SUN'IY INTELLEKTNING CHEGARASI VA TASNIFI: NAZARIY-TAHLILIY SHARH

Aminov Asadbek Narimon o'g'li
Urganch davlat pedagogika instituti akademik litseyi
Informatika va AT fani o'qituvchisi
doi:https://doi.org/10.5281/zenodo.23029543

Annotatsiya. Maqolada sun'iy intellekt (SI) tushunchasi, uning rivojlanish bosqichlari, turli asoslar bo'yicha tasniflari hamda imkoniyatlari chegaralari tizimli ko'rib chiqiladi. SI qobiliyat darajasi, vazifaviy xususiyati, texnologik yondashuvi, qo'llanish sohasi va xavf darajasi bo'yicha tasniflanadi. Chegaralar texnik (ma'lumotga bog'liqlik, gallyutsinatsiya, zaifliklar, izohlanuvchanlik, resurs sarfi), axloqiy-huquqiy (xolislik, maxfiylik, mas'uliyat, tartibga solish) va falsafiy (tushunish, ong) yo'nalishlarda tahlil qilinadi. Natijalar zamonaviy SI tizimlari tor doiradagi vazifalarda yuqori samaradorlikka erishgan bo'lsa-da, umumiy aql, ma'noni chuqur tushunish va mas'uliyat masalalarida inson o'rnini bosa olmasligini ko'rsatadi.
Kalit so'zlar: sun'iy intellekt, tasnif, tor SI, umumiy SI, mashinaviy o'rganish, chuqur o'rganish, generativ SI, gallyutsinatsiya, izohlanuvchan SI, SI etikasi.

ПРЕДЕЛЫ И КЛАССИФИКАЦИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ТЕОРЕТИКО-АНАЛИТИЧЕСКИЙ ОБЗОР

Аннотация. В статье систематически рассматриваются понятие искусственного интеллекта (ИИ), этапы его развития, классификации по различным основаниям и границы его возможностей. ИИ классифицируется по уровню способностей, функциональным признакам, технологическому подходу, области применения и уровню риска. Границы анализируются в технических (зависимость от данных, галлюцинации, уязвимости, объяснимость, ресурсоёмкость), этико-правовых (справедливость, конфиденциальность, ответственность, регулирование) и философских (понимание, сознание) аспектах. Результаты показывают, что современные системы ИИ эффективны в узких задачах, но не заменяют человека в вопросах общего интеллекта, глубокого понимания смысла и ответственности.
Ключевые слова: искусственный интеллект, классификация, узкий ИИ, общий ИИ, машинное обучение, глубокое обучение, генеративный ИИ, галлюцинации, объяснимый ИИ, этика ИИ.

THE LIMITS AND CLASSIFICATION OF ARTIFICIAL INTELLIGENCE: A THEORETICAL-ANALYTICAL REVIEW

Abstract. The article systematically examines the concept of artificial intelligence (AI), the stages of its development, its classifications on several bases, and the limits of its capabilities. AI is classified by level of capability, functional characteristics, technological approach, field of application and level of risk. The limits are analysed in technical (data dependence, hallucination, vulnerabilities, explainability, resource consumption), ethical-legal (fairness, privacy, accountability, regulation) and philosophical (understanding, consciousness) dimensions. The results show that modern AI systems perform well in narrow tasks but cannot replace humans with respect to general intelligence, deep understanding of meaning and accountability.

Keywords: artificial intelligence, classification, narrow AI, artificial general intelligence, machine learning, deep learning, generative AI, hallucination, explainable AI, AI ethics.

KIRISH

So‘nggi o‘n yillikda sun’iy intellekt (SI) ilmiy laboratoriyalardan chiqib, iqtisodiyot va kundalik hayotning ajralmas qismiga aylandi. Tibbiy tasvirlarni tahlil qilish, oqsil tuzilmasini bashorat qilish, mashinaviy tarjima, moliyaviy firibgarlikni aniqlash, ta’limda shaxsiylashtirilgan yordamchilar kabi yo‘nalishlarda SI tizimlari amaliy natija bermoqda. Ayniqsa, 2022-yil oxirida ommaviy foydalanishga chiqarilgan generativ tizimlar texnologiyaga bo‘lgan qiziqishni keskin oshirdi.

Shu bilan birga, SI haqidagi tasavvurlar ko‘pincha noaniq yoki mubolag‘ali. Bir tomondan uni "hamma narsaga qodir aql" deb baholash, ikkinchi tomondan oddiy "murakkab kalkulyator" deb qarash uchraydi. Muammo shundaki, "sun’iy intellekt", "mashinaviy o‘rganish", "chuqur o‘rganish" va "generativ SI" tushunchalari tez-tez aralashtiriladi, tasniflar esa turli asoslarga tayanadi, imkoniyatlar chegaralari haqidagi ma’lumotlar turli manbalarda tarqoq holda beriladi.

