

KLASA: 007-01/24-03/02  
URBROJ: 2170-137-01-24-21  
Rijeka, 23. siječnja 2024.

Na temelju članka 34. Statuta Sveučilišta u Rijeci (KLASA: 030-01/23-01/05, URBROJ: 2170-137-01-23-2 od 2. listopada 2023. godine) Senat Sveučilišta u Rijeci na svojoj 91. sjednici održanoj dana 23. siječnja 2024. godine donosi sljedeću

## ODLUKU

### I.

Senat donosi Politiku korištenja alata umjetne inteligencije na Sveučilištu u Rijeci koja čini sastavni dio ove Odluke.

### II.

Ova Odluka stupa na snagu danom donošenja.



REKTORICA

prof. dr. sc. Snježana Prijjić-Samaržija

#### DOSTAVITI:

1. Sastavnicama Sveučilišta u Rijeci,
2. Pismohrani.

**UNIVERSITAS STUDIORUM FLUMINENSIS  
SVEUČILIŠTE U RIJECI**

**UNIRI**



**POLITIKA KORIŠTENJA ALATA UMJETNE INTELIGENCIJE NA  
SVEUČILIŠTU U RIJECI**

**Rijeka, siječanj 2024.**

Na temelju članka 34. stavka 1. Statuta Sveučilišta u Rijeci (KLASA: 030-01/23-01/05, URBROJ: 2170-137-01-23-2 od 2. listopada 2023. godine), Senat Sveučilišta u Rijeci na svojoj 91. sjednici održanoj dana 23. siječnja 2024. godine donosi

## POLITIKU KORIŠTENJA ALATA UMJETNE INTELIGENCIJE NA SVEUČILIŠTU U RIJECI

### UVOD

**Umjetna inteligencija (UI)**, kao jedan od pokretača tranzicije u digitalno društvo i gospodarstvo, ubrzano se razvija mijenjajući sva područja života i rada, stvarajući preduvjete za napredak kvalitete života kao i za globalni prosperitet i ekonomski razvoj, ali postavljajući i nemale izazove – što sve u svojim dokumentima evidentira i Vijeće Europe. Riječ je o vrlo prilagodljivoj tehnologiji koja poprima različite oblike te pronalazi nove primjene pa su sveučilišta dužna pratiti razvoj UI, omogućiti njeno odgovorno i primjereno korištenje, ali i osvijestiti kontekst mogućih izazova ove disruptivne i transformativne tehnologije.

Znanost i visoko obrazovanje pozvani su poticati korištenje dostupnih i razvoj novih alata i sustava UI koji će biti kompatibilni s **Ciljevima održivog razvoja Ujedinjenih nacija (UN SDGs)** koji, zajedno s digitalizacijom i otpornim jedinstvenim tržištem, čine jedno od **temeljnih načela Ugovora Europske Unije (EU)** i jedan od prioritarnih ciljeva politika Unije. Ističe se tu posebno nužnost sistematizacije postojećih spoznaja o načinu na koji sustavi UI mogu doprinijeti ostvarivanju ciljeva održivog razvoja, ali i sustavni pristup osiguravanju održivosti samih sustava UI.

Podržavajući napredak znanosti i širenja znanja usvajanjem inovativnih tehnologija i modela njihove odgovorne primjene, **Sveučilište u Rijeci zalaže se za odgovorni, etički, otvoreni, transparentni i inovativni pristup upotrebi alata UI i drugih naprednih digitalnih tehnologija.** U tom je smislu **Vijeće časti Sveučilišta u Rijeci** na svojoj sjednici održanoj 18. travnja 2023. godine izdalo *Priopćenje o odgovornoj uporabi alata umjetne inteligencije* u kojem se ističe potreba za konstruktivnom raspravom o odgovornoj, etičkoj i transparentnoj upotrebi alata UI s osobljem i studentima Sveučilišta te pozdravljaju svi institucijski i individualni naponi na odgovornoj upotrebi alata UI u učenju i poučavanju, uključujući njihovu dopuštenu, kritičku i informiranu upotrebu uz poštovanje najviših etičkih načela. Nastavno na tu izjavu, **Sveučilište ovim policy dokumentom, usvojenim na Senatu Sveučilišta, detaljnije razrađuje osnovne pojmove, definira ciljeve korištenja alata UI u nastavi i istraživanju, navodi temeljna načela njihovog korištenja na Sveučilištu, definira aktivnosti Sveučilišta na ovome području te dionike uključene u te aktivnosti.** Time Sveučilište iskazuje svoju odlučnost da maksimalno iskoristi prednosti i potencijale novih tehnologija temeljenih na UI, istovremeno uzimajući u obzir najviše etičke standarde kroz promicanje kontinuiranog obrazovanja i razvijanja kritičkog prosuđivanja svih uključenih dionika.

Imajući u vidu ubrzani rast i razvoj i tehnološku evoluciju UI i na UI temeljenih digitalnih i drugih tehnologija, Sveučilište u Rijeci će svoju Politiku korištenja alata umjetne inteligencije redovito, a najmanje jednom godišnje, revidirati i ažurirati.

