

Inovacije u pretraživanju informacija: uloga generativne umjetne inteligencije

Ljutić, Iva

Undergraduate thesis / Završni rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zadar / Sveučilište u Zadru**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:162:090656>

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

Download date / Datum preuzimanja: **2024-11-06**



Sveučilište u Zadru
Universitas Studiorum
Jadertina | 1396 | 2002 |

Repository / Repozitorij:

[University of Zadar Institutional Repository](#)



Sveučilište u Zadru

Odjel za informacijske znanosti

Sveučilišni prijediplomski studij

Informacijske znanosti



Zadar, 2024.

Sveučilište u Zadru
Odjel za informacijske znanosti
Sveučilišni prijediplomski studij
Informacijske znanosti

Inovacije u pretraživanju informacija: uloga generativne
umjetne inteligencije

Završni rad

Studentica:

Iva Ljutić

Mentor:

Doc. dr. sc. Ante Panjkota

Komentorica:

Dr. sc. Nikolina Peša Pavlović

Zadar, 2024.



Izjava o akademskoj čestitosti

Ja, **Iva Ljutić**, ovime izjavljujem da je moj **završni** rad pod naslovom **Inovacije u pretraživanju informacija: uloga generativne umjetne inteligencije** rezultat mojega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na izvore i radove navedene u bilješkama i popisu literature. Ni jedan dio mojega rada nije napisan na nedopušten način, odnosno nije prepisan iz necitiranih radova i ne krši bilo čija autorska prava.

Izjavljujem da ni jedan dio ovoga rada nije iskorišten u kojem drugom radu pri bilo kojoj drugoj visokoškolskoj, znanstvenoj, obrazovnoj ili inoj ustanovi.

Sadržaj mojega rada u potpunosti odgovara sadržaju obranjenoga i nakon obrane uređenoga rada.

Zadar, rujan 2024.

Sadržaj

1. Uvod	1
2. Pretraživanje informacija	3
2.1. Osnovni koncepti pretraživanja informacija	3
2.2. Komponente sustava za pretraživanje informacija i interakcija sustava s korisnicima	5
2.3. Pretraživanja informacija u digitalnom dobu i generativna umjetna inteligencija	7
3. (Generativna) umjetna inteligencija	9
3.1. Osnovni principi rada (generativne) umjetne inteligencije	10
3.2. Alati (generativne) umjetne inteligencije	12
3.3. Područja primjene (generativne) umjetne inteligencije	13
4. Razvoj (generativne) umjetne inteligencije	16
5. Primjena generativne umjetne inteligencije u pretraživanju informacija	19
5.1. Načini primjene generativne umjetne inteligencije u pretraživanju informacija .	20
5.2. Usporedba mogućnosti generativne umjetne inteligencije s klasičnim sustavima za pretraživanja informacija	22
5.3. Informacijsko preopterećenje i generativna umjetna inteligencija	24
6. Prednosti i izazovi primjene generativne AI u pretraživanju informacija	25
7. Primjeri sustava za pretraživanje temeljenih na generativnoj inteligenciji	27
8. Zaključak	33
9. Literatura	34

Sažetak

Predmet rada je analiza inovacija u pretraživanju informacija s posebnim fokusom na ulogu generativne umjetne inteligencije. Cilj je istražiti kako generativna umjetna inteligencija utječe na tradicionalne sustave pretraživanja te identificirati ključne prednosti i izazove koje donosi. Motivacija za proučavanje ovog predmeta leži u rastućoj složenosti i količini dostupnih informacija u digitalnom dobu, što stvara potrebu za učinkovitijim i preciznijim sustavima pretraživanja. Generativna umjetna inteligencija nudi nove mogućnosti personaliziranog pretraživanja, sažimanja sadržaja i rješavanja problema informacijskog preopterećenja, što je čini potencijalno ključnom tehnologijom u informacijskom dobu. Metodologija rada temelji se na analizi i prikazu relevantne literature te usporedbi između tradicionalnih sustava pretraživanja i novih modela koji koriste generativnu umjetnu inteligenciju. Kako bi dobili uvid u suvremene trendove, u radu su korištene studije slučaja i tehnološki alati. Rezultati pokazuju da generativna umjetna inteligencija značajno poboljšava korisničko iskustvo kroz napredne metode personalizacije, automatizacije i analize podataka, ali ipak izazovi ostaju u etičkom segmentu, sigurnosti i pristranosti algoritama. U konačnici, generativna umjetna inteligencija ne samo da unapređuje pretraživanje informacija, već i redefinira način na koji korisnici pristupaju informacijama, pružajući dublje, preciznije rezultate. Ovo otvara prostor za daljnje inovacije, ali zahtijeva i odgovorno korištenje kako bi se minimizirali potencijalni rizici.

Ključne riječi: Pretraživanje informacija, generativna umjetna inteligencija, strojno učenje, veliki jezični modeli, sustavi za pretraživanje informacija

Popis korištenih kratica

AI Artificial Intelligence = umjetna inteligencija

GAI Generative Artificial Intelligence = generativna umjetna inteligencija

GAN Generative Adversarial Networks = generativna kontradiktorna mreža

GDR Generative Document Retrieval = generativno dohvaćanje dokumenata

IIR Intelligent Information Retrieval = inteligentno pretraživanje informacija

IR Information Retrieval = pretraživanje informacija

LLM Large Language Model = veliki jezični model

MIT Massachusetts Institute of Technology = privatno sveučilište u Cambridgeu

RAG Retrieval Augmented Generation = generiranje s proširenim dohvaćanjem

SLIP Symmetric List Processor = programski jezik

1. Uvod

U digitalnom dobu pretraživanje informacija postalo je svakodnevna aktivnost, kako u poslovnom tako i u privatnom okruženju. „Tradicionalni pristupi upravljanja velikim zbirkama informacija potječu iz discipline knjižničarstva (Sanderson i Croft 2012).“ Iako je upravljanje informacijama izvorno bilo u domeni knjižničarstva, s razvojem tehnologije, pretraživanje informacija postalo je dostupno široj populaciji putem modernih web tražilica. „Danas, ne samo da su knjižničari uključeni u aktivnosti pronalaženja informacija, već se stotine milijuna ljudi svakodnevno bave IR-om kada koriste web tražilice. Vjeruje se da je pronalaženje informacija dominantan oblik pristupa informacijama (Geeks for Geeks 2020).“ S obzirom na stalno rastući broj dostupnih informacija, postavlja se pitanje kako učinkovito pronaći i odabrati relevantne informacije. Jedno od mogućih rješenja je razvoj naprednih sustava pretraživanja koji koriste generativnu umjetnu inteligenciju za predviđanje i personalizaciju rezultata, omogućujući korisnicima brži pristup preciznijim informacijama. U tom kontekstu, generativna umjetna inteligencija igra sve važniju ulogu u stvaranju inovativnih pristupa pretraživanju informacija.

Hambarde i Proenca u radu *Information retrieval: recent advances and beyond* ističu kako su tradicionalni pristupi pretraživanju informacija često ograničeni učinkovitošću i relevantnošću rezultata te ukazuju na probleme s kojima se suočavaju korisnici, a to su najčešće informacijska preopterećenost te preopterećeni i nejasni rezultati pretraživanja (Hambarde i Proenca 2023). Kako bi se ti problemi prevladali, generativna umjetna inteligencija donosi nove mogućnosti i alate koji mijenjaju način na koji se pristupa pretraživanju informacija. Sukladno tomu, Hersh u radu *Search Still Matters: Information Retrieval in the Era of Generative AI* naglašava da su sustavi za pretraživanje informacija naširoko korišteni alati za traženje informacija u različitim znanstvenim područjima, a pojava generativne umjetne inteligencije uvelike je utjecala na daljnji razvoj tih sustava. U današnje vrijeme su sustavi pretraživanja, kao što su Google ili Bing za opće pretraživanje weba, prošireni mogućnostima umjetne inteligencije te se pretraživanje interneta i drugih kolekcija podataka ili bibliografskih zapisa promijenilo na način koji mnogi nisu niti očekivali (Hersh 2023). Ova transformacija omogućila je korisnicima da dobiju preciznije i relevantnije rezultate u kraćem vremenskom roku, pritom smanjujući informacijsko preopterećenje i povećavajući zadovoljstvo (i iskustvo) korisnika.

Međutim, uz sve prednosti koje donosi, kao što su precizniji i relevantniji rezultati u kraćem vremenskom roku, integracija generativne umjetne inteligencije u sustave za

pretraživanje informacija otvara niz novih pitanja. Kako generativni modeli postaju sve složeniji, potrebno je dublje razumijevanje njihovih mogućnosti, ali i ograničenja kako bi se osiguralo da njihova primjena korisnicima donese maksimalnu korist. Pritom, kao što Hersh ističe, izazov nije samo u tehničkom napretku, već i u balansiranju između kvalitete i brzine pretraživanja te u rješavanju etičkih pitanja koja se pojavljuju u kontekstu automatizacije i personalizacije rezultata pretraživanja.

Stoga se, osobito pojačano u novije doba, usporedno s tehnološkim razvojem sustava umjetne inteligencije otvaraju mnogobrojna etička i pravna pitanja u vezi s mogućnostima, opravdanosti i ograničenjem primjene umjetne inteligencije (Hrvatska enciklopedija 2024).

U ovom kontekstu, iznimno je važno ne samo razumjeti kako generativna umjetna inteligencija dodatno unapređuje procese pretraživanja informacija, već i identificirati potencijalne rizike i izazove povezane s njezinom primjenom. Njezina integracija u klasične sustave pretraživanja informacija, kao i primjena u srodnim područjima, donosi nova saznanja koja mogu značajno oblikovati budućnost informacijskih znanosti. Svrha ovog rada je istražiti kako generativna umjetna inteligencija utječe na proces pretraživanja informacija kroz napredne tehnike prepoznavanja uzoraka i automatizirano stvaranje sadržaja.

Ciljevi rada su: istražiti modele generativne umjetne inteligencije i njihove mogućnosti u razumijevanju i generiranju teksta općenito. Nadalje, kroz pregled postojeće literature utvrditi načine primjene generativne umjetne inteligencije u pretraživanju i pronalaženju informacija, što podrazumijeva dohvaćanje dokumenata, odgovaranje na pitanja, konverzacijske agente i sustave preporuka. Zatim, usporediti mogućnosti generativne umjetne inteligencije s klasičnim sustavima u zadacima dohvaćanja dokumenata, rangiranja po relevantnosti, ekstrakciji informacija te identificirati izazove i etičke probleme koji se javljaju uz integraciju generativne umjetne inteligencije u sustave za pretraživanje informacija.

Rad se sastoji od osam poglavlja. Nakon uvoda, slijedi poglavlje o osnovama i ulozi pretraživanja informacija u pronalasku relevantnih informacija među velikim količinama dostupnih. Sljedeća dva poglavlja su posvećena generativnoj umjetnoj inteligenciji uključujući povijesni razvoj i osnovne principe rada ove tehnologije. Zatim će se predstaviti načini primjene generativne umjetne inteligencije u pretraživanju informacija te usporediti mogućnosti generativne umjetne inteligencije s klasičnim sustavima u standardnim zadacima pretraživanja

informacija. Nakon toga, slijede poglavlja o prednostima i izazovima primjene generativne umjetne inteligencije u pretraživanju informacija, kao i o rješavanju problema informacijskog preopterećenja s pomoću generativne umjetne inteligencije. Zaključak donosi sažetak ključnih nalaza rada.

2. Pretraživanje informacija

2.1. Osnovni koncepti pretraživanja informacija

Pretraživanje informacija predstavlja ključnu aktivnost koja se stalno razvija kako bi zadovoljila korisničke potrebe za pristupom potrebnim informacijama. U tom kontekstu, razumijevanje osnovnih definicija pomaže u prepoznavanju važnosti i složenosti ovog procesa te u osiguravanju učinkovitog pristupa potrebnim informacijama. Sukladno tomu, pretraživanje informacija je:

traženje i izdvajanje nekog podatka iz velikoga skupa pohranjenih podataka, primjerice iz teksta, proračunske tablice, baze podataka, zbirke, kataloga, kazala, sekundarnoga dokumenta i slično. Kod računalnih sustava provodi se zadavanjem upita obično jedne ključne riječi ili više njih (Hrvatska enciklopedija 2024).

U kontekstu digitalnog doba, pretraživanje informacija je definirano kao: „proces pronalaženja informacija koje su pohranjene na računalu odnosno proces pronalaženja informacija koje su dostupne na internetu (Cambridge English Dictionary 2024).” Navedene definicije odražavaju važnost prilagodbe pretraživanja informacija na suvremene tehnologije i metode pristupa informacijama. S obzirom na korisnike, pretraživanje informacija se fokusira na metode i tehnike pronalaženja i izdvajanja relevantnih informacija koje zadovoljavaju specifične korisničke potrebe (Cooper 1984). Dakle, sustavi za pretraživanje mogu koristiti različite pristupe kako bi zadovoljili korisničke potrebe za informacijom.

Korisnik može pružiti sustavu relevantne informacije na nekoliko načina. On može izvršiti zadatak eksplicitne povratne informacije, izravno birajući dokumente s popisa rezultata ili zadatak implicitne povratne informacije, gdje sustav pokušava procijeniti interese korisnika koristeći relevantne dokumente u zbirci (Picariello i Rinaldi 2008).

U kontekstu pretraživanja informacija, važno je razumjeti ulogu i funkciju sustava za pretraživanje informacija. Dok je pretraživanje informacija proces koji obuhvaća metode za pronalaženje informacija, sustavi za pretraživanje informacija su alati koji omogućuju i unapređuju taj proces. Prema članku *What is information retrieval* sustavi za pretraživanje informacija su softverski programi koji obavljaju zadatke organizacije, pohrane, dohvaćanje i evaluacije informacija iz dokumenata, baza podataka i repozitorija znanja. Sukladno tomu, sustavi za pretraživanje informacija odabiru i rangiraju informacije koje su potrebne korisnicima odnosno informacije koje su korisnici zatražili u obliku upita. Drugim riječima, sustavi za pretraživanje informacija analiziraju korisničke upite kako bi identificirali i pružili korisnicima najrelevantnije rezultate odnosno informacije (Geeks for Geeks 2020).

