uključeni u sam rad modela i njegove rezultate, oni bi ih trebali nadzirati i korigirati jer sam model nema nikakve

etičke principe niti je osjetljiv na kulturološki kontekst u kojemu bi se rezultati njegova rada mogli pojaviti. Iako možda djeluje banalno, ovo je vrlo važna smjernica jer društvo je već dobro antropomorfiziralo generativnu umjetnu inteligenciju i precijenilo njezine sposobnosti dok u isto vrijeme ne obraća dovoljno pozornosti na društvene probleme koje pristranost podataka za treniranje može produbiti (Bennett, 2023).

S druge strane pretpostavka da bi ljudi uklonili neetičke rezultate koje model ponudi zapravo je utopijska. Pokušaji da se kolegiji o etici uključe u studijske programe čija je specijalnost računarstvo i razvoj umjetne inteligencije zasad nisu dobro prihvaćani (prema Nourbakhsh, 2021), a to što tvrtke sastavljaju, donose ili potpisuju etičke regulative ne znači ujedno da su one u skladu s općim dobrom i pravednošću: “Neusklađenost javnih i privatnih vrijednosti vječno je iskušenje za korporacije, koje bi mogle skrenuti s etičkog smjera kako bi privilegirale privatne nad javnim interesima. (...) Ne možemo očekivati da njihova samoregulacija služi bilo kojoj svrsi izvan njihovih internih sustava vrijednosti” (Ibid, str. 44). Što se pak tiče državnih ili međunarodnih dokumenata, oni mogu biti previše općeniti da bi poslužili bilo čemu osim slobodnoj interpretaciji jer mnogi postupci, ovisno o ideološkim, političkim i filozofskim stavovima mogu biti prikazani kao “korisni za opće dobro” (Whittlestone, Nyrup, Alexandrova, Cave, 2019, str. 197). U svakome slučaju autori se zalažu za stvaranje formalnih standarda koji bi jasno određivali čega se proizvođači i korisnici modela trebaju držati, ali shvaćaju da je detaljniju, jasnu, a s druge strane univerzalnu regulativu teško osmisliti.

Provedena istraživanja i intervjui (Agbese, Mohanani, Khan, Abrahamsson, 2023; Pant, Hoda, Spiegler, Tantithamthavom, Turhan, 2024) s osobama na različitim pozicijama odgovornosti unutar tvrtki koje razvijaju modele umjetne inteligencije pokazuju da postoji svijest o pitanju etike u razvoju tehnologije, ali i svojevrsno svjesno zanemarivanje. Smatra se da su etički principi previše apstraktni da bi ih se moglo implementirati u kod kao matematičku vrijednost ili operaciju te da, kada bi to i bilo moguće, oni bi varirali od društva do društva ili čak od pojedinca do pojedinca, što bi značilo da se i uz najveći napor i dobru volju ne bi mogli zadovoljiti svi standardi (Pant i sur., 2024). Proučavanje i prilagođavanje ovog aspekta svakoga modela značajno bi produljilo vrijeme razvoja samoga modela, što tvrtkama nije isplativo, a jedan ispitanik izravno dodaje: “Ovisi o tržištu. Sve se svodi na potrošačeve želje i ako njih ove

stvari ne zabrinjavaju i ne traže od vas da strogo uzimate u obzir etička pitanja u razvoju proizvoda, ako oni nisu spremni platiti to, onda ja nisam spreman graditi takav model i zahtijevati iste stvari” (Agbese i sur., 2023, str. 66).

