Универзитет у Београду

Факултет организационих наука

Комитет за развој и примену генеративне вештачке интелигенције у акадамеском окружењу

Београд, јун, 2024.

Одлуком број 05-01 бр. 3/156 од 28.12.2023, а на основу члана 58. Статута Факултета организационих наука, формиран је Комитет за развој и примену генеративне вештачке интелигенције у акадамеском окружењу (Комитет) са задатком да изради документ о коришћењу генеративне вештачке интелигенције, који треба да обухвати етичке, правне и техничке аспекте коришћење алата вештачке интелигенције, као и смернице за њихову примену у настави, истраживању и административном раду.

Чланови Комитета подносе Наставно-научном већу ФОН-а предлог

***Општих смерницa***

***за коришћење генеративне вештачке интелигенције***

***на Факултету организационих наука***

**Уводне напомене**

**Опште смернице за коришћење генеративне вештачке интелигенције (ГВИ) на****Факултету организационих наука (Смернице)** су развијене ради пружања оквира за одговорно коришћење ГВИ технологија код студената, наставног и ненаставног особља. Смернице имају за циљ да се унапреде академски интегритет, транспарентност и етичко понашање у истраживању, наставним пословима, стручним пословима и администрацији.

**Кључ успешне имплементације** смерница лежи у континуираном образовању, отвореној комуникацији и сарадњи. Смернице дефинишу оквир за етичко коришћење ГВИ, као и коришћење засновано на правним нормама имајући у виду одговорност према свим учесницима у академском и ширем друштвеном окружењу.

**На овај начин, Факултет организационих наука (ФОН) се залаже** да се технолошке иновације користе у складу са високим етичким стандардима и принципима, доприносећи тако побољшању образовног процес и истраживачког рада уз поштовање академског интегритета.

**Опште одредбе:**

**1.**      **Регулаторни оквир**

**Основ** за доношење "Смернице за коришћење генеративне вештачке интелигенције на Универзитету у Београду - Факултету организационих наука" утемељен је на правним, стратешким и етичким оквирима (регулаторни оквир):

*Правни оквир:* У Србији, као земљи која тежи чланству у Европској унији, постоји снажан напор да се законодавство усклади са ЕУ законима, посебно у областима које обухватају дигитализацију, информационе технологије, заштиту података о личности и вештачку интелигенцију. У развоју и примени LLM модела, као и будућих технолошких решења заснованих на вештачкој интелигенцији, имаћемо у виду ЕУ правни оквир и настојаћемо да га се придржавамо у мери у којој је то могуће и сврсисходно, узимајући у обзир специфичности опен соурце модела и даља дешавања у домену регулативе

Овај процес усклађивања укључује прилагођавање националних закона, као и развој нових подзаконских аката који подржавају примену тих закона.

Прави оквир који је узет у обзир приликом израде Смерница се састоји од:

о   Закона о високом образовању („Сл. гласник РС“ бр. 88/17, 73/18, 27/18, 67/19, 6/20, 11/21, 67/21, 67/21 и 76/23);

о   Закона о науци и истраживањима(„Сл. гласник РС“ бр. 49/19);

о   Закона о ауторском и сродним правима („Сл. гласник РС“ бр. 104/09, 99/11, 119/12, 29/16 и 66/19);

о   Закона о заштити података о личности („Сл. гласник РС“ бр. 87/18);

о   и други релевантни закони.

***Стратешки оквир***

Стратешки оквир за развој Смерница се ослања на неколико кључних стратегија и иницијатива које је Влада Републике Србије усвојила. Ове стратегије пружају смернице и циљеве за ширу примену информационих технологија, дигитализацију, иновације и развој вештачке интелигенције. Три стратегије које су релевантне за имплементацију Смерница су:

о   Стратегија развоја информационог Стратегија развоја информационог друштва и информационе безбедности у Републици Србији за период од 2021. до 2026. године

о   Стратегија развоја вештачке интелигенције за период 2020–2025

о   Национална платформа за вештачку интелигенцију

*Етички оквир*

Етички кодекси и стандарди који се примењују на академско особље и студенте Универзитета у Београду представљају етички оквир за смернице о одговорном коришћењу ГВИ, укључујући поштовање академског интегритета и транспарентности (Табела 1).

