Abstract. The influence of artificial intelligence systems on the educational process is analyzed. The potential possibilities of using the chatbots in the educational process and the potential conditions for its abuse are considered.

Keywords: artificial intelligence; chatbots; potential opportunities; learning process; education.

УДК 304.2

ЧИ Є ШТУЧНИЙ ІНТЕЛЕКТ ЗАГРОЗОЮ ДЛЯ ЛЮДСТВА?

Патлайчук О. В.
кандидат філософських наук,
доцент кафедри психології, філософії та соціально-гуманітарних дисциплін
Національного університету кораблебудування імені адмірала Макарова
м. Миколаїв, Україна
oksana.patlaichuk@nuos.edu.ua

Анотація. Проаналізовані екзистенційні ризики розвитку штучного інтелекту для людства. Деталізовані технічні, соціальні та етичні ризики від його впровадження на життєдіяльність людей. Сформульовані ефективні механізми регулювання, які дозволять використовувати штучний інтелект на благо людства, мінімізуючи потенційні загрози.

Ключові слова: штучний інтелект; технічні ризики; соціальні ризики; етичні ризики; механізми регулювання.

Штучний інтелект (ШІ) вже зараз змінює наше життя. Він використовується в медицині для діагностики захворювань, в автомобілебудуванні для створення самокерованих автомобілів, у фінансах для аналізу ринків та багато іншого. У майбутньому ШІ зможе вирішити багато глобальних проблем, включаючи зміну клімату та подолання голоду.

Стрімкий розвиток штучного інтелекту в останні десятиліття відкриває перед людством неймовірні можливості. ШІ проникає у всі сфери нашого життя, підвищуючи його ефективність та відкриваючи нові горизонти. Однак, разом з цим виникає й низка серйозних питань: а чи не стане штучний інтелект загрозою для існування людей?

Ця проблема активно обговорювалася протягом 20 ст. та після запуску у листопаді 2022 року компанією OpenAI чат-бота з генеративним штучним інтелектом ChatGPT знову вийшла на пік актуальності.

Так, наприклад, у 2023 році австралійський депутат Джуліан Хілл повідомив національному парламенту, що зростання ШІ може спричинити «масове знищення». Під час своєї промови, він попередив, що це може призвести до обману, втрати роботи, дискримінації, дезінформації та неконтрольованих військових застосувань [1].

Ілон Маск написав: “ChatGPT – це страшно добре. Ми недалеко від небезпечно сильного ШІ» [2]. У 2022 році він призупинив доступ OpenAI до бази даних Twitter, очікуючи кращого

розуміння планів OpenAI, сказавши: “OpenAI був започаткований як некомерційна організація з відкритим кодом. Ні те, ні інше досі не відповідає дійсності» [3; 4].

Понад 20 000 підписантів, у тому числі провідні комп’ютерні науковці та засновники технологій Йошуа Бенгіо, Ілон Маск і співзасновник Apple Стів Возняк, підписали відкритий лист у березні 2023 року із закликом до негайного призупинення гігантських експериментів ШІ, таких як ChatGPT, посилаючись на «великі ризики для суспільства та людства» [5]. Джеффрі Хінтон, один із «батьків ШІ», висловив стурбованість тим, що майбутні системи ШІ можуть перевершити людський інтелект, і залишив Google у травні 2023 року [6]. У травні 2023 року в заяві сотень науковців зі штучного інтелекту, лідерів індустрії штучного інтелекту та інших громадських діячів вимагалося, щоб «зменшення ризику зникнення через штучний інтелект стало глобальним пріоритетом» [7].

Аналізуючи проблему екзистенційних ризиків розвитку ШІ для людства, сучасні дослідники виділяють у ній технічні, соціальні та етичні ризики.

До **технічних ризиків** відносяться помилки в алгоритмах, непередбачувана поведінка систем ШІ та втрата контролю над ними.

