RAQAMLI KONTENT VA SUN’IY INTELLEKT TOMONIDAN YARATILGAN MATERIALLARNING AXLOQIY TA’SIRINI BAHOLASH: NOXOLISLIK, DEZINFORMATSIYA VA QADRIYATLAR XAVFI

Turdaliyev Kamronbek Ilhomjon o‘g‘li
Farg‘ona davlat texnika universiteti
Axborot tizimlari va texnologiyalari kaferdasi assistenti
Isoqjonov Shohjaxon Bahodirjon o‘g‘li
Farg‘ona davlat texnika universitet talabasi
Tolibjonova Shohista Ulug‘bek qizi
Farg‘ona davlat texnika universitet talabasi
https://doi.org/10.5281/zenodo.23510879

Annotatsiya. *Mazkur maqolada raqamli kontent hamda sun’iy intellekt texnologiyalari yordamida yaratilayotgan matn, tasvir, audio va videomateriallarning axloqiy ta’siri, ularning jamiyatdagi axborot almashinuvi, ijtimoiy ishonch va qadriyatlar tizimiga ko‘rsatadigan ta’siri ilmiy-nazariy jihatdan tahlil qilinadi. Sun’iy intellekt tizimlarida mavjud ma’lumotlar va algoritmik qarorlar bilan bog‘liq noxolislik, sintetik kontent orqali dezinformatsiyaning tarqalishi, shaxsiy identifikatsiyani soxtalashtirish hamda madaniy va ma’naviy qadriyatlarning buzilishi asosiy tadqiqot yo‘nalishlari sifatida o‘rganiladi.*

Kalit so‘zlar: *sun’iy intellekt, raqamli kontent, generativ sun’iy intellekt, axloqiy xavf, algoritmik noxolislik, dezinformatsiya, deepfake, axborot ishonchliligi, madaniy qadriyatlar, raqamli savodxonlik.*

Kirish. Raqamli texnologiyalarning jadal rivojlanishi natijasida axborotni yaratish, qayta ishlash va tarqatish mexanizmlari tubdan o‘zgarmoqda. Xususan, generativ sun’iy intellekt tizimlarining matn, tasvir, ovoz, video va dasturiy kodlarni inson ishtirokisiz yoki minimal ishtirokida ishlab chiqarish imkoniyatlari kengayib, raqamli kontentning hajmi, shakli va undan foydalanish ko‘lamini sezilarli darajada oshirmoqda. Ushbu texnologiyalar ta’lim, ilm-fan, jurnalistika, davlat boshqaruvi, biznes va ijodiy faoliyat uchun yangi imkoniyatlarni yaratishi bilan birga, axborotning haqqoniyligi, mualliflik mas’uliyati, ijtimoiy adolat va inson qadr-qimmatini himoya qilish bilan bog‘liq murakkab axloqiy muammolarni ham yuzaga chiqarmoqda.

Mazkur muammoning dolzarbligi shundaki, sun’iy intellekt yordamida yaratilgan kontent tashqi ko‘rinishi, til uslubi va mazmuniy tuzilishi jihatidan inson tomonidan tayyorlangan materiallardan deyarli farq qilmasligi mumkin. Natijada foydalanuvchi axborotning manbasini, yaratilish usulini va uning haqiqiylik darajasini mustaqil aniqlashda qiyinchiliklarga duch keladi. Xalqaro sun’iy intellekt xavfsizligi hisobotida umumiy maqsadli sun’iy intellekt tizimlari orqali soxta matn, tasvir va videolarni katta hajmda yaratish, ularni jamoatchilik fikriga ta’sir ko‘rsatish maqsadida qo‘llash imkoniyatlari qayd etilgan. Shu bilan birga, bunday kontentning

jamiyatga umumiy ta’siri va uzoq muddatli oqibatlarini baholashda ilmiy noaniqliklar mavjudligi ham ta’kidlangan [1].

Sun’iy intellektning axloqiy ta’sirini faqat noto‘g‘ri axborot ishlab chiqarish bilan cheklash mumkin emas. Muammo algoritmlarning muayyan ijtimoiy guruhlar, tillar va madaniy qarashlarni qay darajada ifodalashi, foydalanuvchining mustaqil fikrlashiga ta’siri, shaxsiy ma’lumotlarning himoyasi hamda texnologik tizimlar oldidagi javobgarlikning taqsimlanishi bilan ham uzviy bog‘liq. UNESCO tomonidan ishlab chiqilgan generativ sun’iy intellektdan ta’lim va ilmiy tadqiqotlarda foydalanish bo‘yicha tavsiyalarda ham inson manfaatlari, ma’lumotlar maxfiyligi, tenglik, madaniy xilma-xillik va insonning mustaqil faoliyat yuritish imkoniyatlarini himoya qilishga alohida e’tibor qaratilgan [2].