## 5.2 Etika razvoja

UNESCO nadalje u svome prijedlogu skreće pažnju na baze podataka na kojima se modeli treniraju. Već je bilo riječi o načinima i mogućnostima prikupljanja podataka te kako oni ne bi trebali biti “ukradeni” njihovim autorima (v. poglavlje 4.1 *Pitanje autorskoga prava na skupljene podatke*). Druga je važna stavka sigurnost podataka kada se radi o osjetljivim podacima, ali treba uzeti u obzir i namjerne napade putem podataka koji se u danome dokumentu nazivaju trovanje podataka, manipulacija ulaznih vrijednosti i zaključivanje o članstvu (2023, str. 18). Prvi napad odnosi se na namjerno izmjenjivanje rada modela tako da mu se pruže novi podaci koji nisu trebali biti korišteni kao dio baze za treniranje. Druga se vrsta napada odnosi na umetanje zlonamjernog sadržaja kako bi se prevario sustav, a treći se odnosi na mogućnost povezivanja nekoliko specifičnih individualnih podataka kako bi se dobio profil korisnika koji je trebao od početka do kraja rada biti anonimn. Ono na što UNESCO želi skrenuti pozornost jest činjenica da podaci mogu biti pristrani te zbog njih model može dati neprimjerene rezultate, ali jednako tako ti rezultati mogu biti namjerno izazvani, odnosno uvjetovani namjernim dodavanjem ili prilagođavanjem podataka za treniranje (zato se i zalažu za isključivo odgovornost čovjeka za rad modela).

## 5.3 Etika korištenja

Uvjeti i pravila korištenja javno dostupnih modela generativne umjetne inteligencije sadržavaju napomene o savjesnom korištenju. Neke od tih napomena zabranjuju korištenje tehnologije u svrhu ilegalnih ili štetnih aktivnosti, a poseban je naglasak stavljen na prisvajanje ponuđenih odgovora ili prikazivanje generiranih rezultata kao djela ljudskoga rada (detaljnije u <https://openai.com/policies/terms-of-use/>). Ove dvije zabrane, posebice druga, usko su povezane s principom transparentnosti:

“(ova vrsta) odgovornosti drugačije je naravi od primjerice one softverskih inženjera ili tehnoloških kompanija koje koriste umjetnu inteligenciju za stvaranje alata za naš svakodnevni život. Umjetnici koji rade s umjetnom inteligencijom moraju prepoznati ovu odgovornost, ali vjerojatno i mi kao publika možemo podijeliti ovaj teret tako što ćemo istražiti i razumjeti kontekst u kojemu nastaju djela umjetne inteligencije” (Bennett, 2023, str. 55).

Možemo reći da se radi o razlici između crteža i isprintane preslike za koju ne bi nikako imalo smisla tvrditi da je nacrtana rukom. Možda je umjetnikova misao jedna, ali tehnike su različite pa ako možemo gledati na umjetnu inteligenciju kao alat, ne vidimo razloga da je skrivamo od publike kao što ne bismo skrivali tehniku ulja na platnu. Što se književnosti tiče, ona već sama po sebi uključuje vremenski angažman i pretpostavku da je autor s posebnom mišlju ili inspiracijom izabrao korištene riječi ili su one pak posljedica jedinstvenoga stila i truda. Razumljivo je da se čitatelji mogu osjećati na neki način prevareno ako se ispostavi da je tekst rezultat matematičkog procesa, pogotovo kada se to od njih skrivalo. Razumijevanje cijeloga procesa stvaranja djela s pomoću modela generativne umjetne inteligencije može otkloniti strah od tehnologije i spriječiti nesporazume među korisnicima ili onima kojima se djela predstavljaju (UNESCO, 2023).

Spomenuti osjećaj prevare proizlazi upravo iz toga što su ljudi naviknuti stvarati i promatrati umjetnost kao prikazivanje doživljaja u “pojačanome obliku” (Jiang i sur., 2023, str. 365). Taj prikaz doživljaja uvijek u nekoj mjeri u vidu ima i publiku kojoj će se djelo obratiti, odnosno predviđanje kako bi djelo moglo utjecati na publiku ili kako autor želi da djelo utječe na publiku. Autori u istome članku ističu da je upravo ta namjera ono što čini umjetnost, a model nema i ne može imati namjeru, već samo krajnji cilj, izlaznu vrijednost koju (više ili manje uspješno) može ponuditi čovjeku koji ima namjeru. Također umjetnost prikazuje “apstraktne interpretacije između reprezentativnih i imaginarnih subjekata, tema i, naravno, osobnih osjećaja i iskustava” na temelju kojih se ljudi međusobno povezuju i pokušavaju razumjeti (Ibid, str. 366).