## CILJEVI KORIŠTENJA ALATA UI NA SVEUČILIŠTU U RIJECI

**Alati temeljni na UI** (cf. popis osnovnih pojmova na kraju ovoga dokumenta) su računalni programi ili sustavi koji koriste tehnike i algoritme iz područja UI kako bi izvršavali zadatke koji obično zahtijevaju ljudsku inteligenciju. Primjeri alata UI uključuju modele strojnog učenja, *chatbotove*, analitičke alate za podatke velikog opsega, softvere za prepoznavanje slika i govora, sustave za upravljanje podacima, i mnoge druge. Alati UI imaju veliku ulogu u automatizaciji procesa, poboljšanju produktivnosti i pružanju rješenja za složene zadatke koji zahtijevaju analizu i zaključivanje temeljeno na velikim količinama podataka. Ti alati mogu obavljati različite zadatke, uključujući prepoznavanje uzoraka, obradu prirodnog jezika, planiranje, donošenje odluka, izradu preporuka, analizu podataka i još mnogo toga te se stoga mogu koristiti u sektorima kao što su zdravstvo, financije, proizvodnja, prijevoz, marketing, obrazovanje i mnogi drugi. Ovisno o njihovoj složenosti i svrsi, alati UI mogu biti implementirani kao samostalni programi ili mogu biti integrirani u veće aplikacije ili sustave.

Uz nadzor i stručno vođenje obučениh istraživača i nastavnika, alati UI se, uz odgovorno i etičko korištenje, danas tako sve opsežnije koriste, te potpomažu i ubrzavaju procese i postupke, i u obrazovnim te u procesima istraživanja i inovacija.

### U nastavi se alati UI već šire koriste posebno pri:

- koncipiranju i razvoju strukture i sadržaja kurikuluma pojedinih kolegija ili cijelih studijskih programa;
- pripremi ispitnih pitanja;
- korištenju *chatbotova* i drugih alata generativne UI kao virtualnih asistenata u nastavi (primjerice za personaliziranu pomoć, odgovaranje na pitanja ili pri identifikaciji resursa i materijala za učenje), uz razvoj odgovarajućih digitalnih blizanaca;
- pisanje studenskih radova, pri čemu se mora inzistirati da studenti jasno navedu i citiraju alate UI koji su korišteni i za koje dijelove radova su korišteni, te da se tako dobiveni dijelovi koriste posebno za razvoj kritičkog mišljenja i kritičkog propitivanja pouzdanosti informacija i rezultata generiranih sustavima i alatima UI.

### Alati UI su ključna omogućavajuća tehnologija koja se u istraživanju i inovacijama ekstenzivno koristi kao pomoć za:

- „rudarenje“, analizu, sažimanje i sintetiziranje informacija i velikih količina podataka iz literature ili kod njihovog prevođenja;
- *brainstorming* ideja, prepoznavanju uzoraka i korelacija te predlaganje primjerenih metodologija rada;
- pisanje i uređivanje (stil, gramatika, lektoriranje) znanstvenih radova i projektnih prijava, posebno istraživačima kojima, primjerice, engleski nije materinji jezik;
- recenziranje radova i projekata;
- pisanje programskih kôdova, generiranje skupova podataka, slika, vizualizacija, simulacija, 3D modela i sl.;
- optimizaciju i automatizaciju eksperimentalnih postavki;
- optimizaciju elemenata kompleksnih sustava ili automatiziranu klasifikaciju (npr. umjetničkih kompozicija);
- predlaganje medicinskih dijagnoza, dizajniranje proteina, kodiranje genskih sekvencija, traženje rješenja za otpornost na antibiotike i mnoge druge primjene u biomedicini i zdravstvu;
- unapređenje vremenskih prognoza te razvoj inovativnih materijala;

- olakšavanje i ubrzanje administrativnih poslova;
- komunikaciju i promociju znanosti.

### **Prednosti upotrebe alata UI u nastavi i istraživanju su pritom:**

- Povećanje učinkovitosti, produktivnosti i primjene inovativnih metoda u znanstvenom istraživanju i nastavi: alati UI mogu olakšati rutinske zadatke poput pretraživanja relevantnih izvora informacija iz zadanog područja te pronalaženje radova naprednim sustavima preporuka i njihove preliminarnu usporedbu i analize, što povećava produktivnost obrazovnih i istraživačkih procesa i omogućuje studentima i nastavnicima da se usmjere na kreativnije i složenije zadatke. Uz to, alati generativne UI mogu pomoći kod predlaganja adekvatnih formulacija tekstova za znanstvene radove ili projektne prijave kao i kod uređivanja predmetnih tekstova. Doista, mnogi znanstveni časopisi često već imaju objavljenu politiku korištenja alata UI gdje odobravaju njihovu primjenu za unapređivanje teksta, za izradu i/ili izmjenu slika, grafova i vizualnih elemenata te za upravljanje bibliografskim referencama – sve uz obavezno navođenje aktivnosti i dijela sadržaja radova za koje su korišteni alati UI te uz jasnu odgovornost samih autora za konačni sadržaj, a posebno za znanstvene postavke i zaključke radova. S druge bi strane izdavači ali i institucije koje financiraju istraživanja morali jasno i otvoreno deklarirati kada oni koriste alate UI pri selekciji projekata odnosno radova ili pri njihovom recenziranju.
- Brza analiza velikih količina podataka: UI može analizirati i izvlačiti zaključke iz velikih skupova podataka u različitim domenama i područjima znanosti brže i preciznije od ljudi. Doista, algoritmi strojnog i dubokog učenja omogućuju prepoznavanje uzoraka u velikim količinama podataka, klasifikaciju i grupiranje podataka, predviđanje i slične automatizirane postupke koji mogu ubrzati znanstvene analize i dovesti do novih otkrića i spoznaja. Pritom je važno da se koriste odgovarajući podatkovni skupovi koji neće biti pristrani i dovesti do razvoja pogrešnih modela i zaključaka. Naravno, poželjno je da podatkovni skupovi budu i otvoreno dostupni i strukturirani u skladu s FAIR principima otvorene znanosti (tj. da budu vidljivi, dostupni, interoperabilni i ponovno upotrebivi). Sve to je korisno u mnogim sektorima pa bi trebalo omogućiti nastavnicima i studentima da više vremena posvete kreativnoj interpretaciji nego rutinskoj obradi podataka.
- Personalizacija i preporuke: UI može pružiti ciljane i personalizirane preporuke i usluge temeljene na korisničkim preferencijama i povijesti pretraživanja, što poboljšava kvalitetu odgovora i konkretizira rezultate.
- Rješavanje složenih problema: alati UI mogu rješavati složene probleme u područjima poput medicinske dijagnostike, istraživanja i inženjerstva, potičući interdisciplinarni i timski pristup rješavanju problema.
- Stalna dostupnost: UI *chatbotovi* i virtualni asistenti mogu pružiti usluge i podršku korisnicima tijekom cijelog dana i svakog dana u tjednu, što otvara potencijal za unapređenje brojnih usluga.