Tadqiqotda tizimli tahlil, ilmiy manbalarni qiyosiy o‘rganish, konseptual umumlashtirish va xavflarni tasniflash usullaridan foydalanildi. Tahlil uchun UNESCOning generativ sun’iy intellekt bo‘yicha tavsiyalari, NISTning generativ sun’iy intellekt xavflarini boshqarishga oid metodologik hujjati hamda 2025-yilgi Xalqaro sun’iy intellekt xavfsizligi hisobotida keltirilgan ilmiy xulosalar asos qilib olindi [1–3]. Shuningdek, Yevropa Ittifoqining sun’iy intellekt to‘g‘risidagi qonunchiligida belgilangan shaffoflik talablarining mazkur muammo bilan bog‘liq jihatlari o‘rganildi [4].

Tadqiqot doirasida axloqiy xavflar to‘rtta o‘zaro bog‘liq yo‘nalishda ko‘rib chiqildi: algoritmik noxolislik va kamsitish, dezinformatsiya va sintetik kontentning ishonchliligi, madaniy-ma’naviy qadriyatlarning ifodalanishi hamda mas’uliyat va shaffoflik. Ushbu yo‘nalishlar texnologiyaning o‘ziga xos xususiyatlari bilan cheklanmay, uning amaliy qo‘llanish muhiti, foydalanuvchilarning axborot savodxonligi va institutsional nazorat mexanizmlari bilan birgalikda tahlil qilindi.

Sun’iy intellekt tizimlarining axloqiy jihatdan baholanishida algoritmik noxolislik markaziy o‘rin tutadi. Mazkur hodisa tizim natijalarining ayrim ijtimoiy guruhlar, tillar, madaniyatlar yoki qarashlarga nisbatan muntazam ravishda nomutanosib bo‘lishini anglatadi. Noxolislikning kelib chiqishi ko‘pincha o‘quv ma’lumotlarining tarkibi, ma’lumotlarni tanlash mezonlari, belgilash jarayonidagi insoniy qarashlar, model arxitekturasi va tizimdan foydalanish sharoitlari bilan bog‘liq. Agar modelni o‘qitishda ayrim hududlar yoki ijtimoiy guruhlar haqidagi ma’lumotlar yetarli bo‘lmasa, u ushbu guruhlarga tegishli savollarga yuzaki, noto‘liq yoki stereotiplarga asoslangan javoblar berishi mumkin.

2025-yilgi Xalqaro sun’iy intellekt xavfsizligi hisobotida umumiy maqsadli sun’iy intellekt tizimlari irq, jins, yosh, nogironlik, madaniyat va siyosiy qarashlar bilan bog‘liq noxolisliklarni aks ettirishi mumkinligi ko‘rsatilgan. Bunday holatlar resurslarni taqsimlashda tengsizlik, stereotiplarning mustahkamlanishi va ayrim guruhlar nuqtayi nazarining e’tibordan chetda qolishiga olib kelishi ehtimoli mavjud

[1]. Shu sababli algoritmik noxolislikni faqat dasturiy xato sifatida emas, balki ijtimoiy adolat va teng imkoniyatlar bilan bog‘liq muammo sifatida baholash lozim.

Sun’iy intellekt, dezinformatsiya va sintetik kontent ishonchliligi

Generativ sun’iy intellektning rivojlanishi axborot yaratish xarajatlarini kamaytirib, kontentni qisqa vaqt ichida katta auditoriyaga yetkazish imkoniyatini kengaytirdi. Ushbu imkoniyatlardan ta’limiy materiallar, ilmiy izohlar, tarjimalar va axborot xizmatlarini rivojlantirishda samarali foydalanish mumkin. Biroq xuddi shu texnologik imkoniyatlar yolg‘on xabarlar, soxta hujjatlar, qalbakilashtirilgan audioyozuvlar va real voqealarga o‘xshash videolarni ishlab chiqarish uchun ham qo‘llanishi mumkin. Ayniqsa, muayyan shaxsning ovozi yoki qiyofasini taqlid qiluvchi deepfake materiallar axborotning haqiqiyligini tekshirish jarayonini murakkablashtiradi.

Dezinformatsiya tushunchasini oddiy xato ma’lumotdan farqlash muhim. Noto‘g‘ri ma’lumot har doim ham ataylab tarqatilmaydi, dezinformatsiya esa odatda odamlarni chalg‘itish yoki muayyan maqsadga erishish uchun bila turib tarqatiladigan yolg‘on axborotni anglatadi. Sun’iy intellekt ikkala holatda ham vositachi bo‘lishi mumkin: u tasodifan noto‘g‘ri ma’lumot yaratishi yoki ataylab chalg‘ituvchi material tayyorlash uchun ishlatilishi mumkin. Shuning uchun kontentni baholashda uning noto‘g‘riligini aniqlash bilan birga, manbasi, yaratilish sharoiti, tarqatilish usuli va ehtimoliy zararini ham hisobga olish zarur.

Raqamli kontentning axloqiy ta’sirini baholash modeli. Tadqiqot natijalarini umumlashtirish asosida raqamli kontentning axloqiy xavflarini baholash uchun o‘zaro bog‘langan to‘rtta yo‘nalishni qamrab oluvchi konseptual model taklif etiladi. Ushbu model xalqaro tavsiyalar va xavflarni boshqarish yondashuvlarining mazkur maqola doirasida amalga oshirilgan mualliflik umumlashtirishidir; u hali empirik tekshiruvdan o‘tgan universal standart sifatida talqin qilinmasligi kerak.