U radu *Information retrieval: recent advances and beyond* autori naglašavaju primarni cilj pretraživanja informacija, a to je: prepoznavanje i dohvaćanje informacija koje su relevantne korisničkim upitima. Za korisnike je važno i rangiranje rezultata prema relevantnosti odnosno prema sličnosti s njihovim upitom (Hambarde i Proenca 2023). Bez obzira na razvoj tehnologije i utjecaj generativne umjetne inteligencije, cilj sustava za pretraživanje informacija uvijek ostaje isti, a tu tezu potvrđuje i sljedeći citat: „Jedna od najvažnijih komponenti sustava za pretraživanje informacija je korisnik: cilj sustava za pretraživanje informacija je zadovoljiti informacijske potrebe korisnika (Picariello i Rinaldi 2008).“ Pretraživanje informacija i sustavi za pretraživanje informacija su razvijeni s ciljem optimizacije korisničkog iskustva pri rješavanju specifičnih zadataka i problema te pri donošenju informiranih odluka i zaključaka.

Također, pretraživanje informacija omogućuje korisnicima dublji uvid u područja i teme njihovog interesa, kao i praćenje najnovijih informacija iz različitih područja (Hersh 2023). Danas, učinkovitost pretraživanja informacija na internetu ovisi o programskim sustavima poznatim kao tražilice koje omogućuju pretraživanje informacija provjeravajući cijeli dostupan dio interneta. Tražilice koriste razne algoritme za sustavno pregledavanje interneta i pronalaženje relevantnih informacija. Postupak pretraživanja tražilicama uključuje analizu i indeksiranje stranica s pomoću malih programa koji traže ključne riječi i vraćaju jedinstvene adrese stranica i poveznice na srodne dokumente. Dobro uspostavljeni sustavi redovito ažuriraju svoju bazu podataka kako bi osigurali da korisnici dobiju najnovije i najrelevantnije informacije (Hrvatska enciklopedija 2024).

Sustavi za pretraživanje informacija su posrednici između informacija i korisnika te se smatra kako se takvi sustavi mogu procijeniti na temelju tri kriterija, a to su: „prikladnost

sustava u smislu specifičnih IR (eng. *information retrieval*) zadataka za koje će se koristiti, učinkovitost izvedbe zadataka sustava i opseg u kojem sustav zadovoljava informacijske potrebe svojih korisnika (Al-Maskari i Sanderson 2010)“. Zadovoljstvo korisnika se smatra kriterijem uspješnosti i učinkovitosti sustava jer, kao što je spomenuto na početku poglavlja, glavni cilj sustava za pretraživanje informacija je zadovoljiti korisničke potrebe za informacijom. Nakon što su prikazani osnovni koncepti pretraživanja informacija, sljedeća potpoglavlja se usmjeravaju na analizu osnovnih komponenti koje čine temelj učinkovitih sustava za pretraživanje informacija te na interakciju korisnika i sustava za pretraživanje informacija.

2.2. Komponente sustava za pretraživanje informacija i interakcija sustava s korisnicima

Sukladno članku *What is information retrieval* osnovne komponente sustava za pretraživanje informacija koje su ključne za procese pretraživanja i dohvaćanja relevantnih informacija korisniku iz različitih baza podataka su: korisnik (eng. *user*) sustav za pretraživanje informacija (eng. *information retrieval system*), problem (eng. *problem*), zastupljenost (eng. *representation*), upit (eng. *query*), povratna informacija (eng. *feedback*), uspoređivanje (eng. *matching*), organizacija datoteka (eng. *file organisation*), dohvaćen objekt (eng. *retrieved object*) te prikupljanje informacija (eng. *acquisition*).

Proces pretraživanja informacija u sustavu za pretraživanje započinje kada korisnik izrazi informacijski zahtjev odnosno potrebu za određenom informacijom. Ovaj informacijski zahtjev transformira se u upit koji korisnik unosi u sustav za pretraživanje. Upit može biti jednostavan, sastavljen od nekoliko ključnih riječi ili složeniji, kada korisnik upotrebljava operatore za preciznije pretraživanje (npr. Booleovi operatori AND, OR i NOT).

Sustav za pretraživanje informacija obrađuje korisnikov upit tako da analizira njegove pojmove kako bi što preciznije razumio korisnikove potrebe. Nakon toga, upit se pretvara u niz karakteristika koje sustav koristi za usporedbu s dokumentima ili podacima pohranjenima u svojoj bazi. U ovoj fazi ključnu ulogu ima koncept zastupljenosti, koji se odnosi na način pohrane i prikaza dokumenata unutar sustava. Dokumenti mogu biti predstavljeni putem ključnih riječi, metapodataka ili drugih atributa koji olakšavaju njihovo pretraživanje (Belkin 1996).

Nakon što sustav interpretira upit i pretvori ga u dohvaćeni objekt, provodi se usporedba između upita i dokumenata pohranjenih u internoj bazi podataka. Ova usporedba koristi razne algoritme s ciljem identificiranja dokumenata koji su najrelevantniji za korisnikov zahtjev. U suvremenim sustavima, ova faza može uključivati napredne tehnike poput obrade prirodnog jezika i algoritama strojnog učenja (više o tome u sljedećem potpoglavlju), kako bi se korisnikovi upiti što preciznije razumjeli i analizirali. Relevantni rezultati pretraživanja prikazuju se korisniku u obliku rangiranog popisa. Pri tome se koristi mehanizam povratnih informacija, gdje korisnik označava koji su rezultati korisni, a sustav te informacije koristi za optimizaciju budućih pretraga. Tako, sustav kontinuirano prilagođava proces pretraživanja kako bi poboljšao točnost rezultata (Belkin 1996).

Prije nego što sustav za pretraživanje informacija može izvršiti pretraživanje, nužno je provesti dva ključna procesa: prikupljanje informacija i organizaciju podataka. Prikupljanje informacija podrazumijeva prikupljanje informacija iz različitih izvora, poput dokumenata, baza podataka, web stranica ili drugih izvora informacija. Nakon toga slijedi organizacija informacija, pri čemu se prikupljene informacije pohranjuju u jednu ili više baza podataka. Organizacija može biti sekvencijalna, gdje su dokumenti pohranjeni redosljedom kojim su uneseni, ili inverzna, gdje se dokumenti indeksiraju prema ključnim riječima, čime se omogućuje brže pretraživanje. Ovaj cjelokupni proces, od stvaranja upita do prikaza rezultata, temelj je funkcionalnosti sustava za pretraživanje informacija te omogućuje korisnicima pristup relevantnim informacijama u skladu s njihovim zahtjevima (Geeks for Geeks 2020). Obrnuto odnosno inverzno indeksiranje pruža nasumičan, ali sustavan pristup predstavljanju dokumenata, osiguravajući raznoliku pokrivenost sadržaja (Kuo et al. 2024)

Nadalje, interakcija između korisnika i sustava za pretraživanje informacija temelji se na specifičnom informacijskom zahtjevu koji korisnik unosi u sustav u obliku upita. Proces pretraživanja informacija je prethodno opisan, a prilikom interakcije sustava i korisnika ključne su baze podataka te pregledavanje rezultata. Baza podataka (eng. *data base*) je „organizirana zbirka logički povezanih, pretraživih i međusobno ovisnih podataka (informacija), pohranjena u nekom od računalno čitljivih medija (Hrvatska enciklopedija 2024).“

U ranim sustavima dokumenti u bazama podataka opisivali su se skupom ključnih riječi koje su služile kao sažetak sadržaja. Suvremeni sustavi koriste puni tekst dokumenta, čime se povećava preciznost pretraživanja jer svaki pojam u dokumentu može doprinijeti pronalasku relevantnih informacija. Kako bi se smanjila količina nepotrebnih informacija i olakšalo

pretraživanje, često se primjenjuju tekstualne operacije poput uklanjanja zaustavnih riječi, koje nemaju značajnu informacijsku vrijednost (npr. veznici). Time se smanjuje složenost dokumenta, omogućujući algoritmima učinkovitije pretraživanje i brže pronalaženje rezultata. Nakon dohvaćanja relevantnih informacija iz baze podataka, korisnik pregledava (eng. *browsing*) rezultate kako bi pronašao one koji najbolje odgovaraju njegovom informacijskom zahtjevu.

Sukladno tome, interakcija između korisnika i sustava za pretraživanje informacija stavlja korisnika u središte, s naglaskom na optimizaciju procesa pretraživanja putem naprednih metoda dohvaćanja informacija i tekstualnih operacija koje pojednostavljaju i ubrzavaju pronalaženje relevantnih informacija. Tako, sustav korisnicima pruža preciznije i relevantnije odgovore na njihove informacijske zahtjeve.

Budući da su sustavi za pretraživanje informacija osmišljeni kako bi zadovoljili informacijske potrebe korisnika, razumijevanje korisnika postaje ključan element tih sustava. Razumijevanje korisničkih namjera i modeliranje njihovog ponašanja dva su glavna aspekta istraživanja u području IR-a, a značajan napredak postignut je posljednjih desetljeća (Ai et al. 2023).

Također, „istraživači u području pronalaženja informacija kontinuirano razvijaju nove pristupe kako bi optimizirali performanse sustava, osobito s naglaskom na sustave učenja za rangiranje rezultata (Hambarde i Proenca 2023).“ S napretkom novih tehnologija, korisničke potrebe u procesu pretraživanja informacija značajno su evoluirale. U tradicionalnim sustavima za pretraživanje, korisnici su često morali precizno formulirati svoje upite i ručno pregledavati rezultate kako bi došli do željenih informacija. Suvremeni sustavi koriste napredne algoritme strojnog učenja i obrade prirodnog jezika, omogućujući učinkovitiju interpretaciju korisničkih zahtjeva i prilagođavanje rezultata pretraživanja specifičnim potrebama svakog pojedinca (Hersh 2023). Ova promjena u dinamici interakcije između korisnika i sustava za pretraživanje informacija, kao i integracija novih tehnologija poput generativne umjetne inteligencije, omogućava poboljšanje procesa pretraživanja kroz preciznije i personalizirane rezultate, što će biti detaljnije obrađeno u sljedećem potpoglavlju kao i u ostatku rada.

2.3. Pretraživanja informacija u digitalnom dobu i generativna umjetna inteligencija

U današnjem digitalnom dobu, pretraživanje informacija postaje sve zahtjevnije zbog eksponencijalnog rasta količine dostupnih podataka pogotovo polu-strukturiranih i

nestrukturiranih. Iako se može činiti da je koncept pretraživanja informacija relativno nov, njegova povijest je duga i složena. Razumijevanje suvremenih metoda pretraživanja zahtijeva kratki uvid u tradicionalne pristupe informacijama koji su prethodili digitalnoj eri.

Duga povijest pretraživanja informacija ne započinje s razvojem weba. Tek u posljednjem desetljeću, web tražilice su postale sveprisutne, a pretraživanje je postalo integrirano u strukturu desktop i mobilnih operativnih sustava. Prije široke upotrebe tražilica, sustavi za pretraživanje informacija korišteni su u komercijalnim i obavještajnim aplikacijama još od 1960-ih godina. Najraniji računalni sustavi za pretraživanje izgrađeni su kasnih 1940-ih. Kao i kod mnogih računalnih tehnologija, mogućnosti sustava za pretraživanje rasle su s povećanjem brzine procesora i kapaciteta pohrane. Razvoj takvih sustava također odražava brz napredak u prelasku s ručnih pristupa temeljenih na knjižnicama na automatske metode (Sanderson i Croft 2012).

Tradicionalni pristupi pretraživanju informacija su često ograničeni učinkovitošću i relevantnošću rezultata pretraživanja, posebice u kontekstu porasta sadržaja u digitalnim oblicima. Povezano s tim, povećavaju se problemi s kojima se korisnici suočavaju, a to su najčešće: informacijska preopterećenost, nejasni ili preopćeniti rezultati pretraživanja, nedostatak konteksta, personalizacije pretraživanja i slično (Hambarde i Proenca 2023). Povezano s tim, generativna umjetna inteligencija mijenja načine korištenja digitalnih sadržaja tako da korisnici mogu koristiti glasovne naredbe ili multimedijske sadržaje za pretraživanje. Nadalje, generativna umjetna inteligencija se može koristiti za rješavanje složenih problema pretraživanja i pronalaska informacija što uvelike otvara put razvoju novih mogućnosti. Ova tehnologija olakšava pretragu složenih i višeznačnih informacija te pruža relevantne rezultate čak i s nedostatkom jasnih korisničkih upita (Feuerriegel et al. 2024). S obzirom na ogromnu količinu generiranih informacija, postavlja se pitanje kako učinkovito obraditi i analizirati informacije. U ovom kontekstu, umjetna inteligencija i strojno učenje također igraju ključnu ulogu.

Danas količina podataka koju generiraju i ljudi i strojevi daleko nadmašuje sposobnost ljudi da apsorbiraju, interpretiraju i donose složene odluke na temelju tih podataka. Umjetna inteligencija i strojno učenje čine osnovu za praktički sve buduće složene odluke, kako u životu tako i u proizvodnji (CAD/CAM Group 2023).

Pretraživanje informacija i umjetna inteligencija predstavljaju obećavajući spoj dviju tehnologija koje su danas od velike važnosti. Pronalaženje informacija je aktivnost dobivanja relevantnih informacija iz različitih izvora. Proces pretraživanja se može znatno poboljšati korištenjem mogućnosti umjetne inteligencije (AI Glossary 2024). Povezano s tim, inteligentni informacijski sustavi odnosno inteligentni sustavi za pretraživanje informacija mogu osigurati učinkovitiji pristup informacijama (Jones 1991).

Dalje u radu će biti prikazana primjena generativne umjetne inteligencije u pretraživanju informacija uključujući načine primjene alata generativne umjetne inteligencije u sustave za pretraživanje informacija, usporedbu mogućnosti generativne umjetne inteligencije s klasičnim sustavima za pretraživanja informacija te poveznicu između informacijskog preopterećenja i generativne umjetne inteligencije. Sljedeće poglavlje prikazuje kako generativna umjetna inteligencija mijenja način na koji pristupamo stvaranju i interpretaciji sadržaja, proširujući mogućnosti koje su bile nezamislive u ranijim sustavima obrade informacija jer bi u budućnosti generativna umjetna inteligencija mogla značajno unaprijediti pretraživanje informacija.