## 6. Zaključak

Napredak koji velikom brzinom ostvaruje umjetna inteligencija je neporeciv. Združivanje povratnih neuronskih mreža s velikim bazama podataka koji se mogu obrađivati postavilo je temelje za posve nov smjer kretanja razvoja ove tehnologije. Mogućnost nelinearnoga kretanja kroz baze podataka, primjećivanje i određivanje vlastitih kategorija i principa po kojima će se podaci klasificirati osiguralo je modelima umjetne inteligencije da operiraju na razini mnogo višoj od pukoga reagiranja na upit unaprijed zadanim odgovorima. Analiziranje, razumijevanje i korištenje prirodnoga jezika promijenilo je načine na koje korisnici komuniciraju s modelom, pretražuju i organiziraju informacije.

Modeli generativne umjetne inteligencije mogu biti trenirani na bilo kojemu jeziku, prevoditi tekstove ili ponuditi solidnu pomoć prilikom prevođenja, ali pokazuju znakove nesnalaženja s analizom teksta koja nadilazi leksičku razinu, odnosno modeli su zasad neosjetljivi na kulturološke kontekste koji su utjecali na stvaranje teksta te ih tako ne mogu niti reproducirati. Veća im je mana to što će na svaki upit ponuditi rješenje, iako ono ne mora nužno postojati ili je ono negativno. Istrenirani da reflektiraju statistički najvjerojatnije podatke, modeli ne mogu ponuditi odgovor koji će odražavati nesigurnost modela u odgovor. Iako je stilistika također jedna od problematičnih razina, noviji modeli već pokazuju znakove uspješnoga kopiranja stila autora, a u neknjiževnim tekstovima mogu proizvesti zadovoljavajuću strukturu (naprimjer strukturu novinskoga članka). S obzirom na to da model mora proći proces obrade podataka i treniranja na njima, on ne može uvijek ponuditi odgovore koji odražavaju trenutno stanje realnoga konteksta, nego uvijek malo zaostaju. Također, čak i kada bi imali konstantan priljev svježih podataka, oni bi ih mogli isključivo reflektirati, a ne bi mogli ponuditi subjektivnu reakciju kakvu u čovjeku izazivaju zbivanja oko njega.

Kako bi se modeli nastavili razvijati u pozitivnome pravcu, tvrtke su počele raditi upravo na tomu osiguravanju priljeva podataka, što se ponekad kosi s autorskim pravima te različite organizacije ovom problemu pristupaju na različite načine. Predloženi programi isplate naknade autorima za korištenje njihovih djela u treniranju umjetne inteligencije uvelike bi usporili cijeli proces razvoja tako da tvrtke počinju u svojim uvjetima korištenja platformi i društvenih mreža uključivati stavku da korištenjem platforme korisnik pristaje na to da tvrtka koristi njegove



podatke za treniranje umjetne inteligencije. Generativna umjetna inteligencija još nije uspjela samostalno stvoriti kvalitetno književno djelo pa je tako teško govoriti o budućnosti rješavanja autorskoga prava na takve radove, ali pretpostavlja se da će ono pripasti vlasniku modela ili osobi koja s vlasnikom ima uređen poseban ugovor o autorskim pravima.