Tadqiqotning maqsadi SI tushunchasi va tasniflarini tizimlashtirish hamda uning texnik, axloqiy-huquqiy va falsafiy chegaralarini aniqlashdan iborat. Maqsadga erishish uchun quyidagi vazifalar belgilandi:

– SI tushunchasining zamonaviy ta’riflarini va rivojlanish bosqichlarini yoritish;

– SI’ni qobiliyat darajasi, vazifaviy xususiyati, texnologik yondashuvi, qo‘llanish sohasi va xavf darajasi bo‘yicha tasniflash;

– SI’ning texnik, axloqiy-huquqiy va falsafiy chegaralarini tahlil qilish;

– chegaralarni yumshatish yo‘nalishlarini va inson bilan hamkorlik istiqbollarini baholash.

Tadqiqotda nazariy sharh, qiyosiy tahlil va tizimlashtirish usullari qo‘llanildi. Asosiy manbalar sifatida yetakchi darsliklar, ilmiy jurnal maqolalari, xalqaro standartlar va me’yoriy hujjatlardan foydalanildi.

1. SUN’IY INTELLEKT TUSHUNCHASI VA RIVOJLANISH BOSQICHLARI

1.1. Zamonaviy ta’riflar

SI yagona, hamma tomonidan qabul qilingan ta’rifga ega emas. S. Rassell va P. Norvig sohadagi ta’riflarni to‘rt yondashuvga guruhlaydi: inson kabi fikrlash, inson kabi harakat qilish, oqilona fikrlash va oqilona harakat qilish; zamonaviy tadqiqotlarda oxirgi yondashuv, ya’ni maqsadga erishish uchun eng maqbul harakat qiladigan "oqilona agent" tushunchasi asosiy hisoblanadi [1].

Me’yoriy sohada ta’riflar amaliy tus oladi. Iqtisodiy hamkorlik va taraqqiyot tashkiloti (OECD) SI tizimini kiruvchi ma’lumotdan aniq yoki yashirin maqsadlar asosida xulosa chiqarib, bashorat, kontent, tavsiya yoki qaror kabi natijalar hosil qiladigan va real yoki virtual muhitga ta’sir ko‘rsata oladigan mashinaviy tizim sifatida belgilaydi [2]. Xuddi shu yondashuv ISO/IEC 22989 standartida [3] hamda Yevropa Ittifoqining SI to‘g‘risidagi reglamentida [4] turli darajada aks etgan. Bu ta’riflarning umumiy jihati shuki, ular SI’ni "aql" emas, balki muayyan avtonomiya darajasida ishlaydigan tizim sifatida talqin qiladi.

Tushunchalarning o‘zaro nisbatini nisbatan oddiy ko‘rinishda ifodalash mumkin: SI eng keng yo‘nalish bo‘lib, mashinaviy o‘rganish uning ma’lumotlardan o‘rganishga asoslangan qismi, chuqur o‘rganish mashinaviy o‘rganishning ko‘p qatlamli neyron tarmoqlariga tayanadigan turi, generativ SI esa yangi kontent yaratuvchi modellar sinfidir. Shu sababli har bir generativ model SI hisoblanadi, ammo har bir SI tizimi generativ emas.

1.2. Rivojlanish bosqichlari

SI tarixi yuksalish va sovish davrlarining almashinuvidan iborat. 1950-yilda A. Tyuring "mashina fikrlay oladimi?" degan savolni ilmiy muhokamaga olib chiqdi va bu savolni amaliy sinovga aylantiruvchi "taqlid o‘yini" g‘oyasini taklif qildi [5]. "Sun’iy intellekt" atamasi

1955-yilda tayyorlangan Dartmut yozgi tadqiqot loyihasi taklifida qo‘llanilib, 1956-yilgi seminar sohaning rasmiy tug‘ilishi hisoblanadi [6]. Dastlabki optimizm 1970-yillarda amaliy natijalar kutilganidan past bo‘lgani sababli so‘ndi va moliyalashtirish qisqardi ("birinchi SI qishi"); bunda Britaniyadagi Laythill hisoboti muhim rol o‘ynagan [7]. 1980-yillarda ekspert tizimlar tijoriy qo‘llanildi, biroq ularning murakkabligi va moslashuvchanligi cheklanganligi ikkinchi sovish davriga olib keldi.

2010-yillarda katta ma’lumotlar to‘plamlari, grafik protsessorlar quvvati va yangi algoritmlar uyg‘unlashib, chuqur o‘rganish davrini boshladi. 2012-yilda chuqur konvolyutsion tarmoq ImageNet tanlovida tasvirlarni tanish xatosini sezilarli kamaytirdi [8]. 2017-yilda taklif etilgan transformer arxitekturasi [9] keyingi yirik til modellari uchun asos bo‘ldi. Asosiy bosqichlar 1-jadvalda umumlashtirilgan.