### **TEMELJNA NAČELA KORIŠTENJA ALATA UI NA SVEUČILIŠTA U RIJECI**

Uz navedene velike koristi upotrebe alata UI, postoje i nesumnjivi izazovi i rizici njihova korištenja, kao što su:

- oslanjanje na ponavljajuće uzorke bez njihovog razumijevanja te netransparentnost dobivenih informacija i rezultata;

- povećana mogućnost plagiranja te fabriciranja i manipulacije rezultatima znanstvenih radova;
- nemogućnost ponavljanja rezultata (pitanje reproducibilnosti);
- nepostojanje poveznice s izvorom korištenih informacija;
- pristranost (engl. *bias*) te greške („halucinacije“) u podacima i algoritmima;
- mogućnost umnažanja lažnih informacija i krivotvorina s posljedičnom erozijom povjerenja javnosti u sustav znanosti i visokog školstva;
- neriješenost pitanja vezanih uz intelektualno vlasništvo i autorska prava;
- pitanja vezana uz moguće narušavanje privatnosti (engl. *privacy*) i sigurnost podataka;
- mogućnost produljivanja regionalnih i nacionalnih nejednakosti u dostupnosti znanstvene infrastrukture zbog cijene i energetske zahtjevnosti povezanih uz razvoj i korištenje sustava i alata UI, ali i dostupnosti podataka za treniranje modela UI.

U težnji za maksimizacijom koristi od upotrebe alata UI, uz istovremenu minimizaciju potencijalnih rizika njihove primjene, osnovna je postavka da je za korištenje umjetne inteligencije i kvalitetu i pouzdanost rezultata toga korištenja odgovoran čovjek, odnosno nastavnik i student Sveučilišta koji koristi alate UI. To znači da su sami korisnici alata UI odgovorni za nadzor i tumačenje podataka i rezultata dobivenih korištenjem tih alata, autorstvo radova, seminara i drugih nastavnih i znanstvenih publikacija i njihovu recenziju te formulaciju znanstvenih hipoteza i ciljeva istraživanja.

#### Korištenje alata UI na Sveučilištu u Rijeci mora se stoga temeljiti na:

- principima pravednosti, pravičnosti, poštenja i javne odgovornosti (engl. *accountability*), a što implicira i demokratsko sudjelovanje u razvoju i korištenju alata UI uz ispitivanje, debatu i kontrolu te povratne informacije korisnika;
- poštivanju zakona, podzakonskih akata, propisa i standarda koji reguliraju predmetnu materiju, posebno Usklađenih pravila o umjetnoj inteligenciji, odnosno Akta o umjetnoj inteligenciji – *AI Act*) Europske komisije usvojenog u prosincu 2023. godine, te očuvanje temeljnih prava;
- osobnoj odgovornosti, čestitosti, savjesti i posvećenosti istini korisnika – nastavnika i studenata;
- principima etičnog i pouzdanog korištenja fokusiranog na ljude i njihove vrijednosti;
- točnosti te tehničkoj (uključivo kibernetičkoj) sigurnosti i robusnosti, uz izbjegavanje pristranosti pomoću provjere u odnosu na pouzdane uzorke, testiranje primjereno svrsi korištenja, standardizirano certificiranje pouzdanosti alata UI te sljedivost modela i sustava UI;
- upravljivosti te vanjskoj neovisnoj i nepristranoj kontroli uz mogućnost argumentiranog osporavanja rezultata alata UI;
- istraživanju, stručnom i nastavnom radu koji teži izvrsnosti temeljenoj na:
  - sadašnjoj i budućoj sigurnosti i održivosti života, zdravlja i dobrobiti ljudi, okoliša i imovine (uključujući skrb za kulturnu, povijesnu i arheološku baštinu) te društvenim potrebama i interesima;
  - najvišim standardima stručnosti;
  - principima akademskog integriteta te integriteta i dostojanstva pojedinaca – istraživača i/ili predmeta istraživanja i profesije;
  - ljudskoj i znanstvenoj autonomiji, objektivnosti, točnosti, nepristranosti, skromnosti i dostojanstvu;
  - transparentnosti (uključujući transparentnost i sljedivosti izvora financiranja) i otvorenosti