3. (Generativna) umjetna inteligencija

„Umjetna inteligencija je dio računalne znanosti (računarstva) koji se bavi razvojem sposobnosti računala da obavljaju zadaće za koje je potreban neki oblik inteligencije (Hrvatska enciklopedija 2024).” Prema Europskom parlamentu umjetna inteligencija je:

sposobnost nekog uređaja da oponaša ljudske aktivnosti poput zaključivanja, učenja, planiranja i kreativnosti. Umjetna inteligencija omogućuje tehničkim sustavima percipiranje okruženja, uzimanje u obzir onog što vide i rješavanje problema kako bi postigli neki cilj (Europski parlament 2023).

Umjetna inteligencija je tehnologija koja računalima i strojevima omogućuje simulaciju ljudske inteligencije i sposobnosti rješavanja problema (IBM 2023). Umjetna inteligencija je interdisciplinarno područje koje se bavi razvojem računalnih sustava s karakteristikama ljudskog razmišljanja poput učenja, planiranja, razumijevanja jezika i prepoznavanja objekata. Cilj umjetne inteligencije je stvaranje strojeva i programa koji mogu obavljati zadatke poput donošenja odluka, prepoznavanja govora, vizualnog prepoznavanja i prevođenja jezika (Tech Target 2024). Zadnjih godina sve više uočavamo korištenje alata umjetne inteligencije u svakodnevnom životu, posebice u umjetnosti gdje s pomoću umjetne inteligencije nastaju nova

umjetnička djela (Thompson i Clark 2017). Sukladno tomu, jedna od najznačajnijih vrsta umjetne inteligencije je generativna umjetna inteligencija.

Generativna umjetna inteligencija grana je umjetne inteligencije koja može stvarati nove sadržaje poput tekstova, slika ili zvuka koji se sve češće više ne mogu razlikovati od ljudske izrade. Generativna umjetna inteligencija ima potencijal transformirati discipline koje se oslanjaju na kreativnost, inovacije i obradu znanja (Feuerriegel et al. 2024).

Kako su tehnike dubokog učenja u umjetnoj inteligenciji brzo napredovale, umjetna inteligencija je mogla učiti iz velikih količina podataka, stvarajući generativnu umjetnu inteligenciju koja može generirati nove podatke (Jang et al. 2024). Jedna od podjela umjetne inteligencije je na slabu i jaku umjetnu inteligenciju. Slabu umjetnu inteligenciju čine svi dosad osmišljeni sustavi čiji rad se bazira na principima umjetne inteligencije dok sustavi koji koriste jaku umjetnu inteligenciju još nisu postignuti (Hrvatska enciklopedija 2024). Slaba umjetna inteligencija je sposobna obavljati specifične zadatke dok bi jaka umjetna inteligencija u potpunosti trebala biti sposobna oponašati čak i nadmašiti ljudsku inteligenciju (IBM 2023).

Jaka umjetna inteligencija ili svjesna umjetna inteligencija se odnosi na stroj koji ne bi bio samo sposoban izvoditi inteligentno ponašanje, nego bi čak dao dojam realne samosvijesti, istinitih osjećaja, i razumijevanja vlastitih rasuđivanja. Slaba umjetna inteligencija ili ograničena umjetna inteligencija ide za tim da se gradi što više autonomnih sustava, algoritama sposobnih za rješavanje problema određenog područja. Na ovoj razini, stroj simulira inteligenciju, djeluje kao da je inteligentan (Singbo 2008).

3.1. Osnovni principi rada (generativne) umjetne inteligencije

Generativna umjetna inteligencija (eng. *Generative Artificial Intelligence* - GAI) se oslanja na modele strojnoga učenja koji oponašaju ljudske neuronske mreže. „Cilj strojnog učenja je konstruirati i stvoriti algoritme i sustave koji se vrlo lako prilagođavaju novim situacijama i koji uče na temelju iskustva, a algoritmi strojnog učenja otkrivaju obrasce i uče kako napraviti predviđanja i dati preporuke. (CAD/CAM Group 2023).“ Sukladno tomu, strojno učenje:

sustavima omogućuje svladavanje velike količine znanja, komunikaciju s čovjekom ili nekim drugim neživim sustavom, učenje na temelju iskustva, donošenje

zaključaka, prilagodljivo ponašanje i složeno planiranje pri čemu su, za razliku od čovjekove sposobnosti da istovremeno obavlja raznovrsne funkcije, današnji inteligentni sustavi još uvijek specijalizirani za užu raspon mogućnosti (Hrvatska enciklopedija 2024).

Osim strojnog učenja (eng. *Machine Learning* - ML) postoji i duboko učenje (eng. *Deep Learning* – DL) koje se razlikuje po vrstama neuronskih mreža i ljudskoj intervenciji. Strojno učenje uključuje različite modele npr. stabla odlučivanja, pravila odlučivanja, regresijske modele, itd., a posebno mjesto zauzimaju tzv. umjetne neuronske mreže: „neuronske mreže s ulaznim slojem, jednim ili dva skrivena sloja i izlaznim slojem (IBM 2023).” Modeli strojnog učenja koriste različite algoritme kako bi naučili obavljati postavljeni zadatak temeljem ulaznih skupova podataka u različitim modovima učenja kao što su nadgledano učenje (eng. supervised learning), nenadgledano učenje (eng. unsupervised learning), pojačano učenje (eng. reinforcement learning), itd. Duboko učenje kao vrlo razvijeno područje strojnog učenja koristi neuronske mreže s ulaznim slojem, stotinama skrivenih slojeva i izlaznim slojem. Važno je napomenuti da duboke neuronske mreže mogu učiti ne samo u nadziranom načinu, već i u raznim drugim načinima, uključujući polunadzirano učenje (eng. semi-supervised learning), samonadzirano učenje (self-supervised learning) itd. Ovi različiti pristupi omogućuju mrežama da se prilagode različitim vrstama podataka i zadataka, smanjujući potrebu za ljudskom intervencijom.

S obzirom na to da modeli dubokog učenja mogu učiti na više načina, kao što je prethodno i spomenuto, to ih čini prikladnima za trening na velikim skupovima podataka koji mogu biti strukturirani, polu-strukturirani i nestrukturirani, a samim tim dobrim kandidatima za rješavanje složenih tipova problema iz različitih domena. Također, ključna komponenta generativne umjetne inteligencije je generativni model koji uči kako distribuirati podatke i koji koristi distribuciju kako bi generirano nove podatke. Prvo se prikuplja velika količina podataka iz neke domene, a zatim se taj model obučava da generira slične podatke. Intuicija koja stoji iza ovog pristupa, prati citat američkog fizičara Richarda Feynmana koji glasi: „Ono što ne mogu stvoriti, ne razumijem.“ (Open AI 2016).

Transformatori kao jedna od prekretnih arhitektura u području DL i veliki jezični modeli zasnovani na derivatima te arhitekture uvelike su utjecali na razvoj generativne umjetne inteligencije. Transformatori omogućuju treniranje na velikim količinama podataka pa veliki jezični modeli izgrađeni od istih omogućuju generativnoj umjetnoj inteligenciji stvaranje novih

sadržaja odnosno strojno prevođenje, generiranje teksta, razumijevanje prirodnog jezika i obradu slika (Tech Target 2024).

Generativna umjetna inteligencija se sastoji od tri radne faze, a to su: obuka za izradu modela, prilagodba modela specifičnoj namjeni te podešavanje, a samim time i implementacija modela. Obuka za izradu modela odvija se na velikim količinama podataka. U hrvatskom jeziku, ovaj pojam se prevodi kao *obuka*, no u daljnjem tekstu koristit ćemo termin *trening*. Navedeni proces je temelj za stvaranje GAI alata, a omogućuju ga modeli dubokog učenja kao i veliki jezični modeli.

Kako bi stvorili temeljni model, stručnjaci treniraju algoritam dubokog učenja na ogromnim količinama relevantnih neobrađenih, nestrukturiranih, djelomično označenih i većinom neoznačenih podataka, kao što su terabajti ili petabajti podatkovnog teksta ili slika ili videa s interneta. Trening najčešće rezultira s dubokom neuronskom mrežom od nekoliko milijardi *parametara* — kodiranih prikaza entiteta, uzoraka i odnosa u podacima — koji mogu samostalno generirati sadržaj kao odgovor na upite (IBM 2023).

Prilagodba modela specifičnoj namjeni se najčešće radi na dva načina, a to su: fino podešavanje ili pojačano učenje s povratnim informacijama od strane ljudskih stručnjaka. Kod finog podešavanja, modelu se dostavljaju označeni podaci specifični namjeni, a kod pojačanog učenja s povratnim informacijama stručnjaci procjenjuju točnost i relevantnost modela unoseći ispravke. Posljednji korak je implementacija modela uključujući dodatne procjene i redovita ažuriranja sustava. Programeri i korisnici redovito procjenjuju rezultate svojih generativnih AI aplikacija i dodatno prilagođavaju model, čak i jednom tjedno, za veću točnost ili relevantnost. Nasuprot tome, sam temeljni model ažurira se puno rjeđe, možda svake godine ili 18 mjeseci. Još jedna opcija za poboljšanje performansi gen AI aplikacije je generiranje proširenog dohvaćanja (RAG), tehnika za proširenje mogućnosti temeljnog modela uz korištenje relevantnih izvora izvan podataka za trening kako bi se poboljšali parametri za veću točnost ili relevantnost (IBM 2023).

3.2. Alati (generativne) umjetne inteligencije

Alat je „primjenski ili korisnički program, skup naredbi i uputa koji omogućuje izvršenje neke zadaće (Hrvatska enciklopedija 2024).“ Hrvatski jezični portal taj pojam, u okvirima računalnih znanosti, definira kao: „računalni program stvoren za obavljanje određene

zadaće, namjenski program, računalni program s određenom namjenom (Hrvatski jezični portal).“ Razvoj alata čiji rad se temelji na principima (generativne) umjetne inteligencije datira uglavnom iz posljednjeg desetljeća, a temelji za stvaranje ovih alata nastaju davnih 1930-ih godina razvojem *mehaničkog mozga* te od tog trenutka sve do danas brzina razvoja u ovom području svakodnevno raste. Unutar okvira generativne umjetne inteligencije postoje alati za generiranje teksta, slika, glazbe, zvuka, prezentacija i kodova (Tech Target 2024).

Stvaranje AI alata omogućeno je zahvaljujući napretku neuronskih mreža i transformatorskih arhitektura. Stvaranje neuronskih mreža i transformatora omogućilo je računalnim programima učinkovitije kodiranje i dekodiranje prirodnog jezika, olakšavajući razvoj velikih jezičnih modela. Zbog učinkovitosti koju stvaraju transformatori, veliki jezični modeli sada imaju podatke od stotina milijardi riječi, što im omogućuje pristup velikom korpusu suvremenog ljudskog znanja i umjetnosti (Carroll i Borycz 2024).

Pojava alata umjetne inteligencije kao što su *Codex* ili *Chat GPT* otvorila je novu eru pomoćnih alata za programere. Također, procjenjuje se da je učinkovitost AI alata na visokoj razini. Na primjer, Codex, koji ima 12 milijardi parametara modela i obučen je na 54 milijuna GitHub softverskih repozitorija, riješio je preko 70% od 164 problema programiranja Pythona sa 100 primjera. Chat GPT je posebno imao golem utjecaj u raznim domenama. Sa 100 milijuna aktivnih posjetitelja u prva dva mjeseca dostupnosti, Chat GPT je postao najbrže rastući AI alat u povijesti (Kuhail et al. 2024).

3.3. Područja primjene (generativne) umjetne inteligencije

Danas umjetna inteligencija može obavljati razne zadatke puno brže i preciznije od ljudi, bez obzira na područje primjene. Alati umjetne inteligencije posebno su korisni u obradi velikih količina podataka, što čini umjetnu inteligenciju ključnim faktorom u različitim sektorima, od marketinga i obrazovanja do dizajna proizvoda i sličnih područja (Tech Target 2024). Na primjer, „znanstvenici istražuju kako koristiti umjetnu inteligenciju za analizu velike količine zdravstvenih podataka, kako bi se identificirali obrasci koji mogu dovesti do novih medicinskih otkrića i poboljšanja dijagnostike na individualnoj razini (Europski parlament 2023).“

Također, različite online trgovine koriste chatbotove, konverzacijske agente ili virtualne asistente za rješavanje upita i pružanje podrške korisnicima. Ovi alati, zahvaljujući mogućnostima obrade prirodnog jezika, mogu brzo odgovoriti na pitanja poput statusa

narudžbi, detalja o povratu proizvoda ili pravilima povrata. Chatbotovi i slični alati omogućuju neprekidnu podršku korisnicima, brže odgovore na često postavljana pitanja, oslobađanje vremena zaposlenicima za složenije zadatke i pružanje dostupnije usluge korisnicima (IBM 2023).

Umjetna inteligencija u širokoj je upotrebi za pružanje personaliziranih preporuka, na primjer na temelju prethodnih pretraga i kupnji ili drugih oblika ponašanja na internetu. Umjetna inteligencija iznimno je važna u trgovini, kod optimizacije proizvoda, planiranja inventara, logistike itd. (Europski parlament 2023).

Primjena umjetne inteligencije u odnosima s javnošću omogućava preciznije ciljanje strateških publika te povećava interaktivnost i pažnju medija. Umjetna inteligencija olakšava transkripciju govora u tekst, prevođenje audio i tekstualnih datoteka na više jezika, praćenje autentičnosti videozapisa, analizu sentimenta te automatsko generiranje sadržaja putem alata za pisanje. Sve ove funkcionalnosti povećavaju učinkovitost u odnosima s javnošću, optimizirajući komunikacijske procese (Tomić et al. 2022). „Napredne tehnologije poput umjetne inteligencije, strojnog učenja, digitalnih platformi i softverskih rješenja pružaju neprocjenjivu konkurentsku prednost i pomažu u predviđanju budućih projekata (CAD/CAM Group 2023).“

Nadalje, u marketingu, umjetna inteligencija pomaže u razumijevanju ponašanja potrošača putem analize velikih količina podataka, predviđanja budućih odluka te optimizacije marketinških aktivnosti. Tehnike poput prediktivne analitike omogućuju kreiranje personaliziranih ponuda i bolje korisničke usluge. Osim toga, umjetna inteligencija poboljšava odnose s kupcima, upravljanje prodajnim osobljem i omogućuje precizno istraživanje tržišta, smanjujući rizik kod donošenja poslovnih odluka (Arbutina i Nakić 2024).