1-jadval. Sun’iy intellekt rivojlanishining asosiy bosqichlari

Yil	Voqea	Ahamiyati	Manba
1950	Tyuringning "Computing Machinery and Intelligence" maqolasi	Mashina intellekti savoli va taqlid o‘yini (Tyuring testi) g‘oyasi	[5]
1956	Dartmut seminari	Sohaning mustaqil ilmiy yo‘nalish sifatida shakllanishi	[6]
1958	Rozenblattning perseptroni	Dastlabki o‘rganuvchi neyron model	[10]
1973	Laythill hisoboti	Moliyalashtirishning qisqarishi, "birinchi SI qishi"	[7]
1980-y.	Ekspert tizimlarning tijoriy qo‘llanishi	Qoidaga asoslangan SI’ning yuksalishi va keyingi sovish	–
1997	Deep Blue G. Kasparovni mag‘lub etdi	Maxsus SI tizimi shaxmatda jahon chempionidan ustun keldi	[11]
2012	AlexNet ImageNet tanlovida	Chuqur o‘rganish davrining boshlanishi	[8]
2016	AlphaGo Li Sedolni mag‘lub etdi	Chuqur neyron tarmoq va mustahkamlovchi o‘rganish uyg‘unligi	[12]
2017	Transformer arxitekturasi	Zamonaviy yirik til modellari asosi	[9]
2020	GPT-3 (175 mlrd parametr)	Kam misol bilan o‘rganish qobiliyati	[13]
2021	AlphaFold 2	Oqsil fazoviy tuzilmasini yuqori aniqlikda bashorat qilish	[14]
2022	Generativ suhbat tizimlarining ommalashuvi	SI’ning keng auditoriyaga chiqishi	–

2. SUN’IY INTELLEKT TASNIFI

SI’ni yagona mezon bo‘yicha tasniflash mumkin emas, chunki turli mezonlar turli savollarga javob beradi: tizim nimaga qodir, u qanday ishlaydi, qayerda qo‘llaniladi va qanday xavf tug‘diradi. Quyida beshta asosiy mezon ko‘rib chiqiladi.

2.1. Qobiliyat darajasiga ko‘ra

Eng keng tarqalgan tasnif tizimlarni qobiliyat darajasi bo‘yicha uch guruhga bo‘ladi (2-jadval).

2-jadval. Sun'iy intellektning qobiliyat darajasiga ko'ra tasnifi

Tur	Ta'rifi	Hozirgi holati
Tor (zaif) SI, Narrow AI	Aniq bir vazifa yoki tor vazifalar to'plamini bajarishga mo'ljallangan, boshqa sohaga o'z-o'zidan o'ta olmaydigan tizim	Barcha amaliy tizimlar (tavsiya tizimlari, nutqni tanish, tibbiy tasvir tahlili, shaxmat dasturlari)
Umumiy SI, AGI	Istalgan intellektual vazifani inson darajasida o'rganib bajara oladigan, bilimni sohalar orasida erkin ko'chira oladigan tizim	Nazariy maqsad; ta'rifi va amalga oshish muddati bo'yicha yagona fikr yo'q
O'ta aql, ASI	Deyarli barcha sohalarda eng yaxshi insondan ustun bo'lgan taxminiy tizim	Mavjud emas; falsafiy muhokama va xavfsizlik tadqiqotlari mavzusi

Umumiy SI (AGI) tushunchasi tizimning moslashuvchanligi va bilimni ko'chirish qobiliyatiga urg'u beradi, biroq uning aniq mezonlari borasida tadqiqotchilar o'rtasida kelishuv yo'q [15]. O'ta aql esa asosan uzoq muddatli xavflar va boshqaruv muammolari kontekstida muhokama qilinadi [16]. Shuning uchun so'nggi yillarda "bor yoki yo'q" mantig'i o'rniga darajalarga asoslangan yondashuvlar taklif qilinmoqda. Masalan, Morris va hammualliflari tizimlarni ikki o'lchov, ya'ni ish sifati (rivojlanmoqda, malakali, ekspert, virtuoz, insondan ustun) va umumiylik (tor yoki umumiy) bo'yicha baholashni tavsiya etadi [17]. Bunday yondashuv AGI'ga erishish jarayonini bosqichma-bosqich o'lchash imkonini beradi.

2.2. Vazifaviy xususiyatiga ko'ra

A. Xintse tasnifi tizimlarni xotira va "ongni modellashtirish" darajasiga qarab to'rt turga ajratadi [18]:

– **reaktiv mashinalar** xotirasiz bo'lib, faqat joriy holatga javob beradi (Deep Blue);

– **cheklangan xotirali tizimlar** o'tmish ma'lumotlaridan foydalanadi (haydovchiga yordam tizimlari, chatbotlar, tavsiya tizimlari);

– **"ongni tushunuvchi" tizimlar (Theory of Mind)** boshqalarning niyat, e'tiqod va his-tuyg'usini modellashtira olishi kerak; hozircha tadqiqot bosqichida;

– **o'zini anglovchi tizimlar** o'z holatini anglaydi; mavjud emas va nazariy tushuncha sifatida muhokama qilinadi.

A. Kaplan va M. Xaynlayn esa SI'ni "intellekt turi" bo'yicha uchga bo'ladi: tahliliy (kognitiv intellekt, ma'lumotdan xulosa chiqaradi), inson ruhiyatidan ilhomlangan (kognitiv va hissiy intellektni birlashtiradi) va insonsifat (kognitiv, hissiy hamda ijtimoiy intellektni qamrab oladi) tizimlar; bugungi amaliyot asosan birinchi toifaga to'g'ri keladi [19].