Bez obzira na negativne prognoze, uglavnom neutemeljene, književna djelatnost neće zamrijeti. Čovjek je po prirodi kreativno biće i bez obzira na okolnosti ima potrebu stvarati. Kroz svoje stvaranje čovjek razumije svijet oko sebe, održava tradiciju i kulturu ili joj prkosi te pokušava u umjetnosti ovjekovječiti svoje postojanje. Korištenje generativne umjetne inteligencije može biti vrlo pogodno za autore i poticajno za njihovu inspiraciju, to jest modeli mogu biti dobri pomoćni alati prilikom nadilaženja kreativne blokade ili pregledavanja i uređivanja teksta. Jedina prijetnja u industriji zabave i umjetnosti su tvrtke kojima cilj neće biti kvalitetan, već brzo dovršen i jeftin proizvod. Svejedno, čak i kada bi nakladničke kuće imale stručnjake za generiranje tekstova i prestale objavljivati djela koja su napisali ljudi, pisaca bi i dalje bilo te bi se oni koristili drugim načinima dijeljenja svojih radova, kao što je to već sada slučaj.

## 7. Literatura

- Agbese, M., Mohanani, R., Khan, A. A. i Abrahamsson, P. (2023). Implementing AI ethics: making sense of ethical requirements. *Proceedings of the International Conference on Evaluation and Assessment in Software Engineering*, 62–71.
- Ajdagulova, A. R. (2023). Osobnosti tekstov, sgenerirovannyh iskusstvennym intellektom. *Vestnik Baškirkoskogo gosudarstvennogo pedagogičeskogo universiteta im. M. Akmully*, 4(72), 154–156.
- Artamonova, M. V., Mambetov A. A. i Tulina, E. V. (2023). Čat-bot kak instrument v rabote perevodčika. *Litera*. 8, 235–253.
- Artenij, L. S. (2019). Iskusstvennyj intellekt v avtorskom prave. *Vestnik nauki i obrazovanija*, 7(61). 76–81.
- Aydoğdu Çelik, M. (2023). Death of the author: a survey on artificial intelligence in literature. *Journal of communication science researches*, 3(2), 142–154.
- Barsegjan, A. A., Kuprijanov, M. S., Stepanenko, V. V. i Holod, I. I. (2007a). Analiz tekstovoj informacii – Text Mining. U Kondukova, E. (ur.) *Tehnologii analiza dannyh: Data Mining, Visual Mining, Text Mining, OLAP* (str. 194–224). Sankt Peterburg: BHV-Peterburg.
- Barsegjan, A. A., Kuprijanov, M. S., Stepanenko, V. V. i Holod, I. I. (2007b). Intelektual'nyj analiz dannyh. U Kondukova, E. (ur.) *Tehnologii analiza dannyh: Data Mining, Visual Mining, Text Mining, OLAP* (str. 58–91). Sankt Peterburg: BHV-Peterburg.
- Bennett, S. J. (2023). AI, ethics, and the role of the artist. *The New Real Magazine*, 1, 54–55.
- Bergson, H. (2001). *Tvorčeskaja evoljucija*. Moskva: Knižnyj klub Terra.
- Bol'šakova, E. I. (2017). Komp'juternaja lingvistika: zadači, podhody, resursy. U *Avtomatičeskaja obrabotka tekstov na estestvennom jazyke i analiz dannyx* (str. 7–31). Moskva: NIU VŠE.

Bol'shakova, E. I. i Efremova, N. E. (2017). Izvlečenje informacii iz tekstov: portret napravlenija. U *Avtomatičeskaja obrabotka tekstov na estestvennom jazyke i analiz dannyx* (str. 83–127). Moskva: NIU VŠE.

*Characters*. (bez dat.) Sven Hassel. *Official site*. Preuzeto 15. travnja 2024. s <https://www.svenhassel.com/characters/>.

De Cremer, D., Morini Bianzino, N. i Falk, B. (2023). *How Generative AI Could Disrupt Creative Work*. Preuzeto 7. ožujka 2024. s <https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work>.

Ducato, R. i Strowel, A. (2021). Ensuring Text and Data Mining: Remaining Issues With the EU Copyright Exceptions and Possible Ways Out. *CRIDES Working Paper Series*, 1/2021, 322–337.

Elam, M. (2023). Poetry will not optimize; or, what is literature to AI? *American Literature*, 95(2), 281–303.