 **Националне етичке смернице**

о   [Етичке смернице за развој, примену и употребу поуздане и одговорне вештачке интелигенције](https://www.ai.gov.rs/tekst/sr/586/eticke-smernice.php)

о   [Упитник за проверу испуњености услова из Етичких смерница](https://www.ai.gov.rs/tekst/sr/706/upitnik-za-proveru-ispunjenosti-uslova-iz-etickih-smernica.php)

**Општи акти и етички кодекси Универзитета и ФОН-а**

о   **Статут и општи акти**Универзитета у Београду, омогућавају доношење специфичних политика и смерница за употребу дигиталних технологија и ресурса. Статут Универзитета у Београду је основни правни акт Универзитета који дефинише организацију, надлежности и функционисање универзитета. У Статуту се налазе одредбе о академском интегритету, правима и обавезама студената, поступцима у случају кршења правила и осталим документима садржаним у Табели 1.

о   Статут ФОН-а

о   Правилници и процедуре који детаљно описују поступке, обавезе и права студената и наставника

**Глобални ресурси за етику ВИ:** Међународна тела и организације као што су ЕУ, УНЕСЦО и ОЕЦД развила су оквире за етички развој АИ. Ови оквири се баве питањима као што су поштење, одговорност и транспарентност.

о   Етичке смернице за вештачку интелигенцију од поверења (ЕУ) (<https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>)

о   Смернице за особље о коришћењу онлајн доступних генеративних модела вештачке интелигенције (ЕУ)  (<https://www.politico.eu/wp-content/uploads/2023/06/01/COM-AI-GUIDELINES.pdf>)

о   Живе смернице за  одговорну употребу генеративенвештачке интелигенције у истраживању (ЕУ)  (<https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en>)

о   Препорука УНЕСЦО о етици вештачке интелигенције ([https://unesdoc.unesco.org/ark:/48223/pf0000380455](https://unesdoc.unesco.org/ark%3A/48223/pf0000380455))

о   ОЕЦД принципи АИ (<https://oecd.ai/en/ai-principles>)

Комбиновањем постојећих прописа Универзитета у Београду са ширим етичким оквирима, развијене су  Смернице које промовишу одговорно коришћење ГВИ алата на ФОН-у.

У регулаторном смислу, Општа смерница се у највећој мери ослања на документа релевантна за академски интегритет и етичке смернице на Универзитету у Београду, националне смернице за вештачку интелигенцију и опште акте  ФОН-а (Табела 1).

Табела 1. Листа документа релевантних за академски интегритет и етичке смернице на високошколским установама у Србији

| **Документ** | **Година** |
| --- | --- |
| Основе за Кодекс о академском интегритету на високошколским установама у Републици Србији | 2016 |
| Кодекс професионалне етике Универзитета у Београду | 2016, 2021 |
| Кодекс понашања у научноистраживачком раду Универзитета у Београду | 2016 |
| Правилник о поступку утврђивања неакадемског понашања у изради писаних радова | 2024 |
| Правилник о поступку провере оригиналности докторских дисертација које се бране на Универзитету у Београду | Још увек предлог |
| Правилник о полагању испита и оцењивању на испиту | 2007, 2012, 2014 |
| Етичке смернице за развој, примену и употребу поуздане и одговорне вештачке интелигенције. Национална платформа за вештачку интелигенцију | 2024 |
| Кодекс понашања у научноистраживачком раду | 2018 |
| Правилници ФОН-а | - |

**2.**     **Основне дефиниције**

* **Генеративна вештачка интелигенција** (ГВИ) је технологија вештачке интелигенције (ВИ) која аутоматски генерише садржај на основу упита (промптова) написаних у облику разговора на природном језику (Holmes& Миао, 2023).
* **Кодекс професионалне етике Универзитета у Београду**(Кодекс) донет је у циљу очувања достојанства професије, унапређивања моралних вредности, заштите вредности знања и подизања свести о одговорности свих чланова универзитетске заједнице на Универзитету у Београду (Универзитет). Кодекс се примењује у професионалном и јавном деловању свих чланова универзитетске заједнице: наставног и научног особља, студената, ненаставног особља, као и лица која нису чланови академске заједнице Универзитета а учествују у раду Универзитета (Кодекс професионалне етике Универзитета у Београду, 2016).