U obrazovanju, umjetna inteligencija ima potencijal značajno transformirati sustav u bliskoj budućnosti, modernizirajući ga i povećavajući njegovu kvalitetu. Primjena umjetne inteligencije može ubrzati digitalizaciju obrazovanja, automatizirati kreiranje kurikuluma, prepoznati slabosti u obrazovnom sustavu te smanjiti socijalne nejednakosti među učenicima. Iako umjetna inteligencija neće zamijeniti učitelje, ona će unaprijediti proces učenja, omogućujući personalizirano i učinkovitije obrazovanje (Varga i Ružić 2024).

Zatim, u radu *Primjena umjetne inteligencije u analizi konkurencije* se istražuje kako umjetna inteligencija djeluje kao snažan alat u različitim industrijama, omogućujući

prikupljanje, analizu i korištenje podataka iz izvora poput društvenih mreža, online recenzija, novinskih članaka i financijskih izvještaja za dobivanje uvida o konkurenciji. Korištenjem povijesnih podataka i algoritama strojnog učenja, umjetna inteligencija pomaže tvrtkama predvidjeti tržišne trendove, prilagoditi strategije i identificirati nove prilike na tržištu (Ribić et al. 2024).

Sukladno tomu, rad pod nazivom *Uloga umjetne inteligencije u stvaranju medijskog sadržaja* obrađuje primjenu umjetne inteligencije u kreiranju medijskog sadržaja, s naglaskom na tehnologije poput jezičnih, slikovnih, glasovnih i video alata. Umjetna inteligencija značajno transformira medije i komunikaciju, a s daljnjim tehnološkim napretkom, očekuje se da će i dalje mijenjati načine stvaranja i distribucije sadržaja. Rad također govori o korištenju umjetne inteligencije u novinarstvu, društvenim mrežama te drugim oblicima medija. Istaknuto je kako tehnologije umjetne inteligencije poboljšavaju proizvodnju sadržaja automatizacijom pisanja, uređivanja, prepoznavanja lažnih vijesti i pružanjem personaliziranih korisničkih iskustava (Bebić 2023)

Generativna umjetna inteligencija značajno je transformirala proces razvoja softverskih rješenja, omogućujući brže i učinkovitije stvaranje koda. Alati poput GitHub Copilot, Devin (cognition.ai) i Cursor AI (cursor.com) koriste napredne modele umjetne inteligencije za predlaganje koda, automatizaciju zadataka i ubrzanje procesa razvoja softvera. Ovi alati ne samo da poboljšavaju produktivnost programera, već i omogućuju razvoj sofisticiranih rješenja s minimalnim ručnim kodiranjem. Na primjer, GitHub Copilot koristi umjetnu inteligenciju kako bi programerima predlagao odgovarajući kod u realnom vremenu, što smanjuje vrijeme razvoja i poboljšava točnost. Također, Cursor AI je poznat po svojoj primjeni u upravljanju i optimizaciji složenih razvojnih projekata kroz napredne generativne modele, što programerima omogućuje lakše upravljanje velikim količinama podataka te brže testiranje i implementaciju koda. Ovi alati predstavljaju budućnost softverskog razvoja, gdje se umjetna inteligencija koristi ne samo za kodiranje već i za dizajn, testiranje i održavanje softverskih sustava. Navedeni alati dodatno naglašavaju sposobnost generativne umjetne inteligencije da izvrši zadatke brže i preciznije, neovisno o području primjene. Očekuje se da će njezin utjecaj u budućnosti nastaviti rasti kako se tehnologija razvija, pružajući značajnu konkurentsku prednost onima koji je implementiraju (Coutinho et al. 2024).

U sljedećem poglavlju je prikazan povijesni pregled (generativne) umjetne inteligencije. Analizirane su ključne prekretnice u razvoj u ove tehnologije od njezinih početaka do današnje napredne primjene.

4. Razvoj (generativne) umjetne inteligencije

Generativna umjetna inteligencija predstavlja jednu od ključnih tehnologija 21. stoljeća, omogućujući stvaranje novih sadržaja korištenjem sofisticiranih algoritama. Njena primjena proteže se od automatiziranog generiranja slika, glazbe i teksta, pa sve do stvaranja kompleksnih virtualnih okruženja i interakcije s korisnicima. Iako je razvoj generativne AI dugotrajan proces, eksponencijalni rast postignut je u posljednjim desetljećima zahvaljujući napretku u modelima strojnog učenja. Na primjer, iako su prvi oblici generativne AI započeli još 1960-ih godina s jednostavnim chatbotovima, tek je uvođenje generativnih kontradiktornih mreža (GAN-ova) 2014. godine omogućilo stvaranje uvjerljivih slika, videozapisa i zvukova gotovo nerazlučivih od stvarnosti (Tech Target 2024).

Osim toga, generativna umjetna inteligencija se nije razvijala brzo, već je rezultat višedesetljetnog napora u istraživanju umjetne inteligencije. Današnje mogućnosti ove tehnologije temelje se na dugoj povijesti interdisciplinarnog istraživanja koje obuhvaća područja kao što su kognitivne znanosti, matematika i filozofija. Razvoj umjetne inteligencije, od prvih teorijskih koncepata do naprednih algoritama, progresivno je premošćivao jaz između ljudskog i strojno generiranog mišljenja.

Zamisli o umjetnoj inteligenciji pojavljuju se u mitovima, legendama i književnosti od davnih vremena sve do danas, često izazivajući nelagodu oko mogućnosti da stroj intelektualno nadjača čovjeka. Razvoj umjetne inteligencije započeo je pojavom elektroničkih računala u drugoj polovici 20. st., pri čemu su pristupi i tehnike usvojeni iz raznih disciplina, a napose onih koje se među ostalima bave istraživanjima i mogućnostima načina ljudskoga mišljenja: kognitivne znanosti, matematike, logike, filozofije, psihologije i lingvistike (Hrvatska enciklopedija 2024).

Povijest razvoja umjetne inteligencije može se podijeliti u nekoliko ključnih faza koje vode od klasične umjetne inteligencije prema generativnoj. Početak seže u 1932. godinu kada inženjer Georges Artsrouni osmišljava stroj nazvan *mehanički mozak*, dizajniran za prevođenje

s jednog jezika na drugi s pomoću mehaničkog računala. Pravi temelj umjetne inteligencije postavlja Alan Turing 1950. godine objavom rada *Computing Machinery and Intelligence* u kojem predlaže Turingov test kao način mjerenja sposobnosti stroja da oponaša inteligentno ponašanje ljudi. Umjetna inteligencija kao formalno područje istraživanja dobiva prepoznatljivost 1956. godine na konferenciji u Dartmouthu, koju su organizirali John McCarthy, Marvin Minsky, Nathaniel Rochester i Claude Shannon. Tada je termin *umjetna inteligencija* službeno prihvaćen, označivši početak ovog novog znanstvenog polja (Tech Target 2024).

Narednih desetljeća klasična umjetna inteligencija polako se razvija. Jedan od važnih trenutaka je 1957. kada lingvist Noam Chomsky objavljuje *Sintaktičke strukture*, pružajući uvid u gramatička pravila potrebna za raščlanjivanje i generiranje rečenica u prirodnom jeziku. Godine 1966., Joseph Weizenbaum s MIT-a kreira Elizu, prvi chatbot koji simulira razgovore sa psihoterapeutom, pružajući iluziju razumijevanja korisničkih pitanja.

Eliza se temeljila na postupku podudaranja uzoraka dajući iluziju da razumije postavljena pitanja. Temeljila se na skriptama napisanim u programskom jeziku za procesiranje lista SLIP koji je 1960. godine također razvio Weizenbaum. Weizenbaum je skriptu napravio kao parodiju, a odgovori Elize su često zbunjivali korisnike, ostavljajući im dojam da Eliza stvarno daje inteligentne odgovore (Stipaničev et al. 2021).

Iako su odgovori Elize bili daleko od stvarne inteligencije, ovaj program pokazuje mogućnosti računalne interakcije s ljudima. Eliza je time postala rani primjer testnog slučaja za Turingov test. „Eliza je bila jedan od prvih chatbotova. Bio je to i rani testni slučaj za Turingov test, test sposobnosti stroja da pokaže inteligentno ponašanje jednako ljudskom ili se ne razlikuje od njega. Prema današnjim standardima Eliza vrlo brzo pada ako joj postavite nekoliko složenih pitanja (New Jersey Institute of Technology 2024).” Nadalje, 1968. godine Terry Winograd razvija *SHRDLU*, rani računalni program koji razumije prirodni jezik.

Tijekom 1980-ih i kasnije, umjetna inteligencija dobiva zamah s važnim tehničkim naprecima. Judea Pearl 1985. godine uvodi uzročnu analizu putem Bayesovih mreža, što dovodi do novih metoda generiranja sadržaja u specifičnom stilu. Godine 1986. Michael Jordan postavlja temelje za modernu primjenu rekurentnih neuronskih mreža, dok 1989. Yann LeCun, Yoshua Bengio i Patrick Haffner demonstriraju kako se konvolucijske neuronske mreže mogu

koristiti za prepoznavanje slika. Ovi napreci otvaraju put ka sofisticiranijim modelima koji omogućuju dublje razumijevanje i analizu podataka.

Generativna umjetna inteligencija počinje značajnije dobivati na važnosti početkom 2000-tih, kada istraživači sa Sveučilišta u Montrealu objavljuju rad *A Neural Probabilistic Language Model*, u kojem predlažu metode za modeliranje jezika korištenjem neuronskih mreža. Ovaj rad postavlja temelje za daljnji razvoj tehnologija za obradu prirodnog jezika. Godine 2011. Apple lansira Siri, prvog komercijalnog glasovnog osobnog asistenta, što označava početak integracije generativnih sposobnosti umjetne inteligencije u svakodnevni život korisnika.

Posebno važan trenutak u razvoju generativne umjetne inteligencije događa se 2014. godine, kada Ian Goodfellow razvija generativne kontradiktorne mreže (GAN-ovi), koje omogućuju stvaranje novih sadržaja na temelju zadanih uzoraka. Razvoj ovih modela ubrzava daljnja istraživanja, pa 2017. godine Googleovi istraživači razvijaju koncept transformatora (Attention Is All You Need), revolucionarnu arhitekturu koja omogućuje obradu prirodnog jezika na način koji je promijenio područje umjetne inteligencije.

Transformatorski modeli, poput GPT-a, postaju temelj naprednih sustava za generiranje teksta. Godine 2021. OpenAI predstavlja DALL-E, sustav koji može generirati slike iz tekstualnih upita, dok 2022. godine istraživači iz Runway Research i Stability AI objavljuju Stable Diffusion, alat za automatsko generiranje slikovnog sadržaja iz tekstualnih podataka (Tech Target).

Ovaj prijelaz od klasične umjetne inteligencije prema generativnoj označava važnu prekretnicu u razvoju tehnologije, gdje sustavi više nisu samo sposobni za analizu i prepoznavanje podataka, već mogu generirati nove informacije, slike, tekstove i druge sadržaje s razinom sofisticiranosti koja se približava ljudskoj kreativnosti. Također, ovaj napredak otvara vrata širokom spektru primjena, a jedna od najznačajnijih je upravo pretraživanje informacija. U sljedećem poglavlju razmotrit će se kako generativna umjetna inteligencija doprinosi unapređenju ovog procesa, omogućujući korisnicima da ne samo pronalaze informacije, već da interaktivno generiraju relevantne odgovore i sadržaje.

5. Primjena generativne umjetne inteligencije u pretraživanju informacija

Iz prethodnih sekcija vidljiv je razvoj generativne umjetne inteligencije, kao i porast zastupljenosti sustava generativne umjetne inteligencije u različitim djelatnostima. Slično revoluciji koju su ti sustavi donijeli u ostalim područjima i područje pretraživanja informacija doživjelo je značajne promjene od kojih su neke tek u eksperimentalnoj fazi, a druge su već pokazale dobrobiti koje nude, ali i izazove. Tijekom 2022. godine zabilježen je rast generativnog pretraživanja informacija na dva načina: generativno dohvaćanje dokumenata koji odgovaraju korisničkim upitima te generiranja odgovora nastalog kao rezultat korisničkog upita. Umjetno generirani odgovor korisniku može pružiti potrebne informacije ili ga može uputiti na izvore informacija (Bénédict et al. 2023).

Trenutne i buduće mogućnosti sustava za pretraživanje informacija koji koriste generativnu umjetnu inteligenciju proizašle su iz sve većih mogućnosti velikih jezičnih modela, posebno njihove sposobnosti da izvršavaju zadatke prirodnog jezika, uključujući sažimanje, razumijevanje i proširenje upita. Većina sustava zamjenjuje upit i rangiranu listu razgovorom i pisanom sintezom informacija (Alaofi et al. 2024).

Sustavi za inteligentno pretraživanje informacija su tehnološki alati koji olakšavaju izdvajanje, procjenu i prezentaciju potrebnih informacija. Ovakvi sustavi koriste metode generativne umjetne inteligencije za pojednostavljivanje dohvaćanja informacija pružajući kontekstualizirane informacije korisnicima, a ne samo rezultate prikazane u obliku ključnih riječi (AI Glossary 2024).

Trenutno istraživanje u generativnom IR-u primarno je usmjereno prema dva glavna cilja, a to su: generativno pronalaženje dokumenata (eng. *Generative Document Retrieval* - GDR) koje uključuje pronalaženje dokumenata generiranjem njihovih identifikatora i pouzdano generiranje odgovora što podrazumijeva izravno generiranje odgovora usmjerenih na korisnika kroz strategije koje povećavaju njihovu pouzdanost (Li et al. 2024).

Belkin predlaže da se inteligencija u IR sustavima treba promatrati kroz prizmu interakcije između korisnika, sustava i informacijskih izvora. Inteligentni IR sustavi ne bi trebali samo dohvaćati informacije, već trebaju omogućiti korisnicima da aktivno upravljaju informacijskim problemima, prilagođavajući sustav njihovim potrebama kroz interaktivne

procesu (Belkin 1996). Nadalje, Belkin u radu istražuje koncept inteligencije u IR sustavima, postavljajući pitanje čiju inteligenciju bi takvi sustavi trebali upotrebljavati, inteligenciju od korisnika ili sustava. Osim toga, zalaže se za pristup koji je više usmjeren na korisnika gdje se sustav prilagođava promjenjivim potrebama te kontekstu korisnika i njegovim zahtjevima. Belkin ističe važnost razumijevanja korisnikove namjere i konteksta za poboljšanje učinkovitosti IR sustava postavljajući temelje za razvoj personaliziranog pretraživanja i odgovaranja na korisničke upite (Belkin 1996).