Epstein, Z., Hertzmann, A., Herman, L., Mahari, R., Frank, M. R., Groh, M., ... Russakovsky, O. (2023). Art and science of generative AI: a deeper dive. *Science Perspectives*. Preuzeto 11. srpnja 2024. s <https://arxiv.org/abs/2306.04141>.

Gabora, L. i Kaufman, S. B. (2010). Evolutionary Approaches to Creativity. U Kaufmann, J. i Stenberg, R. (ur.), *The Cambridge handbook of creativity*. Cambridge: Cambridge University Press.

Gaikwad, S. V., Chaugule, A. i Patil, P. (2014). Text mining methods and techniques. *International Journal of Computer Applications*, 85(17), 42–45.

Griewank, A. (2012). *Who Invented the Reverse Mode of Differentiation*. Preuzeto 15. ožujka 2024. s <https://www.semanticscholar.org/paper/Who-Invented-the-Reverse-Mode-of-Differentiation-Griewank/8ff0c546aff84566635f0a9a2e01feb3d6588c1c>.

Holdsworth, J. (2024). *What is NLP?* Preuzeto 11. srpnja 2024. s <https://www.ibm.com/topics/natural-language-processing>.

Hwang, A. H. (2022). Too late to be creative? AI-empowered tools in creative processes. *CHI '22: AC; CHI Conference on human factors in computing systems*. Preuzeto 11. srpnja 2024. s <https://doi.org/10.1145/3491101.3503549>.

Jackson, J. (2002). Data mining; a conceptual overview. *Communications of the Association for Information Systems*, 8, 266–297.

Jiang, H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., ... Gebru, T. (2023). AI art and its impact on artist. *AAAI/ACM Conference on AI, Ethics and Society*, 363–374.

Karpathy, A. (2015). The Unreasonable Effectiveness of Recurrent Neural Networks. *Andrej Karpathy blog*. Preuzeto 8. ožujka 2024. s <https://karpathy.github.io/2015/05/21/rnn-effectiveness/>.

Khosrowshahi, G. (10. studenog 2023). *Generative AI and Human Creativity*. Preuzeto 7. ožujka 2024. s <https://www.youtube.com/watch?v=CnDn CZTRVBc>.

Klyšinskij, E. S. i Sanin, A. S. (2017). Morfoložičeskij analiz tekstov. U *Avtomatičeskaja obrabotka tekstov na estestvennom jazyke i analiz dannyx* (str. 31–83). Moskva: NIU VŠE.

*Kyodo*. (2. veljače 2024). *Japan author sparks debate after revealing she used AI in book that won top award*. Preuzeto 20. travnja 2024. s <https://www.scmp.com/news/asia/east-asia/article/3250875/japan-author-sparks-debate-after-revealing-she-used-ai-novel-won-top-award>.

Lawton, G. (2023). *How generative AI is changing creative work*. Preuzeto 7. ožujka 2024. s <https://www.techtarget.com/searchenterpriseai/feature/How-generative-AI-is-changing-creative-work>.

Lihačev, E. V. (2023). Problema konstruirvanija jumorističeskikh vyskazaniy generativnym iskusstvennym intellektom (na materiale nemeckogo jazyka). *Učenyje zapiski Krymskogo federal'nogo universiteta imeni V. I. Vernadskogo. Filologičeskie nauki. Naučnyj žurnal*. 9(75), 92–100.

Lukašević, N. V. (2017). Avtomatičeskie metody analiza tonal'nosti. U *Avtomatičeskaja obrabotka tekstov na estestvennom jazyke i analiz dannyx* (str. 127–195). Moskva: NIU VŠE.

Mahmud, B. U. i Fong, B. (2022). A study of human-AI symbiosis for creative work: recent developments and future directions in deep learning. *ACM Transactions on Multimedia Computing, Communications and Applications*. Preuzeto 11. srpnja 2024. s [https://www.researchgate.net/publication/362319289\\_A\\_Study\\_of\\_Human-AI\\_Symbiosis\\_for\\_Creative\\_Work\\_Recent\\_Developments\\_and\\_Future\\_Directions\\_in\\_Deep\\_Learning](https://www.researchgate.net/publication/362319289_A_Study_of_Human-AI_Symbiosis_for_Creative_Work_Recent_Developments_and_Future_Directions_in_Deep_Learning).