о    **Академски интегритет**дефинисан Законом о високом образовању представља један од принципа високог образовања. Према овом закону, на академском интегритету заснивају се: очување достојанства професије, унапређивање моралних вредности, заштита вредности знања и подизање свести о одговорности свих чланова академске заједнице и унапређење људских права и слобода (Закон о високом образовању, 2017 до 2023). Основе за Кодекс о академском интегритету на високошколским установама у Републици Србији, (24. октобар 2016), донео је Национални савет за високо образовање Републике Србије.

о    **Кодекс понашања у научноистраживачком раду** (2018) дефинише „најтеже облике превара у науци као: плагирање, лажно ауторство, измишљање и кривотворење резултата и аутоплагирање, што се односи на сва објављена научна дела (докторска дисертација, рад у часопису, монографија, поглавља у монографијама и зборницима и сл.) и академске радове (додали аутори Смерница) (Кодекс понашања у научноистраживачком раду, 2018):

о    **Плагирање** је представљање туђих идеја или туђег рада, у целини или деловима, без навођења изворног ауторства или изворника, односно противзаконито присвајање туђих интелектуалних творевина и научних резултата и њихово приказивање као својих, као и:

-          дословно преузимање текста другог аутора, односно копирање из електронских или штампаних извора, са српског или страног језика, у деловима или целости, без навођења имена аутора и извора из којег је текст преузет, као и без јасног обележавања преузетог дела;

-          препричавање или сажимање текста другог аутора из електронских или штампаних извора, са српског или страног језика, у деловима или целости, без одговарајућег навођења имена аутора и извора из којег је текст преузет, као и без јасног обележавања препричаног дела;

-          представљање идеја других аутора као својих, без одговарајућег навођења имена аутора, односно извора из кога је текст преузет.

о   **Лажно ауторство** је навођење као аутора лица које није учествовало у изради научног дела.

о   **Измишљање и кривотворење резултата** је деловање којим се манипулише предметом, опремом или процесом истраживања, са циљем да резултати научног истраживања буду намерно подешени или тенденциозно протумачени

о   **Аутоплагирање** је поновно објављивање свог раније објављеног научног дела или за другу сврху искоришћеног дела као новог и оригиналног”. (Кодекс понашања у научноистраживачком раду, 2018).

 **Парафразирање** се у светлу овог кодекса сматра као плагијат у случају да се примењује као „препричавање или сажимање текста другог аутора из електронских или штампаних извора, са српског или страног језика, у деловима или целости, без одговарајућег навођења имена аутора и извора из којег је текст преузет, као и без јасног обележавања препричаног дела” (Кодекс понашања у научноистраживачком раду, 2018).

* **Академски радови**: есеји, семинарски радови, завршни радови на свим нивоима студија, стручни и научни радови.
* **Образовне активности** су ближе одређене Законом о високом образовању (2017 до 2023)
* **Научно-истраживачке активности** су ближе одређене Законом о науци и истраживањима (2019).
* **Административне активности** студената, наставног i vannastavnog особља обухватају све оне задатке и процедура које су неопходне за правилно функционисање ФОН-а, али које не спадају директно у домен образовних или научно-истраживачких активности.
* **Промт**: Упит (питање) који корисник шаље ГВИ у очекивању одговора на природном језику

3.     **Дозвољена употреба ГВИ**

Факултет организационих наука снажно се залаже да своје студенте и запослене уведе у свет дозвољене и одговорне употребе ГВИ, уз строго поштовање етичких принципа.

Дефинисање дозвољене употребе ГВИ на факултету захтева темељно разматрање етичких, образовних, технолошких и правних аспеката. Дозвољена употреба ГВИ има за циљ да успостави равнотежу између технолошког напретка иновација у образовном процесу са очувањем академског интегритета и заштитом права појединаца. Таквим приступом, ФОН ће максимално искористити потенцијал ГВИ за унапређење процеса образовања и истраживања, уз очување темељних вредности и принципа академске заједнице.

**Делокруг употребе ГВИ:**Алати ГВИ се могу користити у научно-истраживачком раду, наставним и стручним пословима, као и у административним пословима.