5.1. Načini primjene generativne umjetne inteligencije u pretraživanju informacija

Generativna umjetna inteligencija mijenja načine korištenja digitalnih sadržaja tako da korisnici mogu koristiti glasovne naredbe ili multimedijske sadržaje kako bi izvršili pretraživanje. Nadalje, generativna umjetna inteligencija se može koristiti za rješavanje složenih problema pretraživanja i pronalaska informacija što uvelike otvara put razvoju novih mogućnosti. Ova tehnologija olakšava pretragu složenih i višeznačnih informacija te pruža relevantne rezultate čak i u situacijama s nedostatkom jasnih upita (Feuerriegel et al. 2024).

Sa značajnim napretkom jezičnih modela, generativno pretraživanje se pojavilo kao novi smjer u području pretraživanja informacija. Za razliku od tradicionalnih metoda dohvaćanja temeljenih na indeksu, generativno dohvaćanje odnosno pretraživanje oslanja se na unaprijed uvježbane generativne modele za izravno generiranje identifikatora dokumenata povezanih s upitom, čime se postiže end-to-end pretraživanje i pronalaženje (Li et al. 2024).

Generativno dohvaćanje dokumenata (engl. *Generative Document Retrieval* – GDR) je najznačajniji primjer utjecaja generativne umjetne inteligencije na pretraživanje informacija. Alati generativne umjetne inteligencije u kratkom vremenskom intervalu mogu pretražiti velike količine podataka. Umjesto prikaza relevantnih rezultata i izvora informacija, GAI modeli najčešće izvlače podatke iz relevantnih izvora odnosno dokumenata te korisnicima pružaju gotove informacije i objašnjenja (Alaofi et al. 2024). Dohvaćanje dokumenata nije jedini ključni pristup koji je doprinio generativnom pretraživanju. Srodan pristup je generiranje proširenog dohvaćanja (eng. *Retrieval Augmented Generation*- RAG), koji spaja prednosti dohvaćanja i generiranja informacija.

Dok GDR koristi modele za pretraživanje dokumenata i njihovo generativno predstavljanje, RAG integrira generativni model s pretraživanjem kako bi prvo dohvatili

relevantne informacije, a zatim ih koristili za generiranje odgovora. Time RAG smanjuje potrebu za pohranjivanjem ogromnih količina podataka unutar modela, uz omogućavanje detaljnijih i točnijih odgovora (IBM 2023). Odgovaranje na pitanja putem chatbotova je još jedan primjer utjecaja generativne umjetne inteligencije na pretraživanje informacija.

Sljedeća generacija chatbota s generativnim AI mogućnostima nudit će poboljšanu funkcionalnost s razumijevanjem uobičajenog jezika i složenih upita, sposobnošću prilagodbe korisnikovom stilu razgovora i korištenjem empatije prilikom odgovaranja na pitanja korisnika. Samoučeći generativni AI chatbotovi izgrađeni na konverzacijskoj AI platformi neprestano se i automatski poboljšavaju. Oni koriste algoritme koji automatski uče iz prošlih interakcija kako najbolje odgovoriti na pitanja i poboljšati usmjerenje tijekom razgovora (IBM 2021).

Citirani tekst iz 2021. godine opisuje napredne mogućnosti chatbotova s generativnom umjetnom inteligencijom, uključujući prilagodbu korisničkom jeziku i stilovima komunikacije te sposobnost učenja iz prošlih interakcija. Iako su ove funkcionalnosti u to vrijeme predstavljale značajan napredak, važno je istaknuti da je u području generativne umjetne inteligencije tri godine izrazito dug period. Danas već svjedočimo znatno naprednijim sposobnostima generativnih AI sustava, koji ne samo da automatski uče iz interakcija, već koriste i sofisticirane modele za složeniju obradu jezika, kreiranje konteksta, prilagodbu različitim korisnicima i implementaciju emocionalne inteligencije.

U kontekstu primjene generativne umjetne inteligencije u pretraživanju informacija, sustavi preporuka predstavljaju ključnu tehnologiju za personalizaciju korisničkog iskustva. Njihov cilj nije samo predviđanje budućih interesa korisnika na temelju prošlih aktivnosti, već i filtriranje velikog broja informacija kako bi se korisnicima omogućilo relevantno pretraživanje. U tom smislu, generativna umjetna inteligencija dodaje novu dimenziju: umjesto da se oslanja isključivo na analizu podataka o prethodnim odabirima, moderni AI sustavi mogu generirati nove informacije i preporuke kroz razumijevanje složenih obrazaca ponašanja korisnika.

Sustavi za preporuku s jedne strane uče model korisnika kroz njegove aktivnosti, a s druge strane uče model proizvoda ili usluge na temelju interakcija s poznatim, već modeliranim korisnicima. Nakon toga sustavi sparuju korisnike i proizvode na

temelju naučenih modela, te korisnicima nude one proizvode za koje pretpostavljaju da će se baš njima svidjeti (Stipaničev et al. 2021).

Sustavi preporuka postupno su postali sastavni dio našeg svakodnevnog života. Trgovci pažljivo prate naše kupovne navike kako bi nam preporučili slične proizvode i tako povećali svoju prodaju. Društvene mreže analiziraju naše kontakte kako bi nam predložile nove prijatelje i potaknule nas na veću interakciju s platformom dok internetske radio stanice pamte pjesme koje smo preskočili kako bi nam u budućnosti pružile personaliziraniji glazbeni doživljaj (Lü et al. 2012).

Nadalje, konverzijski agenti predstavljaju naprednu tehnologiju koja dodatno unapređuje korisničku interakciju s informacijskim sustavima. Ovi agenti, stvoreni za oponašanje dijaloga nalik ljudskom, omogućuju korisnicima ne samo da pretražuju informacije, već i da kroz prirodne i dinamične razgovore dolaze do relevantnih sadržaja. Generativna umjetna inteligencija omogućuje konverzijskim agentima da nadmaše klasične metode pretraživanja, jer mogu razumjeti složene upite i pružiti personalizirane odgovore koji se temelje na korisničkim potrebama.

Budućnost konverzijskih agenata u pretraživanju informacija leži upravo u napretku generativne AI tehnologije, koja uključuje poboljšano razumijevanje prirodnog jezika, emocionalnu inteligenciju i sposobnost vođenja dinamičnih interakcija. Umjesto statičnih upita, korisnici će moći voditi interaktivne razgovore s agentima koji će ih voditi kroz proces pretraživanja i preporučivati im sadržaje na temelju kontekstualnog razumijevanja njihovih potreba. Ovakav napredak predstavlja inovaciju u pretraživanju informacija jer omogućuje dublju personalizaciju i efikasnije pronalaženje relevantnih informacija u stvarnom vremenu (Casheekar et al. 2024).

5.2. Usporedba mogućnosti generativne umjetne inteligencije s klasičnim sustavima za pretraživanja informacija

Sustavi za pretraživanje informacija razvijali su se kroz desetljeća, prelazeći od jednostavnih modela podudaranja ključnih riječi prema sofisticiranijim algoritmima koji koriste umjetnu inteligenciju. Tradicionalni sustavi pretraživanja prvenstveno su se oslanjali na izravno podudaranje pojmova između korisničkog upita i sadržaja dokumenata. Suvremeni sustavi, potpomognuti umjetnom inteligencijom, pružaju dublje razumijevanje ne samo korisničkih upita, nego i konteksta u kojem su ti upiti postavljeni.

Jedan od ključnih trenutaka u evoluciji pretraživanja bilo je uvođenje velikih jezičnih modela (LLM), temeljenih na Transformer arhitekturi. Ovi modeli, poput onih korištenih u naprednim sustavima kao što je New Bing, pokazali su izvanrednu sposobnost razumijevanja prirodnog jezika te generiranja sadržaja na temelju kontekstualnog učenja. LLM modeli ne samo da poboljšavaju proces dohvaćanja informacija, već omogućuju kondenziranje podataka iz više izvora u sažete i jasne odgovore. Time korisnicima značajno olakšavaju navigaciju kroz informacijsko preopterećenje koje često prati digitalno doba (Zhu et al. 2024).

S uvođenjem naprednih algoritama dubokog učenja, sustavi za pretraživanje postali su učinkovitiji i precizniji. Naime, umjesto jednostavnog podudaranja ključnih riječi, sustavi sada prepoznaju semantičke odnose i kontekst, što omogućuje dohvaćanje relevantnijih informacija (Hambarde i Proenca 2023). Ovaj pomak posebno je važan u pretraživanju velikih i heterogenih baza podataka, gdje klasični pristupi često podbacuju zbog količine i složenosti podataka.

Druga ključna inovacija u ovom području je generativna umjetna inteligencija, primjerice Chat GPT. Za razliku od tradicionalnih sustava koji se fokusiraju na dohvaćanje postojećih podataka, generativni modeli stvaraju nove informacije. Ovaj prijelaz značajno mijenja dinamiku pretraživanja informacija jer korisnici ne dobivaju samo rezultate temeljene na pretraživanim podacima, već i personalizirane, sažete odgovore na svoje upite (Carroll i Borycz 2024).

Ipak, iako su LLM modeli i generativna umjetna inteligencija donijeli velike inovacije, izazovi ostaju. Unatoč širokoj dostupnosti, tehnologija umjetne inteligencije za mnoge korisnike ostaje *crna kutija*. Ograničeno razumijevanje kako ti sustavi funkcioniraju može dovesti do nepovjerenja u dobivene rezultate. Studije pokazuju da korisnici koji bolje razumiju tehnologiju umjetne inteligencije postižu bolje rezultate u pretraživanju i povećavaju svoju produktivnost (Jang et al. 2024).

Sustavi za pretraživanje temeljeni na umjetnoj inteligenciji pružaju naprednije i kontekstualno relevantnije rezultate u usporedbi s tradicionalnim pristupima. Generativna umjetna inteligencija dodatno poboljšava performanse pretraživanja omogućujući bolje razumijevanje složenih podataka. Međutim, točnost, autoritativnost i kontekstualizacija pretraženih informacija ostaju izazovi, osobito u akademskom okruženju. Stoga, iako LLM

modeli znatno unapređuju proces pretraživanja, postoji potreba za daljnjim istraživanjima i poboljšanjima kako bi se osigurala veća točnost i pouzdanost rezultata (Hersh 2023).

5.3. Informacijsko preopterećenje i generativna umjetna inteligencija

U suvremenom digitalnom okruženju, informacijsko preopterećenje postalo je značajan izazov s kojim se suočavaju korisnici. Pojava raznolike i velike količine informacija, uz brzi razvoj tehnologija, otežava korisnicima prepoznavanje relevantnog sadržaja. Generativna umjetna inteligencija nudi potencijalna rješenja za ovaj problem, omogućujući korisnicima da efikasnije pretražuju i analiziraju informacije. Ova tehnologija može poboljšati proces filtriranja i prilagodbe informacija prema specifičnim potrebama, čime se smanjuje informacijsko preopterećenje. U nastavku se razmatra informacijsko preopterećenje kroz prizmu citata koji ukazuje na njegove glavne aspekte.

Informacijsko preopterećenje je postalo veliki izazov u digitalnom dobu, u kojem su ljudi okruženi ogromnim količinama informacija iz različitih izvora. Brzi razvoj tehnologije i širenje platformi društvenih medija i produkcija online sadržaja doveli su do stalnog protoka informacija i pristupa bez presedana velikoj količini informacija. U rješavanju problema informacijske preopterećenosti, preporučuje se niz alata, tehnika i tehnologija za učinkovito upravljanje priljevom dodatnih informacija za korisnike. Među njima, dva su privukla značajnu pozornost u postojećoj literaturi: korištenje inteligentnih agenata za filtriranje informacija i implementacija personaliziranih sustava preporuka (Shahrzadi et al. 2024).

Informacijsko preopterećenje smanjuje sposobnost učinkovitog funkcioniranja, što može dovesti do pogrešnih odluka kao i do nemogućnosti donošenja odluka. Glavni alat za sprečavanje informacijskog preopterećenja je ograničavanje izvora informacija, a strategije poput isključivanja obavijesti i ograničavanja vremena za računalom mogu biti vrlo učinkovite. Tehnološka rješenja ovog problema uključuju implementaciju alata generativne umjetne inteligencije (Tech Target 2024).

Koristeći AI alate i tehnike, pojedinci se mogu učinkovito kretati golemom količinom informacija koje su im dostupne, izvući vrijedne uvide i naposljetku donositi odluke utemeljene na dobrim informacijama. Inteligentni sustavi upravljanja informacijama igraju ključnu ulogu u pojednostavljivanju prioritizacije informacija i pružanju korisnih filtara korisnicima. Korištenjem mehanizama

automatskog skeniranja, semantičkog filtriranja i upozoravanja, korisnici su osnaženi da se s lakoćom kreću kroz ogroman priljev informacija (Shahrzadi et al. 2024).

6. Prednosti i izazovi primjene generativne AI u pretraživanju informacija

S obzirom na brzi napredak tehnologije i promjene u korisničkim potrebama, ključno je razumjeti prednosti i izazove koje generativna umjetna inteligencija donosi u području pretraživanja informacija. Ova tehnologija može značajno unaprijediti preciznost i relevantnost rezultata pretraživanja, omogućujući korisnicima pristup personaliziranim i kontekstualno prilagođenim informacijama. Ipak, uz te prednosti javljaju se i izazovi, kao što su etička pitanja, rizik od pristranosti te potreba za osiguravanjem kvalitete informacija. Kroz analizu ovih aspekata, ovo poglavlje doprinosi boljem razumijevanju uloge generativne umjetne inteligencije u transformaciji pretraživanja informacija i postavlja temelj za daljnje istraživanje i primjenu inovativnih rješenja u tom području.

Generativna umjetna inteligencija ne samo da poboljšava kvalitetu odgovora na specifične korisničke upite, već također omogućava sažimanje složenih informacija i pojednostavljivanje procesa stvaranja sadržaja. Ove sposobnosti su od ključne važnosti za optimizaciju korisničkog iskustva i učinkovitost pretraživanja, posebno u doba informacijskog preopterećenja. Razumijevanje ovih prednosti pomaže u identificiranju načina na koje generativna umjetna inteligencija može unaprijediti pristup informacijama i donošenje odluka. Neke od potencijalnih prednosti implementacije generativne umjetne inteligencije u pretraživanje informacija uključuju poboljšanje odgovora na specifične korisničke upite, sažimanje složenih informacija, pojednostavljivanje procesa stvaranja sadržaja i slično (Tech Target 2024).