McGibbon, S. i Solomon, N. (2024). AI and creative work. *Harvard Data Science Review. Special issue 5: Grappling with the generative AI revolution*. Preuzeto 11. srpnja 2024. s <https://hdsr.mitpress.mit.edu/pub/p6q84rww/release/1>.

Mil'gizin, I. E. i Baeva, L. V. (2017). K voprosu o kreativnosti v nejrosetjah iskusstvennogo intelekta. *Filosofskie problemy informacionnyh tehnologij i kiberprostranstva*. 1(13), 62–72.

Moawad, H. (2023). *What part of the brain controls speech?* Preuzeto 20. ožujka 2024. s <https://psychcentral.com/health/what-part-of-the-brain-controls-speech>.

Murel, J. i Kavlakoglu, E. (2023). What are stemming and lemmatization? Preuzeto 17. ožujka 2024. s <https://www.ibm.com/topics/stemming-lemmatization>.

Nourbakhsh, I. R. (2021). AI ethics: a call to faculty. *Communications of the ACM*, 64(9), 43–45.

*OpenAI*. (2022). Introducing ChatGPT. Preuzeto 9. ožujka 2024. s <https://openai.com/blog/chatgpt>.

Pandžić, I. (2015). Oblikovanje korjenovatelja za hrvatski jezik. *Časopis Instituta za hrvatski jezik i jezikoslovlje*, 41(2), 301–327.

Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C. i Turhan, B. (2024). Ethics in the age of AI: an analysis of AI practitioners' awareness and challenges. *ACM Transactions on Software and Methodology*, 33(3).

Plotnikova, A. M. (2023). Nejroset' kak ključevoe slovo tekuščego momenta. *Filologičeskij klass*. 2(28), 45–54.

Prijedlog Akta o umjetnoj inteligenciji Europske unije prema Uredbi Europskog parlamenta i Vijeća o utvrđivanju usklađenih pravila o umjetnoj inteligenciji. (2024). Preuzeto 20. travnja 2024. s <https://artificialintelligenceact.eu/the-act/>.

*Računalno jezikoslovlje*. (bez dat.) U *Struna. Hrvatsko strukovno nazivlje*. Institut za hrvatski jezik i jezikoslovlje. Preuzeto 12. travnja 2024. s <http://struna.ihj.hr/naziv/racunalno-jezikoslovlje/51811/#naziv>.

Rafner, J., Beaty, R. E., Kaufman, J. C., Laubart, T. i Sherson, J. (2023). Creativity in the age of generative AI. *Nature Human Behaviour*, 7, str. 1836–1838.

Schwartz, O. (2018). Competing Versions of AI. *Digital Culture and Society*, 4(1), 87–105.

Senftleben, M. (2023). Generative AI and Author Remuneration. *IIC – International Review of Intellectual Property and Competition Law*, 54, 1535 – 1560.

Šuman, S. (2021). Pregled metoda obrade prirodnih jezika i strojnog prevođenja. *Zbornik Veleučilišta u Rijeci*, 9(1), 371–384.

Tel'pov, R. E. i Larcina, L. S. (2023). Tipovye različija estestvennyh i sgenerirovannyh nejronnoj set'ju tekstov v kvantitativnom aspekte. *Naučnij dialog*. 7(12), 47–65.

Torkington, S. (2023). *How might generative AI change creative jobs?* Preuzeto 7. ožujka 2024. s <https://www.weforum.org/agenda/2023/05/generative-ai-creative-jobs/>.

Toubia, O. (10. studenog 2023). *Generative AI and Human Creativity*. Preuzeto 7. ožujka 2024. s <https://www.youtube.com/watch?v=CnDn CZTRVBc>.