- u odnosu na izvore podataka korištenih za treniranje modela UI te u odnosu na načine kako sustavi UI donose odluke;
- zaštiti osobnih i povjerljivih podataka uz informirani pristanak na istraživanje i sudjelovanje u njemu;
  - zaštiti privatnosti, intimnosti te društvene, kulturne i jezične različitosti, što uključuje (ali nije ograničeno na) zabranu korištenja sustava i alata UI za društveno rangiranje, automatsko profiliranje ponašanja te masovno nadziranje;
  - uključenosti, nediskriminaciji i solidarnosti među ljudima i generacijama, a što uključuje i minimizaciju razlika u dostupnosti resursa i infrastrukture za korištenje alata UI;
  - omogućavanju ponovljivosti i objašnjivosti korištenih postupaka, a što u nekim slučajevima implicira i konzekventnu potrebu utvrđivanja poštenja i moralnosti korištenih algoritama UI;
  - poticanju otvorenosti inovacija i dijeljenja znanja;
  - uspostavi instrumenata podizanja UI pismenosti u društvu;
  - nultoj toleranciji prema svim oblicima prijevare i plagiranja;
  - težnji ka održivom ekonomskom, društvenom te razvoju okoliša (tj. bez sadašnje ili buduće biološke, fizičke ili moralne opasnosti za ljude, životinje, biljke, okoliš i resurse);
  - publiciranju, uređivanju i recenziranju/vrednovanju koji se temelje na:
    - integritetu pri određivanju i navođenju svih autora rada, što implicira potrebu da se svako korištenje alata UI treba jasno navesti, da u svim radovima treba biti jasno koji su dijelovi dobiveni korištenjem UI te na kojim se izvorima oni temelje kao i da se kritički vrednuju tako dobiveni podaci i rezultati;
    - objektivnosti, poštenju, argumentiranosti i povjerljivosti;
  - drugim principima etičnosti koji moraju uključivati i:
    - pitanje stručne (profesionalne) odgovornosti u odnosu prema javnosti te samoj struci, a što uključuje i skrb za daljnji razvoj, unapređenje i diseminaciju;
    - pitanja intelektualnog vlasništva;
    - pitanja legitimnosti rezultata istraživanja;
    - izbjegavanje sukoba interesa;
  - objektivnoj i istinitoj komunikaciji uz osvješćivanje i stvaranje povjerenja javnosti u rezultate korištenja alata UI te uz poticanje sinergije i suradnje svih zainteresiranih dionika, a što uključuje i promociju ekspertize na području UI u javnim i državnim službama;
  - jasnoj, eksplicitnoj i nepobitnoj definiciji odgovornosti, koja ni u kojem slučaju ne može biti prebačena s čovjeka na stroj ili alat UI;
  - međunarodnoj suradnji i koordinaciji, posebno na području normiranja i regulacije UI;
  - razvoju metoda procjene rizika i istinitom javnom iskazivanju tih rizika;
  - po potrebi, principu predostrožnosti (*no safety- no usage*);
  - mjerama opreza koje uključuju:
    - mjere zaštite i dovoljno dugog nadzora utjecaja na zdravlje, javnu sigurnost i okoliš;
    - predviđanje mogućih negativnih posljedica korištenja alata UI te upotrebe mjera za njihovo izbjegavanje;
    - zahtjev da svaka prijava za financiranje istraživanja sadrži i procjenu rizika uz zabranu financiranja istraživanja koja bi mogla dovesti do kršenja općih prava i temeljnih etičkih načela.

## AKTIVNOSTI NA PODRUČJU UI NA SVEUČILIŠTU U RIJECI

Sveučilište u Rijeci će:

- Poticati razvoj ljudskih potencijala na području UI, posebno kroz radionice i druge strukturirane i podržane obrazovne aktivnosti kojima će se razvijati digitalna i informacijska pismenost nastavnika i studenata s ciljem učinkovite, odgovorne i etičke upotrebe alata UI, uz povezani razvoj kritičkog mišljenja i ukazivanje na potrebu pridržavanja gore definiranim načelima korištenja alata UI. Okvir za pismenost na području UI će pritom obuhvatiti osnovne softverske i hardverske alate, tehnike procjene rezultata dobivenih korištenjem alata UI, komunikaciju sa sustavima AI i *chatbotovima*, upotrebu alata UI u suradnji s drugima, stvaranje sadržaja korištenjem alata UI, sigurnost te pravne i etičke aspekte korištenja alata UI kao i specifična znanja za određena područja primjene alata UI.
- Poticati, i kroz model znanstvenika povratnika, privlačenje istraživača s ekspertizom na području korištenja alata UI.
- Poticati razvoj metoda poučavanja i vrednovanja koji koriste UI te pripadajući razvoj inovativnih i interdisciplinarnih kurikuluma i kolegija, uz ukazivanje na potrebno kritičko razmišljanje i razumijevanje.
- Poticati, i kroz podršku aktivnostima Centra za umjetnu inteligenciju i kibernetičku sigurnost, aktivnostima projekta EDIH Adria te podršku održavanju i daljnjem razvoju istraživačke, a posebno digitalne infrastrukture Sveučilišta, aktivnosti istraživanja i inovacija na području UI i povezanim interdisciplinarnim područjima, a posebno projekte i primjenu alata UI u suradnji s dionicima iz javnog i poslovnog sektora, uz posljedični razvoj regionalnog inovacijskog ekosustava.
- Poticati osnivanje *startup* i *spin-off* tvrtki nastavnika i studenata Sveučilišta temeljenih na primjeni UI.
- Pratiti utjecaj UI na razvoj istraživačkih karijera, posebno mladih znanstvenika (doktoranada i poslijedoktoranada) i akademsku čestitost.
- Inzistirati na fleksibilnom pristupu i pronalaženju ravnoteže između upotrebe novih digitalnih tehnologija te odgovorne i etičke upotrebe alata UI. Doista, ti alati se moraju promatrati kao pomoć u istraživanju i nastavi, a nikako kao zamjena za nastavnike i istraživače. Nastavnicima je pritom prepušten izbor žele li u nastavi koristiti alate UI te u kojoj mjeri upotreba alata UI pridonosi postizanju ili poboljšavanju ishoda učenja.
- Na gore navedenim načelima korištenja alata UI, podržavati razvoj i implementaciju etičkog i pravnog okvira kao i normi kvalitete za primjenu UI te poticati otvorenu i transparentnu javnu raspravu o prednostima i izazovima upotrebe UI.
- Na principima otvorene znanosti i otvorenih inovacija, zalagati se za dostupnost velikih (posebno lokalno relevantnih) količina podataka pod ingerencijom državne i javne uprave akademskim istraživačkim skupinama u svrhu razvoja i istraživanja na području UI, te se zalagati za razvoj javne digitalne infrastrukture.
- Usklađeno s temeljnim ljudskim pravima i Ciljevima održivog razvoja Ujedinjenih nacija, aktivno poticati međunarodnu suradnju i globalni uključivi pristup UI i pripadajućoj infrastrukturi, posebno u okviru sveučilišnih mreža (YUFE; YERUN i druge) u kojima je Sveučilište aktivni dionik.
- Osnovati Savjetodavno tijelo koje će pratiti razvoj alata UI i njihovih implikacija na nastavne i istraživačke procese na Sveučilištu te koje će upravi Sveučilišta i Senatu predlagati reviziju i