2.3. Texnologik yondashuviga ko'ra

Texnologik nuqtayi nazardan SI bir necha yondashuvga bo'linadi (3-jadval). Mashinaviy o'rganishda uch asosiy paradigma ajratiladi: nazoratli o'rganish (javoblari ma'lum misollar asosida), nazoratsiz o'rganish (belgilanmagan ma'lumotdagi qonuniyatlarni topish) va mustahkamlovchi o'rganish (muhit bilan o'zaro ta'sirda mukofot orqali o'rganish) [20], [21].

3-jadval. Sun'iy intellektning texnologik yondashuvlari

Yondashuv	Mohiyati	Afzalliklari	Cheklovlari
Ramziy (qoidaga asoslangan) SI	Inson yozgan mantiqiy qoidalar va bilimlar bazasi	Shaffoflik, natijani tushuntirish oson	Qoidalarni qo'lda yozish, o'zgarishga moslashuvchanlikning pastligi

Yondashuv	Mohiyati	Afzalliklari	Cheklovlari
Mashinaviy o‘rganish	Qonuniyatlarni ma’lumotdan avtomatik topish	Moslashuvchanlik, qoidalarni yozish shart emas	Ma’lumot sifati va hajmiga bog‘liqlik
Chuqur o‘rganish	Ko‘p qatlamli neyron tarmoqlar	Tasvir, nutq, matndagi yuqori aniqlik	Ko‘p ma’lumot va hisoblash quvvati, izohlash qiyinligi
Generativ SI	Yangi matn, tasvir, ovoz, kod yaratuvchi modellar	Ko‘p vazifali, tabiiy til orqali boshqariladi	Gallyutsinatsiya, tekshirish zarurati
Gibrid (neyro-ramziy) SI	Neyron tarmoqlar va ramziy mantiq uyg‘unligi	O‘rganish va mantiqiy xulosa chiqarishni birlashtirish	Tadqiqot bosqichida, amaliy qo‘llanishi cheklangan

Chuqur o‘rganish [22] tasvir va nutqni tanishda, transformer asosidagi yirik modellar esa matn bilan ishlashda yutuqlarga erishdi. Bunday modellarni bir vaqtning o‘zida ko‘plab vazifalar uchun moslashtirish mumkin bo‘lgani uchun ular "poydevor modellar" (foundation models) deb ataladi [23]. Amalda bu yondashuvlar ko‘pincha bir tizim ichida birga qo‘llaniladi.

2.4. Qo‘llanish sohasiga ko‘ra

Qo‘llanish sohasi bo‘yicha tasnif SI’ning amaliy yo‘nalishlarini ko‘rsatadi:

– **tabiiy tilni qayta ishlash**: mashinaviy tarjima, matn tahlili, savol-javob, suhbat tizimlari;

– **kompyuter ko‘rishi**: tibbiy tasvirlar tahlili, yuzni tanish, sifat nazorati, yo‘l belgilarini aniqlash;

– **nutqni qayta ishlash**: ovozni matnga aylantirish, ovozli yordamchilar, sintez;

– **robototexnika va avtonom tizimlar**: sanoat robotlari, dronlar, avtonom transport;

– **qaror qabul qilishni qo‘llab-quvvatlash va bashoratlash**: kredit skoringi, talabni bashorat qilish, hosildorlik prognozi, tavsiya tizimlari;

– **ilmiy hisoblash**: biologiya, kimyo va materialshunoslikda modellashtirish (masalan, oqsil tuzilmasini bashorat qilish [14]).

2.5. Xavf darajasiga ko‘ra

SI tizimlari keng qo‘llanila boshlagach, tasnifning yangi mezoni, ya’ni xavf darajasi paydo bo‘ldi. Yevropa Ittifoqining SI to‘g‘risidagi reglamenti tizimlarni to‘rt xavf toifasiga ajratib, ularga turli talablar qo‘yadi [4] (4-jadval). AQShning Milliy standartlar va texnologiya instituti (NIST) esa tashkilotlarga SI xavflarini aniqlash, o‘lchash va boshqarish uchun ixtiyoriy tavsiya ramkasi taklif qiladi [24].

4-jadval. Yevropa Ittifoqi reglamenti bo‘yicha xavf toifalari

Xavf darajasi	Mohiyati	Misollar
Qabul qilib bo‘lmaydigan xavf	Asosiy huquqlarga zid bo‘lgani uchun taqiqlanadigan amaliyotlar	Hokimiyat tomonidan fuqarolarni ijtimoiy reytinglash, odamlarni yashirin manipulyatsiya qiluvchi tizimlar
Yuqori xavf	Qat’iy talablar (xavfni boshqarish, ma’lumot sifati, inson nazorati, hujjatlashtirish) qo‘yiladi	Ta’lim va ishga qabul qilishda, kredit baholashda, muhim infratuzilmada, huquq-tartibotda qo‘llanadigan tizimlar