Tredinnick, L. i Laybats, C. (2023). Black-box creativity and generative artificial intelligence. *Business Information Review*, 40(3), str. 98–102.

UNESCO. (2023). *Ethical impact assessment. A tool of the recommendation on the ethics of artificial intelligence*. Preuzeto 12. srpnja 2024. s <https://www.unesco.org/ethics-ai/en/eia>.

Uredba predsjednika Ruske Federacije O razvoju umjetne inteligencije u Ruskoj Federaciji (2019/2024). № 490. [Ukaz Prezidenta Rossijskoj Federacii O razvitii iskusstvennogo intellekta

v Rossijskoj Federaciji] Preuzeto 11. ožujka 2024. s

<http://www.kremlin.ru/acts/bank/44731/page/1>.

Washington, J. W. (2023). *The Impact of Generative Artificial Intelligence on Writer's Self-Efficacy: A Critical Literature Review*. Preuzeto 7. ožujka 2024. s <https://ssrn.com/abstract=4538043>.

Van Heerden, I. i Bas, A. (2021), AI as author – bridging the gap between machine learning and literary theory. *Journal of Artificial Intelligence Research*, 71, 175–189.

Wang, H., Wu, H., He, Z., Huang, L. i Church, K. W. (2022). Progress in machine translation. *Engineering*, 18, 143–153.

Weiss, E. A. (1992). Biographies: Eloge: Arthur Lee Samuel (1901-1990). *IEEE Annals of the History of Computing*, 14(3), 55–69.

*Wernickeovo područje*. (bez dat.) U *Struna. Hrvatsko strukovno nazivlje*. Institut za hrvatski jezik i jezikoslovlje. Preuzeto 20. ožujka 2024. s

<http://struna.ihj.hr/naziv/wernickeovo-podrucje/25968/>.

Whittlestone, J., Nyrupe, R., Alexandrova, A. i Cave, S. (2019). The role and limits of principles in AI ethics: towards a focus on tensions. *AAAI/ACM Conference on AI, Ethics, and Society*, 195–200.

Zakon o autorskom pravu i srodnim pravima NN 111/21. Preuzeto 20. travnja 2024. s <https://www.zakon.hr/z/106/Zakon-o-autorskom-pravu-i-srodnim-pravima>.

## Popis slika

Slika 1. Sastavnice procesa *data mining*. *Data Mining Process Vector Illustration Concept*.

[Slika] (bez dat.) Preuzeto 11. srpnja 2024. s

<https://www.shutterstock.com/image-vector/data-mining-banner-web-icon-vector-2304502967>.

Str. 8.

Slika 2. Sastavnice procesa *text mining*. *Text mining Concept Icons*. [Slika] (bez dat.) Preuzeto 11. srpnja 2024. s

<https://www.shutterstock.com/image-vector/text-mining-concept-icons-written-resources-2178507869>. Str. 10.

Slika 3. Slijed obrade podataka tijekom *text mininga*. Str. 12.

Slika 4. Povratna neuronska mreža. *Neural Network Algorithm*. [Slika] (bez dat.) Preuzeto 11. srpnja 2024. s

<https://www.shutterstock.com/image-illustration/tech-icon-neural-network-algorithm-stock-1747591643>. Str. 15.

Slika 5. Snimka zaslona; *ChatGPT*. Str. 23.

Slika 6. Snimka zaslona; *ChatGPT*. Str. 23.

Slika 7. Snimka zaslona; *Microsoft Copilot*. Str. 24.

Slika 8. Snimka zaslona; *Microsoft Copilot*. Str. 25.

Slika 9. Snimka zaslona; *ChatGPT*. Str. 28.

Slika 10. Snimka zaslona; *ChatGPT*. Str. 28.