ažuriranje Politike korištenja alata umjetne inteligencije na Sveučilištu u Rijeci. Savjetodavno će tijelo imenovati rektor Sveučilišta a činit će ga:

- prorektor zadužen za digitalizaciju (predsjednik);
- prorektor zadužen za inovacije (zamjenik predsjednika);
- predstojnik Centra za umjetnu inteligenciju i kibernetičku sigurnost;
- predstavnik Sveučilišne knjižnice stručan na području otvorene znanosti;
- ekspert o pitanjima etičnosti upotrebe UI (npr. predsjedavajući Etičkog savjeta projekta EDIH Adria);
- predstavnik Vijeća časti Sveučilišta;
- dva eksperta za područje UI;
- djelatnik rektorata Sveučilišta zadužen za administrativne podršku aktivnostima Savjetodavnog tijela;
- po potrebi se u članstvo mogu priključiti i predstavnici dionika te eksperti izvan sveučilišne zajednice.

## DIONICI KORIŠTENJA ALATA UI NA SVEUČILIŠTU U RIJECI

Dionici korištenja alata UI u gore navedenim okvirima su uprava Sveučilišta u Rijeci, uprave sastavnica Sveučilišta, nastavnici i istraživači, administrativno i stručno osoblje i studenti Sveučilišta te partneri Sveučilišta u regionalnom inovacijskom ekosustavu kao i međunarodni partneri Sveučilišta. Svi ti dionici u svojem djelokrugu trebaju brinuti za odgovorno, etično, otvoreno i transparentno korištenje alata UI uz striktno identificiranje sadržaja za koje su korišteni alati UI te uz jasnu odgovornost samih korisnika za konačne rezultate tih aktivnosti.

**Uprava Sveučilišta u Rijeci** usmjerava i nadzire aktivnosti na području UI na Sveučilištu definirane ovim *policy* dokumentom u skladu s gore definiranim načelima, skrbi za razvoj tih aktivnosti i načela te, putem opisanog Savjetodavnog tijela, razmatra i Senatu Sveučilišta predlaže ažuriranja i revizije ove Politike.

**Uprave sastavnica Sveučilišta** predlažu svojim fakultetskim odnosno akademskim vijećima donošenje strateških smjernica i politika za korištenje alata UI na svojim institucijama sukladno znanstvenim i umjetničkim područjima u kojima djeluju te usmjeravaju i nadziru operativne aktivnosti vezane uz te dokumente.

**Nastavnici i istraživači Sveučilišta** aktivno sudjeluju u radionicama i drugim obrazovnim aktivnostima kojima grade i unapređuju svoje vještine korištenja alata UI te koriste te alate u nastavnoj i istraživačkoj djelatnosti pridržavajući se načela njihovog korištenja definiranih ovom Politikom i odgovarajućih dokumenata svojih sastavnica.

**Administrativno i stručno osoblje Sveučilišta** aktivno unapređuje svoje vještine korištenja alata UI te koristi same alate sukladno ovoj Politici i strateškim smjernicama svojih sastavnica.

**Studenti Sveučilišta u Rijeci** obrazuju se za korištenje alata UI te ih koriste u svojim studijskim i drugim aktivnostima uz konzekventnu primjenu načela definiranih ovom Politikom i uz uvažavanje preporuka sastavnica Sveučilišta na kojima studiraju.

**Partneri Sveučilišta u regionalnom inovacijskom ekosustavu iz gospodarstva, lokalne ili područne samouprave, civilnog sektora i nevladinih udruga te šire zajednice** surađuju s djelatnicima i studentima Sveučilišta na provedbi aktivnosti koje uključuju korištenje alata UI

doprinosеći, i na principima peterostruke inovacijske uzvojnice (engl. *quintuple innovation helix framework*), pametne specijalizacije te dvosmjernog i cirkularnog transfera znanja, regionalnom, nacionalnom i međunarodnom ekonomskom i društvenom razvoju. Na takvim aktivnostima djelatnici i studenti Sveučilišta aktivno surađuju i s **međunarodnim partnerima Sveučilišta**, posebno u okviru sveučilišnih mreža (YUFE; YERUN i druge) u kojima je Sveučilište aktivni dionik.

## OSNOVNI POJMOVI

### Umjetna inteligencija (UI)

- U referentnim izvorima nema općeprihvaćene i sveobuhvatne definicije pojma umjetne inteligencije. Mogu se pronaći vrlo jednostavne definicije koje o umjetnoj inteligenciji govore kao o „inteligentnim strojevima” ili o „strojevima koji djeluju na način koji se smatra inteligentnim”, do složenijih definicija kako se radi o grani znanosti koja obuhvaća računalne tehnologije koje su nadahnute načinom na koji ljudi upotrebljavaju živčani sustav i tijela da bi osjećali, učili, prosuđivali, i djelovali. Većina se definicija ipak slaže da se radi o strojevima i softveru koji različitim razinama autonomije prilagođavaju i unapređuju izvršenje svojih zadaća temeljeći se na sposobnosti da za zadani skup ciljeva koje odredi čovjek generiraju izlazne rezultate kao što su sadržaj, predviđanja, preporuke ili odluke koji utječu na okolinu s kojom su u fizičkoj ili digitalnoj interakciji. Tu onda spadaju pristupi strojnog učenja (što obuhvaća nadzirano, nenadzirano i podržano učenje pomoću raznih metoda, uključujući duboko učenje), pristupi temeljeni na logici i na znanju (uključujući prezentaciju znanja, induktivno (logičko) programiranje, baze znanja, module za induktivno i deduktivno (simboličko) zaključivanje i stručni sustavi) te statistički pristupi, Bayesijska procjena kao i metode pretraživanja i optimizacije.