Xavf darajasi	Mohiyati	Misollar
Cheklangan xavf	Shaffoflik majburiyatlari: foydalanuvchi SI bilan muloqot qilayotganini bilishi kerak	Suhbat tizimlari, sintetik kontent
Minimal xavf	Maxsus talablar qo‘yilmaydi	Spam filtrlari, video o‘yinlardagi SI

3. SUN’IY INTELLEKTNING CHEGARALARI

3.1. Texnik chegaralar

Ma’lumotga bog‘liqlik. Mashinaviy o‘rganish tizimlarining sifati o‘qitish ma’lumotlarining sifati, to‘liqligi va vakilligiga bog‘liq. Ma’lumotlardagi xato yoki noxolislik modelga o‘tadi. Bundan tashqari, tizim amalda o‘qitilgan taqsimotdan farq qiladigan ma’lumotga duch kelsa (taqsimot siljishi), aniqligi keskin pasayishi mumkin.

Gallyutsinatsiya. Generativ til modellari ishonchli va ravon ko‘rinadigan, biroq noto‘g‘ri yoki manbasiz ma’lumot berishi mumkin; bu hodisa tadqiqotlarda "gallyutsinatsiya" deb ataladi [25]. Sababi shundaki, model faktlarni tekshirmaydi, balki o‘qitish ma’lumotlaridagi qonuniyatlarga ko‘ra ehtimolli davomni hosil qiladi. Uni kamaytirish uchun tashqi manbadan ma’lumot qidirib, javobni shunga asoslash (retrieval-augmented generation) kabi usullar qo‘llanadi [26], biroq muammo to‘liq hal etilmagan, shuning uchun muhim qarorlarda natijani tekshirish shart.

Zaifliklar. Neyron tarmoqlar inson sezmaydigan darajada kichik o‘zgartirilgan kirish ma’lumotiga (adversar misollarga) noto‘g‘ri javob berishi mumkinligi ko‘rsatilgan [27], [28]. Bu xavfsizlikka sezgir sohalarda (avtonom transport, biometriya) jiddiy xavf tug‘diradi.

Umumlashtirish va sog‘lom fikr. Chuqur o‘rganish tizimlari o‘qitilgan sharoitdan tashqarida ko‘pincha xato qiladi, ular sabab-oqibat munosabatlarini emas, ko‘proq statistik bog‘liqliklarni o‘zlashtiradi [29]. Inson uchun oddiy bo‘lgan jismoniy va ijtimoiy sog‘lom fikr mashina uchun qiyin, aksincha, inson uchun murakkab hisob-kitoblar oson: bu Moravek paradoksi deb ataladi [30].

Izohlanuvchanlik. Chuqur neyron tarmoqlar qarorining sababini aniq izohlab bermaydi ("qora quti"). Tibbiyot, huquq va moliya kabi yuqori xavfli sohalarda bu ayniqsa muammo: ba’zi tadqiqotchilar bunday sohalarda dastlabdan izohlanadigan modellardan foydalanishni tavsiya etadi [31]. Buning uchun izohlanuvchan SI (XAI) yo‘nalishi rivojlanmoqda [32].

Resurs sarfi. Yirik modellarni o‘qitish katta hisoblash quvvati va elektr energiyasini talab qiladi, bu esa atrof-muhitga ta’sir qiladi va katta tashkilotlar bilan cheklangan resursli tadqiqotchilar o‘rtasida tengsizlik yaratadi [33], [34]. Model hajmi va ma’lumot oshgani sari sifat ham oshishi kuzatilgan bo‘lsa-da [35], bu qonuniyat cheksiz davom etishi kafolatlanmagan va iqtisodiy hamda energetik jihatdan cheklanadi. Texnik chegaralar 5-jadvalda umumlashtirilgan.

5-jadval. Sun’iy intellektning asosiy texnik chegaralari

Chegara	Mohiyati	Oqibati	Yumshatish yo‘llari
Ma’lumotga bog‘liqlik	Sifat ma’lumotning to‘liqligi va vakilligiga bog‘liq	Noxolis yoki noto‘g‘ri natijalar	Ma’lumotni tekshirish, monitoring, qayta o‘qitish
Gallyutsinatsiya	Ishonchli ko‘rinadigan noto‘g‘ri javob	Dezinformatsiya, noto‘g‘ri qaror	Manbaga asoslash, inson nazorati, faktlarni tekshirish

Chegara	Mohiyati	Oqibati	Yumshatish yo‘llari
Zaifliklar	Kichik o‘zgarishlarga sezgirlik	Xavfsizlik buzilishlari	Mustahkam o‘qitish, sinovdan o‘tkazish
Umumlashtirish	Notanish vaziyatda xato qilish	Ishonchsizlik nostandart holatlarda	Ko‘p xil ma’lumot, insonni jarayonga qo‘shish
Izohlanuvchanlik	Qaror sababini tushuntira olmaslik	Ishonch va mas’uliyat muammosi	XAI usullari, izohlanadigan modellar
Resurs sarfi	Katta hisoblash va energiya talabi	Xarajat, ekologik ta’sir, tengsizlik	Samarali arxitektura, kichik modellar

3.2. Axloqiy va huquqiy chegaralar

Xolislik va kamsitish. Ma’lumotlardagi tarixiy noxolisliklar modelga o‘tib, ishga qabul qilish, kredit berish, huquq-tartibot kabi sohalarda adolatsiz natijalarga olib kelishi mumkin [36]. Keng tarqalgan misollardan biri: tijoriy yuz tahlili tizimlarida xato darajasi teri rangi to‘qroq ayollar uchun sezilarli yuqori bo‘lgani aniqlangan [37]. Yirik til modellari ham o‘qitish matnlaridagi stereotiplarni takrorlashi mumkin [38].