Prema članku *What is Information Retrieval with AI?* iz 2024. godine generativna umjetna inteligencija donosi niz prednosti sustavima za pronalaženje informacija kao što su: *Poboljšana preciznost:* Umjetna inteligencija povećava preciznost rezultata pretraživanja razumijevanjem konteksta iza kompleksnih korisničkih upita i isporukom točnijih rezultata. *Brže donošenje odluka:* pronalaženje informacija vođeno umjetnom inteligencijom nudi brži pristup podacima u stvarnom vremenu koji ubrzava procese poslovnog odlučivanja. *Smanjenje ljudskih pogrešaka:* Automatiziranjem podataka i procesa analize, umjetna inteligencija smanjuje mogućnost ljudskih pogrešaka i nedosljednosti, čime se povećava učinkovitost i

točnost. *Poboljšano korisničko iskustvo*: Intuitivni i inteligentni sustavi značajno poboljšavaju korisničko iskustvo nudeći personalizirane preporuke na temelju ponašanja i preferencija korisnika. *Bolji uvidi u procese*: Kroz prepoznavanje uzoraka i prediktivnu analitiku, umjetna inteligencija može pružiti vrijedne uvide za učinkovito donošenje odluka i planiranje strategije (AI Glossary 2024).

Prema članku *From Matching to Generation: A Survey on Generative Information Retrieval* tradicionalni IR sustavi oslanjali su se na podudaranje ključnih riječi, dok generativni modeli omogućuju semantičko pretraživanje, poboljšavajući točnost pronalaska relevantnih informacija. Prednosti uključuju mogućnost odgovaranja na složene upite, učenje bez ili s vrlo malo dodatnih podataka te širenje pretraživanja izvan teksta na slike i videozapise (Li et al. 2024).

Uz prednosti, prikazani su i potencijalni izazovi implementacije generativne umjetne inteligencije u sustave za pretraživanje informacija: *Priistranost algoritama*: budući da modeli umjetne inteligencije uče iz podataka, mogli bi nenamjerno naslijediti pristranosti koje postoje u podacima o obuci što bi moglo utjecati na pravednost i neutralnost sustava u pronalaženju informacija. *Složena implementacija*: Dizajniranje i implementacija sustava za pretraživanje informacija temeljenog na umjetnoj inteligenciji može biti složeno i zahtjeva značajnu stručnost u umjetnoj inteligenciji i srodnim tehnologijama. *Visoki troškovi*: Troškovi razvoja i održavanja sustava temeljenih na umjetnoj inteligenciji mogli bi biti financijski problem za mala i srednja poduzeća. *Redovita ažuriranja sustava*: Kako se tehnologija umjetne inteligencije brzo razvija, potrebna su stalna ažuriranja i održavanje sustava kako bi se iskoristio njezin puni potencijal. *Sigurnosni problemi*: sustavi umjetne inteligencije zahtijevaju okvir kibernetičke sigurnosti za zaštitu osjetljivih informacija od potencijalnih prijetnji i kršenja, što može biti posebno izazovno s obzirom na stalni i nekontrolirani porast informacija (AI Glossary 2024).

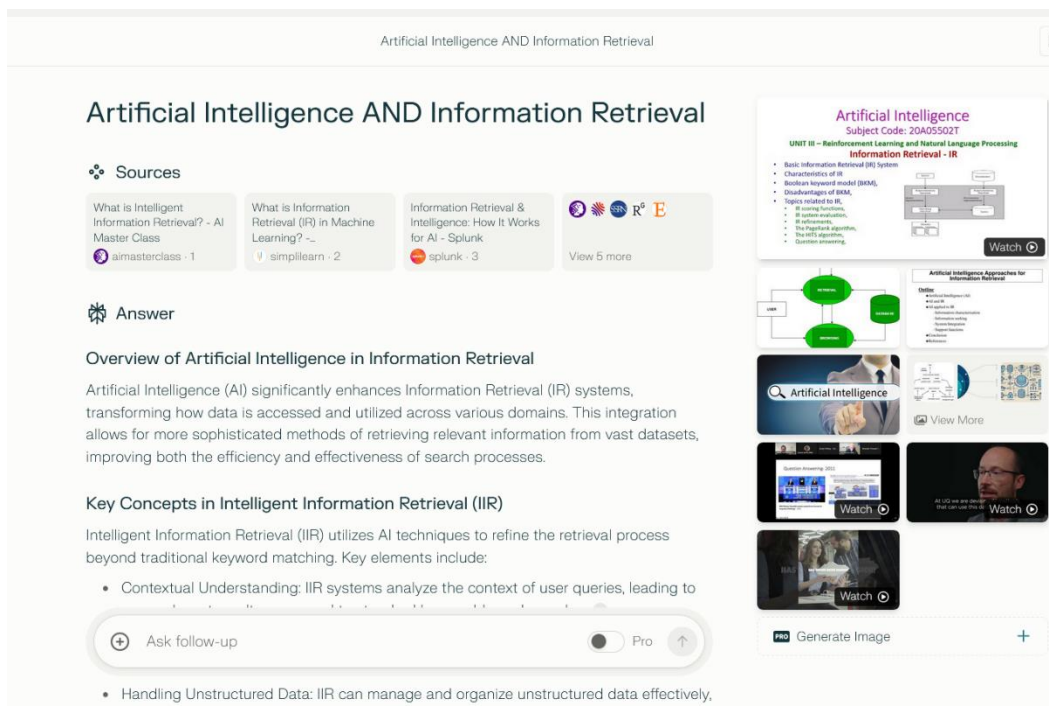
Iako prednosti upućuju na sve veći tehnološki napredak, izazovi poput pristranosti, sigurnosti i održivosti sustava naglašavaju veću potrebu za promišljanjem o odgovornoj primjeni AI tehnologija. Ovi aspekti vode prema ključnim etičkim pitanjima i budućim smjerovima razvoja koji će biti presudni u definiranju načina na koji će generativna umjetna inteligencija oblikovati pretraživanje informacija. Iako porast primjene generativne umjetne inteligencije može donijeti poboljšanja unutar sustava za pretraživanje informacija, moguće je

da umjetna inteligencija nenamjerno učini više štete nego koristi ako se ne razvije i ne primijeni na etičan i odgovoran način (McGovern et al. 2022).

7. Primjeri sustava za pretraživanje temeljenih na generativnoj inteligenciji

U današnje vrijeme, sustavi za pretraživanje informacija temeljeni na umjetnoj inteligenciji postaju sve prisutniji i sofisticiraniji, pružajući korisnicima mogućnost pristupa relevantnim informacijama na način koji nadilazi tradicionalne sustave za pretraživanje. Dok klasični sustavi, poput Googlea, uglavnom nude liste linkova kao odgovore na korisničke upite, novi AI alati kao što su Perplexity AI, Elicit, Consensus, SciSpace i Scite pružaju kontekstualizirane, sažete i personalizirane odgovore. Ovi sustavi koriste napredne tehnike obrade prirodnog jezika kako bi korisnicima omogućili preciznije i učinkovitije pretraživanje informacija u različitim područjima. Razumijevanje razlika između ovih naprednih AI i klasičnih sustava ključno je za optimizaciju procesa pronalaska relevantnih informacija koje su namijenjene korisnicima.

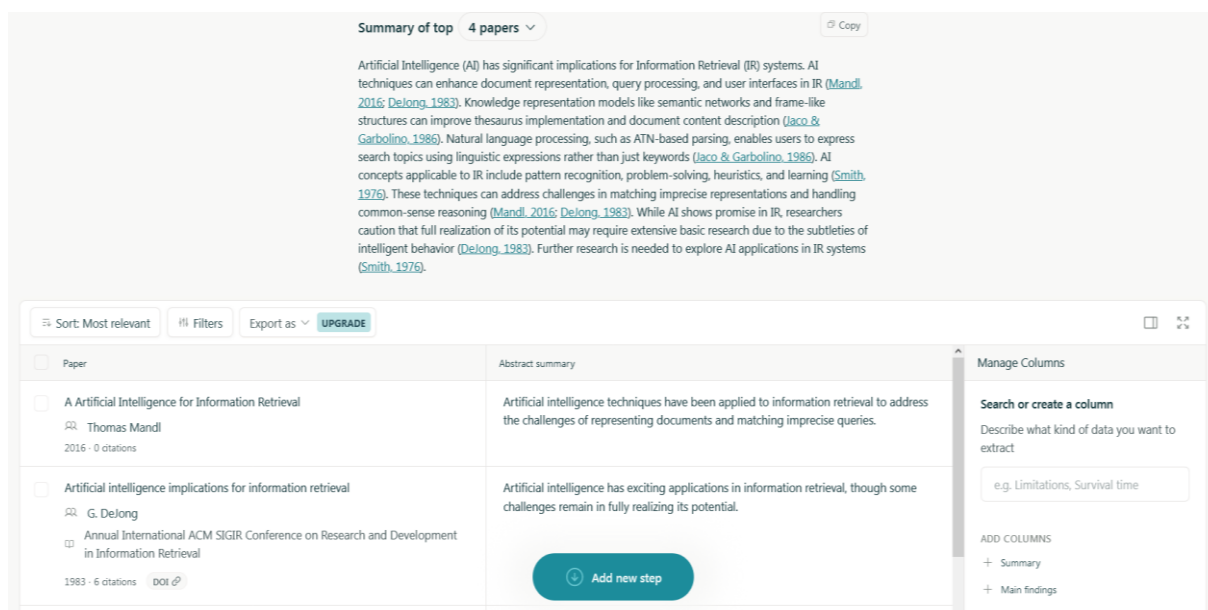
Perplexity AI je napredni alat za pretraživanje temeljen na modelima umjetne inteligencije koji omogućava korisnicima da postavljaju kompleksna pitanja i dobiju koherentne, kontekstualne odgovore. Njegova snaga leži u mogućnosti da sažima informacije iz različitih izvora te da direktno odgovara na pitanja, za razliku od klasičnih sustava za pretraživanje koji uglavnom nude linkove na stranice s potencijalnim odgovorima. Uz to, Perplexity AI koristi modeliranje prirodnog jezika kako bi kreirao relevantnije rezultate te ima i mogućnost praćenja konteksta kroz više ponavljanja korisničkih upita, što omogućuje složenije pretrage i preciznije odgovore (Perplexity AI 2024).



Slika 1: Perplexity AI

Na slici 1 je prikaz sučelja Perplexity AI-a s upitom “Artificial Intelligence AND Information Retrieval”. Rezultati pretraživanja jasno prikazuju sažetak relevantnih informacija o ulozi umjetne inteligencije u sustavima za pretraživanje informacija. U desnom dijelu sučelja nalaze se povezani upiti i teme kao i poveznice za detaljnije istraživanje dok je u glavnom dijelu zaslona prikazan tekst koji objedinjuje više izvora informacija koji su navedeni na samom vrhu. Sučelje je jednostavno i fokusirano na sadržaj što omogućuje korisnicima brzo razumijevanje teme kroz sažetke umjesto prikaza samo poveznica. Ovakav pristup omogućuje korisnicima učinkovitije pretraživanje informacija i dobivanje relevantnih odgovora u kratkom vremenu čime se optimizira proces istraživanja.

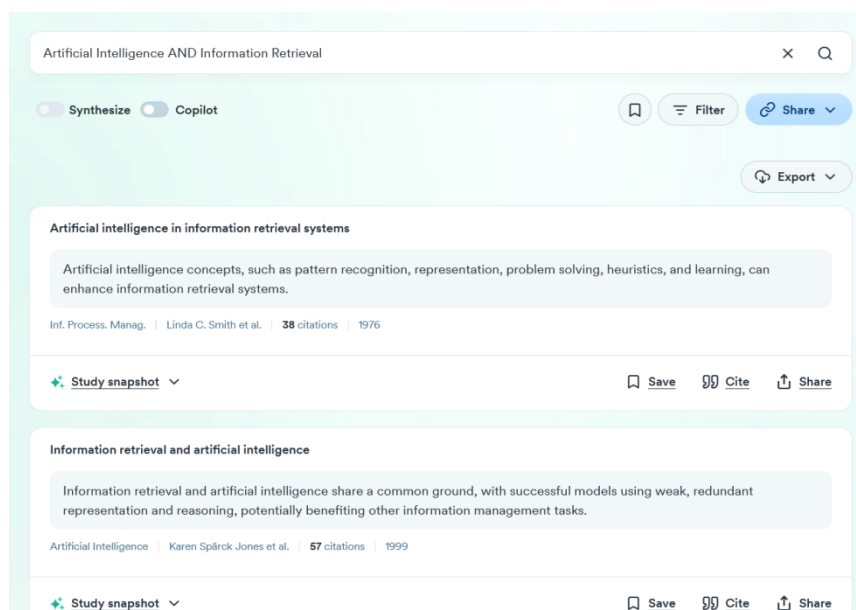
Elicit je alat za znanstvena istraživanja koji koristi umjetnu inteligenciju za analizu i ekstrakciju relevantnih informacija iz znanstvene literature. Fokusiran je na pomoć istraživačima, omogućujući brzo pregledavanje rezultata znanstvenih studija, a često ga koriste u kontekstu sustavnih pregleda literature. Alat omogućava i filtriranje rezultata prema hipotezama, metodologijama ili zaključcima, što klasični sustavi većinom ne nude. Njegova glavna prednost je sposobnost da ubrza proces traženja specifičnih znanstvenih informacija, a dodatno omogućava i prilagođavanje pretrage prema određenim znanstvenim temama odnosno područjima istraživanja (Elicit 2024).



Slika 2: Elicit

Slika 2 prikazuje sučelje Elicita-a s upitom “Artificial Intelligence AND Information Retrieval”. Rezultati pretraživanja prikazuju sažetke znanstvenih radova, uključujući hipoteze, metodologije i zaključke što korisnicima olakšava pregledavanje specifičnih informacija. Ovo omogućava istraživačima da brzo identificiraju ključne informacije i uspoređuju različite studije. Perplexity AI se fokusira na generiranje kontekstualnih odgovora na kompleksna pitanja sa sažetkom iz različitih izvora dok Elicit omogućuje detaljnu analizu znanstvenih radova s naglaskom na metodologije i hipoteze. Elicit je posebno koristan za sustavne preglede literature, a Perplexity AI bolje odgovara općim upitima i brzim pretragama. Također, Elicit nudi filtriranje rezultata prema specifičnim znanstvenim kriterijima, što nije glavna značajka Perplexity AI-a.