## Primjena i problematika umjetne inteligencije u književnosti

### Sažetak

Modeli generativne umjetne inteligencije svakoga dana postaju sve dostupniji poslovnim i privatnim korisnicima te se, kako eksperimentalno tako i s ciljem kreiranja proizvoda za krajnjeg korisnika, počinju koristiti i u stvaranju književnih djela. Stavovi stručnjaka iz područja računalnih znanosti, filologije i raznih umjetničkih područja podijeljeni su u pogledu budućeg utjecaja umjetne inteligencije na stvaralaštvo. Cilj je ovoga rada analizirati načine na koje modeli generativne umjetne inteligencije obrađuju i generiraju prirodni jezik kako bi se odredile granice njihovih mogućnosti i potencijalne prijetnje autorskome radu, pravima i etičkim obvezama prema krajnjim korisnicima.

U radu se donosi pregled razvoja tehnologije koju se naziva umjetnom inteligencijom te njezini glavni mehanizmi obrade podataka, odnosno *data mining* i *text mining*, te povratne neuronske mreže čije je nelinearno obrađivanje podataka omogućilo stvaranje kvalitetnijih izlaznih vrijednosti. Kako je fokus istraživanja stavljen na književnost, posebno se posvećuje pažnja računalnom jezikoslovlju te radu s prirodnim jezikom u smislu mogućnosti generiranja, prevođenja, analize te stilskih obilježja teksta koje model umjetne inteligencije može ponuditi. Postupci obrade, razumijevanja i stvaranja na prirodnome jeziku popraćeni su analizama i usporedbama s načinima na koje uče ljudi, usvajaju i koriste jezik te kako nastaju ideje i razmatranjima što je to zapravo kreativnost. Neovisno o kvaliteti ponuđenih rezultata javlja se problem autorskoga prava na podatke koji se koriste za treniranje te generirane tekstove. Trenutno se pokušavaju osmisliti pravedna rješenja podjele prava ili kompenzacije izvorima i korisnicima, a to uključuje i etička pitanja pravednog korištenja, zloupotrebe modela ili rezultata.

Modeli generativne umjetne inteligencije sposobni su ponuditi tekstove koji su jezično smisleni, ali umjetnički bezvrijedni. Brzina generiranja teksta profitabilna je ako čitateljska publika ne zahtijeva autorsku intenciju, što je rijetko slučaj, a cijeli proces zadavanja cilja, usmjeravanja i korigiranja rezultata još uvijek zahtijeva ljudski nadzor.

**Ključne riječi:** generativna umjetna inteligencija, književnost, text mining, kreativnost, autorsko pravo

## **Application and issues of artificial intelligence in literature**

### **Summary**

Models of generative artificial intelligence are becoming more and more available to business and private users every day and, both experimentally and with the aim of creating products for the end user, they are also starting to be used in the creation of literary works. The views of experts from the fields of computer science, philology and various artistic fields are divided regarding the future impact of artificial intelligence on creativity. The aim of this paper is to analyze the ways in which generative artificial intelligence models process and generate natural language in order to determine the limits of their possibilities and potential threats to author's work, rights and ethical obligations towards end users.

The paper provides an overview of the development of technology called artificial intelligence and its main mechanisms of data processing, namely data mining and text mining, and recurrent neural networks whose non-linear data processing enabled the creation of higher quality output. As the focus of the research is on literature, special attention is paid to computational linguistics and work with natural language in terms of the possibility of generation, translation, analysis and stylistic features of the text that the artificial intelligence model can offer. The processes of processing, understanding and creation in natural language are accompanied by analysis and comparisons with the ways in which people learn, adopt and use language, as well as how ideas are created and considerations of what creativity actually is. Regardless of the quality of the offered results, the problem of copyright arises for the data used for training and the generated texts. Currently, attempts are being made to devise fair solutions for the distribution of rights or compensation to sources and users, and this includes ethical issues of fair use, misuse of models or results.

Generative artificial intelligence models are capable of offering texts that are linguistically meaningful but artistically worthless. The speed of text generation is profitable if the readership does not require the author's intention, which is rarely the case, and the whole process of setting the goal, directing and correcting the results still requires human supervision.

**Keywords:** generative artificial intelligence, literature, text mining, creativity, copyright