Maxfiylik. SI tizimlari katta hajmdagi, jumladan shaxsiy ma’lumotlarni talab qiladi. Ma’lumotni to‘plash, saqlash va qayta ishlash shaxsiy hayot daxlsizligi hamda ma’lumotlarni himoya qilish qonunchiligi bilan to‘qnashishi mumkin.

Mas’uliyat. SI tizimi xato qilganda mas’uliyatni kim ko‘taradi: ishlab chiquvchimi, joriy etuvchi tashkilotmi yoki foydalanuvchimi? Tizim qarori izohlanmasa, bu savolga javob berish yanada murakkablashadi. Shu sababli yuqori xavfli sohalarda inson nazorati va hujjatlashtirish talablari qo‘yilmoqda [4].

Mualliklik huquqi va soxta kontent. Generativ modellar o‘qitish ma’lumotlarida himoyalangan asarlardan foydalanishi va yaratilgan kontentning mualliklik maqomi haqidagi savollarni keskinlashtirdi. Real odamlarning ovozi va tasvirini yuqori aniqlikda taqlid qiluvchi soxta kontent (deepfake) dezinformatsiya va firibgarlik xavfini oshirmoqda.

Tartibga solish. Xalqaro darajada bir necha yondashuv shakllangan. UNESKO 2021-yilda SI etikasi bo‘yicha tavsiyanomani qabul qildi [39], OECD tamoyillari SI’ni ishonchli va inson huquqlarini hurmat qiladigan tarzda rivojlantirishni tavsiya etadi [2], Yevropa Ittifoqi esa xavfga asoslangan yagona majburiy reglamentni joriy etdi [4]. Ko‘plab mamlakatlar, jumladan O‘zbekiston ham, SI’ni rivojlantirish va tartibga solish bo‘yicha strategik hamda me’yoriy hujjatlarni shakllantirmoqda.

3.3. Falsafiy chegaralar

Falsafiy nuqtayi nazardan eng chuqur savol mashina haqiqatan "tushunadi"mi yoki tushunganga o‘xshaydimi degan masalaga borib taqaladi. Tyuring testi mashinaning xatti-harakatini inson xatti-harakatidan farqlab bo‘lmasligini o‘lchov deb hisoblaydi [5]. Unga qarshi Dj. Serlning "Xitoy xonasi" fikriy tajribasi ramzlarni qoidaga ko‘ra to‘g‘ri qayta ishlash hali ma’noni tushunish degani emasligini ko‘rsatishga urinadi [40]. H. Dreyfus esa insoniy bilim katta qismi ochiq qoidalarga keltirilmaydigan, tajriba va jismoniy mavjudlikka bog‘liq bilim ekanini ta’kidlab, ramziy SI’ni tanqid qilgan [41].

Zamonaviy til modellari haqida ham shunga o‘xshash bahs davom etmoqda: E. Bender va A. Koller faqat matn shakliga asoslanib o‘qitilgan model ma’noni o‘zlashtira olmasligini asoslashga urinadi [42]. Ong va sub’ektiv tajriba masalasi, ya’ni Chalmers "ongning qiyin muammosi" deb atagan savol [43], ham ochiq qolmoqda. Hozirgi tizimlarda ong borligini

tasdiqlovchi ilmiy asos yo‘q, biroq masala yakuniy hal etilmagan va bahs falsafa, nevrologiya va kompyuter fanlari kesishmasida davom etmoqda.

4. MUHOKAMA: CHEGARALARNI KENGAYTIRISH YO‘NALISHLARI VA ISTIQBOLLAR

4.1. Chegaralarni kengaytirish yo‘nalishlari

Tadqiqotchilar yuqorida ko‘rsatilgan chegaralarni bir necha yo‘l bilan kengaytirishga urinmoqda. Birinchidan, model hajmi, ma’lumot va hisoblash quvvatini oshirish (masshtablash) ko‘p vazifalarda sifatni oshirgan [35], biroq bu yo‘l resurs va ma’lumot mavjudligi bilan cheklanadi. Ikkinchidan, matn, tasvir va ovozni birgalikda qayta ishlovchi ko‘p modalli modellar tizimning dunyoni idrok etish doirasini kengaytirmoqda. Uchinchidan, tashqi ma’lumot manbalariga ulash va vositalardan (qidiruv, hisob-kitob, ma’lumotlar bazasi) foydalanish gallyutsinatsiyani kamaytirishga xizmat qiladi [26]. To‘rtinchidan, neyron tarmoqlarning o‘rganish qobiliyatini ramziy mantiqning aniqligi bilan birlashtiruvchi neyro-ramziy yondashuvlar sog‘lom fikr va izohlanuvchanlik muammolarini yumshatish umidini beradi.