Корисници алата ГВИ (корисници): **студенти, наставно и ненаставно особље ФОН-а.**

**Дозвољено коришћење ГВИ**

Да би користили ГВИ у образовним, научно-истраживачким и административним активностима, кључно је доследно поштовање **општих принципа дозвољеног коршићења ГВИ**, који укључују:

* **ГВИ се посматра као алат односно трећа страна**. Иако није реч о човеку, свака помоћ у образовном и научно-истраживачком процесу која долази од ГВИ мора бити експлицитно наведена. Када су у питању административни задаци, овај захтев може бити ублажен, под условом да њихово извршавање не доводи до проблема везаних за етичност и академски интегритет.

* **Транспарентност**у коришћењу ГВИ је од суштинског значаја за интегритет наставно-научног процеса и истраживања, а истовремено је важна за градњу поверења јавности у технологију. Имплементација следећих корака би могла повећати транспарентност и веродостојност коришћења ГВИ:

о    *Објашњење методологије*: Требало би јасно навести како и зашто је ГВИ коришћен у истраживачком процесу, укључујући детаље о томе како су модели тренирани, који су подаци коришћени, и како су резултати интерпретирани. Ово може помоћи у разумевању потенцијалних ограничења и предности примене ГВИ.

о    *Указивање потенцијала за пристрасност*: алати ГВИ производе одговоре на основу улазних захтева и података упућених од стране људи (промптова), које могу садржати друштвене пристрасности и стереотипе, које заузврат могу бити реплициране у одговору алата ГВИ.

о    *Декларација извора података*: Јасно навођење извора података који су коришћени за тренирање модела ГВИ може помоћи у идентификацији потенцијалних пристрасности и нетачности. Такође, то може омогућити другим истраживачима да оцене квалитет и релевантност коришћених података.

о    *Декларација о квалитету добијених резултата*: подаци и информације садржани унутар алата ГВИ добијају се из широког спектра извора, укључујући оне који су слабо референцирани или нетачни. Слично, нејасни налози или информације могу бити погрешно тумачени од стране алата ГВИ и произвести нетачне, небитне или застареле информације. То значи да одговорност за тачност информација генерисаних овим алатима, када се преносе у други контекст, лежи на кориснику.

о    *Правила о коришћењу*: Објављивање смерница или правила о коришћењу резултата генерисаних помоћу ГВИ може помоћи у смањењу злоупотребе и нетачном тумачењу. Такође, то може помоћи корисницима да разумеју како најбоље да користе и интерпретирају информације генерисане од стране АИ.

о    *Ревизија и валидација*: Подстицање ревизије и валидације резултата добијених помоћу ГВИ модела може допринети побољшању прецизности и поузданости. Ревизија од стране стручњака такође може помоћи у идентификовању пристрасности и нетачности у излазима вештачке интелигенције.

о    Транспарентност и одговорност у развоју: Коначно, организације које развијају ГВИ алате требало би да буду транспарентне у вези са својим методама развоја, укључујући детаље о тимовима који стоје иза развоја, њиховим етичким смерницама, и како се баве проблемима пристрасности и тачности.

* **Академски интегритет:** Коришћење ГВИ је дозвољено искључиво уз поштовање базичних принципа академског интегритета дефинисаних у складу са Основама за Кодекс о академском интегритету на високошколским установама у Републици Србији (2016) и Кодексом понашања у научноистраживачком раду (2018).
* **Поштовање ауторских права и интелектуалне својине:** Употреба ГВИ се врши уз поштовање ауторских права, у складу са Законом о ауторском и сродним правима (2009 до 2019).
* **Поштовање приватности и заштите података о личности:**Приликом коришћења ГВИ мора се водити рачуна о правима грађана утврђеним Законом о заштити података о личности (2018).
* **Поштовање етичких принципа:**УпотребаГВИ се вршиу складу са начелима и етичким правилима у наставном и научном раду наведеним у Кодексу професионалне етике Универзитета у Београду (2016):
* **Принцип чувања пословне и службене тајне:**Приликом употребе ГВИ захтева се**уздржавање од радњи**које би могли угрозити пословање Факултета и партнера, повредом дужности поверљивости која проилзлази из законског оквира или уговорних обавеза.