Birinchi yo‘nalish — haqqoniylik va kelib chiqish shaffofligi. Bunda kontentdagi faktlarning ishonchli manbalar bilan tasdiqlanishi, yaratilish usuli, mualliflik va tahrirlash tarixi hamda sun’iy intellektdan foydalanilganligi haqidagi ma’lumotlarning ochiqligi tekshiriladi. Ikkinchi yo‘nalish — xolislik va inklyuzivlik. Unda turli ijtimoiy guruhlar, tillar va madaniyatlar tasviridagi nomutanosiblik, stereotiplar hamda kamsituvchi mazmun ehtimoli baholanadi. Uchinchi yo‘nalish — qadriyatlar va ijtimoiy zarar. Bu bosqichda materialning inson sha’ni, madaniy xilma-xillik, shaxsiy hayot daxlsizligi va ijtimoiy ishonchga ehtimoliy ta’siri ko‘rib chiqiladi. To‘rtinchi yo‘nalish — javobgarlik va inson nazorati. Unda kontent uchun mas’ul tomonlarning aniqligi, inson tomonidan tekshirilganligi, xatolarni tuzatish va zararli materialga munosabat bildirish mexanizmlari baholanadi.

Raqamli kontent va sun’iy intellekt tomonidan yaratilgan materiallarning axloqiy ta’sirini baholash zamonaviy axborot jamiyatining muhim ilmiy-amaliy

muammolaridan biri hisoblanadi. Tadqiqot shuni ko‘rsatadiki, algoritmik noxolislik, dezinformatsiya va qadriyatlar bilan bog‘liq xavflar alohida hodisalar bo‘lishiga qaramay, ular o‘zaro chambarchas bog‘liq bo‘lib, kontentning ishonchliligi, ijtimoiy adolat va insonning axborotga nisbatan mustaqil munosabatiga birgalikda ta’sir ko‘rsatishi mumkin. Shu bois sun’iy intellektning axloqiy xavfsizligini faqat texnik himoya vositalari yoki avtomatik aniqlash algoritmlari orqali ta’minlash yetarli emas.

Tadqiqot asosida raqamli kontentni baholashda haqqoniylik va kelib chiqish shaffofligi, xolislik va inklyuzivlik, qadriyatlar va ijtimoiy zarar, javobgarlik va inson nazoratini qamrab oluvchi to‘rt yo‘nalishli konseptual model taklif etildi. Ushbu yondashuv sun’iy intellektdan foydalanish jarayonida xavflarni tizimli aniqlash, ularning ehtimoliy oqibatlarini baholash va tegishli himoya choralarini belgilash uchun metodologik asos bo‘lib xizmat qilishi mumkin.

Amaliy jihatdan ta’lim va ilmiy muassasalarda sun’iy intellekt yaratgan materiallarni tekshirish tartibini joriy etish, raqamli va media savodxonlikni rivojlantirish, kontentning kelib chiqishini imkon qadar ochiq ko‘rsatish, algoritmlarni muntazam auditdan o‘tkazish va inson nazoratini saqlab qolish tavsiya etiladi. Ayniqsa, o‘zbek tilidagi raqamli muhitda lingvistik aniqlik, mahalliy madaniy kontekst va ishonchli manbalardan foydalanish masalalarini alohida o‘rganish muhimdir. Keyingi tadqiqotlarda taklif etilgan modelni real kontent namunalari asosida sinovdan o‘tkazish, baholash mezonlarining ishonchliligini aniqlash va turli xavf omillari o‘rtasidagi bog‘liqlikni empirik ma’lumotlar yordamida tekshirish maqsadga muvofiqdir.

Foydalanilgan adabiyotlar

1. International AI Safety Report. *International AI Safety Report 2025*. 2025. URL: https://internationalaisafetyreport.org/publication/international-ai-safety-report-2025
2. Miao, F., Holmes, W. *Guidance for Generative AI in Education and Research*. UNESCO, Paris, 2023. DOI: https://doi.org/10.54675/EWZM9535
3. Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST AI 600-1, 2024. DOI: https://doi.org/10.6028/NIST.AI.600-1
4. European Parliament and Council of the European Union. *Regulation (EU) 2024/1689 of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj
5. Chandra, B., Dunietz, J., Roberts, K., Lee, Y., Fontana, P., Awad, G. *Reducing Risks Posed by Synthetic Content: An Overview of Technical Approaches to Digital Content Transparency*. NIST AI 100-4, 2024. URL: https://www.nist.gov/publications/reducing-risks-posed-synthetic-content-overview-technical-approaches-digital-content
6. European Commission. *Guidelines on Transparency Obligations for Providers and Deployers of Certain AI Systems*. 2026. URL: https://digital-strategy.ec.europa.eu/en/policies/guidelines-ai-transparency-obligations