### Alati UI (engl. *AI tools*)

- Naziv za sve alate koji koriste tehnologije UI. Trenutno su tu u fokusu alati koji se temelje na generativnoj umjetnoj inteligenciji te omogućuju kreiranje novog sadržaja kao što su, primjerice, tekstovi u zadanom stilu, programski kôd, vizualizacije i realistične slike, umjetničke slike, glazba, video, 3D modeli, simulacije i sl.

### Duboko učenje (engl. *deep learning*)

- Vrsta strojnog učenja koja se temelji na primjeni umjetnih neuronskih mreža s više slojeva. Algoritmi dubokog učenja omogućuju učenje i pronalaženje značajki velikih skupova podataka, odnosno prepoznavanje obrazaca i uzoraka koji se tu javljaju. Primjenjuje se u različitim zadacima kao što su npr. preporučivanje sadržaja, dijagnostika na temelju medicinskih slika, analiza genetskih podataka, identifikacija rizika za određene bolesti, razvoj novih materijala i dr.

### Chatbot

- Računalni program, odnosno aplikacija koja može komunicirati s korisnicima putem tekstualnog razgovora. Koristi se za odgovaranje na pitanja i pružanje informacija u različitim domenama. Pored sustava koji su razvijani za specifične, uže domene, kao što su, primjerice, davanje informacija o proizvodima ili uslugama, razvijaju se i *chatbotovi* kao što su ChatGPT i Google Bard koji imaju sposobnost razgovora o bilo kojoj temi.

### **Generativna UI (engl. *generative AI*)**

- Grana UI koja, kao odgovor na tekstualni upit putem pogodnog sučelja, automatizirano stvara sadržaj kao što su tekst, slike, video, programski kôd, glazba i dr. Generativna umjetna inteligencija koristi algoritme dubokog učenja i neuronske mreže kako bi, na temelju postojećih podatkovnih skupova, mogla kreirati nove sadržaje. Generativni modeli mogu stvarati raznovrsne i kreativne sadržaje kao što su npr. tekstovi različitih stilova, kvalitetne slike i fotografije na zadanu temu, realistične simulacije ili glazbu koja zvuči kao da ju je skladao čovjek.

### **Objašnjivost (engl. *explainability*)**

- Kod UI se objašnjivost definira kao sposobnost razumijevanja i tumačenja odluka, ponašanja i rezultata koje generiraju sustavi UI, kao što su modeli strojnog ili dubokog učenja. Objašnjivost je ključna u kontekstu sustava UI jer omogućuje korisnicima da shvate kako i zašto određeni model donosi odluke. Objašnjivost je pojam sličan interpretabilnosti (tumačenje), ali razlika je u tome što se interpretabilnost odnosi na razumijevanje unutarnjeg djelovanja modela UI, dok objašnjivost uključuje pružanje ljudski razumljivih objašnjenja za odluke koje model donosi. Interpretabilan model omogućuje razumijevanje kako model dolazi do svojih predviđanja, dok objašnjiv model pruža objašnjenja zašto je model donio određenu odluku. Objašnjivost je važan aspekt etičkog razvoja i primjene UI, posebno u situacijama gdje sustavi UI donose odluke važne za ljude.

### **Opća UI (engl. *artificial general intelligence, AGI*)**

- Razina UI koja je sposobna za razumijevanje i primjenu znanja u širokom spektru različitih zadataka. Opća UI ima sposobnost djelovanja u bilo kojoj domeni i može se nositi s različitim zadacima bez potrebe za posebnim prilagodbama ili programiranjem za svaki novi zadatak. Takvi sustavi su sposobni razumjeti, učiti, primjenjivati znanje i donositi odluke na način koji je sličan ljudskoj inteligenciji.

### **Podaci velikog opsega (engl. *big data*)**

- Veliki heterogeni skupovi podataka iz različitih izvora koji se brzo mijenjaju i rastu. Analize podataka velikog obujma presložena je da bi je mogli provesti ljudi bez pomoći računalnih programa. Iz toga se razloga koriste algoritmi strojnog učenja i duboke neuronske mreže, što omogućuje pronalaženje obrazaca, uzoraka i relacija u podacima, te izvođenje novog znanja.

### **Ponovljivost/reproducibilnost (engl. *reproducibility*)**

- Kod UI se pojam odnosi na mogućnost drugih istraživača da reproduciraju i potvrde rezultate istraživanja i eksperimente koji su provedeni korištenjem UI. U kontekstu strojnog učenja je tako važno da se može ponoviti eksperiment te da drugi istraživači mogu koristiti iste podatke, modele i metode kako bi provjerili rezultate. To je ključno kako bi se osigurala valjanost i pouzdanost rezultata te kako bi se stvorila transparentnost i povjerenje u istraživanja temeljena na upotrebi UI.