Consensus je dizajniran kao alat za pretraživanje znanstvenih radova i stručne literature s fokusom na davanje odgovora temeljenih na znanstvenom konsenzusu (sporazumu). Umjesto da jednostavno nudi članke, Consensus pokušava sažeti glavne zaključke iz relevantnih istraživanja i prikazati koje su informacije podržane u većini slučajeva. To pomaže znanstvenicima odnosno korisnicima da brže donesu informirane odluke, oslanjajući se na relevantne i pouzdane izvore. U usporedbi s klasičnim pretraživačima, Consensus pruža jasniji uvid u stav znanstvene zajednice o određenom pitanju. Istovremeno omogućuje i brzo prepoznavanje neslaganja među istraživanjima, što olakšava analizu kontradiktornih rezultata (Consensus 2024).



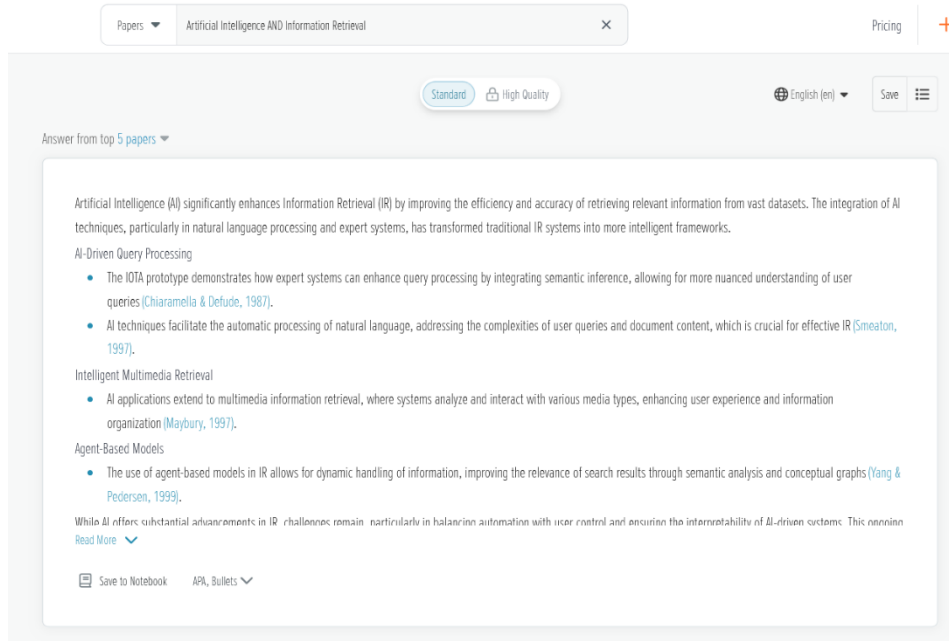
Slika 3: Consensus

Prikazano je sučelje Consensus-a kao i upit “Artificial Intelligence AND Information Retrieval”. Ovaj alat umjetne inteligencije sažima glavne zaključke iz relevantnih istraživanja naglašavajući koje su informacije podržane od strane većine znanstvenika. Ovaj pristup omogućava korisnicima brže donošenje informiranih odluka oslanjajući se na provjerene izvore i prepoznajući nesuglasice među istraživanjima. Elicit se fokusira na detaljnu analizu znanstvenih radova i pruža specifične informacije kao što su hipoteze, metodologije i zaključci dok Consensus pruža općenitiji pregled konsenzusa unutar znanstvene zajednice. Perplexity AI nudi koherentne odgovore na kompleksna pitanja sažimanjem iz različitih izvora, ali ne naglašava konsenzus ili nesuglasice u istraživanju. Njegova snaga leži u brzini i lakoći pretrage, a ne u analizi znanstvenih stavova.

Iako sva tri alata koriste umjetnu inteligenciju, razlikuju u pristupu i ciljevima: Elicit je orijentiran prema detaljnoj analizi i istraživanju znanstvene literature, Consensus se fokusira na kratke sažetke analizu konsenzusa u znanstvenoj zajednici, a Perplexity AI pruža brze i kontekstualizirane odgovore s naglaskom na širu primjenu.

SciSpace fokusiran je na olakšavanje razumijevanja i stvaranja znanstvenih dokumenata. Alat koristi AI kako bi analizirao znanstvene članke i pružio sažetke te olakšao čitanje kroz složene tekstove. Također, omogućuje korisnicima da pišu vlastite radove s automatskim formatiranjem prema standardima publikacija. Za razliku od klasičnih sustava za pretraživanje, SciSpace je orijentiran na interakciju sa sadržajem (pisanje, analiza, komentari)

i poboljšava razumijevanje znanstvenih radova s pomoću AI alata. Osim toga, nudi funkcije poput automatskog označavanja ključnih elemenata istraživanja, što pomaže u bržem pregledu relevantnih informacija i poboljšava surađivanje među istraživačima (SciSpace 2024).

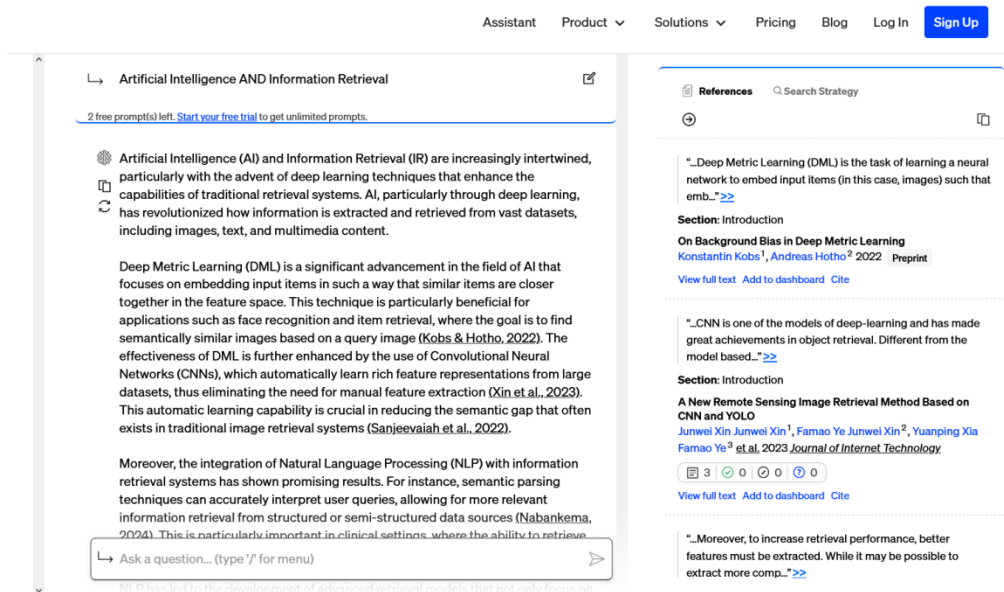


Slika 4: SciSpace

Slika 4 prikazuje upit “Artificial Intelligence AND Information Retrieval” u SciSpace-u koji donosi rezultate uključujući članke koji se bave raznim aspektima primjene umjetne inteligencije u pretraživanju informacija. Sučelje prikazuje istaknute dijelove znanstvenih radova poput naslova, sažetka i uvoda. SciSpace ima napredne značajke poput automatskog formatiranja citata i izdvajanja ključnih riječi što korisnicima olakšava pretraživanje kao i odabir najrelevantnijih rezultata. Ovaj AI alat pomaže u bržem razumijevanju i pregledu ključnih elemenata složenih znanstvenih radova, pružajući jasan i intuitivan prikaz bitnih informacija.

Scite je alat koji koristi umjetnu inteligenciju za analiziranje znanstvenih radova i pružanje informacija temeljenih na citiranjima. Njegova ključna prednost je u analizi kako su određeni radovi citirani – jesu li citati pozitivni, negativni ili neutralni. Ovaj alat omogućava korisnicima da brzo shvate utjecaj i valjanost nekog rada. U usporedbi s klasičnim sustavima za pretraživanje, Scite omogućava dublji uvid u kredibilitet radova kroz citate, što pomaže u procjeni kvalitete literature. Dodatno, omogućuje vizualizaciju citata kroz vrijeme i prikazuje

trendove u citiranju, što olakšava praćenje razvoja određenih teorija ili hipoteza kroz godine (Scite 2024).



Slika 5: Scite

Slika 5 prikazuje sučelje alata Scite s rezultatima pretrage za upit “Artificial Intelligence AND Information Retrieval”. Prikazani su dijelovi znanstvenih radova uz analizu citata. Desna strana sučelja vizualizira trendove citiranja kroz vrijeme, naglašavajući promjene u načinu na koji je tema citirana tijekom posljednjih godina. Svaki rad uključuje sažetak kao i kontekst u kojem je citiran što pomaže u procjeni relevantnosti i utjecaja rada. Prikazani su i vizualni indikatori koji pokazuju kredibilitet rada na temelju njegovih citata.

Alati temeljeni na umjetnoj inteligenciji razlikuju se od klasičnih sustava za pretraživanje po nekoliko ključnih aspekata. Prvo, koriste napredne metode dubinske analize za interpretaciju i sažimanje sadržaja, omogućujući brži pristup relevantnim informacijama, umjesto da korisnicima jednostavno pružaju popis linkova odnosno rezultata pretraživanja. Nadalje, umjesto da se oslanjaju isključivo na ključne riječi, ovi sustavi nude kontekstualizirane odgovore na pitanja ili sažetke znanstvenih radova. Posebno su korisni jer su specijalizirani za određene znanstvene i istraživačke domene, dok su klasični sustavi generalizirani. Osim toga, AI alati za pretraživanje često pružaju interaktivnost i personalizaciju, nudeći preporuke i analize prilagođene potrebama korisnika, što je rijetko prisutno u tradicionalnim pretragama.

8. Zaključak

Zaključno, generativna umjetna inteligencija donosi značajne promjene u području pretraživanja informacija, otvarajući nove mogućnosti za pronalaženje, analizu i generiranje sadržaja. Tradicionalni sustavi pretraživanja, iako dugo učinkoviti, suočavaju se s izazovima u digitalnom dobu, osobito zbog ogromne količine dostupnih informacija. Informacijsko preopterećenje i nedostatak preciznosti u rezultatima pretraživanja sve su češći problemi. Generativna umjetna inteligencija, kroz napredne algoritme strojnog učenja i obrade prirodnog jezika, pruža rješenja koja omogućuju preciznije i kontekstualno relevantnije rezultate. Time se pretraživanje informacija ne svodi samo na dohvaćanje dokumenata već i na generiranje personaliziranih odgovora temeljenih na korisnikovim potrebama i upitima.

Jedna od ključnih prednosti generativne umjetne inteligencije je sposobnost kreiranja sažetih i personaliziranih informacija, što korisnicima olakšava pronalaženje relevantnih informacija bez potrebe za ručnim filtriranjem velikih količina informacija. Sustavi temeljeni na generativnim modelima omogućuju dublje razumijevanje korisničkih upita, pružajući relevantne odgovore čak i u situacijama kada su upiti složeni ili nejasni. Osim toga, napredni modeli poput generativnih kontradiktornih mreža (GAN-ova) i velikih jezičnih modela (LLM) omogućuju generiranje novih informacija, čime se pretraživanje informacija značajno unapređuje.

Iako generativna umjetna inteligencija donosi mnoge prednosti, važno je osvijestiti izazove koje ona nosi. Etika i sigurnost postaju ključne teme, osobito u kontekstu potencijalne pristranosti modela, širenja dezinformacija i zaštite privatnosti korisnika. Generativni modeli mogu proizvesti informacije koje, iako uvjerljive, nisu uvijek točne ili provjerene, što postavlja pitanja o odgovornosti i kontroli nad generiranim sadržajem. Također, iako je brzina pretraživanja značajno povećana, važno je osigurati da kvaliteta rezultata ne pati zbog ubrzanih procesa generalizacije odgovora.

Budućnost pretraživanja informacija neizbježno će se kretati u smjeru daljnje integracije generativne umjetne inteligencije. Ova tehnologija ima potencijal revolucionirati način na koji pristupamo informacijama, pružajući korisnicima brže, preciznije i prilagođenije odgovore. Ipak, ključno je osigurati ravnotežu između napretka tehnologije i očuvanja etičkih standarda kako bi se omogućila sigurna i odgovorna primjena generativnih sustava. Kroz daljnji razvoj i istraživanje, generativna umjetna inteligencija imat će ključnu ulogu u unapređenju

pretraživanja informacija, osiguravajući korisnicima učinkovitiji i jednostavniji pristup relevantnim informacijama u sve složenijem digitalnom okruženju.

9. Literatura

Ai, Qingyao, Ting Bai, Zhao Cao, Yi Chang, Jiawei Chen, Zhumin Chen, Zhiyong Cheng, i sur. 2023. „Information Retrieval Meets Large Language Models: A Strategic Report from Chinese IR Community“. *AI Open* 4:80–90. <https://doi.org/10.1016/j.aiopen.2023.08.001>.

AI Chat for Scientific PDFs | „SciSpace“. 2024. Pristupljeno 23. rujan 2024. <https://typeset.io>.

Alaofi, Marwah, Negar Arabzadeh, Charles L. A. Clarke, i Mark Sanderson. 2024. „Generative Information Retrieval Evaluation“. *arXiv*. Pristupljeno: 21. kolovoz 2024. <https://doi.org/10.48550/arXiv.2404.08137>.

Alat. *Hrvatski jezični portal*. Pristupljeno: 20. kolovoz 2024. https://hjp.znanje.hr/index.php?show=search_by_id&id=fVdjUA%253D%253D.

Al-Maskari, Azzah, i Mark Sanderson. 2010. „A Review of Factors Influencing User Satisfaction in Information Retrieval“. *Journal of the American Society for Information Science and Technology* 61 (5). Pristupljeno: 5. rujan 2024. <https://doi.org/10.1002/asi.21300>.