Shu bilan birga, bu yo‘nalishlarning hech biri chegaralarni to‘liq bartaraf etmasligini ta’kidlash zarur. Masalan, masshtablash sifatni oshirsa-da, gallyutsinatsiya, xolislik va mas’uliyat muammolarini o‘z-o‘zidan hal qilmaydi.

4.2. Inson va SI hamkorligi

Chegaralarni hisobga olgan holda, SI’ni insonni almashtiruvchi emas, uning imkoniyatini kuchaytiruvchi vosita sifatida qo‘llash maqbul yondashuv hisoblanadi. Amaliyotda bu "inson jarayonda" (human-in-the-loop) tamoyilida ifodalanadi: SI katta hajmdagi ma’lumotni tez qayta ishlaydi va variantlar taklif qiladi, yakuniy qaror va mas’uliyat esa insonda qoladi. Tibbiyotda SI shifokorga tasvirlarni tahlil qilishda yordam beradi, ta’limda o‘qituvchi ish yukini yengillashtiradi va individual yondashuvni kuchaytiradi, qishloq xo‘jaligida hosildorlikni bashoratlash va resurslarni tejashga xizmat qiladi, moliyada firibgarlikni aniqlash tezligini oshiradi. Ammo har bir holatda natijani tekshirish, ma’lumot maxfiyligini ta’minlash va tizim cheklovlarini foydalanuvchiga ochiq bildirish zarur.

Foydalanuvchilar uchun amaliy tavsiyalar quyidagilardan iborat: SI natijasini, ayniqsa faktlar, raqamlar va havolalarni, mustaqil manbadan tekshirish; muhim qarorlarni faqat SI’ga topshirmaslik; shaxsiy va maxfiy ma’lumotni tizimga kiritishdan oldin foydalanish shartlarini o‘rganish; SI bilan yaratilgan kontentda foydalanish faktini kerakli hollarda ochiq ko‘rsatish.

4.3. Tadqiqotning cheklovlari

Maqola nazariy-tahliliy sharh xarakterida bo‘lib, yangi eksperimental ma’lumot keltirmaydi. SI sohasi juda tez o‘zgargani uchun aniq tizimlar va me’yoriy hujjatlarga oid ma’lumotlar tezda eskirishi mumkin. Tasniflar shartli bo‘lib, bir tizim bir nechta toifaga birdaniga kirishi mumkin. Kelgusi tadqiqotlarda aniq sohalar (tibbiyot, ta’lim, moliya) kesimida chegaralarni miqdoriy baholash va milliy me’yoriy muhitni qiyosiy o‘rganish maqsadga muvofiq.

XULOSA

Olib borilgan tahlil quyidagi xulosalarga kelish imkonini berdi:

– SI yagona ta’rifga ega emas; zamonaviy ta’riflar (OECD, ISO/IEC, Yevropa Ittifoqi) uni muayyan avtonomiya darajasida ishlaydigan, kiruvchi ma’lumotdan natija hosil qiluvchi mashinaviy tizim sifatida talqin qiladi.

– SI’ni qobiliyat darajasi, vazifaviy xususiyati, texnologik yondashuvi, qo‘llanish sohasi va xavf darajasi bo‘yicha tasniflash mumkin; bu tasniflar bir-birini to‘ldiradi va bir tizim bir nechta toifaga kirishi mumkin.

– Bugungi barcha amaliy tizimlar tor SI toifasiga kiradi; umumiy SI va o‘ta aql hozircha nazariy tushunchalar bo‘lib qolmoqda.

– SI chegaralari texnik (ma'lumotga bog'liqlik, gallyutsinatsiya, zaifliklar, izohlanuvchanlik, resurs sarfi), axloqiy-huquqiy (xolislik, maxfiylik, mas'uliyat, tartibga solish) va falsafiy (tushunish, ong) turlarga bo'linadi.

– Chegaralarni kengaytirish yo'nalishlari mavjud, biroq ular muammolarni to'liq bartaraf etmaydi; shu bois SI'dan inson nazoratida, aqlni kuchaytiruvchi vosita sifatida foydalanish maqbul.

Tasniflarni bilish texnologiyaga real baho berishga, chegaralarni tushunish esa undan xavfsiz va samarali foydalanishga yordam beradi.