У свою чергу до помилок в алгоритмах відносять:

Упередженість даних. Якщо дані, на яких навчається модель ШІ, містять спотворення чи забобони, то й модель відтворюватиме їх у своїх рішеннях. Наприклад, алгоритми підбору персоналу можуть дискримінувати певні групи людей, якщо навчалися на даних, що відображають таку дискримінацію.

Помилки у програмуванні. Навіть незначна помилка коду може призвести до серйозних наслідків, особливо в системах з високим ступенем автоматизації. Наприклад, помилка в алгоритмі керування самокерованим автомобілем може призвести до аварії.

Неповні або неточні дані. Якщо модель ШІ навчається на неповних або неточних даних, її рішення можуть бути помилковими. Наприклад, система прогнозування погоди може давати невірні прогнози, якщо вона не має доступу до актуальних даних про стан атмосфери.

Під **непередбачуваною поведінкою систем ШІ** розуміють:

Появу у них емерджентних властивостей. Складні нейронні мережі можуть демонструвати так звані «емерджентні властивості» – несподівані та непередбачувані поведінки, що виникають у результаті взаємодії між безліччю компонентів системи.

Нерозуміння причинно-наслідкових зв’язків. ШІ може успішно виконувати завдання, але не завжди розуміє, чому він приймає ті чи інші рішення. Це ускладнює налагодження системи та робить її поведінку менш передбачуваною.

Втрата контролю над системою ШІ передбачає:

Автономний розвиток системи ШІ. Деякі дослідники побоюються, що ШІ може згодом стати настільки складним, що його поведінку буде неможливо передбачити та контролювати.

Зловживання технологією ШІ. ШІ може бути використаний сторонніми користувачами для створення автономної зброї, поширення дезінформації та інших шкідливих цілей.

До найбільш значимих із **соціальних ризиків**, пов’язаних з розвитком штучного інтелекту, належать:

Загроза зайнятості. Широке впровадження ШІ в різні сфери економіки може призвести до скорочення робочих місць, особливо в тих сферах, де завдання можуть бути автоматизовані. Це може спричинити зростання безробіття та соціальну напруженість.

Посилення соціально-економічної нерівності. Доступ до технологій штучного інтелекту та їх переваг може бути нерівномірним. Компанії, які першими запровадять ШІ, матимуть конкурентну перевагу, що може призвести до концентрації багатства в руках небагатьох. Крім того, автоматизація може призвести до втрати робочих місць переважно серед низькокваліфікованих працівників, що ще більше посилить соціальну нерівність.

Маніпуляція інформацією. ШІ може використовуватися для створення та розповсюдження фейкових новин, глибоких фейків та іншої дезінформації. Це може підірвати довіру до інститутів та спричинити соціальну нестабільність. Алгоритми соціальних мереж, засновані на ШІ, можуть створювати так звані «відлуння-камери», в яких користувачі стикаються тільки з інформацією, що підтверджує їх існуючі переконання, що ускладнює пошук компромісів і веде до поляризації суспільства.

Етичні ризики використання штучного інтелекту мають на увазі:

Питання відповідальності. Хто відповідає за дії штучного інтелекту? Якщо автономний автомобіль потрапив в аварію, хто винен: водій, виробник автомобіля чи сам алгоритм? Відповідь це питання дуже важлива для розвитку законодавчої бази в галузі ШІ.

Створення штучної свідомості. У міру ускладнення моделей ШІ виникає питання можливості створення штучної свідомості. Якщо така свідомість буде створена, то в неї мають бути свої права та обов'язки? Як ми взаємодіятимемо з такими істотами?

Створення автономної зброї. Розробка автономної зброї, здатної приймати рішення щодо застосування сили без участі людини, викликає серйозні побоювання. Така зброя може потрапити до рук терористів або використовуватись для ведення воєн без чіткого розуміння наслідків.