Ankica, Arbutina i Nakić Stanislav. „PRIMJENA UMJETNE INTELIGENCIJE U DIGITALNOM MARKETINGU.“ *Oeconomicus* 1, br. 23 (2024): 11-21. Pristupljeno: 22. rujan 2024. <https://hrcak.srce.hr/320058>

Belkin, Nicholas. 1996. „Intelligent Information Retrieval: Whose Intelligence?“ *Ingénierie Des Systèmes D'information*. Pristupljeno: 28. kolovoz 2024. https://www.researchgate.net/publication/240053046_Intelligent_Information_Retrieval_Whose_Intelligence

Bénédict, Gabriel, Ruqing Zhang, i Donald Metzler. 2023. „Gen-IR @ SIGIR 2023: The First Workshop on Generative Information Retrieval.“ *arXiv*. Pristupljeno: 20.

kolovoz 2024. <http://arxiv.org/abs/2306.02887>.

Bebić, Domagoj. „Uloga umjetne inteligencije u stvaranju medijskog sadržaja.“ *Suvremene teme* 14, br. 1 (2023): 47-59. Pristupljeno: 24. rujan 2024. <https://doi.org/10.46917/st.14.1.3>

CADCAM Group. 2023. „Što su umjetna inteligencija i strojno učenje i zašto su važni?“ Pristupljeno: 20. kolovoz 2024. <https://www.cadcam-group.eu/hr/knowledge/sto-su-umjetna-inteligencija-i-strojno-ucenje-i-zasto-su-vazni/>.

Cambridge English Dictionary. 2024. „Information Retrieval.“ Pristupljeno: 20. kolovoz 2024. <https://dictionary.cambridge.org/dictionary/english/information-retrieval>.

Carroll, Alexander J., i Joshua Borycz. 2024. „Integrating large language models and generative artificial intelligence tools into information literacy instruction“. *The Journal of Academic Librarianship* 50 (4). Pristupljeno: 21. kolovoz 2024. <https://doi.org/10.1016/j.acalib.2024.102899>.

Casheekar, Avyay, Archit Lahiri, Kanishk Rath, Kaushik Sanjay Prabhakar, i Kathiravan Srinivasan. 2024. „A Contemporary Review on Chatbots, AI-Powered Virtual Conversational Agents, ChatGPT: Applications, Open Challenges and Future Research Directions“. *Computer Science Review* 52 (5). Pristupljeno: 22. kolovoz 2024. <https://doi.org/10.1016/j.cosrev.2024.100632>.

Cooper, W. S. 1984. „Bridging the gap between AI and IR“. U *Proceedings of the 7th annual international ACM SIGIR conference on Research and development in information retrieval*. Pristupljeno: 21. kolovoz 2024. <https://dl.acm.org/doi/pdf/10.5555/636805.636822>.

Coutinho, Mariana, Lorena Marques, Anderson Santos, Marcio Dahia, Cesar Franca, i Ronnie de Souza Santos. 2024. „The Role of Generative AI in Software Development Productivity: A

Pilot Case Study “. arXiv. PRISTUPLJENO: 27. RUJAN 2024.
<https://doi.org/10.48550/ARXIV.2406.00560>.

Consensus AI-Powered Academic Search Engine. 2024. Consensus: AI Search Engine for Research. Pristupljeno 23. rujan 2024. <https://consensus.app/>.

Elicit: The AI Research Assistant. 2024. Pristupljeno 23. rujan 2024. <https://elicit.com/>.

Feuerriegel, Stefan, Jochen Hartmann, Christian Janiesch, i Patrick Zschech. 2024. „Generative AI.“ *Business & Information Systems Engineering* 66 (1). Pristupljeno: 20. kolovoz 2024. doi:10.1007/s12599-023-00834-7.

Grace, Katja, Harlan Stewart, Julia Fabienne Sandkühler, Stephen Thomas, Ben Weinstein-Raun, i Jan Brauner. 2024. „Thousands of AI Authors on the Future of AI“. *arXiv*. Pristupljeno: 22. kolovoz 2024. <https://arxiv.org/abs/2401.02843v2>.

Hambarde, Kailash A., i Hugo Proenca. 2023. „Information Retrieval: Recent Advances and Beyond.“ Pristupljeno: 20. kolovoz 2024.
doi:10.48550/ARXIV.2301.08801.

Hersh, William R. 2023. „Search Still Matters: Information Retrieval in the Era of Generative AI.“ *arXiv*. Pristupljeno: 20. kolovoz 2024.
<http://arxiv.org/abs/2311.18550>.

Hrvatska enciklopedija, mrežno izdanje. 2024. „Pretraživanje informacija.“
Leksikografski zavod Miroslav Krleža. Pristupljeno: 20. kolovoz 2024.
<https://enciklopedija.hr/clanak/pretrazivanje-informacija>.

Hrvatska enciklopedija, mrežno izdanje. 2024. „Umjetna inteligencija.“
Leksikografski zavod Miroslav Krleža. Pristupljeno: 20. kolovoz 2024.
<https://enciklopedija.hr/clanak/umjetna-inteligencija>.

Hrvatska enciklopedija , mrežno izdanje. 2024. „Baza podataka.“ Leksikografski zavod Miroslav Krleža. Pristupljeno 9. rujan 2024. <https://www.enciklopedija.hr/clanak/baza-podataka>.

Jang, Soobin, Haeyoon Lee, Yujin Kim, Daeho Lee, Jungwoo Shin, i Jungwoo Nam. 2024. „When, What, and How Should Generative Artificial Intelligence Explain to Users?“ *Telematics and Informatics* 93 (9). Pristupljeno: 20. kolovoz 2024. <https://doi.org/10.1016/j.tele.2024.102175>.

Jones, Karen Sparck. 1991. „The Role of Artificial Intelligence in Information Retrieval.“ *Journal of the American Society for Information Science* 42 (8). Pristupljeno: 20. kolovoz 2024. [doi:10.1002/\(SICI\)1097-4571\(199109\)42:8<558::AID-ASI4>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-4571(199109)42:8<558::AID-ASI4>3.0.CO;2-4).

Kuhail, Mohammad Amin, Sujith Samuel Mathew, Ashraf Khalil, Jose Berengueres, i Syed Jawad Hussain Shah. 2024. „Will I Be Replaced? Assessing ChatGPT’s Effect on Software Development and Programmer Perceptions of AI Tools“. *Science of Computer Programming* 235 (7). Pristupljeno: 20. kolovoz 2024. <https://doi.org/10.1016/j.scico.2024.103111>.

Kuo, Tzu-Lin, Tzu-Wei Chiu, Tzung-Sheng Lin, Sheng-Yang Wu, Chao-Wei Huang, i Yun-Nung Chen. 2024. „A Survey of Generative Information Retrieval“. arXiv. Pristupljeno: 23. rujan 2024. <https://doi.org/10.48550/arXiv.2406.01197>.

Li, Xiaoxi, Jiajie Jin, Yujia Zhou, Yuyao Zhang, Peitian Zhang, Yutao Zhu, i Zhicheng Dou. 2024. „From Matching to Generation: A Survey on Generative Information Retrieval“. Pristupljeno: 27. kolovoz 2024. <https://doi.org/10.48550/ARXIV.2404.14851>.

Lü, Linyuan, Matúš Medo, Chi Ho Yeung, Yi-Cheng Zhang, Zi-Ke Zhang, i Tao Zhou. 2012. „Recommender Systems“. *Physics Reports* 519 (1). Pristupljeno: 22. kolovoz 2024. <https://doi.org/10.1016/j.physrep.2012.02.006>.

McGovern, Amy, Imme Ebert-Uphoff, David John Gagne II, i Ann Bostrom. 2022. „Why We Need to Focus on Developing Ethical, Responsible, and Trustworthy Artificial Intelligence Approaches for Environmental Science“. *Environmental Data Science* 1 (1). Pristupljeno: 27. kolovoz 2024. <https://doi.org/10.1017/eds.2022.5>.

New Jersey Institute of Technology. 2024. „Eliza, a chatbot therapist“. Pristupljeno 20. kolovoz 2024. <https://web.njit.edu/~ronkowitz/eliza.html>.

Open AI. 2016. „Generative models.“ Pristupljeno: 20. kolovoz 2024. <https://openai.com/research/generative-models>.

Picariello, Antonio i Rinaldi Antonio M. 2008. „A User-Centered Approach for Information Retrieval“. Pristupljeno: 15. rujan 2024. <https://doi.org/10.4018/978-1-60566-144-5.ch009>.

Perplexity. 2024. Pristupljeno: 23. rujan 2024. <https://www.perplexity.ai/>

Ribić, Damir, Nikolina Medvedović i Domagoj Glumac. „Primjena umjetne inteligencije u analizi konkurencije.“ *ET²eR – ekonomija, turizam, telekomunikacije i računarstvo VI*, br. 1 (2024): 1-10. Pristupljeno: 24. rujan 2024. <https://doi.org/10.70077/et2er.6.1.1>

Sanderson, Mark i W. Bruce Croft. 2012. „The History of Information Retrieval Research“. *Proceedings of the IEEE* 100. Pristupljeno: 9. rujan 2024. <https://doi.org/10.1109/JPROC.2012.2189916>.

Shahrzadi, Leila, Ali Mansouri, Mousa Alavi, i Ahmad Shabani. 2024. „Causes, Consequences, and Strategies to Deal with Information Overload: A Scoping Review“. *International Journal of Information Management Data Insights* 4 (2). Pristupljeno: 23. kolovoz 2023. <https://doi.org/10.1016/j.jjime.2024.100261>.

Singbo, Odilon Gbenoukpo. „Umjetna inteligencija u suvremenom biokibernetičkom svijetu.“ *Spectrum*, br. 3-4 (2008): 55-63. Pristupljeno: 20. kolovoz 2024. <https://hrcak.srce.hr/38568>.

Stipaničev, Darko , Ljiljana Šerić, i Maja Braović. „Uvod u umjetnu inteligenciju”. Split: Sveučilište u Splitu, 2021.

Što je umjetna inteligencija i kako se upotrebljava? *Europski parlament, članak 2023.*

Pristupljeno: 20. kolovoz 2024.

https://www.europarl.europa.eu/pdfs/news/expert/2020/9/story/20200827STO85804/20200827STO85804_hr.pdf.

Tech Target. 2024. „What is generative AI? Everything you need to know.”

Pristupljeno: 20. kolovoz 2024.

<https://www.techtarget.com/searchenterpriseai/definition/generative-AI>.

Tech Target. 2024. „What Is Artificial Intelligence (AI)?” Pristupljeno: 20. kolovoz

2024. <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>.

Tech Target. 2024. „The future of AI: What to expect in the next 5 years?”

Pristupljeno: 22. kolovoz 2024.

<https://www.techtarget.com/searchenterpriseai/tip/The-future-of-AI-What-to-expect-in-the-next-5-years>.

Tech Target. 2024. „What is information overload?” Pristupljeno: 28. kolovoz 2024.

<https://www.techtarget.com/whatis/definition/information-overload>

Thompson, E., i Clark, A. 2017. AI in the World of Art: A New Wave of Creativity.

Art and Design Studies, str. 67-82.

Tomić, Zoran, Tomislav Volarić i Đorđe Obradović. „UMJETNA INTELIGENCIJA U ODNOSIMA S JAVNOŠĆU.“ South Eastern European Journal of Communication 4, br. 2 (2022): 7-16. Pristupljeno: 22. rujan 2024. <https://doi.org/10.47960/2712-0457.2022.2.4.7>

Varga, Matija i Dražen Ružić. „UMJETNA INTELIGENCIJA U OBRAZOVANJU.“

Varaždinski učitelj 7, br. 15 (2024): 87-96. Pristupljeno: 22. rujan 2024.

<https://hrcak.srce.hr/316187>

What Is Information Retrieval? 2020. *Geeks for Geeks*. Pristupljeno: 20. kolovoz 2024. <https://www.geeksforgeeks.org/what-is-information-retrieval/>.

What Is Artificial Intelligence (AI)? 2023. *IBM*. Pristupljeno: 20. kolovoz 2024. <https://www.ibm.com/topics/artificial-intelligence>.

What is Information Retrieval with AI? 2024. *AI Glossary*. Pristupljeno 21. kolovoz 2024. <https://www.aimasterclass.com/glossary/information-retrieval-with-ai>.

What Is a Chatbot? 2021. *IBM*. Pristupljeno: 21. kolovoz 2024. <https://www.ibm.com/topics/chatbots>.

What is Intelligent Information Retrieval? 2024. *AI Glossary*. Pristupljeno 27. kolovoz 2024. <https://www.aimasterclass.com/glossary/intelligent-information-retrieval>.

Your AI Research Assistant – Scite. 2024. *Scite.Ai*. Pristupljeno 23. rujan 2024. <https://scite.ai>.

Zhu, Yutao, Huaying Yuan, Shuting Wang, Jiongnan Liu, Wenhan Liu, Chenlong Deng, Haonan Chen, Zheng Liu, Zhicheng Dou, i Ji-Rong Wen. 2024. „Large Language Models for Information Retrieval: A Survey“. *arXiv*. Pristupljeno: 23. rujan 2024. <http://arxiv.org/abs/2308.07107>.

Popis slika

<u>Slika 1: Perplexity AI</u>	28
<u>Slika 2: Elicit</u>	29
<u>Slika 3: Consensus</u>	30
<u>Slika 4: SciSpace</u>	31
<u>Slika 5: Scite</u>	32

Innovations in information retrieval: the role of generative artificial intelligence

Abstract

The subject of the work is the analysis of innovations in information search with a special focus on the role of generative artificial intelligence. The aim is to investigate how generative artificial intelligence affects traditional search systems and to identify the key advantages and challenges it brings. The motivation for studying this subject lies in the growing complexity and amount of information available in the digital age, which creates a need for more efficient and accurate search systems. Generative artificial intelligence offers new possibilities for personalized search, content summarization and solving the problem of information overload, making it a key technology in the information age. The methodology of the work is based on the analysis and presentation of relevant literature and comparisons between traditional search systems and new models that use generative artificial intelligence. Case studies and technological tools were used to gain insights into contemporary trends. The most important results show that generative artificial intelligence significantly improves the user experience through advanced methods of personalization, automation and data analysis, while challenges remain in the ethical segment, security and algorithm bias. The main takeaway is that generative artificial intelligence not only improves information search, but also redefines the way users access information, providing deeper, more accurate results. This opens up space for further innovation, but requires responsible use and development to minimize potential risks.

Keywords: Information retrieval, generative artificial intelligence, machine learning, large language models, information retrieval systems