### **Priistranost (engl. *bias*)**

- Priistranost je sklonost predrasudi u korist ili protiv osobe, stvari ili stava. Iako se očekuje da su sustavi temeljeni na UI logični i ispravni, i oni mogu biti pristrani. Naime, algoritmi strojnog učenja odražavaju priistranost skupova podataka na kojima se treniraju pa mogu usvojiti i iskazivati pristrana rješenja i stavove koji često narušavaju etiku. Primjerice, jezični modeli mogu generirati tekstove koji sadrže predrasude, stereotipe ili diskriminatorne izraze, što je posljedica učenja iz internetskog sadržaja koji može sadržavati takve elemente. Sustavi za

strojno prevođenje mogu, pak, prenositi predrasude i pristranost iz jednog jezika na drugi, što može rezultirati neprikladnim ili uvredljivim prijevodima.

### **Strojno učenje (engl. *machine learning, ML*)**

- Grana umjetne inteligencije koja se bavi razvojem algoritama za učenje na temelju zadanih skupova podataka. Algoritmi strojnog učenja analiziraju podatke na temelju određenih svojstava tih podataka (skupa značajki), te rješavaju različite zadatke bez da su za to izričito programirani. Osnovna ideja strojnog učenja je da računalni sustavi mogu automatski prepoznavati obrasce u podacima i koristiti te obrasce za donošenje odluka i rješavanje problema.

### **Sustavi temeljeni na UI (engl. *AI-based systems*)**

- Softverski ili hardverski sustavi koji se zasnivaju na tehnologijama i algoritmima iz područja UI. Koriste, primjerice, strojno učenje, duboko učenje, metode za obradu prirodnog jezika i slične pristupe analizi velikih količina podataka. Na taj način omogućuju učenje iz podataka, izvođenje zaključaka i generiranje novoga znanja te donošenje odluka. Sustavi temeljeni na UI primjenjuju se u različitim domenama ljudskog djelovanja kao što su medicina, inženjerstvo, financije, logistika, obrazovanje, sigurnost, istraživanje svemira i dr. Za takve sustave koriste se još i nazivi sustav UI ili inteligentni sustavi.

### **Treniranje modela (engl. *model training*)**

- Termin koji se koristi u području UI i strojnog učenja za procese u kojima se algoritmu ili modelu omogućuje učenje iz podataka kako bi mogao obavljati određene zadatke ili donositi odluke. Tijekom procesa treniranja model uči kako prepoznati obrasce u podacima te prilagođava svoje parametre s ciljem postizanja najboljih radnih svojstava.

### **Uska UI (engl. *artificial narrow intelligence, ANI; narrow AI; specialized AI*)**

- Izraz se odnosi na vrstu UI koja je specijalizirana za obavljanje specifičnih domenskih zadataka, kao što su prepoznavanje lica, prevođenje jezika, igranje šaha ili analiza podataka. Uska UI je specijalizirana i konkretno usmjerena na rješavanje jednog ili nekoliko srodnih problema te nema sposobnost općeg razumijevanja ili primjene znanja izvan tih specifičnih zadataka. Sustavi uske UI mogu izvršavati zadatke na temelju unaprijed programiranih pravila ili na temelju velikih količina podataka. Međutim, takvim sustavima nedostaje razumijevanje dubljeg konteksta i svrhe tih zadataka.

### **Usklađivanje UI (engl. *AI alignment*)**

- Fokusira se na osiguravanje da se sustavi UI strukturiraju, razvijaju i treniraju na način koji je u skladu s ciljevima, preferencijama ili etičkim načelima ljudi. Ključni cilj usklađivanja UI je osigurati da sustavi umjetne inteligencije razumiju i djeluju u skladu s onim što ljudi žele ili smatraju etičnim. To se primarno odnosi na usklađivanje sustava UI s ljudskim vrijednostima, uključujući etičke principe te uključuje razvoj sustava UI koji donose odluke i poduzimaju akcije u skladu s moralnim i etičkim smjernicama koje je društvo uspostavilo. Usklađivanje AI ključno je za izgradnju sustava UI koji su ne samo tehnološki napredni, već i sigurni, pouzdani i korisni za društvo.

### **UI i edukacija (engl. *artificial intelligence and education, AI&ED*)**

- Obuhvaća različite poveznice između UI i obrazovanja koje uključuju koncepte kao što su „učenje uz podršku UI“, „učenje o UI“ te „pripremu za UI“. Koncept učenja uz podršku umjetne inteligencije (*“learning with AI”*) naziva se još i umjetna inteligencija za edukaciju (*“artificial intelligence for education”*).

### UI u edukaciji (engl. *artificial intelligence in education, AIED*)

- Akademsko područje istraživanja, začeto u 1980-im godinama, koje prvenstveno istražuje alate UI koji omogućuju podršku učenju (*learning with AI*).

### Veliki jezični modeli (engl. *large language models, LLM*)

- Modeli temeljeni na UI koji imaju mogućnost prezentacije i generiranja novog teksta. Za razvoj velikih jezičnih modela primjenjuju se duboke neuronske mreže koje se treniraju na ogromnim količinama tekstova, pri čemu se koriste deseci milijuna ili čak stotine milijuna parametara. Na temelju velikih količina tekstualnog sadržaja modeli mogu naučiti kako generirati vrlo kompleksne i raznolike tekstove. Također, takvi su se modeli pokazali iznimno dobrima pri rješavanju zadataka iz područja obrade prirodnog jezika, kao što su strojno prevođenje, sažimanje teksta, analiza osjećaja, klasifikacija teksta, prepoznavanje entiteta u tekstu i sl. Primjeri takvih modela jesu GPT-4 ili BERT.