FOYDALANILGAN ADABIYOTLAR

1. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. 4th ed. – Hoboken: Pearson, 2021. – 1136 p.
2. OECD. Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). – Paris: OECD, 2019 (amended 2024).
3. ISO/IEC 22989:2022. Information technology – Artificial intelligence – Artificial intelligence concepts and terminology. – Geneva: ISO, 2022.
4. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // Official Journal of the European Union. – 2024.
5. Turing A.M. Computing Machinery and Intelligence // Mind. – 1950. – Vol. 59, No. 236. – P. 433–460.
6. McCarthy J., Minsky M.L., Rochester N., Shannon C.E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. – 1955.
7. Lighthill J. Artificial Intelligence: A General Survey // Artificial Intelligence: a paper symposium. – London: Science Research Council, 1973.
8. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems 25 (NeurIPS 2012). – 2012. – P. 1097–1105.
9. Vaswani A. et al. Attention Is All You Need // Advances in Neural Information Processing Systems 30 (NeurIPS 2017). – 2017. – P. 5998–6008.
10. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain // Psychological Review. – 1958. – Vol. 65, No. 6. – P. 386–408.
11. Campbell M., Hoane A.J., Hsu F. Deep Blue // Artificial Intelligence. – 2002. – Vol. 134, No. 1–2. – P. 57–83.
12. Silver D. et al. Mastering the game of Go with deep neural networks and tree search // Nature. – 2016. – Vol. 529. – P. 484–489.
13. Brown T.B. et al. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems 33 (NeurIPS 2020). – 2020. – P. 1877–1901.
14. Jumper J. et al. Highly accurate protein structure prediction with AlphaFold // Nature. – 2021. – Vol. 596. – P. 583–589.
15. Goertzel B. Artificial General Intelligence: Concept, State of the Art, and Future Prospects // Journal of Artificial General Intelligence. – 2014. – Vol. 5, No. 1. – P. 1–48.
16. Bostrom N. Superintelligence: Paths, Dangers, Strategies. – Oxford: Oxford University Press, 2014. – 328 p.
17. Morris M.R. et al. Levels of AGI: Operationalizing Progress on the Path to AGI. – arXiv:2311.02462, 2023.
18. Hintze A. Understanding the four types of AI, from reactive robots to self-aware beings // The Conversation. – 2016. – 14 November.

19. Kaplan A., Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence // Business Horizons. – 2019. – Vol. 62, No. 1. – P. 15–25.
20. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge, MA: MIT Press, 2016. – 775 p.
21. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. 2nd ed. – Cambridge, MA: MIT Press, 2018. – 526 p.
22. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. – 2015. – Vol. 521. – P. 436–444.
23. Bommasani R. et al. On the Opportunities and Risks of Foundation Models. – arXiv:2108.07258, 2021.
24. NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0): NIST AI 100-1. – Gaithersburg: NIST, 2023.
25. Ji Z. et al. Survey of Hallucination in Natural Language Generation // ACM Computing Surveys. – 2023. – Vol. 55, No. 12. – Article 248.
26. Lewis P. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems 33 (NeurIPS 2020). – 2020. – P. 9459–9474.
27. Szegedy C. et al. Intriguing properties of neural networks // International Conference on Learning Representations (ICLR). – 2014.
28. Goodfellow I.J., Shlens J., Szegedy C. Explaining and Harnessing Adversarial Examples // International Conference on Learning Representations (ICLR). – 2015.
29. Marcus G. Deep Learning: A Critical Appraisal. – arXiv:1801.00631, 2018.
30. Moravec H. Mind Children: The Future of Robot and Human Intelligence. – Cambridge, MA: Harvard University Press, 1988. – 214 p.
31. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead // Nature Machine Intelligence. – 2019. – Vol. 1, No. 5. – P. 206–215.
32. Arrieta A.B. et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI // Information Fusion. – 2020. – Vol. 58. – P. 82–115.
33. Strubell E., Ganesh A., McCallum A. Energy and Policy Considerations for Deep Learning in NLP // Proceedings of the 57th Annual Meeting of the ACL. – 2019. – P. 3645–3650.
34. Patterson D. et al. Carbon Emissions and Large Neural Network Training. – arXiv:2104.10350, 2021.
35. Kaplan J. et al. Scaling Laws for Neural Language Models. – arXiv:2001.08361, 2020.
36. Mehrabi N., Morstatter F., Saxena N., Lerman K., Galstyan A. A Survey on Bias and Fairness in Machine Learning // ACM Computing Surveys. – 2021. – Vol. 54, No. 6. – Article 115.
37. Buolamwini J., Gebru T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification // Proceedings of Machine Learning Research. – 2018. – Vol. 81. – P. 77–91.
38. Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? // Proceedings of ACM FAccT 2021. – 2021. – P. 610–623.
39. UNESCO. Recommendation on the Ethics of Artificial Intelligence. – Paris: UNESCO, 2021.
40. Searle J.R. Minds, Brains, and Programs // Behavioral and Brain Sciences. – 1980. – Vol. 3, No. 3. – P. 417–424.

41. Dreyfus H.L. What Computers Can’t Do: The Limits of Artificial Intelligence. – New York: Harper & Row, 1972. – 259 p.
42. Bender E.M., Koller A. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data // Proceedings of the 58th Annual Meeting of the ACL. – 2020. – P. 5185–5198.
43. Chalmers D.J. Facing up to the problem of consciousness // Journal of Consciousness Studies. – 1995. – Vol. 2, No. 3. – P. 200–219.