о    Етичка употреба АИ: Промовише етичку употребу технологије, посебно у светлу све веће забринутости о утицају АИ на приватност и сигурност. Јасне смернице штите не само институцију и њене партнере, већ и појединце чији се подаци обрађују.

о    Спречавање злоупотребе: У свету где су подаци нови облик валуте, заштита кључних информација од неауторизованог приступа, крађе или злоупотребе постаје примарни приоритет за сваку организацију.

о    Одржавање поверења: Институције које успешно заштите поверљиве информације стичу и одржавају поверење својих студената, особља, и пословних партнера. Ово је посебно важно у образовном сектору, где су интегритет и репутација од суштинског значаја.

о    Укључивање овог принципа у смернице за употребу ГВИ одражава преданост факултета сигурности, етици, и законским обавезама, постављајући снажне темеље за одговорно коришћење АИ технологија. Ово не само да штити факултет и његове партнере, већ и промовише културу одговорности и транспарентности међу корисницима ГВИ.

* **Аутономија наставног рада научно-истраживачког рада**. Као и до сада, одговорност наставника је да обезбеди испуњење циљева и исхода учења дефинисаних картоном предмета, без обзира да ли се, и у којој мери, користи ГВИ,  уз поштовање моралних и правних норми и начела.

Принцип аутономије наставног и научно-истраживачког рада у вези са дефинисањем нивоа и детаља дозвољене употребе ГВИ је веома важан и користан приступ из више разлога:

о    Специфичност дисциплина: Свака академска дисциплина има своје посебности, које могу захтевати прилагођавање употребе технологија као што је ГВИ. Дајући наставницима могућност да дефинишу дозвољене употребе ГВИ, омогућава се флексибилност која узима у обзир те специфичности.

о    Експертиза наставника: Наставници су експерти у својим областима и најбоље разумеју потребе и могућности својих предмета у контексту ГВИ. Овај приступ користи њихово стручно знање за оптимизацију коришћења ГВИ у настави и истраживањима.

о    Иновативност у настави: Аутономија може подстаћи наставнике да иновативно размишљају о коришћењу ГВИ у својим дисциплинама, што може довести до нових и побољшаних метода поучавања и истраживања.

о    Одговорност и разумевање: Наставници који активно учествују у дефинисању смерница за употребу ГВИ вероватније ће развити дубоко разумевање потенцијалних предности и ризика технологије, чиме се повећава одговорност у њеном коришћењу.

о    Придржавање општих принципа: Овај приступ омогућава прилагођавање уз истовремено одржавање усклађености са општим принципима и стандардима факултета, што гарантује да ће све употребе ГВИ подржавати академски интегритет, етику, транспарентност, доступност, једнакост и инклузивност.

* **Доступност, једнакост и инклузивност:**Осигурати да све технологије генеративне вештачке интелигенције (ГВИ) буду једнако доступне свим корисницима, уз активну промоцију инклузивности и обезбеђивање једнаких образовних шанси за све студенте

**4.**     **Недозвољене употребе ГВИ**

Недозвољене употребе ГВИ представљају активности које се не уклапају у оквире дефинисаних општих принципа дозвољене употребе ГВИ. Ове активности не само да подразумевају кршење општих принципа дозвољеног коришћења ГВИ, већ могу значајно утицати на кредибилитет и интегритет академске и шире заједнице.

У наставку се налази неискључива листа недозвољене употребе ГВИ:

* **Кршење правила академског интегритета**, што укључује:
	+ *Плагијаризам*: коришћење туђих идеја, речи или радова без одговарајућег признавања или цитирања.
	+ *Лажно ауторство*: приписивање ауторства особи која није допринела раду или тврдње о ауторству над радом који није индивидуално креиран.
	+ *Измишљање и кривотворење података и резултата*: стварање неистинитих података или алтерација постојећих података и резултата како би се погрешно представила истраживања.
	+ *Генерисање и ширење лажних садржаја*: употреба ГВИ за креирање и дистрибуцију информација које су обмањујуће или неистините, потенцијално ширећи дезинформације.
* **Кршење етичких норми**, што подразумева активности које нарушавају моралне и етичке кодексе понашања унутар и изван академске заједнице, укључујући, али не ограничавајући се на, дискриминацију, узнемиравање и злоупотребу података.
* **Кршење ауторских права**, што се односи на неовлашћено коришћење заштићеног садржаја без дозволе власника ауторских права, што може укључивати репродукцију, дистрибуцију, извођење или приказивање таквог садржаја.
* **Уношење у промпт било којих података или садржаја који представљају пословну или службену тајну** или на било који начин могу угрозити пословање неког привредног друштва или појединца. Ово укључује, али није ограничено на, осетљиве информације које, ако би биле откривене, могле би нанети штету појединцима, организацијама или националној безбедности.
* **Било које активности које компромитују етичке смернице и академски интегритет**, укључујући поступке који подривају основне вредности поштења, транспарентности, фер плеја и одговорности које су кључне за одржавање повјерења и интегритета у академском и професионалном окружењу, као што су:

о    *Употреба ГВИ за стварање дискриминишућег садржаја*: Коришћење ГВИ алата за креирање садржаја који промовише или подстиче дискриминацију на основу расе, пола, сексуалне оријентације, религије, националности, инвалидитета или било које друге заштићене категорије.

о    *Манипулација оценама или академским списима*: Коришћење ГВИ алата за неовлашћену измену оцена, академских списа или сертификата са циљем побољшања академског статуса или стицања незаслужених академских квалификација.

о    *Стварање и ширење деепфаке садржаја*: Употреба ГВИ за креирање високо реалистичних, али лажних аудиовизуелних материјала који могу бити искоришћени за обману, клевету или узнемиравање појединаца.

о    *Аутоматизовано генерисање злонамерног софтвера или хакерских алата*: Употреба ГВИ у сврху креирања софтвера дизајнираног да наруши безбедност информационих система, украде осетљиве информације, или на други начин наруши интегритет и приватност појединаца и организација.

о    *Неовлашћено прикупљање или анализа података*: Употреба ГВИ алата за прикупљање, обраду или анализу података без сагласности власника тих података, што може укључивати неовлашћено скенирање мрежа или система за прикупљање података.

о    *Генерисање садржаја који подрива академску дискусију или критичко мишљење*: Употреба ГВИ за генерисање радова, есеја или истраживања која су намерно дизајнирана да обману, изазову контроверзе без основе или на други начин наруше интегритет академског дијалога.

* Други случајеви употребе ГВИ који су недозвољени у складу са правним и етичким оквиром.

Дефинисање и јасно разграничење недозвољених употреба ГВИ од суштинског је значаја за одржавање високих стандарда академског и професионалног понашања, као и за заштиту права и добробити свих укључених страна. Овим се осигурава да иновације које ГВИ доноси буду коришћене на етички одговоран начин који промовише интегритет и транспарентност у свим областима примене.

**5.**     **Образовне смернице**

**Учење и вештине**: Охрабривати развој критичког размишљања и дигиталних вештина међу студентима и наставницима кроз коришћење ГВИ.

**Адаптација наставног плана**: Размотрити интеграцију ГВИ у наставне планове и програме како би се студентима пружило практично искуство са овом технологијом.

**6.**     **Имплементација и надзор**

За ефикасну имплементацију смерница, важно је:

* Укључивање заинтересованих страна: Укључите студенте и ненаставно особље ФОН-а у развој и ревизију смерница како бисте осигурали да буду релевантне и практичне.
* Обука у циљу правилног информисања и јачања свести о Смерницама међу студентима и особљем (наставно и vannastavno).
* Успостављање механизма за надзор и ревизију како би се осигурало да се смерница адекватно примењује и ажурира у складу са технолошким развојем и променама у регулаторним оквирима.

*Будуће акције:* Упознавање наставника и студената ФОН-а са садржајем документа релевантних за академски интегритет и етичке смернице на високошколским установама у Србији, укључујући и специфичне правилнике Универзитета у Београду, као и националне смернице за вештачку интелигенцију (Табела 1.).

**7.**     **Казнена политика**

Факултет ће осигурати да се одступања од дозвољење употребе ГВИ, дефинисано овим Смерницама процесуира и санкционише у складу са Правилником о дисциплинској одговорности студената Универзитета у Београду, третирајући недозвољено коришћење ГВИ као тежу повреду дисциплине.