### Virtualni asistent (engl. *virtual assistant*)

- Računalni program, odnosno aplikacija koja se uglavnom temelji na UI, a koja korisnicima pruža podršku u obavljanju različitih zadataka uključujući upite pretraživanja, postavljanje podsjetnika, rezervacije, praćenje informacija i sl. Mogu pristupiti i pretraživati ogromne količine podataka kako bi korisnicima pružili informacije i rješenja. Često pružaju personalizirane informacije i preporuke temeljene na povijesti korisničkih interakcija. Slični su *chatbotovima*, ali često imaju širu funkcionalnost te mogu voditi višesmjernu komunikaciju s korisnikom. Virtualni asistenti su sposobni za komunikaciju s korisnicima putem različitih sučelja i komunikacijskih kanala, uključujući tekstualne, govorne i vizualne interakcije.

## POPIS KORIŠTENIH BIBLIOGRAFSKIH VRELA

Bockting, Claudi L.; van Dis, Eva A. M.; van Rooij, Robert; Zuidema, Willem and Bollen, Johan (2023). Living guidelines for generative AI — why scientists must oversee its use. *Nature*, 622(7984), 693-696. doi: 10.1038/d41586-023-03266-1

Brown, Tom; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared D. et al. (2020). Language models are few-shot learners. u „Advances in Neural Information Processing Systems“, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan and H. Lin (ed.), 33, 1877-1901.

Conroy, Gemma (2023). How ChatGPT and other AI tools could disrupt scientific publishing. *Nature*, 622(7982), 234-236. doi: 10.1038/d41586-023-03144-w

Council of Europe on Artificial Intelligence, <https://www.coe.int/en/web/artificial-intelligence>, a posebno <https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6> s povezanim dokumentima i odlukama Vijeća, pristupljeno 5. siječnja 2024.

Devlin, Jacob; Ming-Wei, Chang; Kenton, Lee and Kristina Toutanova (2018). Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805.

Elsevier: „The use of generative AI and AI-assisted technologies in writing for Elsevier“ - <https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier>, pristupljeno 26. prosinca 2023.

European Commission – DG for Research and Innovation (2023). AI in Science: Harnessing the power of AI to accelerate discovery and foster innovation – Policy Brief. European Union.

doi:10.2777/401605

European Commission High-level expert group on artificial intelligence (2019). Ethics guidelines for trustworthy AI.

European Commission (2021). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM(2021) 206 final, 2021/0106 (COD).

European Parliament - Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies (2019). State of the art and future of artificial intelligence – Brief.

European Research Council (2023). Foresight: Use and impact of Artificial Intelligence in the scientific process. ERC Executive Agency. doi: 10.2828/10694

European Research Council (2023). Current position of the ERC Scientific Council on Artificial Intelligence.

European University Association (2023). Artificial intelligence tools and their responsible use in higher education learning and teaching.

Glenster, Ann Kristin and Gilbert, Sam (2023). Policy brief: Generative AI. Bennett Institute for Public Policy.

Hanson, Brooks; Stall, Shelley; Cutcher-Gershenfeld, Joel; Vrouwenvelder, Kristina; Wirz, Christopher; Rao, Yuhan and Peng, Ge (2023). Garbage in, garbage out: mitigating risks and maximizing benefits of ai in research. Nature, 623(7985), 28-31. doi: 10.1038/d41586-023-03316-8

Holmes, Wayne; Persson, Jen; Chounta, Irene-Angelica; Wasson, Barbara and Dimitrova, Vania (2022). Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law. Council of Europe.

Meštrović, Ana (2023). Umjetna inteligencija u otvorenoj znanosti – novi izazovi pred istraživačima. predavanje na Open Science Café Sveučilišne knjižnice Sveučilišta u Rijeci na temu „AI u kontekstu otvorene znanosti“, 8. prosinca 2023. – dostupno na <https://urn.nsk.hr/urn:nbn:hr:195:319846>, pristupljeno 24. prosinca 2023.

Ministarstvo gospodarstva, poduzetništva i obrta (2019). Nacrt Nacionalnog plana za razvoj umjetne inteligencije.

Montréal Declaration for a Responsible Development of Artificial Intelligence (2018). dostupno na <https://montrealdeclaration-responsibleai.com/the-declaration/>, pristupljeno 24. prosinca 2023.

OECD (2023). Recommendation of the Council on Artificial Intelligence.

OpenAI (2023). GPT4 technical report.

Parrilla, Juan Manuel (2023). ChatGPT use shows that the grant-application system is broken. Nature, 623(7986), 443. doi: 10.1038/d41586-023-03238-5

Partnership on AI - <https://partnershiponai.org/>, pristupljeno 25. prosinca 2023.

Priopćenje Vijeća časti Sveučilišta u Rijeci o odgovornoj uporabi alata umjetne inteligencije (2023).

Russell, Stuart J. and Norvig, Peter (2020). Artificial Intelligence: A Modern Approach – 4<sup>th</sup> ed. Pearson.

UK Government – Department for Science, Innovation and Technology (2023). Policy paper: The Bletchley Declaration by Countries Attending the AI Safety Summit.

UNESCO (2023). Guidance for generative AI in education and research.

University of Cyprus Recommendations for Use of Artificial Intelligence in Educational Processes (2023).

USA Whitehouse (2023). FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.

Van Noorden, Richard and Perkel, Jeffrey M. (2023). AI and science: What 1,600 researchers think. Nature, 621(7980), 672-675. doi: 10.1038/d41586-023-02980-0

World Economic Forum (2023). Generative AI has disrupted education. Here's how it can be used for good – UNESCO.

World Economic Forum (2023). The Presidio Recommendations on Responsible Generative AI.

Zelenika, Saša (2011). Etička načela na području tehničkih (inženjerskih) znanosti, tehnologija te prirodnih znanosti. Sveučilište u Rijeci.



Sveučilište u Rijeci, Hrvatska

KLASA: 030-01/24-01/07

URBROJ: 2170-137-01-24-1

Rijeka, 23. siječnja 2024.