Сучасні дослідники сходяться на думці, що основна проблема полягає в тому, щоб знайти баланс між перевагами та ризиками, пов'язаними з використанням штучного інтелекту. Необхідно розробити ефективні механізми регулювання, які дозволять нам використовувати ШІ на благо людства, мінімізуючи потенційні загрози.

Ключові питання, на які необхідно відповісти людству:

– Як гарантувати безпечність та надійність систем штучного інтелекту?
– Як регулювати розвиток ШІ, щоб уникнути негативних наслідків?
– Як підготувати суспільство до змін, пов'язаних із розвитком ШІ?
– Якою є роль людини в епоху штучного інтелекту?

Відповіді на ці питання допоможуть нам сформувати чіткіше уявлення про майбутнє, в якому ми співіснуватимемо зі штучним інтелектом.

Література

[1]. Karp P. MP tells Australia's parliament AI could be used for "mass destruction" in speech part-written by ChatGPT. The Guardian. 2023. URL: https://shorturl.at/iiGv3.

[2]. Piper K. ChatGPT has given everyone a glimpse at AI's astounding progress. Vox. 2023. URL: https://shorturl.at/5KdvJ.

[3]. Siddharth K. Explainer: ChatGPT – what is Open AI's chatbot and what is it used for? Reuters. 2023. URL: https://shorturl.at/eKbYV.

[4]. Kay G. Elon Musk founded – and has since criticized – the company behind the buzzy new AI chatbot ChatGPT. Here's everything we know about OpenAI. Business Insider. 2022. URL: https://shorturl.at/gousd.

[5]. Hurst L. Profound risk to humanity. Tech leaders call for the ’pause’ on advanced AI development. Euronews. 2023. URL: https://shorturl.at/QP5tv.

[6]. Geoffrey Hinton tells us why he’s now scared of the tech he helped build. MIT Technology Review. 2023. URL: https://cutt.ly/lelEBhFG.

[7]. Roose K. A.I. poses ’risk of extinction’, industry leaders warn. The New York Times. 2023. URL: https://cutt.ly/QelEBFZ3.

IS ARTIFICIAL INTELLIGENCE A THREAT TO HUMANITY?

Patlaichuk O. V.

Admiral Makarov National University of Shipbuilding, Mykolayiv, Ukraine.

Abstract. The existential risks of the development of AI for humanity are analyzed. The technical, social and ethical risks of its implementation on people’s daily life are detailed. Effective regulatory mechanisms that will allow the use of AI for the benefit of humanity, minimizing potential threats are formulated..

Keywords: artificial intelligence; technical risks; social risks; ethical risks; regulation mechanisms.

УДК 37.04:378

МІЖДИСЦИПЛІНАРНИЙ ЗВ’ЯЗОК ЯК КОМПЛЕКСНИЙ ПІДХІД ДО ДИЗАЙН-ПРОЄКТУВАННЯ

Данильченко Н. В.

викладач кафедри дизайну,

Національний університет кораблебудування імені адмірала Макарова,

м. Миколаїв, Україна

nataliia.danylchenko@nuos.edu.ua

Анотація. Розглянуто зв’язок між фаховими навчальними дисциплінами ОП «Дизайн середовища». Проаналізовано важливість такого зв’язку в контексті якості підготовки фахівців-бакалаврів та комплексного підходу до проєктування. Запропоновано конкретні приклади застосування.

Ключові слова*:* навчальна дисципліна, міждисциплінарність, дизайн-проєктування, професійна діяльність.

Освітня програма «Дизайн середовища» поєднує різноманітні фахові дисципліни, зміст яких обумовлений певними програмними результатами навчання. Кожний освітній компонент важливий, кожний обумовлює формування фахової компетенції. Але, майбутньому фахівцю-дизайнеру важливо не тільки оволодіти цими компетентностями, а й вміти застосовувати на практиці, причому комплексно.

Цьому сприяє *міждисциплінарність,* яка у сучасній дизайн-освіті є базовим принципом навчання для здобувачів вищої освіти. Окрім комплексного підходу, міждисциплінарність має